Update Strategies for HMM-Based 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*—Biometric authentication on devices such as smartphones and tablets has increased significantly in the last years. One of the most acceptable and increasing traits is the handwriting signature as it has been used in financial and legal agreements scenarios for over a century. Nowadays, it is frequent to sign in banking and commercial areas on digitizing tablets. For these reasons, it is necessary to consider a new scenario where the number of training signatures available to generate the user template is variable and besides it has to be taken into account the lap of time between them (inter-session variability). In this work we focus on dynamic signature verification. The main goal of this work is to study system configuration update strategies of time functions-based systems such as Hidden Markov Model (HMM) and Gaussian Mixture Models (GMM). Therefore, two different cases have been considered. First, the usual case of having an HMM-based system with a fixed configuration (i.e. Baseline System). Second, an HMM-based and GMM-based systems whose configurations are optimized regarding the number of training signatures available to generate the user template. The experimental work has been carried out using an extended version of the Signature Long-Term database taking into account skilled and random or zero-effort forgeries. This database is comprised of a total of 6 different sessions distributed in a 15-month time span. Analyzing the results, the Proposed Systems achieve an average absolute improvement of 4.6% in terms of EER(%) for skilled forgeries cases compared to the Baseline System whereas the average absolute improvement for the random forgeries cases is of 2.7% EER. These results show the importance of optimizing the configuration of the systems compared to a fixed configuration system when the number of training signatures available to generate the user template increases.**

Keywords*—Biometrics, dynamic signature, system configuration update, time functions-based system, HMM, GMM, Signature Long-Term database

I. Introduction

Due to the technological evolution and the quality improvement of sensors, devices such as smartphones and tablets have experimented a great deployment nowadays [1]. Therefore, the use of these newer devices as biometric authentication systems have begun to be applied in many sectors due to the higher number of advantages compared to traditional ways of authentication (i.e. password and card authentication systems). Handwritten signature is one of the most socially accepted traits as it has been used in financial and legal agreements scenarios for many years [2]. In addition, it is worth noting that signatures are very easy to acquire by means of these devices through stylus or even the finger [3]. For this reason, this paper is focused on dynamic or on-line signature verification systems where information of each instant of the signing process is available.

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). Another important problem related to intrinsic or intra-class variability of signatures is known as *aging* term [4] (i.e. the gradual decrease in a system performance due to the changes suffered by the user's trait along the time). Finally, it is also worth noting to consider sources of extrinsic variability such as the new device interoperability scenario [5] due to the high deployment of devices in the last years.

Traditionally, the number of signatures used to obtain a user's template in the training stage for an on-line signature verification system was being between 3 to 5 signatures [6], [7], [8]. However, due to the higher acceptability of devices in our society nowadays, the number of signatures available per user is rapidly increasing with time. Therefore, the main goal of this work is to analyze the performance and the optimal system configuration update strategies over different time functions-based signature verification systems taking into account the scenario where the number of training signatures available increases with time. It is worth noting that the case proposed in this work is an ideal case since we know that all training signatures are genuine (groundtruth). However, in a real application, when incorporating new signatures as training data there is the possibility of making errors regarding the labelling of the data [9], so the performance of the ideal case proposed in this work will be compared to the real case in future studies.

One of the most well-known and competitive state-of-the-art systems HMM-based system is considered in this work. Basically, the HMM represents a doubly stochastic process governed by an underlying Markov chain with finite number of states and a set of random functions each of which is associated with the output observation of one state [10]. There are many previous studies in which the system proposed for signature verification is based on HMM algorithms [11], [12]. In addition, a GMM-based system which can be seen as a particular case of HMM with only one hidden state is considered in this work. This GMM-based system has been proposed as it has been widely used in other biometric traits such as speech recognition [13] and it has provided a good performance in previous studies related to on-line signature verification [14]. To the best of our knowledge, despite the long amount of studies related to HMM-based dynamic signature verification systems, none of them have analyzed the optimal configuration of the HMM-based system (i.e. number of hidden states (N) and number of Gaussian Mixtures per state (M)) in function of the number of training signatures available in the enrollment stage. Experiments are carried out using an extended version of the on-line Signature Long-Term database [4] in which both skilled and random forgeries cases are considered.

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 an extended version of the Signature Long-Term database [4]. Fig. 1 shows the number of genuine signatures per user and the general time diagram of the different acquisition sessions of it. This database was used in [4] taking into account random forgeries. However, skilled forgeries are considered in this extended version of the database too, which will be made publicly available. This database is comprised of a total of 29 users. The problem of inter-session variability is also considered in this work due to signatures were acquired in 6 different sessions with a 15-month time span, emulating a real scenario like we can find in commercial and banking sectors nowadays. The total number of genuine signatures and skilled forgeries per user are 46 and 10 respectively. 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.

Signatures were captured using a digitizing pen tablet WACOM Intuos3 A6 digitizer at 100 Hz and writing on a paper. The available information of this device is the following: *X* and *Y* pen coordinates, pressure, pen angular orientation (azimuth and altitude angles) and timestamp information. For more information about the Signature Long-Term database see [4].

III. Dynamic Signature Verification System

A. Feature Extraction and Selection

A time functions-based system based on previous studies [8], [15] is considered here. Signals captured by the digitizer are used to extract a set of 23 time functions (see Table I) for each signature. Only time functions related to *X*, *Y* coordinates and pressure information are considered in this work. Time functions related to pen angular orientation (azimuth and altitude angles) have been discarded in order to consider the same set of time functions that we would be able to use in general purpose devices such as tablets and smartphones.

Sequential Forward Feature Selection (SFFS) algorithm [16] is performed in order to obtain a subset of the 23 time functions 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 this work, the HMM-based system used in the experiments is based on [15]. An optimal subset comprised of 9 time functions was chosen using SFFS algorithm.

TABLE I. *Set of time functions 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: α_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

B. Time Functions-Based Signature Verification System

HMM algorithm [17] represents a double stochastic process, governed by an underlying Markov chain, with a finite number of states and random function set that generate symbols or observations each of which is associated with one state. The basic configuration of an HMM-based system (see Fig. 2) is comprised by the following elements:

- Number of hidden states N.
- Number of Gaussian mixtures per state M.

The HMM-based system considered has a left-to-right configuration without skipping state transitions (see Fig. 2). In addition, a GMM-based system which can be seen as a particular case of HMM with only one hidden state is also considered. Similarity scores are computed as the log-likelihood of the target signature (using the Viterbi algorithm) divided by the total number of samples of the signature signal. In order to keep scores between a reasonable range, normalised scores s_n between (0,1) are obtained as $s_n = exp(s(x, C)/30)$, where $s(x, C)$ is the score returned by the HMM algorithm and x and C represent respectively the input signature to verify and the enrolled model of the claimed identity.

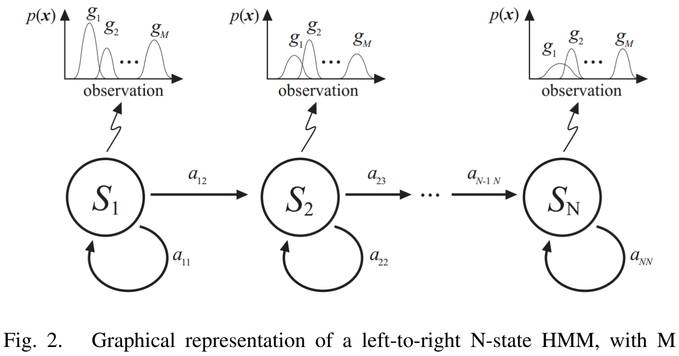


Fig. 2. Graphical representation of a left-to-right N-state HMM, with M Gaussian Mixtures per state. Ref. [18]

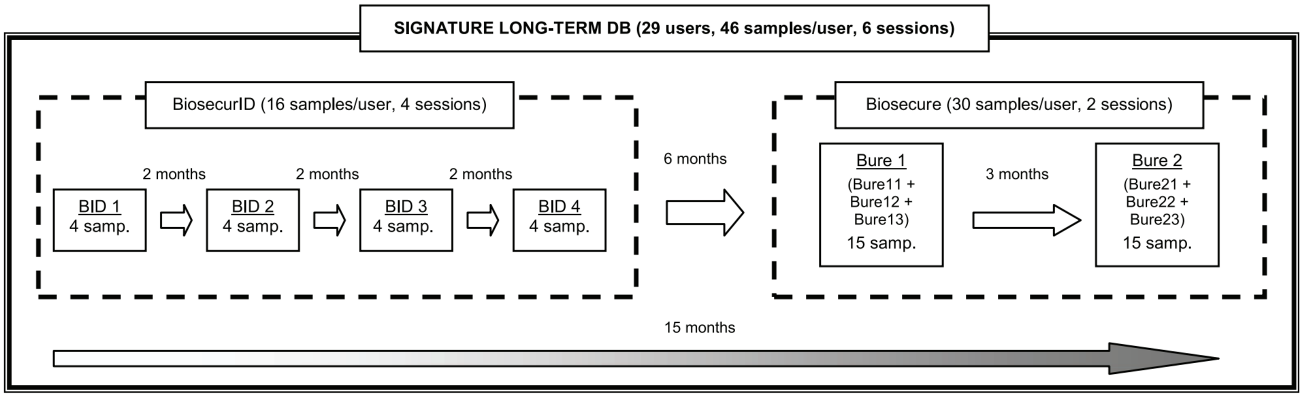

SIGNATURE LONG-TERM DB (29 users, 46 samples/user, 6 sessions)
BiosecurID (16 samples/user, 4 sessions)
2 months
2 months
2 months
BID 1
4 samp.
BID 2
4 samp.
BID 3
4 samp.
BID 4
4 samp.
6 months
Biosecure (30 samples/user, 2 sessions)
Bure 1
(Bure11 + Bure12 + Bure13)
15 samp.
3 months
Bure 2
(Bure21 + Bure22 + Bure23)
15 samp.
15 months

Fig. 1. General time diagram of the different acquisition sessions and number of genuine signatures per user that conform the Signature Long-Term database. Ref. [4]

IV. Experimental Work

A. Experimental Protocol

The main goal of this work is to analyze the optimal configuration of the HMM-based and GMM-based systems regarding the number of training signatures available per user. Therefore, the last 5 genuine signatures (i.e. Bure23 block) of the sixth session are always used for testing. Skilled forgery scores are obtained by comparing training signatures against the 10 available skilled forgeries 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 following experiments have been proposed in order to know how the configuration (i.e. number of hidden states (N) and Gaussian mixtures per state (M)) of the HMM-based and GMM-based systems change with the number of training signatures available in the enrollment stage:

- Exp. A: The first 4 genuine signatures (i.e. BID 1) are used in the enrollment stage.
- Exp. B: The first 16 genuine signatures (i.e. BiosecurID) are used in the enrollment stage.
- Exp. C: The first 31 genuine signatures (i.e. BiosecurID + Bure 1) are used in the enrollment stage.
- Exp. D: The first 41 genuine signatures (i.e. BiosescurID + Bure 1 + Bure21 + Bure22) are used in the enrollment stage. In this last case it is important to highlight that signatures from the same session are used for both training and testing the system, so results can be overoptimistic in this case.

B. Experimental Results

1) Baseline System: In this section, the usual case of having an HMM-based system whose configuration is fixed (i.e. N = 2 and M = 32) [15] is analyzed when the number of training signatures available to generate the user template increases. Table II shows the performance of the Baseline System in terms of the EER(%) for the different experiments quoted in the experimental protocol IV-A.

Analyzing the skilled forgeries cases, the performance of the Baseline System improves when the number of training signatures increases. The results obtained in this section agree with previous studies [4]. However, analyzing the random forgeries cases, it seems that there is a limit in the improvement of the system regarding the number of training signatures. This effect has been reported in previous studies [8] as well.

2) Proposed Systems: In this experiment, the goal is to analyze the optimal configuration of both HMM and GMM Proposed Systems regarding the number of training signatures available to generate the user template. Skilled forgeries case has been chosen as the case to optimize as it is the most challenging case to authenticate. Analyzing the configuration of the HMM-Proposed System, Table III shows the performance (i.e. EER(%)) of the HMM-Proposed System in terms of the number of hidden states (N) and the number of Gaussian mixtures per state (M) for the four experiments.

Some important conclusions can be extracted from the results. First, results show the importance of taking into account different configurations of the HMM-based system regarding the number of training signatures available to generate the user template. In general, when the number of training signatures is low (i.e. Exp. A), the optimal configuration of the HMM-based system tends to have a higher number of Gaussian mixtures than hidden states. However, when the number of training signatures available increases (i.e. Exp. B, C, D), the number of hidden states is higher than the number of Gaussian Mixtures per state. Second, it seems that choosing an optimal configuration of the HMM-based system tends to improve the performance of the system as the number of training signatures increases. This effect is different that we could see in the Sec. IV-B1. Therefore, the higher number of training signatures, the better performance of the system.

TABLE II. Baseline System: Performance of the fixed HMM-based system with N = 2 and M = 32 in terms of the EER(%).

	Exp. A	Exp. B	Exp. C	Exp. D
Skilled Forgeries	16.6	13.1	9.0	7.6
Random Forgeries	8.3	2.8	2.4	3.4

TABLE III. EER(%) FOR DIFFERENT HMM CONFIGURATIONS FOR BOTH SKILLED (TOP) AND RANDOM (BOTTOM) FORGERIES REGARDING THE NUMBER OF TRAINING SIGNATURES AVAILABLE TO GENERATE THE USER TEMPLATE (I.E. EXP. A, B, C AND D). N = NUMBER OF HIDDEN STATES; M = NUMBER OF GAUSSIAN MIXTURE PER STATE.

Exp. A (4 signatures)				Exp. B (16 signatures)			
N	M=2	M=16	M=32	N	M=2	M=16	M=32
2	26.9 8.2	**13.1** **5.5**	16.6 8.3	2	26.2 5.5	11.0 1.6	13.1 2.8
16	17.2 6.9	29.7 23.1		16	11.7 2.1	15.2 2.9	
32	19.3 7.6	37.2 24.8		32	**9.0** **0.7**	13.8 6.2	
64	27.6 12.5			64	12.4 5.5		

Exp. C (31 signatures)				Exp. D (41 signatures)			
N	M=2	M=16	M=32	N	M=2	M=16	M=32
2	24.1 4.8	6.9 3.4	9.0 2.4	2	22.1 4.8	3.4 0.0	7.6 3.4
16	6.2 0.2	9.7 1.4		32	2.1 0.0	3.4 0.0	
32	**4.8** **0.0**	6.2 0.2		64	**1.4** **0.0**	2.1 0.0	
64	5.0 2.1	6.9 3.4		128	14.5 10.3	14.8 10.3	

Regarding the Proposed GMM-Based System, Table IV shows the performance of the system in terms of the EER(%) regarding the number of training signatures available to generate the user template and the configuration of the system (i.e. the number of Gaussian mixtures per state (M)). The results of the GMM-based system are similar to the HMM-based system. It is very important to analyze different configurations of the GMM-based system in terms of the number of training signatures available to generate the user template. In general, we can see that the number of Gaussian mixtures required (M) increases with the number of training signatures available.

Finally, Fig. 3 shows the best performance of the Baseline, HMM and GMM Proposed Systems for the experiments quoted in the experimental protocol (see Sec. IV-A). Some important conclusions can be extracted from Fig. 3. The performance of both HMM and GMM Proposed Systems are much better compared to the Baseline System for both skilled and random forgeries cases. For this reason, hereinafter we evaluate the average performance of the HMM and GMM Proposed Systems compared to the Baseline System pointing out some important differences between the Proposed Systems.

Analyzing the skilled forgeries cases, the Proposed Systems achieves an average absolute improvement of 4.6% compared to the Baseline System. In addition, it is important to highlight that when the number of training signatures increases, there is a higher difference in the performance among the Proposed and Baseline Systems. For example, in the Exp. A there is an absolute improvement of 3.2% in the Proposed Systems compared to the Baseline System whereas in the Exp. C the absolute improvement is 4.6%. Analyzing the random forgeries cases, the Proposed Systems achieves an average absolute improvement of 2.7% compared to the Baseline System. In addition, analyzing the HMM and GMM Proposed Systems for random forgeries case, it is worth noting that when the number of training signatures available is low (i.e. Exp. A) the GMM-Proposed System works better than the HMM-Proposed System whereas this effect is the opposite when the number of training signatures available is higher (i.e. Exp. C and Exp. D).

Finally, the best performance of the system is obtained using an HMM-based system with N=64 and M=2 achieving an EER of 1.4% and 0.0% for skilled and random forgeries, respectively. However, it is important to highlight that in this case signatures from the same session have been used for training and testing the system. Regarding the case of training and testing the system with signatures coming from different sessions, the best performance of the system has been obtained using a GMM-based system with M=128 achieving an EER of 4.1% and 0.7% for skilled and random forgeries, respectively.

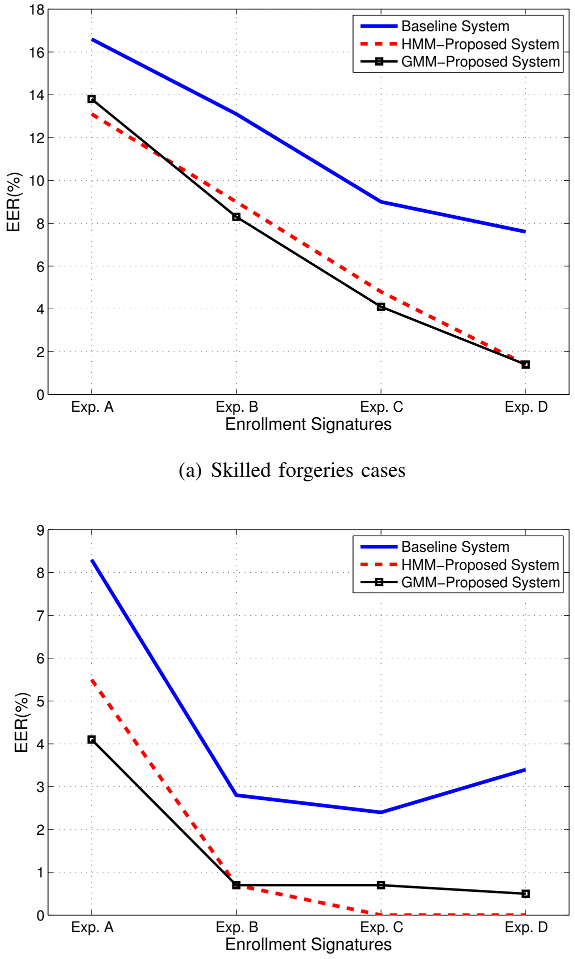
Therefore, the results show the importance of optimizing the configuration of the systems compared to a fixed configuration system when the number of training signatures available to generate the user template increases.

V. CONCLUSION

In this paper, the optimal configuration of time functions-based systems regarding the number of training signatures

TABLE IV. EER(%) FOR DIFFERENT GMM CONFIGURATIONS FOR BOTH SKILLED (TOP) AND RANDOM (BOTTOM) FORGERIES REGARDING THE NUMBER OF TRAINING SIGNATURES AVAILABLE TO GENERATE THE USER TEMPLATE (I.E. EXP. A, B, C AND D). M = NUMBER OF GAUSSIAN MIXTURE PER STATE.

	M=16	M=32	M=64	M=128	M=256	M=512	M=1024
Exp. A	18.6 6.9	**13.8** **4.1**	18.6 6.9				
Exp. B	14.5 3.4	11.0 2.1	8.3 2.1	**8.3** **0.7**	9.0 2.8		
Exp. C	9.0 2.1	4.1 1.4	4.8 0.7	**4.1** **0.7**	4.2 0.8		
Exp. D	5.5 1.6	3.4 0.7	2.8 0.7	2.1 0.2	2.1 0.2	**1.4** **0.5**	4.1 0.7



(a) Skilled forgeries cases

(b) Random forgeries cases

Fig. 3. *Best performance in terms of the EER(%) of the Baseline and Proposed Systems for the 4 different experiments quoted in the experimental protocol (i.e. Sec. IV-A) and for skilled and random forgeries.*

available to generate the user template has been studied for dynamic signature verification. First, the traditional case of having an HMM-based system with a fixed configuration is considered (i.e. Baseline System). Second, an HMM-based and GMM-based systems whose configuration is optimized in terms of the number of training signatures available to generate the user template have been proposed (i.e. HMM-Proposed and GMM-Proposed Systems). The experimental work has been carried out using an extended version of the Signature Long-Term database taking into account skilled and random or zero-effort forgeries. This database is comprised of a total of 6 different sessions distributed in a 15-month time span. The results reported in this work have shown the importance of taking into account different configurations of the systems regarding the number of training signatures available for both HMM-based and GMM-based systems. Both HMM and GMM Proposed Systems have achieved a similar performance in dynamic signature verification. Analyzing the results, the Proposed Systems have achieved an average absolute improvement of 4.6% for skilled forgeries cases compared to the Baseline System whereas the average absolute improvement for the random forgeries cases has been 2.7%. In conclusion, the results show the importance of optimizing the configuration of the systems compared to a fixed configuration system when the number of training signatures available to generate the user template increases. For future work, the system configuration update strategies proposed in this work will be analyzed using different databases. Therefore, due to the lack of databases with a higher number of genuine signatures per user, we will acquire a new database in order to check the performance of the Proposed Systems using a different set of users for development and testing the system. Furthermore, it would be interesting to analyze the system configuration update strategies for the HMM and GMM Proposed Systems regarding the complexity of signatures.

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] F. S. Meng W, Wong DS and Z. J, “Surveying the Development of Biometric User Authentication on Mobile Phones,” *IEEE Communications Surveys Tutorials*, vol. PP, no. 99, pp. 1–1, 2015.

[2] D. Impedovo and G. Pirlo, “Automatic Signature Verification: The State of the Art.” *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, vol. 38, no. 5, pp. 609–635, 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 International Workshop on Biometrics and Forensics (IWBF)*, March 2015.

[4] J. Galbally, M. Martinez-Diaz, and J. Fierrez, “Aging in Biometrics: An Experimental Analysis on On-Line Signature,” *PLOS ONE*, vol. 8, no. 7, p. e69897, July 2013.

[5] 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.

[6] A. K. Jain, F. D. Griess, and S. D. Connell, “On-Line Signature Verification,” *Pattern Recognition*, vol. 35, p. 2002, 2002.

[7] 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 Annual Int. Carnahan Conf. on Security Technology*, September 2015, to appear.

[8] 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.

[9] R. Vera-Rodriguez, J. S. D. Mason, and N. W. D. Evans, “Automatic cross-biometric footstep database labelling using speaker recognition,” in *Proc. International Conference on Biometrics (ICB)*, ser. LNCS, vol. 5558. Springer, June 2009, pp. 503–512.

[10] L. Yang, B. Widjaja, and R. Prasad, “Application of Hidden Markov Models for Signature Verification,” *Pattern Recognition*, vol. 28, no. 2, pp. 161 – 170, 1995.

[11] J. Fierrez-Aguilar, S. Krawczyk, J. Ortega-Garcia, and A. K. Jain, “Fusion of Local and Regional Approaches for On-Line Signature Verification,” in *Proc. Intl. Workshop on Biometric Recognition Systems, IWBRS*, ser. LNCS, vol. 3781. Springer, October 2005, pp. 188–196.

[12] J. Dolfing, E. Aarts, and J. van Oosterhout, “On-Line Signature Verification with Hidden Markov Models,” in *Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on*, vol. 2, Aug 1998, pp. 1309–1312 vol.2.

[13] D. Povey, L. Burget, M. Agarwal, P. Akyazi, K. Feng, A. Ghoshal, O. Glembek, N. Goel, M. Karafiat, A. Rastrow, R. Rose, P. Schwarz, and S. Thomas, “Subspace gaussian mixture models for speech recognition,” in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, March 2010, pp. 4330–4333.

[14] J. Richiardi and A. Drygajlo, “Gaussian Mixture Models for On-line Signature Verification,” in *Proceedings of the 2003 ACM SIGMM Workshop on Biometrics Methods and Applications*, ser. WBMA ’03. New York, NY, USA: ACM, 2003, pp. 115–122. [Online]. Available: http://doi.acm.org/10.1145/982507.982528

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

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

[17] L. R. Rabiner, “A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition,” in *Proc. IEEE 77 (2)*, pp. 257–286.

[18] M. Martinez-Diaz, “Dynamic Signature Verification for Portable Devices,” Master’s thesis, Universidad Autonoma de Madrid, November 2008.