Fusion of Local and Regional Approaches for On-Line Signature Verification

Julian Fierrez-Aguilar[1*], Stephen Krawczyk[2,3], Javier Ortega-Garcia[1], and Anil K. Jain[2]

[1] ATVS, Escuela Politecnica Superior, Universidad Autonoma de Madrid, Avda. Francisco Tomas y Valiente 11, Campus de Cantoblanco, 28049 Madrid, Spain
{julian.fierrez,javier.ortega}@uam.es

[2] Department of Computer Science and Engineering, Michigan State University, East Lansing, MI 48823, USA
jain@cse.msu.edu

[3] Information Technology Division, Naval Research Laboratory, Washington DC, USA
stephen.krawczyk@nrl.navy.mil

Abstract. Function-based methods for on-line signature verification are studied. These methods are classified into local and regional depending on the features used for matching. One representative method of each class is selected from the literature. The selected local and regional methods are based on Dynamic Time Warping and Hidden Markov Models, respectively. Some improvements are presented for the local method aimed at strengthening the performance against skilled forgeries. The two methods are compared following the protocol defined in the Signature Verification Competition 2004. Fusion results are also provided demonstrating the complementary nature of these two approaches.

1 Introduction

The goal of biometrics is to infer the identity of people based on anatomical or behavioral data (e.g., fingerprint, face, palmprint, signature, or voice) [1]. The current interest in biometrics is due to the increasing number of important applications where an automatic assessment of identity is crucial. Within biometrics, automatic signature verification has been an intense research area because of the social and legal acceptance and widespread use of the written signature as a personal authentication method [2], and still is a challenging problem. This is mainly due to the large intra-class variations and, when considering forgeries, small inter-class variations. See Fig. 2 for one example Chinese signature where this effect is evident.

This work deals with on-line signature verification, where the time functions of the dynamic signing process are available for recognition (e.g., position trajectories, or pressure versus time). Different approaches have been considered

* This work has been carried out while J. F.-A. was on a research stay at Michigan State University.

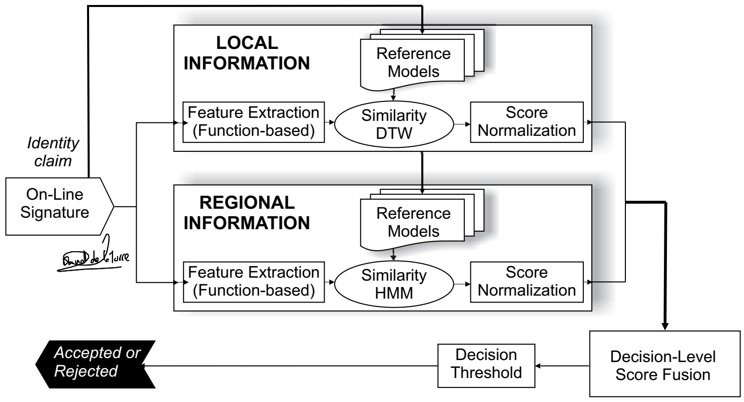

Fig. 1. System model of multilevel signature verification.

in the literature in order to extract relevant information from on-line signature data [3]; they can broadly be divided into: *i*) feature-based approaches, in which a holistic vector representation consisting of a set of global features is derived from the signature trajectories [4–7], and *ii*) function-based approaches, in which time sequences describing local properties of the signature are used for recognition (e.g., position trajectory, velocity, acceleration, force, or pressure) [5–12]. Although recent work shows that feature-based approaches are competitive with respect to function-based methods in some conditions [7], the latter methods have traditionally yielded better results [3].

Function-based approaches can be classified into local and regional methods. In local approaches, the time functions of different signatures are directly matched by using elastic distance measures such as Dynamic Time Warping (DTW) [9–11]. In regional methods, the time functions are converted to a sequence of vectors describing regional properties. One of the most popular regional approaches is the method based on Hidden Markov Models (HMM) [5, 7, 8, 12].

One major research trend in biometric verification is the successful exploitation of the different information levels embodied in the biometric signals [13]. This is usually done by combining the confidences provided by systems that utilize the different information levels in the form of similarity scores [14, 15] (see Fig. 1 for the system model). Examples can be found regarding other biometrics like fingerprint, where the combined use of local (e.g., minutiae-based) and global (e.g, ridge-based) approaches has been shown to improve verification performance [16], or speaker verification, where multilevel approaches are currently the state-of-the-art [17]. Some work on multi-level signature verification is available in [5–7].

The purpose of this work is to compare two representative local and regional approaches from the literature and to fuse them in order to study the benefits of their combination. In order to do so, we adhere to the standard benchmark defined in the First International Signature Verification Competition 2004 (SVC2004) [18]. In the case of the local DTW system, we also present some improvements over published work related to feature extraction and score normalization.

The paper is structured as follows. The system based on local information is described in Sect. 2 with emphasis on the improvements over published work. The regional system is briefly sketched in Sect. 3. The decision-level combination strategies are presented in Sect. 4. Experimental procedure and results are given in Sect. 5. Conclusions are finally drawn in Sect. 6.

2 System Based on Local Information (DTW)

An extension of the signature verification algorithm proposed by Jain et al. [10] was designed and developed. The original algorithm is based on the string matching technique called Dynamic Time Warping (DTW). Modifications of the original system were made to the preprocessing, feature extraction, and matching components. Each of the changes made had the intention of relying more on the temporal information as opposed to the spatial information of the signature. Here, we briefly outline the changes made to the original algorithm. For more details, we refer the reader to [19].

Preprocessing. Originally, the x and y signals of the signature were resampled in order to provide equidistant spacing between points in the signals. When resampling is performed, some temporal information is lost regarding the speed of the signature. In order to strengthen the system against skilled forgeries, we consider the speed of the signature as a feature by not resampling the input signatures. The only preprocessing steps performed are position normalization and stroke concatenation.

Feature Extraction. Three features are extracted at each point in the signature; δx, δy, and p. These features correspond to the change in the x direction, change in the y direction, and the pressure value, respectively. The feature values are normalized using signature-specific z-score normalization.

Matching. The matching algorithm is the standard DTW algorithm. The only differences with respect to the algorithm in [10] are related to the spurious and missing penalties, which we assume are equivalent. We refer to both penalties as the gap penalty G_P. Instead of imposing a constant penalty for missing or spurious points, the penalty is assigned based on the distance of the two feature vectors being compared, multiplied by a constant, i.e., $G_P(i,j) = c \cdot d_E(i,j)$, where $d_E(i,j)$ is the Euclidean distance between the feature vector of point i in the input signature and point j in the reference signature.

Score Normalization. We also added user-dependent score normalization to the new system by using the procedure described by Kholmatov et al. [11]. This method characterizes each matching with a three-dimensional score vector including maximum, minimum, and average distance information to the set of training signatures. This vector is then projected to one dimension with Principle Component Analysis (PCA).

3 System Based on Regional Information (HMM)

The regional system follows closely the system submitted by the authors to SVC 2004 [18], where it was ranked first and second for random and skilled forgeries, respectively (in Task 2, where signature data included position, pen orientation and pressure functions). The main difference of the HMM in the present work is the lack of user-dependent score normalization. A brief sketch of the system is given in this section, for more details we refer the reader to [12].

Feature extraction. Signature trajectories are first preprocessed with position and rotation normalization. The signature is parameterized as a set of 7 discrete-time functions $\{x, y, p, \theta, v, \rho, a\}$, and first-order time derivatives of all of them, totaling 14 discrete functions; p, θ, v, ρ and a stand respectively for pressure, path tangent angle, path velocity magnitude, log curvature radius and total acceleration magnitude. A z-score transformation is finally applied to each discrete-time function.

Matching. Left-to-right Hidden Markov Models with no transition skips between states are used for identity modeling. A reduced number of states and multivariate Gaussian Mixture density for modeling the observations are used.

4 Fusion Strategies

Two theoretical frameworks for combining matchers with application to biometric verification are described in [14] and [15]. More recent approaches are reviewed in [1]. It is now generally agreed that the weighted average is a good way of combining the similarity scores provided by the different systems.

In this work, fusion strategies based on the max, product, and sum rules are compared [15]. Similarity scores given by the local and regional systems are normalized before fusion into the range $[0, 1]$ with logistic functions, in order to overcome the problem of score outliers [20].

5 Experiments

5.1 Database and Experimental Protocol

Not many signature databases are publicly available for research purposes [21]. As a result, the common practice in on-line signature recognition research is to evaluate the proposed recognition strategies on small data sets acquired at individual research laboratories. In this environment, the First International Signature Verification Competition (SVC) was organized in 2004 to provide a common reference for system comparison on the same signature data [18].

Development corpus of the Task 2 in SVC2004 is used in the experiments. The on-line signature data in this case include not only position coordinates (as in Task 1) but also pressure and pen angles at a sampling frequency equal to 100

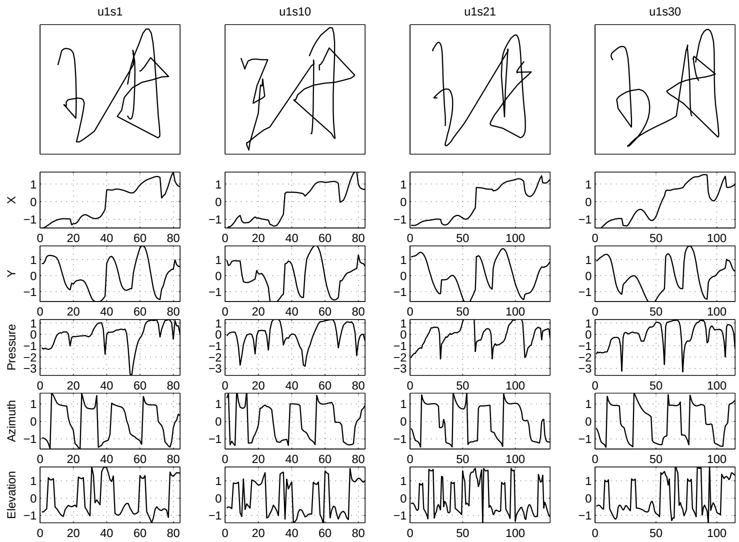

Fig. 2. Signature examples from SVC 2004 corpus. Two genuine signatures (left two columns) and two skilled forgeries (right two columns) are given along with their temporal trajectories for x and y position, pressure and pen angles.

Hz. Pen angles are not used in our study, as they probed to be highly unstable in previous work [7]. The corpus consists of 40 sets of signatures. Each set contains 20 genuine signatures from one subject (acquired in two separate sessions) and 20 skilled forgeries from five other subjects. The signatures are mostly in either English or Chinese. Some examples are depicted in Fig. 2 for one specific signer.

Training signatures are randomly selected from the first acquisition session (5 in the competition, from 3 to 10 in this work). 10 runs of this random selection are conducted. For each user, the 10 genuine signatures of the second session and the 20 skilled forgeries are used for testing. In case of random forgeries, one random genuine signature from each of 20 different subjects randomly selected is used as impostor data. Whenever randomness was involved, the same random sets are used for the two systems under study.

Overall system performances using *a posteriori* user-independent decision thresholds are reported by means of DET plots [22]. Average EER tables for *a posteriori* user-dependent decision thresholds are also given. For more details on *a priori* and *a posteriori* decision thresholding techniques and their application to signature verification, we refer the reader to [12].

Finally, we must also stress the difficulty of the tasks proposed in SVC as compared to previous reported work on signature verification. This is mainly because: *i*) data was acquired by using a pen tablet (WACOM Intuos) with an inkless pen, *ii*) signers were asked to create new signatures different to the ones used in daily life in order to protect their privacy, *iii*) forgers were given not only the shape but also the dynamics of the signatures to forge, and *iv*) training

Table 1. Comparative experiments. Verification performance with 3, 5 and 10 training signatures for *a posteriori* user-independent and user-dependent decision thresholds, both for skilled and random forgeries. Average EER is reported in %

		Local (DTW)			Regional (HMM)		
		3 sign.	5 sign.	10 sign.	3 sign.	5 sign.	10 sign.
skilled	user-indep.	17.02	14.26	11.57	23.69	15.04	11.56
	user-dep.	12.98	10.20	6.76	11.56	8.04	4.50
random	user-indep.	1.83	0.57	0.24	7.64	2.41	0.69
	user-dep.	0.96	0.24	0.24	0.90	0.34	0.00

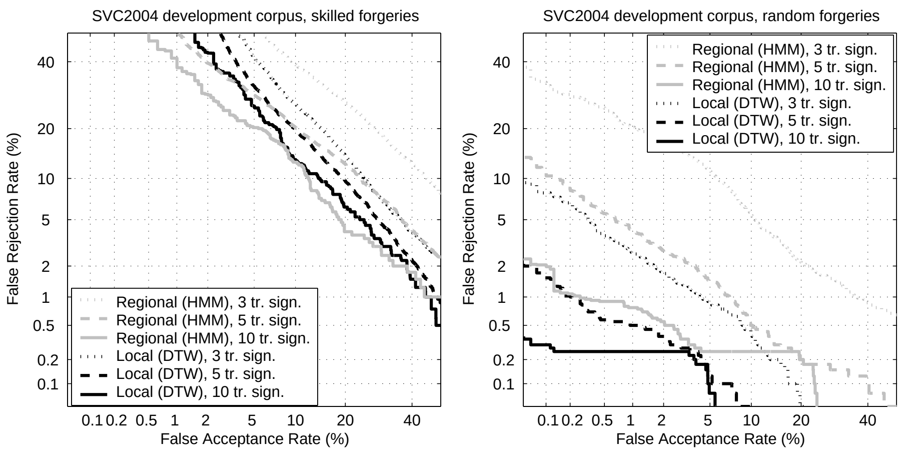

Fig. 3. Verification performance of the local and regional approaches under study on SVC2004 development corpus following SVC2004 protocol.

data is from one session and testing data is from a different session separated from the first session by at least one week.

5.2 Results

Verification performances of local and regional systems are given in Table 1. We observe a progressive improvement in performance as we increase the number of training signatures, especially for the HMM system. Utilizing user-dependent decision thresholds leads to significantly lower error rates than user-independent thresholds for both systems in the case of skilled forgeries. This effect is even more pronounced in the case of random forgeries for the HMM system, whereas it is not present in the DTW system, especially for large training set sizes. This can be explained by the user-dependent score normalization stage introduced in the DTW system, and the lack of user-dependent score normalization in this version of the HMM system. Comparatively, the DTW system outperforms the HMM

Table 2. Fusion experiments. Verification performance with 5 training signatures for *a posteriori* user-independent and user-dependent decision thresholds, both for skilled and random forgeries. Best SVC2004 refers to the best system competing in SVC2004 on this task and dataset. Average EER is reported in %

		Local	Regional	Max(L,R)	Prod(L,R)	Sum(L,R)	Best SVC2004
skilled	u-indep.	14.26	15.04	12.60	10.95	**10.91**	6.90
	u-dep.	10.20	8.04	7.69	7.60	**6.91**	
random	u-indep.	0.57	2.41	1.31	0.44	**0.49**	3.02
	u-dep.	0.24	0.34	0.19	0.14	**0.15**	

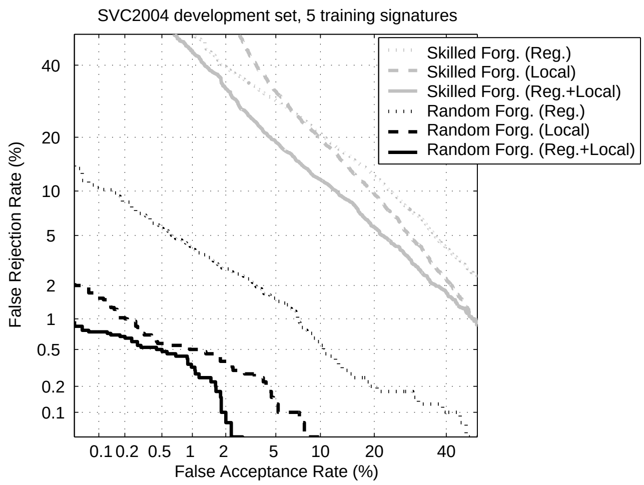

Fig. 4. Verification performance of the local and regional approaches under study as well as their combination on SVC2004 development corpus.

system in almost all test conditions considering user-independent thresholds, especially for low training set sizes. The opposite occurs when employing user-dependent thresholds, especially for large training set sizes.

Detection error trade-off curves for the two competing systems are given in Fig. 3 both for skilled (left) and random forgeries (right).

Fusion results following SVC 2004 protocol are given in Table 2 (5 training signatures). The two systems are shown to provide complementary information for the verification task, which is well exploited with the three reported fusion strategies, especially the sum rule. Results of the SVC2004 winner system [11] following the same experimental protocol are also given for reference. The two studied systems are worse than the SVC2004 winner for skilled forgeries when considered independently, but their combination results in similar performance when considering user-dependent thresholds. In case of random forgeries, the combination of DTW and HMM significantly outperforms the SVC2004 winner.

These results motivate the development of a hybrid system combining the DTW and HMM systems followed by user-dependent score normalization [12].

In Fig. 4 we depict detection error trade-off curves of the individual systems and their combination through the sum rule both for skilled and random forgeries.

6 Conclusions

Function-based approaches for on-line signature verification have been studied. These have been classified into local and regional approaches depending on the feature extraction process. Local and regional representative methods from the literature based on Dynamic Time Warping and Hidden Markov Models, respectively, have been selected and compared following the SVC2004 protocol. In case of the DTW system, some improvements aimed at strengthening the system against skilled forgeries have been introduced.

The enhanced DTW system including user-dependent score normalization appears to be quite robust to the common user-dependencies found in signature verification regarding score distribution miss-alignment between different users, especially for random forgeries when the number of training signatures increases. Comparatively, the DTW system outperformed the HMM system in almost all test conditions considering user-independent thresholds, especially for low training set sizes. The opposite occurred when employing user-dependent thresholds, especially for large training set sizes. The two systems are also shown to provide complementary information which is well exploited with the sum fusion rule.

The difficulty of the tasks proposed in SVC2004 comes from several factors, including: *i*) no visual feedback when writing, *ii*) subjects using invented signatures different to the ones used in daily life, *iii*) skilled forgers imitating not only the shape but also the dynamics, and *iv*) time variability between training and testing (at least one week). These factors affect significantly the performance of signature verification systems and are not well studied in the literature. Future work will be focused on characterizing the behavior of representative methods from the literature to these effects. In order to do so, collection of large biometric databases including these factors under controlled conditions is one of our priorities.

7 Acknowledgements

This work has been supported by Spanish MCYT TIC2003-08382-C05-01 and by European Commission IST-2002-507634 Biosecure NoE projects. J. F.-A. is supported by a FPI scholarship from Comunidad de Madrid.

References

1. Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition. IEEE Trans. on Circuits and Systems for Video Technology **14** (2004) 4–20

2. Plamondon, R., Srihari, S.N.: On-line and off-line handwriting recognition: A comprehensive survey. IEEE Trans. on PAMI **22** (2000) 63–84
3. Plamondon, R., Lorette, G.: Automatic signature verification and writer identification - The state of the art. Pattern Recognition **22** (1989) 107–131
4. Lee, L.L., Berger, T., Aviczer, E.: Reliable on-line human signature verification systems. IEEE Trans. on Pattern Anal. and Machine Intell. **18** (1996) 643–647
5. Kashi, R.S., Hu, J., Nelson, W.L., Turin, W.: On-line handwritten signature verification using Hidden Markov Model features. In: Proc. of ICDAR. (1997) 253–257
6. Zhang, K., Nyssen, E., Sahli, H.: A multi-stage on-line signature verification system. Pattern Analysis and Applications **5** (2002) 288–295
7. Fierrez-Aguilar, J., Nanni, L., Lopez-Penalba, J., Ortega-Garcia, J., Maltoni, D.: An on-line signature verification system based on fusion of local and global information. In: Proc. of AVBPA, Springer LNCS-3546 (2005) 523–532
8. Yang, L., Widjaja, B.K., Prasad, R.: Application of Hidden Markov Models for signature verification. Pattern Recognition **28** (1995) 161–170
9. Nalwa, V.S.: Automatic on-line signature verification. Proceedings of the IEEE **85** (1997) 215–239
10. Jain, A.K., Griess, F., Connell, S.: On-line signature verification. Pattern Recognition **35** (2002) 2963–2972
11. Kholmatov, A., Yanikoglu, B.: Biometric authentication using online signatures. In: Proc. ISCIS, Springer LNCS-3280 (2004) 373–380
12. Fierrez-Aguilar, J., Ortega-Garcia, J., Gonzalez-Rodriguez, J.: Target dependent score normalization techniques and their application to signature verification. IEEE Trans. on Systems, Man and Cybernetics, part C **35** (2005) 418–425.
13. Jain, A.K., Ross, A.: Multibiometric systems. Communications of the ACM **47** (2004) 34–40
14. Bigun, E.S., Bigun, J., Duc, B., Fischer, S.: Expert conciliation for multi modal person authentication systems by Bayesian statistics. In: Proc. of AVBPA, Springer LNCS-1206 (1997) 291–300
15. Kittler, J., Hatef, M., Duin, R., Matas, J.: On combining classifiers. IEEE Trans. on Pattern Anal. and Machine Intell. **20** (1998) 226–239
16. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer (2003)
17. Fierrez-Aguilar, J., Garcia-Romero, D., Ortega-Garcia, J., Gonzalez-Rodriguez, J.: Speaker verification using adapted user-dependent multilevel fusion. In: Proc. of MCS, Springer LNCS-3541 (2005) 356–365
18. Yeung, D.Y., Chang, H., Xiong, Y., George, S., Kashi, R., Matsumoto, T., Rigoll, G.: SVC2004: First International Signature Verification Competition. In: Proc. of ICBA, Springer LNCS-3072 (2004) 16–22
19. Krawczyk, S.: User authentication using on-line signature and speech. Master's thesis, Michigan State University, Dep. of Computer Science and Engineering (2005)
20. Jain, A., Nandakumar, K., Ross, A.: Score normalization in multimodal biometric systems. Pattern Recognition (2005) (to appear).
21. Ortega-Garcia, J., et al.: MCYT baseline corpus: A bimodal biometric database. IEE Proc. Vision, Image and Signal Processing **150** (2003) 395–401
22. Martin, A., et al.: The DET curve in assessment of decision task performance. In: Proc. of EuroSpeech. (1997) 1895–1898