Increasing the Robustness of Biometric Templates for Dynamic Signature Biometric Systems

Ruben Tolosana, Ruben Vera-Rodriguez, Javier Ortega-Garcia and Julian Fierrez
Biometric Recognition Group - ATVS
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 high deployment of devices such as smartphones and tablets and their increasing popularity in our society, the use of biometric traits in commercial and banking applications through these novel devices as an easy, quick and reliable way to perform payments is rapidly increasing. The handwritten signature is one of the most socially accepted biometric traits in these sectors due to the fact that it has been used in financial and legal transitions for centuries. In this paper we focus on dynamic signature verification systems. Nowadays, most of the state-of-the-art systems are based on extracting information contained in the *X* and *Y* spatial position coordinates of the signing process, which is stored in the biometric templates. However, it is critical to protect this sensible information of the users signatures against possible external attacks that would allow criminals to perform direct attacks to a biometric system or carry out high quality forgeries of the users signatures. Following this problem, the goal of this work is to study the performance of the system in two cases: first, an optimal time functions-based system taking into account the information related to *X* and *Y* coordinates and pressure, which is the common practice (i.e. Standard System). Second, we study an extreme case not considering information related to *X*, *Y* coordinates and their derivatives on the biometric system (i.e. Secure System), which would be a much more robust system against attacks, as this critical information would not be stored anywhere. The experimental work is carried out using e-BioSign database which makes use of 5 devices in total. The systems considered in this work are based on Dynamic Time Warping (DTW), an elastic measure over the selected time functions. Sequential Forward Features Selection (SFFS) is applied as a reliable way to obtain an optimal time functions vector over a development subset of users of the database. The results obtained over the evaluation subset of users of the database show a similar performance for both Standard and Secure Systems. Therefore, the use of a Secure System can be useful in some applications such as banking in order to avoid the lost of important user information against possible external attacks.**

***Keywords*—Biometrics, dynamic signature, banking security, feature selection, time functions-based system, e-BioSign**

I. Introduction

User authentication in different services and systems is a critical need in many scenarios nowadays. Biometric recognition systems have many advantages compared to traditional schemes, which are based on what the user knows (passwords, keys, etc.) or what the user has (card, token). In this sense, biometric traits cannot be lost, it is not necessary to memorize them as they are part of ourselves. 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]. 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, it is important to take into account the signature variability problem. 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 interoperability scenario which affect significantly the performance of the system as it can be seen in recent works [2]. Therefore, it is important to take into account this new device interoperability scenario due to the high deployment of devices such as smartphones and tablets and their increasing popularity in our society.

Regarding on-line signature verification, there are two main approaches for feature and time functions 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 [3], [4]. On the other hand, time functions-based systems use the signature time-functions (e.g. *X* and *Y* pen coordinates, pressure, etc.) for verification [5]. Traditionally, time functions-based systems have achieved better recognition performance than feature-based systems [6], [7], [3].

In this work we focus on a time functions-based system. The most common algorithms employed in time functions-based systems are DTW (Dynamic Time Warping) [8], HMM (Hidden Markov Models) [9] [5], NN (Neural Networks) [10] and SVM (Support Vector Machines) [11]. DTW has the advantage that it does not need a previous training of the user models.

The main contribution of the present work is to study the performance of a dynamic signature verification system applied for security applications in the banking and commercial sectors. Once the dynamic information of a signature (i.e. *X* and *Y* coordinates, pressure, etc.) has been acquired by means of devices such as pen tablets or smartphones, all this information is stored in a biometric template which is later used as a representation of the subject identity and deployed for authentication purposes. The problem is this biometric template can contain sufficient information to allow the reconstruction of the original biometric trait, which can be used to perform direct attacks to the biometric system [12], [13], [14].

In order to solve this problem, many efforts have been done trying to protect the biometric template information [15], [16]. However, the performance of the system using this biometric template protection gets worse in some biometric traits. For this reason, the main goal of this work is to study the case of dynamic signature verification system; in this sense we propose a system which does not store any sensible information in the biometric template, thus not allowing criminals to reconstruct the original signature (i.e. X, Y coordinates and their derivatives). Therefore, two different systems have been considered in this work: *i) Standard System* which considers information related to X, Y coordinates and their derivatives in order to know what is the best performance of the system we can achieve; and *ii)* a *Secure System* which does not take into account information related to X, Y coordinates and their derivatives in order to avoid possible reconstruction of the original signature.

A time functions-based system with 23 time functions is considered in the Standard System whereas 17 time functions are considered in the Secure System. DTW algorithm is used to compare the similarity between signatures. Experiments are carried out using e-BioSign database with a total of 70 users. In addition, Sequential Forward Feature Selection (SFFS) has been used as a reliable way to obtain an optimal time-functions subset on a development stage of the system.

The remainder of the paper is organized as follows. Section II describes the database used in the experimental work carried out. Section III describes the time functions-based 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 e-BioSign [17], a new database with information related to dynamic signature and handwriting. e-BioSign is comprised of 5 devices in total, three Wacom devices (DTU-500, DTU-530 and STU 1031) specifically designed to capture dynamic signatures and handwriting, and two Samsung general purpose tablets (Samsung Galaxy Note 10.1 and Samsung ATIV7). For these two Samsung tablets, data is collected using a pen stylus but also the finger in order to take into account a usual mobile scenario where users do not have a pen stylus to sign. In this work, dynamic signatures from all five devices using pen stylus are considered in order to study the performance of the system related to the device quality.

The available information of Wacom tablets using the pen stylus is the following: X and Y pen coordinates, pressure, pen angular orientation (azimuth and altitude angles) and timestamp information. However, in Samsung tablets using the pen stylus just X and Y pen coordinates, pressure and timestamp are available.

Fig. 1 shows an image of the setup used to capture the database, with all five capturing devices. The same capturing protocol was used for all five devices, they were placed on a table and subjects were told to feel comfortable when writing on them, so small rotation of the devices were allowed. Data was collected in two sessions (i.e. multi-session) for 70 subjects with a time gap between session of at least three weeks. For each user, there are a total of 8 genuine signatures and 6 skilled forgeries per device. Two different types of skilled forgeries are considered in the e-BioSign database. In the first session users were allowed to visualize a recording of the dynamic realization of the signature to forge for a few times. In the second session, a paper with the image of the signatures to forge is placed over the devices and they can trace the lines to perform the forgery. For more information about the e-BioSign database see [17].

III. Dynamic Signature Verification System

A. Data Preprocessing Stage

Due to the e-BioSign database is comprised of five different devices and signatures from all five devices are used in the development of the system stage, the first step before comparing the similarity between signatures is to make a data preprocessing stage in order to make signatures from the same user coming from different devices as similar as possible. This stage is similar than the proposed in previous works [2]. Several statistical data normalization techniques have been studied in order to compensate for geometric differences between devices. Finally, the mean and standard deviation normalization has been applied since it achieved the best results. Other normalization techniques were also studied such as max-min or mean normalizations. An additional preprocessing step using interpolation based on splines [18] is necessary due to the difference sampling frequency between devices being 200 Hz the final sampling frequency chosen for all devices.

B. Feature Extraction and Selection

A time functions-based system based on previous works [5], [19] is considered. Only time functions related to X, Y coordinates and pressure are considered in this work. Time functions related to pen angular orientation (azimuth and altitude angles) have been discarded due to this information is not available in Samsung tablets. The number of time functions considered by SFFS algorithm depends on the case to study. On the one hand, analyzing the Standard System (i.e. using X, Y coordinates and first- and second-order derivatives of them), signals captured by the tablets are used to extract a set of 23 time functions (see Table I) for each signature. On the other hand, analyzing the Secure System a total of 17 time functions are considered (i.e. not using X, Y coordinates and first- and second-order derivatives of them). These correspond to the following time functions reported in Table I: 1, 2, 8, 9, 15 and 16.

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 time functions for each system considered in this work improving the performance 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 development of the system stage, signatures from all five devices have been taken into account using

Fig. 1. Acquisition setup for e-BioSign database.

TABLE I. *Set of local features considered in this work.*

#	Feature
1	x-coordinate: x_n
2	y-coordinate: y_n
3	Pen-pressure: z_n
4	Path-tangent angle: θ_n
5	Path velocity magnitude: v_n
6	Log curvature radius: ρ_n
7	Total acceleration magnitude: a_n
8-14	First-order derivate of features 1-7: $\dot{x_n}, \dot{y_n}, \dot{z_n}, \dot{\theta_n}, \dot{v_n}, \dot{\rho_n}, \dot{a_n}$
15-16	Second-order derivate of features 1-2: $\ddot{x_n}, \ddot{y_n}$
17	Ratio of the minimum over the maximum speed over a 5-samples window: v_n^r
18-19	Angle of consecutive samples and first order difference: $\alpha_n, \dot{\alpha_n}$
20	Sine: s_n
21	Cosine: c_n
22	Stroke length to width ratio over a 5-samples window: r_n^5
23	Stroke length to width ratio over a 7-samples window: r_n^7

SFFS algorithm in order to obtain an optimal time functions-based system per case (i.e. Standard and Secure Systems) which works properly for all devices (see Sec. IV-A).

C. Time Functions-Based Signature Verification System

DTW algorithm [8] is used to compare the similarity between time functions from signatures. Scores are obtained as:

$$score = e^{-D/K} \tag{1}$$

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

IV. Experimental Work

A. Experimental Protocol

The first 35 users of the e-BioSign database are used for development and training of the system, while the remaining 35 users are employed for evaluating the system. Two different experimental protocols have been applied depending on the stage to analyze.

On the one hand, in the development and training of the system stage, signatures from all five devices are considered in order to obtain an optimal time functions-based system which works properly for all devices and therefore, it takes into account the problem of device interoperability [2]. A total of 5 genuine signatures (i.e. one genuine signature per device) of the first session are used as a training signatures, whereas the remaining 4 genuine signatures of the second session from all 5 devices (i.e. a total of 20 genuine signatures) are left for testing. Skilled forgery scores are obtained by comparing training signatures against a total of 30 available skilled forgeries signatures (i.e. 6 skilled forgeries signatures $\times$ 5 devices) 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 of all five devices (i.e. 34 users $\times$ 5 devices $\times$ 1 genuine signature = 170 random signatures).

On the other hand, in the validation of the system stage, the same device has been used for training and testing the system in order to analyze the performance of the system applied to security applications in banking and commercial sectors. Therefore, the first 4 genuine signatures of the first session are used as training signatures, whereas the remaining 4 genuine signatures of the second session are left for testing. Skilled forgery scores are obtained by comparing training signatures against the 6 available skilled forgeries signatures for the same user whereas random or zero-effort forgery scores are obtained

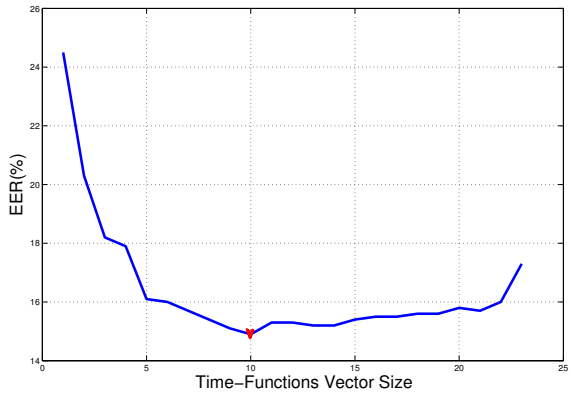
by comparing the training signatures to one genuine signature of the remaining users.

Two different scenarios have been considered, as in [21]: i) an office scenario with a high quality pen tablet specifically designed to acquire signatures (i.e. Wacom device), and ii) a mobile scenario where users sign on their general purpose smartphones or tablets (i.e. Samsung device). In the validation stage of the proposed systems (see Sec. IV-C), one device from the office scenario (i.e. Wacom STU-530) and another one from the mobile scenario (i.e. Samsung ATIV7) are considered in order to analyze the differences between both scenarios.

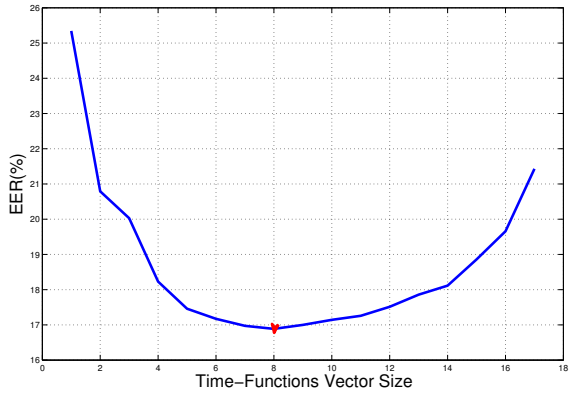
B. Development Experimental Results

Time functions selection is performed on the development set of 35 users. SFFS algorithm has been implemented in order to improve the EER of the system for skilled forgeries case as it is the most challenging case. As the goal of this work is to study the performance of both Standard and Secure Systems, SFFS algorithm has been individually applied per system in order to obtain the best optimal subset of time functions for each case. Fig. 2 shows the verification performance in terms of the size of the optimal time-functions vector selected by the SFFS algorithm for each case. As it can be seen, analyzing the Standard System, a subset of 10 time functions has been obtained as the best optimal time-functions subset whereas for the Secure System the best optimal subset is comprised of 8 time functions.

The performance of the systems in the development stage is slightly better for the Standard System compared to the Secure System (15% and 17% of EER respectively). This is due to the fact that X, Y and their second-order derivative time-functions have been chosen in the optimal time-functions subset in the Standard System.



(a) Standard System



(b) Secure System

Fig. 2. *Verification performance in terms of the size of the optimal time-functions vector selected by the SFFS algorithm.*

C. Validation Experimental Results

In order to validate the implemented systems, we compute the verification performance system on the remaining 35 users of e-BioSign database using the optimal time-functions vectors obtained on the development stage of the system for both Standard and Secure cases. Two devices (Wacom STU-530 and Samsung ATIV7) have been considered in order to analyze the performance of the system over the office and mobile scenarios quoted in Sec. IV-A. The performance of both Standard and Secure Systems is represented using DET plots as shown in Fig. 3. In addition, the EER values for both Standard and Secure Systems and for skilled and random forgeries are also depicted in Table II.

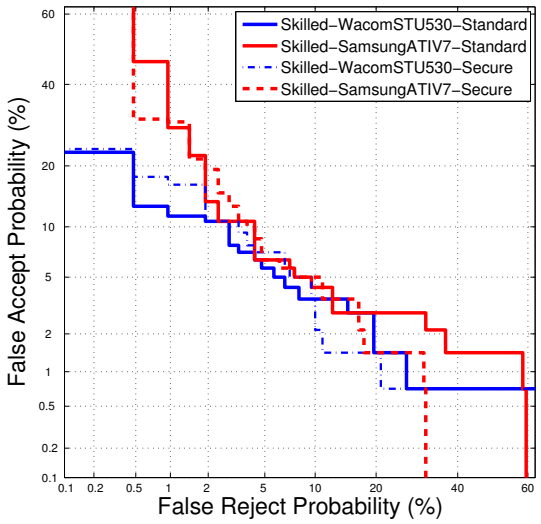
Two important conclusions can be extracted from the results: First, the performance of the system for both office and mobile scenarios (i.e. Wacom STU-530 and Samsung ATIV7) is very similar for the skilled forgeries cases. The Standard System for Wacom STU-530 achieves better results in absolute numbers of 0.7% EER compared to the Standard System for Samsung ATIV7. However, analyzing the Secure System for skilled forgeries cases, the Samsung ATIV7 device achieves better results in absolute numbers of 0.5% EER compared to the Wacom STU-530 device. Furthermore, analyzing the random forgeries cases for both devices we can see that Samsung ATIV7 device achieves a better performance in all cases. Second, the performance of both Standard and Secure Systems are very similar for both devices. Analyzing the skilled forgeries cases, the Standard System for Wacom STU-530 is only 1.2% better in terms of EER compared to the Secure System and the performance of the Samsung ATIV7 is even the same for both Standard and Secure Systems. Finally, analyzing the random forgeries cases, the performance of both systems (i.e. Standard and Secure Systems) for each device are very similar, even achieving better results for the Samsung device. Therefore, in this work we can conclude that the use of a Secure System applying a good time-function extraction and selection algorithm can provide a performance similar to a Standard System. These results open the door to the use of Secure Systems in some applications such as banking in order to avoid the lost of important user information against possible external attacks.

V. Conclusion

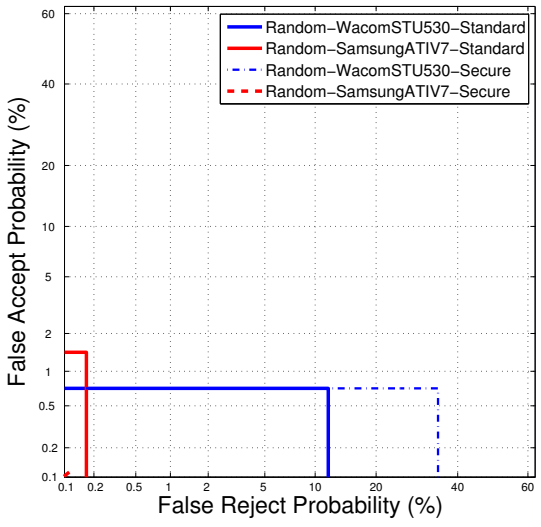
In this paper, two different approaches of time functions-based systems have been studied for dynamic signature verification: Standard System, an optimal time functions-based system taking into account the information related to X and Y coordinates and pressure, which is the common practice and a Secure System, following an extreme case not considering

TABLE II. *System performance in terms of EER (%) on the evaluation set of 35 users using time functions-based systems. Comparison of the results obtained by Standard and Secure Systems.*

	Skilled forgeries		*Random forgeries*	
Device	Standard System	Secure System	Standard System	Secure System
Wacom STU-530	5.7	6.9	0.7	0.7
Samsung ATIV7	6.4	6.4	0.3	0.1



(a) Skilled forgeries case



(b) Random forgeries case

Fig. 3. *DET curves for the proposed time functions-based signature recognition systems on the evaluation set of e-BiosSign for both Standard and Secure Systems.*

information related to X, Y coordinates and their derivatives on the biometric system which would be a much more robust system against attacks, as this critical information would not be stored anywhere. An optimal time-functions vector has been chosen per system (i.e. Standard and Secure Systems) using the SFFS algorithm in the development stage of the system and taking into account the device interoperability problem as signatures from all five devices of e-BioSign database have been considered. The results reported in Sec. IV-C have shown two important conclusions. First, the use of newer general purpose devices (mobile scenario) such as Samsung ATIV7 has shown very robust recognition performance, so they could be reliably used in banking and commercial applications. Second, the use of a Secure System in some applications such as banking applying a good time-functions extraction and selection algorithm can provide a performance similar to a Standard System. In this case sensible information is not stored in the biometric template information, the opposite to a standard case which could allow criminals to reconstruct the original signatures of the users. For future work, it would be interesting to analyze the performance of the Secure System also for mobile scenarios using the finger instead of the pen stylus.

ACKNOWLEDGMENT

This work was supported in part by the Project Bio-Shield (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] R. Plamondon and G. Lorette, "Automatic Signature Verification and Writer Identification - The State of the Art." *Pattern Recognition*, vol. 22, no. 2, pp. 107–131, 1989.

[2] R. Tolosana, R. Vera-Rodriguez, J. Ortega-Garcia, and J. Fierrez, "Preprocessing and Feature Selection for Improved Sensor Interoperability in Online Biometric Signature Verification," *IEEE Access*, vol. 3, pp. 478 – 489, May 2015.

[3] 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. 5th IAPR Intl. Conf. on Audio- and Video-based Biometric Person Authentication, AVBPA*, ser. LNCS, vol. 3546. Springer, July 2005, pp. 523–532.

[4] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia, "Feature-Based Dynamic Signature Verification Under Forensic Scenarios," in *Proc. 3rd International Workshop on Biometrics and Forensics (IWBF)*, March 2015.

[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*, vol. 28, no. 16, pp. 2325–2334, December 2007.

[6] A. Kholmatov and B. Yanikoglu, "Identity Authentication Using Improved Online Signature Verification Method," *Pattern Recognition Letters*, vol. 26, no. 15, pp. 2400 – 2408, 2005.

[7] M. Faundez-Zanuy, "On-Line Signature Recognition Based on VQ-DTW," *Pattern Recognition*, vol. 40, no. 3, pp. 981 – 992, 2007.

[8] H. Sakoe and S. Chiba, "Dynamic Programming Algorithm Optimization for Spoken Word Recognition," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, no. 1, pp. 43–49, 1978.

[9] D. Muramatsu and T. Matsumoto, "An HMM On-Line Signature Verifier Incorporating Signature Trajectories." in *Proc. ICDAR*. IEEE Computer Society, 2003, pp. 438–442.

[10] M. M. Fahmy, "Online Handwritten Signature Verification System Based on DWT Features Extraction and Neural Network Classification," *Ain Shams Engineering Journal*, vol. 1, no. 1, pp. 59 – 70, 2010.

[11] Y. Liu, Z. Yang, and L. Yang, "Online Signature Verification Based on DCT and Sparse Representation," *Cybernetics, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2014.

[12] J. Galbally, A. Ross, M. Gomez-Barrero, J. Fierrez, and J. Ortega-Garcia, "Iris Image Reconstruction From Binary Templates: An Efficient Probabilistic Approach Based on Genetic Algorithms," *Computer Vision and Image Understanding*, vol. 117, no. 10, pp. 1512–1525, October 2013, selected for Elsevier Virtual Issue: Celebrating the Breadth of Biometrics Research.

[13] M. Gomez-Barrero, J. Galbally, A. Morales, M. A. Ferrer, J. Fierrez, and J. Ortega-Garcia, "Inverse Biometrics: A Case Study in Hand Geometry Authentication," in *Proc. IAPR Int. Conf. on Pattern Recognition, ICPR*, November 2012, pp. 1281–1284.

[14] J. Feng and A. Jain, "Fingerprint Reconstruction: From Minutiae to Phase," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 2, pp. 209–223, Feb 2011.

[15] A. K. Jain, K. Nandakumar, and A. Nagar, "Biometric Template Security," *EURASIP J. Adv. Signal Process*, vol. 2008, pp. 113:1–113:17, Jan. 2008. [Online]. Available: http://dx.doi.org/10.1155/2008/579416

[16] M. Gomez-Barrero, C. Rathgeb, J. Galbally, J. Fierrez, and C. Busch, "Protected Facial Biometric Templates Based on Local Gabor Patterns and Adaptive Bloom Filters," in *Proc. IAPR/IEEE Int. Conf. on Pattern Recognition, ICPR*, August 2014, pp. 4483–4488.

[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. 3rd International Workshop on Biometrics and Forensics (IWBF)*, March 2015.

[18] 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. Intl. Conf. on Document Analysis and Recognition, ICDAR*, vol. 2, September 2007, pp. 1113–1117.

[19] M. Martinez-Diaz, J. Fierrez, R. P. Krish, and J. Galbally, "Mobile Signature Verification: Feature Robustness and Performance Comparison," *IET Biometrics*, 2014.

[20] J. N. P. Pudil and J. Kittler, "Floating Search Methods in Feature Selection." *Pattern Recognition Letters*, vol. 15, no. 10, pp. 1119–1125, 1994.

[21] R. Vera-Rodriguez, J. Fierrez, A. Morales, R. Tolosana, and J. Ortega-Garcia, "Automatic Student Authentication Using Dynamic Signature Recognition," in *Proc. CINAIC*, 2015.