Optimal Feature Selection and Inter-Operability Compensation for On-Line Biometric Signature Authentication

Ruben Tolosana, Ruben Vera-Rodriguez, Javier Ortega-Garcia and Julian Fierrez
Biometric Recognition Group - ATVS, Escuela Politecnica Superior
Universidad Autonoma de Madrid
Avda. Francisco Tomas y Valiente, 11 - Campus de Cantoblanco - 28049 Madrid, Spain
{ruben.tolosana, ruben.vera, javier.ortega, julian.fierrez}@uam.es

Abstract

Due to the technological evolution and the increasing popularity of smartphones, people can access an application with many different devices. This device inter-operability is a very challenging problem for biometrics. In this paper we focus on inter-operability device compensation for on-line signature verification. The proposed approach is based on two main stages. The first one is a pre-processing stage where data acquired from different devices are processed in order to normalize the signals in similar ranges. The second one is based on a feature selection of time functions taking into account the inter-operability device comparisons in order to select features which are robust in these conditions. The experimental work has been carried out with Biosecure database using a Wacom tablet (DS2) and a PDA tablet (DS3), and the system developed is based on dynamic time warping (DTW) elastic measure over the selected time functions. The performance of the proposed system is very similar compared to an ideal system. Also, the proposed approach provides average relative improvements for the cases of inter-operability comparisons of 26.5% for random forgeries and, around 14.2% for the case of skilled forgeries comparing the results with the case of having a system specifically tuned for each device, proving the robustness of the proposed approach. These results open the door for future works using devices as smartphones or tablets, commonly used nowadays.

1. 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 [14]. Nowadays, signatures can be easily captured by means of multiple electronic devices (e.g. Pen tablets, PDAs, Grip Pens, Smartphones). For this reason the popularity of this biometric trait has rapidly increased in the last years. However, one of the main challenges in signature verification is related to the signature variability. While signatures from a genuine user differ significantly (high intra-class variability), skilled forgeries could be similar to genuine signatures (low inter-class variability).

Together with this intrinsic variability of signatures, there are sources of extrinsic variability such as the device inter-operability which can affect significantly the recognition performance. For example, due to the increasing deployment of smartphones in the commercial sector to facilitate payments, people can access an application with different devices [17]. For all these reasons, the main goal of this work is to study the performance of the system in an inter-operable case for dynamic signature verification since there are very few works regarding this subject [1, 2]. In addition, it is important to note that the systems used in these related works do not take into account the inter-operability problem in the development phase.

Regarding on-line signature verification, there are two main approaches for feature extraction: feature-based 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 [6, 16]. On the other hand, function-based systems use the signature time functions (e.g. X and Y pen coordinates, pressure, etc.) for verification [5]. Traditionally, function-based systems have achieved better recognition performance than feature-based systems [8, 4, 6].

The most common algorithms employed in function-based systems are DTW (Dynamic Time Warping) [15], HMM (Hidden Markov Model) [11, 5] and NN (Neural Network) [3]. DTW has the advantage that it does not need a previous training.

The main contribution of the present work is to propose an optimal function-based feature vector which can deal with device inter-operability in terms of recognition perfor-

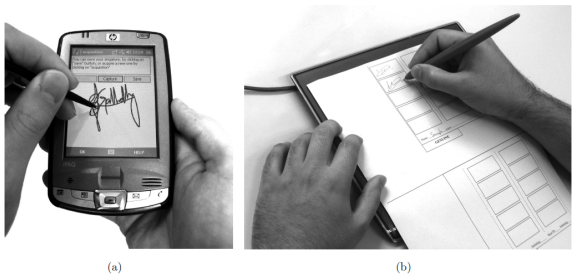


Figure 1. (a) PDA signature capture scenario process in the Biosecure DS3 - Mobile Scenario dataset. (b) Pen tablet capture process in the Biosecure DS2 - Access Control Scenario dataset.

mance. To achieve this, two main stages have been followed in this work: first, a data pre-processing step has been applied in order to reach a high similarity between signatures coming from different devices. After this data pre-processing step, a feature selection phase taking into account inter-operability between devices has been employed, using the Sequential Forward Feature Selection (SFFS) [13] which is one of the best performing methods reported [7]. A function-based system with 21 functions and DTW algorithm are used to compare the similarity between signatures. Experiments are carried out using Biosecure DS2 (pen tablet Wacom) and DS3 (PDA HP) datasets with 120 common users. Finally, an only one function-based system with 7 functions has been considered for all the cases, achieving a good performance for the cases with and without inter-operability.

The remainder of the paper is organized as follows. Section 2 describes the database used in the experimental work carried out. Section 3 describes the function-based signature verification system proposed. Section 4 reports the experimental work. Finally, Section 5 draws the final conclusions and future work.

2. Signature Database

The database used to carry out the experimental work of this paper is Biosecure [12] with datasets DS2 and DS3. DS3 dataset was captured using a PDA HP iPAQ hx2790 with a sampling frequency of 100 Hz, whereas DS2 dataset was captured with a digitizing pen tablet WACOM Intuos3 A6 digitizer at 100 Hz and writing on a paper as can be seen in Fig. 1. A subset of 120 common users in DS2 and DS3 is considered in the experimental work reported due to the goal is to study the effect of the inter-operability of devices.

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. For this reason, in order to make comparable information between DS2 and DS3 datasets, pressure and pen angular orientation have been discarded to focus on the inter-operability performance of the system.

In both datasets (DS2 and DS3), signatures were captured in two separate sessions with a 2 months time gap between them. For each user, there are a total of 30 genuine signatures and 20 skilled forgeries in each dataset. The users had visual access to the dynamics of the signing process of the signatures they had to forge as many times as they wanted.

3. Dynamic Signature Verification System for Inter-operability Device Compensation

This section describes the function-based system and the two main approaches proposed in this work to improve the problem of device inter-operability. First, a data pre-processing step is applied (Sec. 3.1) 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 inter-operability between devices (Sec. 3.2). Finally, the proposed function-based system is studied (Sec. 3.3).

3.1. Data Pre-processing

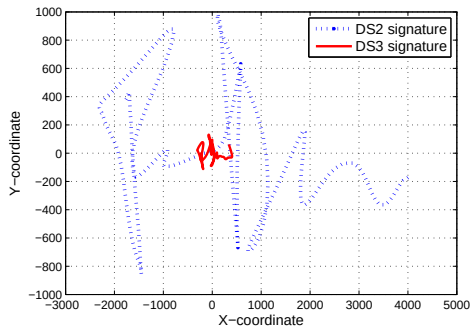
The first stage of the proposed system to compensate device inter-operability is concerned with data pre-processing. Several statistical data normalization techniques have been studied in order to compensate for geometric differences between DS2 and DS3 datasets (see Fig. 2). The different spatial position between signatures is due to the acquisition protocol followed in Biosecure, where in DS2 dataset users had to sign in different boxes on a paper (see Fig. 1(b)) whereas the different size among signatures from DS2 and DS3 could be due to the screen resolution of the devices (see Fig. 2(a)).

In order to improve the performance of the system for the inter-operability case, the mean and standard deviation normalization was applied since it achieved best results. Other normalization techniques were also studied such as max-min or mean normalizations. Fig. 2(b) represents signatures normalised from DS2 and DS3 datasets. An additional pre-processing step using interpolation based on splines [9] is necessary in DS3 dataset in order to correct sampling errors (missing samples).

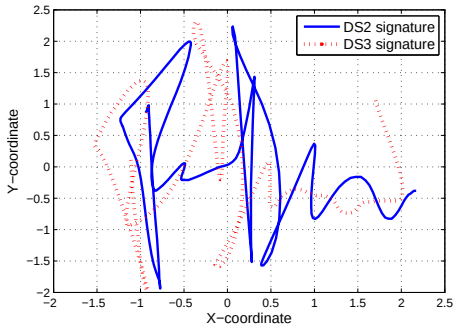
It is also worth noting that information between pen-ups and pen-downs is not recorded by the PDA. Therefore, this information was discarded in DS2 in order to achieve a similar proccesing conditions in both devices.

3.2. Feature Extraction and Selection

A function-based or local feature signature verification system based on previous works [5, 10] is considered in this



(a) Spatial resolution problem between DS2 and DS3 datasets



(b) Signatures from DS2 and DS3 datasets applying mean and standard deviation normalization

Figure 2. *Signatures from DS2 and DS3 datasets.*

work. Signals captured by the digitizer are used to extract a set of 21 time-functions (see Table 1) for each signature.

The second stage of the proposed system is concerned with feature selection. Due to the the low amount of available training data in a signature real case, Sequential Forward Feature Selection (SFFS) algorithm [13] is performed in order to obtain a subset of the 21 local features which improves the performance of the system in terms of 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 chosen as the optimization criterion.

In the proposed approach, in order to achieve a high performance of the system for inter-operability cases, the criterion of this algorithm has been modified taking into account the EER of all possible comparisons for DS2 and DS3 with and without inter-operability (8 cases) at the same time with the goal to obtain an only-one optimal feature vector for all cases (see Sec. 4.2.3).

3.3. Local Signature Verification System

DTW algorithm [15] is used to compare the similarity between time-functions from signatures. Scores are obtained as:

#	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: $\alpha_n, \dot{\alpha_n}$
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

Table 1. *Set of local features considered in this work.*

$$score = e^{-D/K} \quad (1)$$

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

4. Experimental Work

4.1. Experimental Protocol

The first 5 genuine signatures of the first session are used as training signatures, whereas the remaining 15 genuine signatures of the second session are left for testing. Skilled forgery scores are obtained by comparing training signatures against the 20 available skilled forgeries signatures for the same user whereas random or zero-effort forgery scores are obtained by comparing the training signatures to one genuine signature of the remaining users.

The nomenclature used in this work for inter-operability cases (when training and testing signatures come from different capture devices) is denoted as follows:

$$a — b — c$$

Where *a* indicates *skilled* or *random* forgeries case and, *b* and *c* represent the device used for training and testing respectively.

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.

4.2. Development Experimental Results

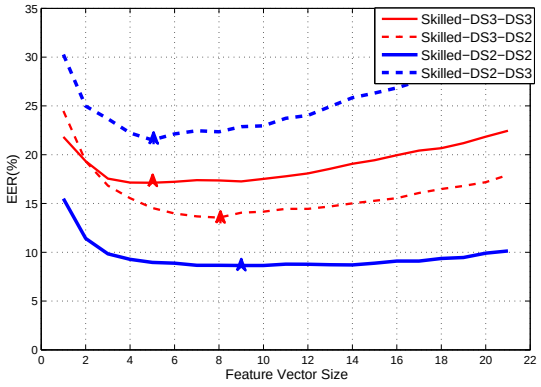
The experiments are structured as follows: first, we evaluate the standard case of having a recognition system tuned specifically for each device, without taking into account inter-operability conditions. Both of them are optimized for skilled forgeries case. In the second experiment we evaluate an ideal case where for each comparison case, a different system is tuned (i.e. eight systems are developed, four for random cases and other four for skilled cases) achieving therefore the best possible performance (although unrealistic). Finally, the system proposed in this work is studied (Experiment 3), where only one system is tuned for all eight possible comparisons where data pre-processing and feature selection have been taken into account to improve the recognition performance for the cases of device inter-operability. All these experiments have employed the development set of 50 users.

It is worth noting that the pre-processing stage of the approach proposed in this work has also been applied in the first two experiments, as the recognition performance was really bad for the inter-operability cases otherwise.

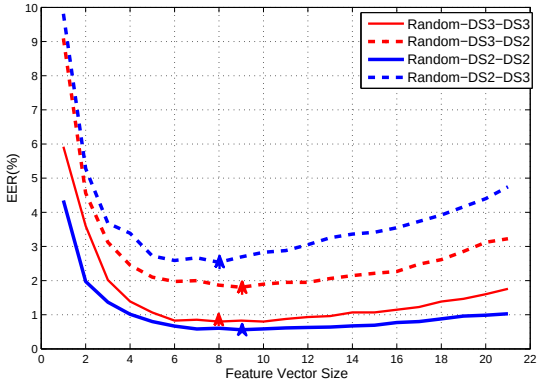
4.2.1 Experiment 1 - Baseline System

In this experiment, the goal is to obtain the performance of a baseline system for inter-operability cases. SFFS algorithm has been implemented in order to improve the individually EER for DS2 and DS3 datasets. In this case we consider two systems, one tuned for DS2 dataset and another one tuned for DS3 datset, and optimized for a skilled forgery case which is the most challenging case. Table 2 represents the performance of this baseline system applying normalization techniques (see Sec. 3.1), since the performance was so bad otherwise.

Analyzing the no inter-operability cases, the performance of the system is better for DS2 compared to DS3 datasets. This is due to the fact that DS2 device (Pen tablet Wacom) is a higher quality device designed for capturing signatures. Analyzing the inter-operability cases, the performance of the system degrades very significantly, especially when it is trained for DS2 device (DS2 - DS3). So, in this experiment we can conclude that training and testing with different devices has a big 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 inter-operability case has been studied in recent works for random forgeries cases [2], but not proposing any system which compensates the inter-operability between different quality devices. 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.



(a) Skilled forgeries cases



(b) Random forgeries cases

Figure 3. *Verification performance in terms of the size of the optimal feature selected by the SFFS algorithm.*

4.2.2 Experiment 2 - Individual Optimized Systems

In this experiment, the goal is to see the best possible performance of the system in an individual optimized case. SFFS algorithm has been individually applied for each comparison case (4 for random and 4 for skilled forgeries). The verification performance in terms of the EER for all the possible values of the optimal feature vector dimensionality is depicted in Fig. 3. Table 2 represents the best EER for individual optimized cases applying the first stage of the proposed approach as we did in the previous experiment. Optimal feature vectors are different for each case as can be seen in Fig. 3, where the number of features selected for every case is depicted with a marker.

The performance of individual optimized system is much better than the baseline system, specially for inter-operability cases. This is due to the fact that the inter-operability case has been taken into account by SFFS algorithm in this individual optimized systems. In addition, it considers 8 different optimal feature vectors (one for each case), so this would not be realistic in a real application.

	Skilled forgeries			*Random forgeries*		
Training vs Testing	Baseline	Individual optimized	Proposed	Baseline	Individual optimized	Proposed
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

Table 2. *System performance in terms of EER (%) on the development set of 50 users using function-based system. Comparison of the results obtained in experiments 1, 2 and 3.*

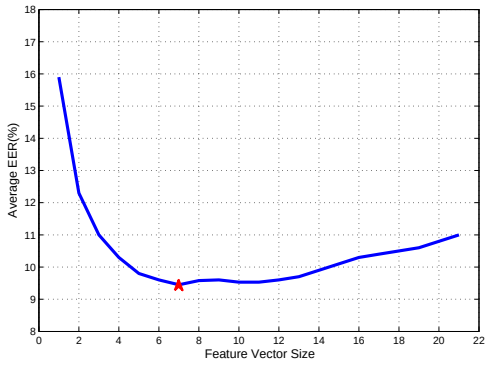


Figure 4. *Average EER (%) of the system in terms of the size of the optimal feature selected by the SFFS algorithm applying the new criterion to optimize.*

Therefore, these results help us to know the best ideal performance we are able to achieve.

It is important to highlight the case when system is trained and tested with DS3 and DS2 devices respectively for skilled forgeries case since the performance of the system is better compared to not having inter-operability (DS3 - DS3). This shows again the low quality of DS3 device compared to DS2 device. Finally, it is important to note that the worst performance of the system is obtained for skilled-DS2-DS3 case, so this is the most challenging case to take into account for the next experiment.

4.2.3 Experiment 3 - Proposed System

In this experiment, the goal is to obtain an optimal feature vector which works satisfactory for all the cases at the same time. 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 (sum of EER of each case) and the lowest EER for skilled-DS2-DS3 case since it is the worst case as we have seen in Sec. 4.2.2. Fig. 4 shows the performance of the system applying SFFS algorithm with the new criterion to evaluate. A subset of 7 features was obtained, in which features relating to the Y-coordinate and velocity are the most important.

The performance of the system for every case using this proposed feature vector is represented in Table 2. These results are just a bit worse compared to the individual optimized system performance. Analyzing the inter-operability case, the proposed system provides an average relative improvement of 14.0% for skilled forgeries and 26.5% 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 4.4% compared to baseline system.

4.3. Validation Experimental Results

To validate the implemented system, we compute the verification performance system on the remaining 70 users of Biosecure datasets using the optimal feature vector obtained on the development phase. The system performance is represented using DET plots as shown in Fig. 5. The EER for individual optimized and proposed systems are shown in Table 3.

As can be seen, the proposed system achieves similar performances compared to the individual optimized system in all cases. It is interesting to note that it even achieves better performance for DS3 - DS2 cases compared to the ideal system. Therefore, this proves the robustness of the proposed feature vector obtained in the development phase.

5. Conclusions and Future work

In this paper, a function-based or local feature system has been proposed for signature verification, specially designed to deal with device inter-operability conditions. Two main stages have been considered in this work. The first one is the pre-processing stage where data acquired from different devices are pre-processed in order to normalize the signals in similar ranges. The second stage is a selection of time functions taking into account the device inter-operability comparisons, in order to select features which are robust in these conditions. This optimal feature vector contains 7 time-functions selected by SFFS algorithm in the development phase applying the new criterion. As can be seen in Sec. 4.3, the performance of the proposed system achieves similar performance compared to an ideal system for all cases. This proves the robustness of the system proposed in this work specially in the cases of device inter-operability which was the main objective of this work. For future work, it would be interesting to see the perfor-

	Skilled forgeries		*Random forgeries*	
Training vs Testing	Individual optimized	Proposed	Individual optimized	Proposed
DS2 - DS2	8.5	9.5	1.3	1.8
DS3 - DS3	14.8	16.2	0.9	1.4
DS2 - DS3	22.0	22.8	2.6	3.1
DS3 - DS2	17.1	16.9	2.4	1.7

Table 3. *System performance in terms of EER (%) on the evaluation set of 70 users using function-based system. Comparison of the results obtained by individual optimized and proposed systems.*

mance of the system using devices with the same quality for inter-operability cases, and also using newer devices such as tablets and smartphones.

6. Acknowledgment

This work was supported in part by the Project BioShield (TEC2012-34881), in part by Cecabank e-BioFirma Contract, in part by the BEAT Project (FP7-SEC-284989) and in part by Catedra UAM-Telefonica.

References

[1] F. Alonso-Fernandez, J. Fierrez-Aguilar, and J. Ortega-Garcia. Sensor interoperability and fusion in signature verification: A case study using tablet pc. In *Proc. IWBRS*, volume 3781 of *LNCS*, pages 180–187. Springer, October 2005.

[2] R. Blanco-Gonzalo, R. Sanchez-Reillo, O. Miguel-Hurtado, and J. Liu-Jimenez. Performance evaluation of handwritten signature recognition in mobile environments. *IET Biometrics*, 2014.

[3] M. M. Fahmy. Online handwritten signature verification system based on dwt features extraction and neural network classification. *in Shams Engineering Journal*, 1(1):59 – 70, 2010.

[4] M. Faundez-Zanuy. On-line signature recognition based on VQ-DTW. *Pattern Recognition*, 40(3):981 – 992, 2007.

[5] J. Fierrez, J. Ortega-Garcia, D. Ramos, and J. Gonzalez-Rodriguez. Hmm-based on-line signature verification: feature extraction and signature modeling. *Pattern Recognition Letters*, 2007.

[6] J. Fierrez-Aguilar, L. Nanni, J. Lopez-Pealba, J. Ortega-Garcia, and D. Maltoni. An on-line signature verification system based on fusion of local and global information. In *Proc. AVBPA*. Springer, 2005.

[7] A. K. Jain and D. E. Zongker. Feature selection: Evaluation, application, and small sample performance. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(2):153–158, 1997.

[8] A. Kholmatov and B. Yanikoglu. Identity authentication using improved online signature verification method. *Pattern Recognition Letters*, 26(15):2400 – 2408, 2005.

[9] 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. ICDAR*, 2007.

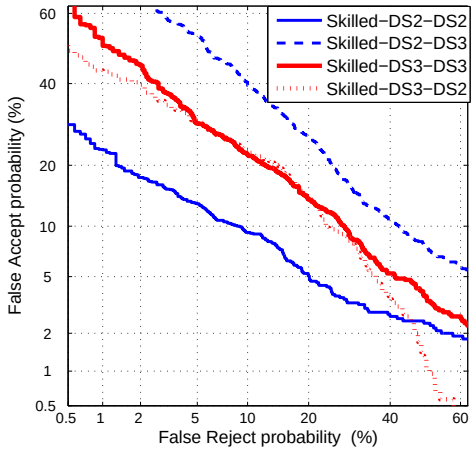
[10] M. Martinez-Diaz, J. Fierrez, R. P. Krish, and J. Galbally. Mobile signature verification: Feature robustness and performance comparison. *IET Biometrics*, 2014.

[11] D. Muramatsu and T. Matsumoto. An hmm on-line signature verifier incorporating signature trajectories. In *Proc. ICDAR*, pages 438–442. IEEE Computer Society, 2003.

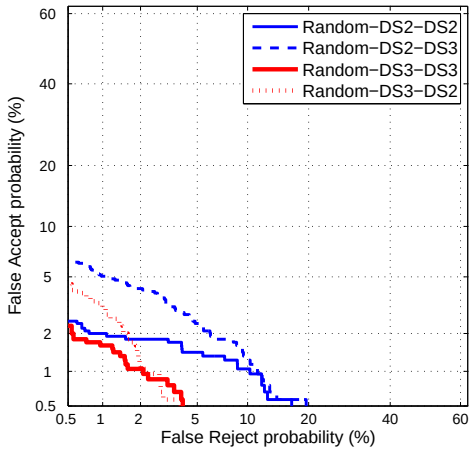
[12] J. Ortega-Garcia, J. Fierrez, et al. The multi-scenario multi-environment biosecure multimodal database (bmdb). *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2010.

[13] J. N. P. Pudil and J. Kittler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15(10):1119–1125, 1994.

[14] R. Plamondon and G. Lorette. Automatic signature verification and writer identification - the state of the art. *P.R.*, 1989.



(a) Skilled forgeries case



(b) Random forgeries case

Figure 5. *DET curves for the proposed function-based signature recognition system on the evaluation set of Biosecure DS2 and DS3, and the device inter-operability cases.*

[15] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE TASSP*, 1978.

[16] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia. Feature-based dynamic signature verification under forensic scenarios. In *Proc. IWBF*, 2015.

[17] 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. IWBF*, 2015.