Received April 15, 2015, accepted May 2, 2015, date of publication May 8, 2015, date of current version May 20, 2015.
Digital Object Identifier 10.1109/ACCESS.2015.2431493

Preprocessing and Feature Selection for Improved Sensor Interoperability in Online Biometric Signature Verification

RUBEN TOLOSANA, RUBEN VERA-RODRIGUEZ, JAVIER ORTEGA-GARCIA, (Senior Member, IEEE), AND JULIAN FIERREZ, (Member, IEEE)
ATVS-Biometric Recognition Group, Escuela Politecnica Superior, Universidad Autonoma de Madrid, Madrid 28049, Spain
Corresponding author: R. Tolosana (ruben.tolosana@uam.es)

This work was supported in part by the Project Bio-Shield under Grant TEC2012-34881, in part by Cecabank e-BioFirma Contract, and in part by Catedra UAM—Telefonica

ABSTRACT Due to the technological evolution and the increasing popularity of smartphones, people can access an application using authentication based on biometric approaches from many different devices. Device interoperability is a very challenging problem for biometrics, which needs to be further studied. In this paper, we focus on interoperability device compensation for online signature verification since this biometric trait is gaining a significant interest in banking and commercial sector in the last years. The proposed approach is based on two main stages. The first one is a preprocessing stage where data acquired from different devices are processed in order to normalize the signals in similar ranges. The second one is based on feature selection taking into account the device interoperability case, in order to select to select features which are robust in these conditions. This proposed approach has been successfully applied in a similar way to two common system approaches in online signature verification, i.e., a global features-based system and a time functions-based system. Experiments are carried out using Biosecure DS2 (Wacom device) and DS3 (Personal Digital Assistant mobile device) dynamic signature data sets which take into account multisession and two different scenarios emulating real operation conditions. The performance of the proposed global features-based and time functions-based systems applying the two main stages considered in this paper have provided an average relative improvement of performance of 60.3% and 26.5% Equal Error Rate (EER), respectively, for random forgeries cases, compared with baseline systems. Finally, a fusion of the proposed systems has achieved a further significant improvement for the device interoperability problem, especially for skilled forgeries. In this case, the proposed fusion system has achieved an average relative improvement of 27.7% EER compared with the best performance of time functions-based system. These results prove the robustness of the proposed approach and open the door for future works using devices as smartphones or tablets, commonly used nowadays.

INDEX TERMS Device interoperability, on-line signature, time functions-based system, global features-based system, fusion, DTW, Biosecure.

I. INTRODUCTION

Handwritten signatures are one of the most socially accepted biometric traits. They have been employed in financial and legal agreements scenarios for over a century [1], [2]. Nowadays, signatures can be easily captured by means of multiple electronic devices (e.g. pen tablets, Personal Digital Assistants (PDAs), grip pens, smartphones, etc) [3]. For this reason the popularity of this biometric trait has rapidly increased in the last years, especially in banking and commercial areas as can be seen in recent events.[1] [2] However, one of the main challenges in signature verification is related to signature variability. While genuine signatures can differ significantly (high intra-class variability), skilled forgeries could be similar to genuine signatures (low inter-class variability). Together with this high intra-class or intrinsic variability of signatures, there are sources of extrinsic variability such as device interoperability which can affect significantly the recognition performance. For example, due to the increasing deployment of smartphones in commercial applications to facilitate payments, people can access an application with different devices [4]. For all these reasons, the main goal of this work is to analyze and improve the system performance in an inter-operable case for dynamic signature verification.

[1] http://www.eab.org/events/program/66, May 2014
[2] http://www.eab.org/events/program/73, October 2014

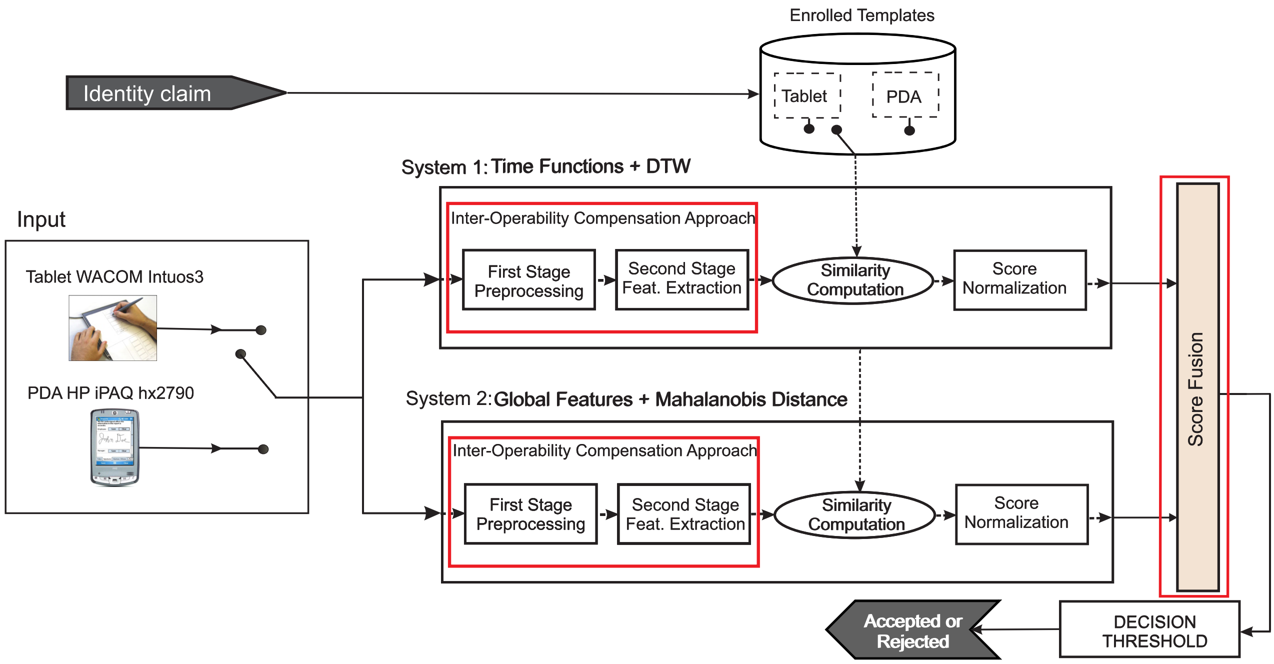


FIGURE 1. Architecture of the proposed strategy to compensate for device interoperability. The proposed approach is comprised on two stages (data preprocessing and feature selection), and is applied to two different recognition systems, one based on time functions and another based on global features. A final fusion of both proposed systems is carried out at the score level, which can further improve the system performance in general, but specially in a device interoperability case. The two proposed stages (consequently, stage 1 and stage 2) considered in this work and the score fusion from both systems are highlighted with red boxes.

Regarding on-line signature verification, there are two main approaches for feature extraction: global features-based systems (commonly known as global systems), which extract global information from the signature (e.g. signature duration, number of pen ups, etc.) in order to obtain a holistic feature vector [5]–[7]. On the other hand, time functions-based systems (commonly known as local systems) use the signature time functions (e.g. X and Y pen coordinates, pressure, etc.) for verification [8]. Traditionally, time functions-based systems have achieved better recognition performance than global features-based systems [5], [9], [10].

The most common algorithms employed in global features-based systems are based on statistical classifiers such as Gaussian Mixture Models [11] or Mahalanobis distance [12] whereas in time functions-based systems are DTW (Dynamic Time Warping) [13], HMM (Hidden Markov Models) [8], [14], NN (Neural Networks) [15] and SVM (Support Vector Machines) [16]. DTW has the advantage that it does not need a previous training of the user models.

To the best of our knowledge, there are very few works focused on the problem of device interoperability for dynamic signature recognition [17]–[19]. In [17], device interoperability is studied employing signatures coming from two different tablet PCs under an access control scenario. Signatures were downsampled to a constant sampling frequency of 100 Hz by using linear interpolation. A time functions-based system based on HMM with 14 discrete-time functions is considered, evaluating the performance of the system taking into account monosensor and multisensor enrolment and fusion of sensors. On the other hand, in a recent study [18], signatures coming from tablet PCs, smartphones and tablets are considered. The performance of the system for interoperability devices is evaluated using a time functions-based system based on DTW with 4 discrete-time functions for random (zero-effort) forgeries. To achieve a higher similarity between signatures coming from different devices, time- and spatial-based preprocessing normalizations are applied. However, apart from the normalization stage, the system used in that work was not specifically designed for compensating the device interoperability problem.

Very recently, in our previous work [19], we proposed a time functions-based system approach specifically designed to improve the device interoperability problem in dynamic signature recognition. This approach achieved an average relative performance improvement of 26.5% for random forgeries and 14.2% for skilled forgeries, comparing the results with the case of having a system specifically adjusted for each device. In this paper we extend this previous work by considering two different systems (global features-based and time functions-based systems) and a fusion of both. Besides, multi-session and both access control and mobile scenarios are considered, emulating realistic operating conditions since in a mobile scenario users had to sign while standing and holding the device in one hand.

Fig. 1 shows the architecture of the proposed strategy to compensate for device interoperability, which is based on two stages (data preprocessing and feature selection). This approach is applied to two common systems in online signature recognition (local and global systems). A final fusion of both proposed systems is carried out at the score level, which can further improve the system performance in general, but specially in a device interoperability case. The two stages followed in the proposed approach are a first data preprocessing step, which is applied in order to reach high similarity between signatures coming from different devices. After this data preprocessing step, a global feature and time function selection phase is proposed in order to further reduce the effect of device interoperability. For this purpose, Sequential Forward Feature Selection (SFFS) [20] algorithm has been used, which is one of the best performing methods reported [21]. In order to compare the similarity between signatures, a total of 100 global features and Mahalanobis distance algorithm are used for the global features based system whereas a total of 21 time functions and DTW algorithm are considered for the time functions based system. Finally, a fusion of both global features based and time functions based systems is performed via weighted sum of the matching scores [22] providing a further significant improvement for the device interoperability problem too. Experiments are carried out using Biosecure DS2 (Wacom pen tablet under access control scenario) and DS3 (HP PDA under mobile scenario) datasets with 120 users which are common to both DS2 and DS3 datasets. These datasets have been used in competitions such as Biosecure Signature Competition Campaign (BSEC 2009) [23] focused on the quality of signatures. Finally, a global system based on fusion of global features-based system with 28 global features and time functions-based system with 7 time functions is proposed for all the comparison cases. Therefore, it is important to highlight that the final proposed system achieves a good performance and works properly for the cases with and without device interoperability.

The rest of the paper is organized as follows. Section II describes the database used in the experimental work carried out. Section III describes the proposed approach in this work based on two main stages and the signature verification system proposed. Section IV reports the experimental work. Finally, Section V draws the final conclusions and future work.

II. SIGNATURE DATABASE

The database used to carry out the experimental work of this paper is Biosecure [24] with dynamic signature datasets DS2 and DS3. The main advantage of these datasets is that the DS2 dataset was captured under access control scenario where users had to sign while sitting, whereas the DS3 dataset considers a mobile scenario where users had to sign while standing and holding the device in one hand, emulating realistic operating conditions. Furthermore, it is important to highlight that intra-class variability problem is also considered, as Biosecure DS2 and DS3 datasets contain two different sessions separated by a 3 month time lap between them. DS3 dataset was captured using a PDA HP iPAQ hx2790 with a sampling frequency of 100 Hz, whereas the DS2 dataset was captured with a digitizing pen tablet WACOM Intuos3 A6 digitizer at 100 Hz and writing simultaneously on a paper sheet as can be seen in Fig. 2. A subset of 120 common users in DS2 and DS3 is considered in the experimental work reported here, as the goal is to study and compensate for the effect of the device interoperability.

(a) (b)

FIGURE 2. (a) Pen tablet capture process in the Biosecure DS2 - Access Control Scenario dataset. (b) PDA signature capture scenario process in the Biosecure DS3 - Mobile Scenario dataset.

The available information in Biosecure DS2 is the following: X and Y pen coordinates, pressure, pen angular orientation (azimuth and altitude angles) and timestamp information. However, in Biosecure DS3 just X and Y pen coordinates and timestamp are available.

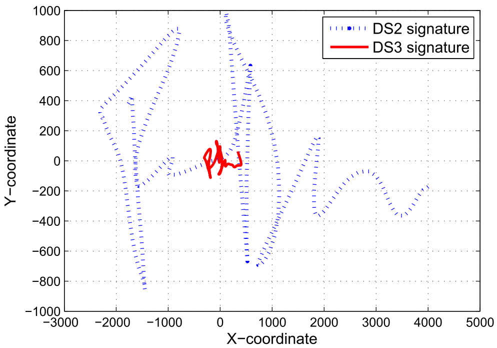
In both datasets (DS2 and DS3), signatures were captured in two separate sessions acquisition (i.e. multi-session) with a 3 month time lap between them. For each user, there are a total of 30 genuine signatures (i.e. 15 genuine signatures per session) and 20 skilled forgeries (i.e. 10 skilled forgeries per session) in each dataset. For the skilled forgeries case, users had visual access to the dynamics of the signing process of the signatures they had to forge.

III. DYNAMIC SIGNATURE VERIFICATION SYSTEM FOR INTEROPERABILITY DEVICE CASES

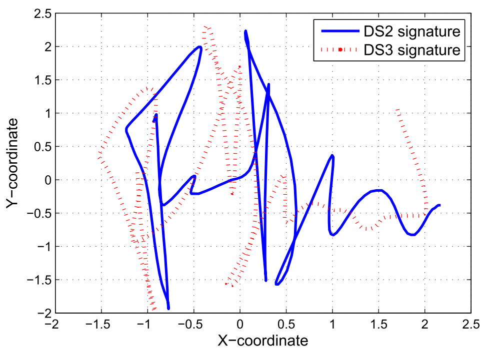
This section describes the system and the two main approaches proposed in this work to improve the problem of device interoperability. First, a data preprocessing step is applied (Sec. III-A) in order to achieve a high similarity between signatures coming from different devices. Second, a new criterion to extract and select features is considered in order to obtain an optimal feature vector taking into account the case of interoperability between devices (Sec. III-B). Finally, the proposed global features-based system (commonly known as global system), time functions-based system (commonly known as local system) and fusion of both systems are studied (Sec. III-C, III-D and III-E).

A. DATA PREPROCESSING STAGE

The first stage of the proposed system to compensate for device interoperability is related to data preprocessing.



(a)



(b)

FIGURE 3. **Signatures from DS2 and DS3 datasets. (a) Spatial resolution difference between DS2 and DS3 datasets. (b) Signatures from DS2 and DS3 datasets applying mean and standard deviation normalization.**

The aim of this first stage is to obtain signatures with the same type of information (i.e. X and Y coordinates, pressure, etc) and time and spatial position standard formats so as to improve the performance of the system in a device interoperability case. Several statistical data normalization techniques were studied in order to compensate for geometric differences between DS2 and DS3 datasets (see Fig. 3). For DS2, there is a different spatial position for the signatures due to the acquisition protocol followed in Biosecure where users had to sign in different boxes on a sheet of paper (see Fig. 2(b)) whereas the different size among signatures from DS2 and DS3 could be due to the screen resolution of the devices (see Fig. 3(a)).

In order to improve the performance of the system for the interoperability case, normalization based on the mean and standard deviation was applied to both systems since it achieved the best results. Fig. 3(b) represents signatures normalized from DS2 and DS3 datasets. An additional preprocessing step using interpolation based on splines [25] is necessary in DS3 dataset in order to correct sampling errors (missing samples).

Furthermore, only X and Y spatial coordinates are considered in this work. Pressure and pen angular orientation have been discarded in order to focus on the interoperability performance of the system due to these information is not provided by DS3 device.

TABLE 1. **Set of time functions considered in this work. Table adapted from [6] and [8].**

#	Feature
1	X-coordinate: x_n
2	Y-coordinate: y_n
3	Path-tangent angle: θ_n
4	Path velocity magnitude: v_n
5	Log curvature radius: ρ_n
6	Total acceleration magnitude: a_n
7-12	First-order derivate of features 1-6: $\dot{x_n}, \dot{y_n}, \dot{\theta_n}, \dot{v_n}, \dot{\rho_n}, \dot{a_n}$
13-14	Second-order derivate of features 1-2: $\ddot{x_n}, \ddot{y_n}$
15	Ratio of the minimum over the maximum speed over a 5-samples window: v_n^r
16-17	Angle of consecutive samples and first order difference: α_n, $\dot{\alpha_n}$
18	Sine: s_n
19	Cosine: c_n
20	Stroke length to width ratio over a 5-samples window: r_n^5
21	Stroke length to width ratio over a 7-samples window: r_n^7

It is also worth noting that information between pen-ups and pen-downs is not recorded by the PDA. Therefore, this information was discarded in DS2 (but not in IV-B1 in order to analyze the proposed data preprocessing stage) in order to achieve similar processing conditions in both devices.

B. FEATURE EXTRACTION AND SELECTION STAGE

The second stage of the proposed system is focused on obtaining a selection of global features and time functions, which are robust to data comparisons with and without device interoperability. In this work, a global features-based (commonly known as global) and time functions-based (commonly known as local) systems with a total of 100 and 21 global features and time functions respectively are considered. Both systems are based on previous works [26], [27]. Table 1 shows the 21 time functions respectively extracted for each signature. Information related to the 100 global features considered in this work has not been shown due to the lack of space. For more details, see Refs. [26], [27].

Due to the the low amount of available training data in a signature real case, Sequential Forward Feature Selection (SFFS) algorithm [20] is performed in order to obtain a subset of the 100 global features and 21 time functions which improves the performance of the system in terms of Equal Error Rate (EER) (%). This technique offers a suboptimal solution since it does not take into account all the possible
feature combinations, although it considers correlations between features. This is the main goal of this algorithm. The EER has been used as optimization criterion.

In the proposed approach, in order to achieve a high performance of the system for interoperability cases, the criterion of this algorithm has been modified, taking into account the EER of all the comparison cases (with and without device interoperability) at the same time with the goal of obtaining a global optimal feature/time function vector for each system (see Sec. IV-B4).

C. GLOBAL FEATURES BASED VERIFICATION SYSTEM

The Mahalanobis distance [28] is used to compare the similarity between a signature and a claimed user model. A user model is created from a training set of signatures. This model is defined as $C = (\mu,\Sigma)$, where μ is a feature vector with the mean of feature vectors extracted from each signature of this user and Σ is a diagonal covariance matrix. The matching score is obtained as the inverse of the Mahalanobis distance between the input signature feature vector x and the claimed user model C:

$$s(x, C) = \left((x-\boldsymbol{\mu})^T(\boldsymbol{\Sigma})^{-1}(x-\boldsymbol{\mu})\right)^{-1/2} \tag{1}$$

If the score $s(x, C)$ is above a specific threshold, the signature is considered genuine. Otherwise, it is rejected by the system.

D. TIME FUNCTIONS-BASED VERIFICATION SYSTEM

DTW algorithm [13] is used to compare the similarity between time functions from signatures. Scores are obtained as:

$$score = e^{-D/K} \tag{2}$$

where D and K represent respectively the minimal accumulated distance and the number of points aligned between two signatures using DTW algorithm.

E. FUSION OF GLOBAL FEATURES-BASED AND TIME FUNCTIONS-BASED SYSTEMS

A fusion of the global and local systems with the optimal global feature/time function vectors proposed in this work (see Sec. IV-B4) is performed via weighted sum of the match scores [29]. Before applying fusion of the systems, global and local scores are normalized in a range [0,1] using tanh-estimators [30]. The fusion score s_f is obtained as:

$$s_f = k \cdot s_g + (1-k) \cdot s_l \tag{3}$$

where s_f is the final score and, s_g and s_l are the match scores of the global and local systems respectively. The fusion weighting coefficient k is heuristically determined taking into account the system performance in terms of EER on the development signature set (see Sec. IV-B5).

IV. EXPERIMENTAL WORK

A. EXPERIMENTAL PROTOCOL

The first 5 genuine signatures of the first session are used as training signatures, whereas the 15 genuine signatures of the second session are left for testing in order to take into account intra-class variability problem. Therefore, the 10 remaining genuine signatures from the first session are not used in our experiments. Skilled forgeries scores are obtained by comparing training signatures against the 20 available skilled forgeries signatures for the same user whereas random (zero-effort) forgeries scores are obtained by comparing the training signatures with one genuine signature of the remaining users. For the global features-based verification system, scores are obtained by comparing signatures against the user model obtained with the first 5 training genuine signatures, while for the time functions-based system, the average score of the five one-to-one comparisons is performed.

The nomenclature used in this work is denoted as follows:

$$a - b - c$$

Where a indicates *skilled* or *random* forgeries cases and, b and c represent the device used for training and testing respectively (DS2 or DS3).

The first 50 users of the selected datasets are used for development and training of the system, while the remaining 70 users are employed for evaluating the system.

B. DEVELOPMENT EXPERIMENTAL RESULTS

Experiments are structured as follows: first (**Experiment 1**), we evaluate the first stage (data preprocessing) of the proposed approach for a standard case of having a recognition system adjusted specifically for each device, without taking into account interoperability conditions. Both of them are optimized for skilled forgeries case as it is the common practise in on-line signature verification (see [27]). In the next experiments (Experiments 2, 3 and 4) the first stage of the proposed approach is applied in order to just evaluate the second stage based on global features and time functions selection techniques. In **Experiment 2**, the system used in the first experiment applying the first data preprocessing stage is considered as the baseline system in order to know the improvement we can achieve applying the second proposed stage in the following experiments. In **Experiment 3**, we evaluate an ideal case where for each system (global features-based system and time functions-based system) and comparison case, a different optimal global feature and time function vectors are adjusted (i.e. eight vectors are developed per system, four for random cases and other four for skilled cases) achieving therefore the best possible performance (although this could be unrealistic). In **Experiment 4**, we propose the case of developing only one system (one for time functions-based system and another for global features-based system) which is adjusted for all possible comparisons with and without device interoperability where both data preprocessing and feature selection stages have been taken into account to improve the recognition performance for the cases of device interoperability.

Finally, a fusion of both systems with the optimal global feature and time-function vectors obtained in experiment 4 is performed via weighted sum of the match scores (**Experiment 5**) providing a further significant improvement for the device interoperability problem. In all these experiments only the development dataset of 50 users has been used.

1) EXPERIMENT 1 - DATA PREPROCESSING STAGE

In this experiment, the aim is to evaluate the need of the first proposed pre-proccesing stage (see Sec. III-A)

TABLE 2. Results for Experiment 1. System performance in terms of EER (%) on the development set of 50 users for global features-based system (top) and time functions-based system (bottom). Comparison of the results with and without applying the first pre-processed stage proposed in this work.

Global System	*Skilled forgeries*	
Training vs Testing	Without stage 1	With stage 1
DS2 - DS2	4.3	4.1
DS3 - DS3	25.3	14.8
DS2 - DS3	30.6	23.5
DS3 - DS2	49.3	46.9
Local System	***Skilled forgeries***	
Training vs Testing	Without stage 1	With stage 1
DS2 - DS2	7.1	8.6
DS3 - DS3	28.6	17.1
DS2 - DS3	45.5	27.3
DS3 - DS2	56.6	17.6

to compensate for the device interoperability problem. To carry out the experiment, systems for both global features and time functions are adjusted not considering device interoperability, which is the common procedure in on-line signature verification. SFFS algorithm has been implemented in order to improve the individually EERs for DS2 and DS3 datasets. In this case we consider two vectors per system (global features-based and time functions-based systems), one adjusted for DS2 dataset and another one fixed for DS3 dataset, and optimized for the skilled forgeries case which is the most challenging case to authenticate. Table 2 shows the performance for both systems. Analyzing the interoperability cases, the performance of the system applying the first proposed stage is significantly better compared to not applying this first preprocessing stage, especially for the time functions-based system. Therefore, the performance of the system for device interoperability cases not taking into account this first preprocessing stage is very poor achieving EER results of around 50% in most cases. This experiment proves the importance of considering this preprocessing stage for a device interoperability case (as was the only case considered in [18]). Analyzing the no interoperability cases, the performance of the system when it is trained for DS2 device (DS2 - DS2) is very similar in both global features-based and time functions-based systems for the case with and without applying the preprocessing stage even the performance is better in the time functions-based system not applying the preprocessing stage as the information between the pen up and the pen down is not removed. On the other hand, the performance of the system when it is trained for DS3 (DS3 - DS3) is much better applying the proposed preprocessing stage, as sampling errors are corrected using interpolation based on splines [25].

Once evaluated the importance of applying this first preprocessing stage in a device interoperability case, the following experiments (Exp. 2, 3 and 4) have been designed to show the improvement of performance of the system applying the second proposed stage. This second stage is based on feature/time function selection techniques in order to select features/time functions which are robust to data comparisons in cases with and without device interoperability.

2) EXPERIMENT 2 - BASELINE SYSTEM

In this experiment, the system previously mentioned applying the first preprocessing stage is considered as the baseline for both global features-based and time functions-based systems which are adjusted not considering interoperability of devices, the common procedure in on-line signature verification. As pointed out in Experiment 1, SFFS algorithm has been implemented in order to improve the individually EERs for DS2 and DS3 datasets and considering two vectors per system (global features-based and time functions-based systems), one adjusted for DS2 dataset and another one adjusted for DS3 dataset, and optimized for a skilled forgeries case which is the most challenging case. This baseline system allows us to know the improvement we can achieve applying the second proposed stage implemented in the following experiments. Table 3 shows the performance of these baseline systems for both global features-based (known as global) and time functions-based (known as local) systems and considering skilled and random forgeries applying the first stage of the proposed approach (see Sec. III-A).

Analyzing the no interoperability cases in Table 3, the performance of the system is much better for DS2 compared to DS3 datasets for both systems. This is due to the fact that DS2 device (pen tablet Wacom) is a higher quality device designed for capturing signatures and besides, in DS3 dataset signatures were captured under a mobility scenario where people had to sign standing and holding the PDA in one hand. Analyzing the interoperability cases, the performance of the system degrades very significantly in both systems, especially when it is trained for DS2 device (DS2 - DS3) where the performance of the system in some cases is even 6 times worse. So, in this experiment we can conclude that training and testing with different devices has a higher impact in the performance, and the critical case is when the quality of the device used for testing is worse than the quality of the device used for training. The performance of the system in an interoperability case has been evaluated in recent works for random forgeries cases [18], but not proposing any solution for compensating the interoperability between different quality devices apart from the preprocessing step. For this reason, the aim of the next experiments is to obtain an optimal feature vector which works satisfactory for all the cases at the same time.

3) EXPERIMENT 3 - INDIVIDUALLY OPTIMIZED SYSTEMS

In this experiment, the goal is to observe which is the best ideal possible performance of both systems in an individually optimized case. It is important to highlight that this would not be realistic in an application. SFFS algorithm has been individually applied to each system (global features-based and time functions-based systems) and for each comparison case (4 for random and 4 for skilled forgeries).

TABLE 3. System performance in terms of EER (%) on the development set of 50 users for global features-based system or global system (top) and time functions-based system or local system (bottom). Comparison of the results obtained in experiments 1, 2 and 3.

Global System	*Skilled forgeries*			*Random forgeries*		
Training vs Testing	Baseline *Exp. 2*	Individually optimized *Exp. 3*	Proposed *Exp. 4*	Baseline *Exp. 2*	Individually optimized *Exp. 3*	Proposed *Exp. 4*
DS2 - DS2	4.1	4.1	8.3	3.5	1.5	4.0
DS3 - DS3	14.8	14.8	20.5	10.8	5.5	8.6
DS2 - DS3	23.5	16.1	20.0	23.3	9.5	13.7
DS3 - DS2	46.9	13.9	21.9	44.4	7.2	13.2
Local System	*Skilled forgeries*			*Random forgeries*		
Training vs Testing	Baseline *Exp. 2*	Individually optimized *Exp. 3*	Proposed *Exp. 4*	Baseline *Exp. 2*	Individually optimized *Exp. 3*	Proposed *Exp. 4*
DS2 - DS2	8.6	8.6	9.3	1.2	0.6	0.9
DS3 - DS3	17.1	17.1	18.1	2.1	0.8	1.5
DS2 - DS3	27.3	21.5	22.9	4.7	2.5	4.3
DS3 - DS2	17.6	13.6	15.7	5.1	1.8	2.9

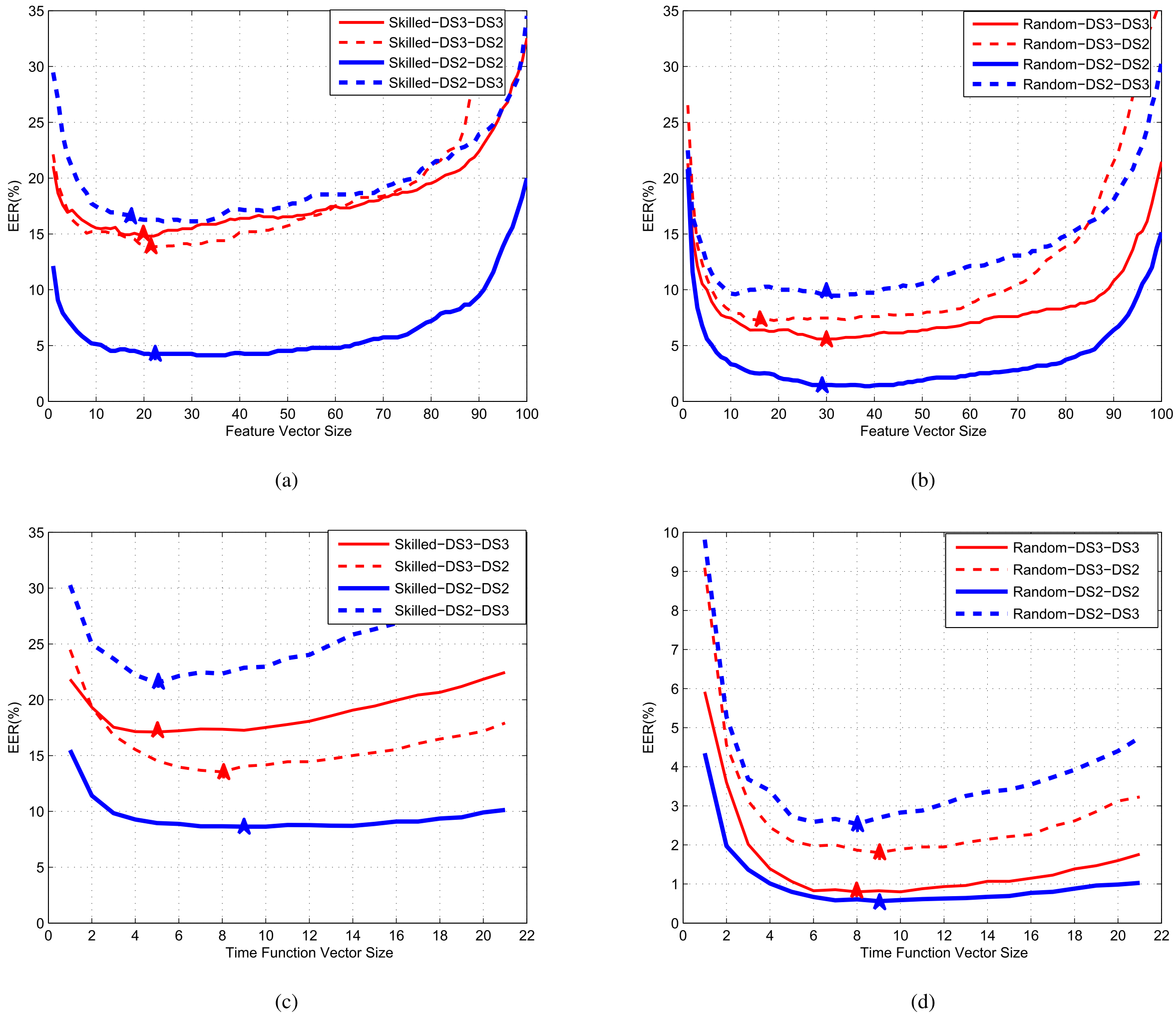
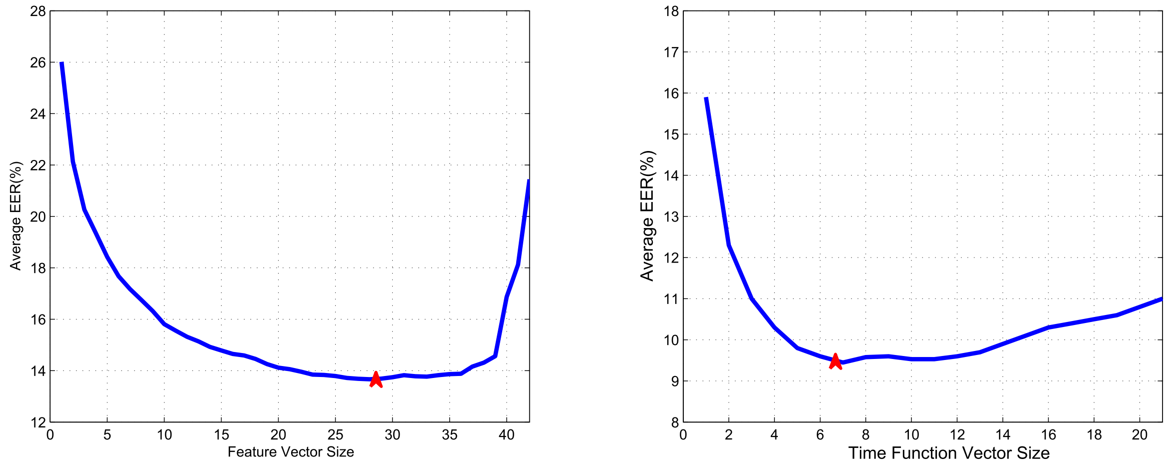


FIGURE 4. Experiment 3: Verification performance in terms of the size of the optimal feature/time function vector selected by the SFFS algorithm. Top: global features-based system or global system cases. Bottom: time functions-based system or local system cases. (a) Global features-based system/Skilled forgeries. (b) Global features-based system/Random forgeries. (c) Time functions-based system/Skilled forgeries. (d) Time functions-based system/Random forgeries.

Verification performance in terms of the EER for all the possible values of optimal feature vector size is depicted in Fig. 4. Table 3 represents the best EER for individually optimized cases applying the first stage of the proposed approach as we did in the previous experiment. Optimal feature/time function vectors are different for each case as



(a) (b)

FIGURE 5. Experiment 4: Average EER (%) of the system in terms of the size of the optimal feature/time function vector selected by the SFFS algorithm applying the new criterion to optimize for global features-based and time functions-based systems. (a) Global features-based system. (b) Time functions-based system.

can be seen in Fig. 4, where the number of features/time functions selected for every case is depicted with a marker.

The performance of individually optimized system is much better compared to the baseline system, specially for interoperability cases. This is due to the fact that the interoperability case has been taken into account by SFFS algorithm in this individually optimized systems. In addition, it considers 16 different optimal feature vectors (one for each case and system), so this would not be realistic in an application. Therefore, these results allow us to know the best ideal performance we would be able to achieve.

It is interesting to highlight in both systems (global features-based and time functions-based systems) the case when the systems are trained and tested with DS3 and DS2 devices (DS3 - DS2) respectively for the skilled forgeries case since the performance of the system is better compared to not having interoperability (DS3 - DS3). This shows again the lower quality of the DS3 device and the mobile scenario considered on DS3 dataset compared to DS2 dataset. Finally, it is important to note that in both systems the worst performance in all cases is obtained for the skilled forgeries DS2 - DS3 case, so this is the most challenging case to take into account for the next experiment.

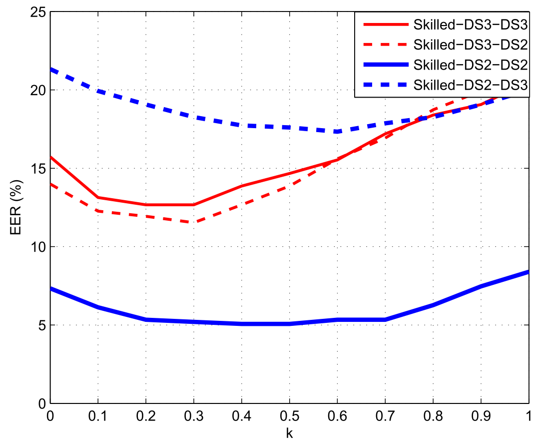
4) EXPERIMENT 4 - PROPOSED SYSTEM

In this experiment, the goal is to obtain an optimal feature/time function vector for each system (global features-based and time functions-based systems) which works satisfactory for all the comparisons cases at the same time (with and without interoperability). To achieve this, the two stages proposed in this work have been applied and therefore the criterion of SFFS algorithm has been modified in order to obtain the lowest total EER (average of the EERs obtained for all the comparison cases) and the lowest EER for the skilled forgeries DS2 - DS3 case since this is the worst case in both systems as it is seen in Sec. IV-B3. Fig. 5 shows the average performance of the system applying SFFS algorithm with the new criterion to evaluate for global features-based and time functions-based systems. A subset of 28 global features and 7 time functions were chosen for global features-based and time functions-based systems respectively, in which features related to the geometry and speed and acceleration are the most important for the global features-based system, whereas time functions related to Y-coordinate and velocity are the best performing for the time functions based system.

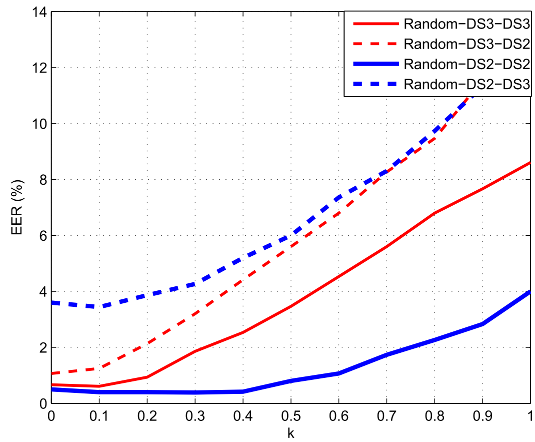
The performance of the system for every case using this proposed feature/time function vector is represented in Table 3. The performance results are slightly worse in general for the proposed approach compared to the individually optimized system, as could be expected. It is worth noting that time functions-based system achieves better results compared to the global features-based system. Therefore, the proposed approach to select features is more robust for a time functions-based system.

Analyzing the interoperability case for global features-based system, the proposed system provides an average relative improvement of 40.5% EER for skilled forgeries and 60.3% EER for random forgeries case compared to the baseline system. Besides, it is important to note that the most challenging case (skilled - DS2 - DS3) has improved in absolute numbers the EER in 3.5% compared to the baseline system.

On the other hand, analyzing the interoperability case for the time functions-based system, the proposed system provides an average relative improvement of 14.0% EER for skilled forgeries and 26.5% EER for random forgeries case compared to the baseline system. In addition, as it occurs for



(a)



(b)

FIGURE 6. Experiment 5: System verification performance on the development set for the fusion of the local and global systems at the score level for different values of the fusion weighting coefficient k. (a) Skilled forgeries case. (b) Random forgeries case.

TABLE 4. Experiment 5: System performance in terms of EER (%) on the development set of 50 users for global features-based system (known as global system), time functions-based system (known as local system) and fusion of both systems (proposed in the Experiment 4).

	Skilled forgeries			*Random forgeries*		
Training vs Testing	Global	Local	Fusion	Global	Local	Fusion
DS2 - DS2	8.3	9.3	5.2	4.0	0.9	0.4
DS3 - DS3	20.5	18.1	12.7	8.6	1.5	1.9
DS2 - DS3	20.0	22.9	18.3	13.7	4.3	4.3
DS3 - DS2	21.9	15.7	11.5	13.2	2.9	3.2

the global system, the performance of the most challenging case (skilled forgeries DS2 - DS3 case) improves in absolute numbers the EER in 4.4% compared to the baseline system.

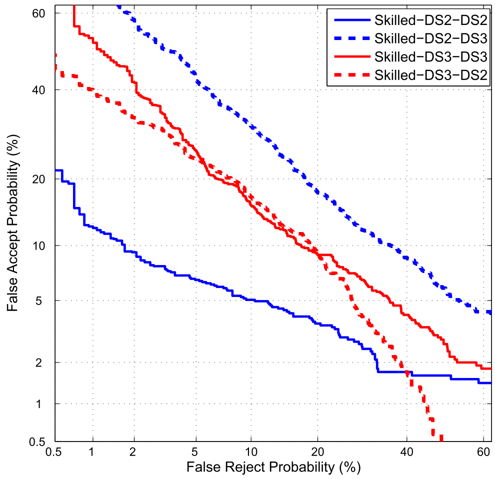
5) EXPERIMENT 5 - FUSION OF THE PROPOSED SYSTEMS

In this experiment, the goal is to achieve a significant improvement for the device interoperability problem applying a fusion of the global features-based and time functions-based systems developed using the proposed optimal feature/time function vectors obtained in Sec. IV-B4. The fusion of systems was performed following the steps of Sec. III-E. The fusion weighting coefficient k was heuristically set by observing the performance of the system in terms of the EER and taking into account all the cases at the same time. Fig. 6 depicts the performance of the fusion system for different values of k. As it can be seen, system performance gets worse for random and skilled forgeries when we choose a high value of k, whereas for a low value of k the performance of the system also gets worse for skilled forgeries cases. For this reason, 0.3 is chosen as the most appropriate value of k, since it gets a good performance for all cases at the same time. Therefore, the time functions-based system outweights the global features-based system in the final score. Table 4 shows the individually performance of global features-based and time functions-based systems (see Sec. IV-B4) and the fusion performance with a fusion weighting coefficient $k = 0.3$. As can be seen, the performance of the proposed fusion system is much better in most of the cases compared to the individually performance of the proposed global features-based and time functions-based systems, especially for skilled forgeries cases where the proposed fusion system provides an average relative improvement of 27.7% EER compared to the time functions-based system.

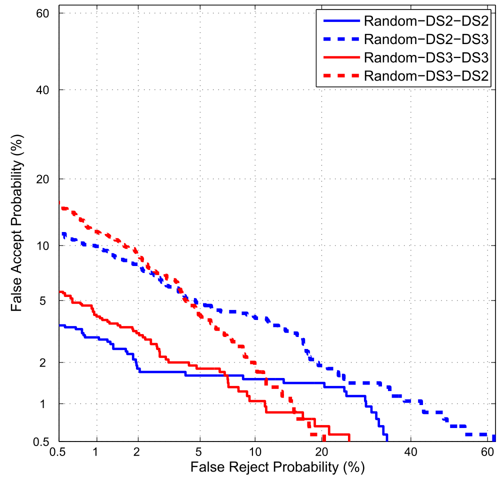
C. VALIDATION EXPERIMENTAL RESULTS

In order to validate the implemented system, we evaluate the verification performance system on the remaining 70 users of Biosecure datasets using the optimal fusion system obtained on the development phase. System performance is represented using DET plots as shown in Fig. 7. The EER for the baseline and proposed systems using a fusion weighting coefficient $k = 0.3$ are shown in Table 5 (in both systems $k = 0.3$ achieves the best performance). Analyzing the interoperability case, the proposed system provides an average relative improvement of 11.0% EER for skilled forgeries and 37.3% EER for random forgeries case compared to the baseline system. Therefore, these results are similar compared to the previous one obtained in the development phase, proving the robustness of the proposed scheme.

IEEE Access The journal for rapid open access publishing



(a)



(b)

FIGURE 7. Validation Results: DET curves for the final signature recognition system based on fusion of the proposed local and global systems on the evaluation set of Biosecure DS2 and DS3, and the device interoperability cases. (a) Skilled forgeries case. (b) Random forgeries case.

TABLE 5. Validation Results: System performance in terms of EER (%) on the evaluation set of 70 users for the fusion of the global and local systems using a weighted sum of scores. Comparison of the results obtained by baseline and proposed systems choosing a value of k equal 0.3 for the fusion.

Fusion of Systems	*Skilled forgeries*		*Random forgeries*	
Training vs Testing	Baseline	Proposed	Baseline	Proposed
DS2 - DS2	7.1	6.2	3.4	2.0
DS3 - DS3	11.3	12.8	2.6	2.7
DS2 - DS3	21.5	18.9	10.6	4.9
DS3 - DS2	14.8	13.4	4.7	4.7

V. CONCLUSION

In this paper, the main goal was to analyze and compensate the very challenging problem of device interoperability for dynamic signature verification due to the large number of electronic devices employed nowadays (e.g. pen tablets, PDAs, grip pens, smartphones, etc) and the increasing deployment of this trait in banking and commercial applications. Besides, it is important to highlight that there are very few works focused on this critical problem. Two scenarios are considered in this work: access control and mobile scenarios, emulating realistic operating conditions since in a mobile scenario users had to sign while standing and holding the device in one hand. Furthermore, it is important to highlight that intra-class variability problem is considered in the experiments too, using Biosecure DS2 and DS3 datasets which contain two different sessions separated by 3 months.

Two main stages are proposed in this work in order to compensate the interoperability problem. The first one is the data preprocessing stage where data acquired from different devices is pre-processed in order to reach a high similarity between signatures coming from different devices. The second stage is a selection of the best features in order to further reduce the effect of device interoperability, selecting features which are robust in these conditions. This proposed approach has been successfully applied to the two main system approaches in on-line signature verification.

Two different systems are considered: a global features-based and time functions-based systems based on previous works with 100 and 21 global features and time functions respectively. The performance of the proposed global features based and time functions based systems applying the two main stages considered in this work have provided an average relative improvement of 40.5% EER and 14.0% EER respectively for skilled forgeries compared to the baseline systems, whereas the relative improvement for random forgeries is 60.3% EER and 26.5% EER respectively compared to the baseline systems.

Finally, the fusion of the proposed global and local systems has been considered giving more weight to the local system. The proposed fusion system has achieved a further significant improvement for the device interoperability problem, especially for skilled forgeries where the proposed fusion system has achieved an average relative improvement of 27.7% EER compared to the best performance of time functions-based system. This proves the robustness of the system proposed specially in the cases of device interoperability which was the main objective of this work.

For future work, it will be interesting to see the performance of the system using devices with the same quality for interoperability cases and also, using newer devices such as tablets and smartphones [3]. Furthermore, it is interesting to study the performance of a dynamic signature verification system applied to security applications, taking into account the extreme case which X and Y coordinates are not used for the system, which would be a much more robust system against attacks [31].

REFERENCES

[1] R. Plamondon and G. Lorette, "Automatic signature verification and writer identification—The state of the art," *Pattern Recognit.*, vol. 22, no. 2, pp. 107–131, 1989.

[2] D. Impedovo and G. Pirlo, "Automatic signature verification: The state of the art," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 38, no. 5, pp. 609–635, Sep. 2008.

[3] R. Vera-Rodriguez, R. Tolosana, J. Ortega-Garcia, and J. Fierrez, "e-BioSign: Stylus- and finger-input multi-device database for dynamic signature recognition," in *Proc. 3rd Int. Workshop Biometrics Forensics (IWBF)*, Mar. 2015.

[4] W. Meng, D. S. Wong, S. Furnell, and J. Zhou, "Surveying the development of biometric user authentication on mobile phones," *IEEE Commun. Surveys Tuts.*, vol. PP, no. 99, p. 1, Dec. 2014.

[5] J. Fierrez-Aguilar, L. Nanni, J. Lopez-Peñalba, J. Ortega-Garcia, and D. Maltoni, "An on-line signature verification system based on fusion of local and global information," in *Proc. 5th IAPR Int. Conf. Audio-Video-Biometric Pers. Authentication (AVBPA)*, vol. 3546. Jul. 2005, pp. 523–532.

[6] M. Martinez-Diaz, "Dynamic signature verification for portable devices," M.S. thesis, Dep. Comput. Sci., Univ. Autónoma Madrid, Madrid, Spain, Nov. 2008.

[7] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia, "Feature-based dynamic signature verification under forensic scenarios," in *Proc. 3rd Int. Workshop Biometrics Forensics (IWBF)*, Mar. 2015.

[8] J. Fierrez, J. Ortega-Garcia, D. Ramos, and J. Gonzalez-Rodriguez, "HMM-based on-line signature verification: Feature extraction and signature modeling," *Pattern Recognit. Lett.*, vol. 28, no. 16, pp. 2325–2334, Dec. 2007.

[9] A. Kholmatov and B. Yanikoglu, "Identity authentication using improved online signature verification method," *Pattern Recognit. Lett.*, vol. 26, no. 15, pp. 2400–2408, 2005.

[10] M. Faundez-Zanuy, "On-line signature recognition based on VQ-DTW," *Pattern Recognit.*, vol. 40, no. 3, pp. 981–992, 2007.

[11] M. Martinez-Diaz, J. Fierrez, and J. Ortega-Garcia, "Universal background models for dynamic signature verification," in *Proc. 1st IEEE Conf. Biometrics, Theory, Appl., Syst. (BTAS)*, Sep. 2007, pp. 1–6.

[12] J. Galbally, J. Fierrez, and J. Ortega-Garcia, "Bayesian hill-climbing attack and its application to signature verification," in *Proc. IAPR Int. Conf. Biometrics (ICB)*, vol. 4642. Aug. 2007, pp. 386–395.

[13] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 26, no. 1, pp. 43–49, Feb. 1978.

[14] D. Muramatsu and T. Matsumoto, "An HMM online signature verifier incorporating signature trajectories," in *Proc. 7th ICDAR*, Aug. 2003, pp. 438–442.

[15] M. M. M. Fahmy, "Online handwritten signature verification system based on DWT features extraction and neural network classification," *Ain Shams Eng. J.*, vol. 1, no. 1, pp. 59–70, 2010.

[16] Y. Liu, Z. Yang, and L. Yang, "Online signature verification based on DCT and sparse representation," *IEEE Trans. Cybern.*, Dec. 2014, doi: 10.1109/TCYB.2014.2375959.

[17] F. Alonso-Fernandez, J. Fiérrez-Aguilar, and J. Ortega-Garcia, "Sensor interoperability and fusion in signature verification: A case study using tablet PC," in *Proc. IWBRS*, vol. 3781. Oct. 2005, pp. 180–187.

[18] R. Blanco-Gonzalo, R. Sanchez-Reillo, O. Miguel-Hurtado, and J. Liu-Jimenez, "Performance evaluation of handwritten signature recognition in mobile environments," *IET Biometrics*, vol. 3, no. 3, pp. 139–146, Sep. 2014.

[19] R. Tolosana, R. Vera-Rodriguez, J. Ortega-Garcia, and J. Fierrez, "Optimal feature selection and inter-operability compensation for on-line biometric signature authentication," in *Proc. IEEE/IAPR Int. Conf. Biometrics (ICB)*, May 2015.

[20] P. Pudil, J. Novovičová, and J. Kittler, "Floating search methods in feature selection," *Pattern Recognit. Lett.*, vol. 15, no. 11, pp. 1119–1125, 1994.

[21] A. Jain and D. Zongker, "Feature selection: Evaluation, application, and small sample performance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 2, pp. 153–158, Feb. 1997.

[22] A. A. Ross, K. Nandakumar, and A. K. Jain, *Handbook of Multibiometrics: Human Recognition Systems*. New York, NY, USA: Springer-Verlag, 2006.

[23] N. Houmani *et al.*, "BioSecure signature evaluation campaign (BSEC'2009): Evaluating online signature algorithms depending on the quality of signatures," *Pattern Recognit.*, vol. 45, no. 3, pp. 993–1003, 2012.

[24] J. Ortega-Garcia *et al.*, "The multiscenario multienvironment BioSecure multimodal database (BMDB)," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 6, pp. 1097–1111, Jun. 2010.

[25] M. Martinez-Diaz, J. Fierrez, M. R. Freire, and J. Ortega-Garcia, "On the effects of sampling rate and interpolation in HMM-based dynamic signature verification," in *Proc. 9th Int. Conf. Document Anal. Recognit. (ICDAR)*, vol. 2. Sep. 2007, pp. 1113–1117.

[26] J. Fierrez-Aguilar, "Adapted fusion schemes for multimodal biometric authentication," Ph.D. dissertation, Dept. Signs Radiocommun. Syst., Univ. Politecnica Madrid, Madrid, Spain, May 2006.

[27] M. Martinez-Diaz, J. Fierrez, R. P. Krish, and J. Galbally, "Mobile signature verification: Feature robustness and performance comparison," *IET Biometrics*, vol. 3, no. 4, pp. 267–277, Dec. 2014.

[28] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 4th ed. Boston, MA, USA: Academic, 2009, p. 4.

[29] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 3, pp. 226–239, Mar. 1998.

[30] A. Jain, K. Nandakumar, and A. Ross, "Score normalization in multimodal biometric systems," *Pattern Recognit.*, vol. 38, no. 12, pp. 2270–2285, 2005.

[31] R. Tolosana, R. Vera-Rodriguez, J. Ortega-Garcia, and J. Fierrez, "Increasing the robustness of biometric templates for dynamic signature biometric systems," in *Proc. 49th Annu. Int. Carnahan Conf. Secur. Technol.*, Sep. 2015.

RUBEN TOLOSANA received the M.Sc. degree in electrical engineering from the Universidad Autonoma de Madrid, in 2014, where he is currently pursuing the Ph.D. degree. In 2014, he joined the Biometric Recognition Group, Universidad Autonoma de Madrid, where he is also collaborating as an Assistant Researcher. He is involved in research on dynamic signature recognition, paying special interest to compensation of device inter-operability, aging problem, and fusion of systems.

RUBEN VERA-RODRIGUEZ received the M.Sc. degree in telecommunications engineering from the Universidad de Sevilla, Spain, in 2006, and the Ph.D. degree in electrical and electronic engineering from Swansea University, U.K., in 2010. The topic of his Ph.D. thesis was on multimodal biometrics with an emphasis on footsteps as a novel biometric. Since 2010, he has been with the Biometric Recognition Group, Universidad Autonoma de Madrid, Spain, first as the recipient of a Juan de la Cierva Post-Doctoral Fellowship from the Spanish Ministry of Innovation and Sciences, and is currently as an Assistant Professor. His research interests include signal and image processing, pattern recognition, and biometrics. He was a recipient of the best paper award at the 4th International Summer School on Biometrics in 2007.

JAVIER ORTEGA-GARCIA received the M.Sc. degree in electrical engineering (Ingeniero de Telecomunicacion) and the Ph.D. (*cum laude*) degree in electrical engineering (Doctor Ingeniero de Telecomunicacion) from the Universidad Politecnica de Madrid, Spain, in 1989 and 1996, respectively, and the master's degree in biometrics. He is currently the Founder and Co-Director of the Biometric Recognition Group. He is a Full Professor with Escuela Politecnica Superior, Universidad Autonoma de Madrid, where he teaches digital signal processing and speech processing. He has authored over 150 international contributions, including book chapters, refereed journal, and conference papers. His research interests are focused on biometrics signal processing: on-line signature verification, speaker recognition, Web-based biometrics, data fusion, and multibiometrics. He was the Chair of the Odyssey: The Speaker Recognition Workshop in 2004, the 3rd IAPR International Conference on Biometrics (ICB) in 2009, and the 6th IAPR ICB in 2013.

JULIAN FIERREZ received the M.Sc. and Ph.D. degrees in telecommunications engineering from the Universidad Politecnica de Madrid, Madrid, Spain, in 2001 and 2006, respectively. From 2007 to 2009, he was a Visiting Researcher with Michigan State University, East Lansing, MI, USA, under a Marie Curie Fellowship. Since 2002, he has been with the Biometric Recognition Group, Universidad Politecnica de Madrid, and the Universidad Autonoma de Madrid, Madrid, since 2004, where he is currently an Associate Professor. He has been and is actively involved in European projects focused on biometrics. His research interests and areas of expertise include signal and image processing, pattern recognition, and biometrics, with an emphasis on signature and fingerprint verification, multibiometrics, biometric databases, and system security. He was a recipient of a number of distinctions for his research, including the Best Ph.D. Award in computer vision and pattern recognition by the IAPR Spanish Liaison from 2005 to 2007, the Motorola Best Student Paper at ICB in 2006, the EBF European Biometric Industry Award in 2006, the IBM Best Student Paper at ICPR in 2008, and the EURASIP Best Ph.D. Award in 2012.