Quality Analysis of Dynamic Signature Based on the Sigma-Lognormal Model

Javier Galbally, Julian Fierrez, Marcos Martinez-Diaz
ATVS - Biometric Recognition Group
Universidad Autonoma de Madrid, EPS
C/ Francisco Tomas y Valiente, 11. 28049 Madrid. SPAIN
{*javier.galbally,julian.fierrez,marcos.martinez*}*@uam.es*

Réjean Plamondon
École Polytechnique de Montréal
P.O. Box 62079, Station Centre-Ville
Montréal H3C 3A7. CANADA
rejean.plamondon@polymtl.ca

***Abstract*—An analysis of the quality of on-line handwritten signatures is carried out based on the Sigma-Lognormal model. In the study, two main issues are addressed from a kinematic perspective of humanly-produced movements. On the one hand, what makes some signatures perform better than others in automatic signature verification systems, and on the other hand if that information may be used as a quality measure in order to predict the expected performance of a given sample. Experiments were carried out on the MCYT database and show the high potential of certain kinematic features for signature quality assessment.**

***Keywords*-Signature recognition; Quality assessment; Kinematic Theory of rapid human movements; Sigma-Lognormal model**

I. Introduction

Due to the fact that biometrics, as an automatic means of human recognition, constitutes a relatively novel field of research [1], most efforts undertaken by the different parties involved in the development of this technology (researchers, industry, evaluators, etc.) have been mainly (but not exclusively) directed to the improvement of its performance (i.e., finding ways to obtain lower error rates)[2], [3]. This has left partially uncovered other important aspects involved in the complex biometric recognition problem.

In particular, it has not been until recently when biometric quality assessment has emerged in the biometric community as a primary field of research, as a consequence of the concern arisen after the poor performance observed in different biometric systems on certain pathological samples [4]. Different studies have proved that biometric systems performance is heavily affected by the quality of the input signals, and that even the best systems worldwide struggle in the presence of noisy samples [5], [6].

One of the main reasons that has led to a slow start in biometric quality research is the intrinsic difficulty to clearly define the term *quality* in the biometric context. Recent standardization efforts [7] have established that biometric sample quality can be considered from three different points of view, namely: i) *character*, ii) *fidelity*, and iii) *utility*. This last characteristic refers to the impact of the individual biometric sample on the overall performance of a biometric system and it is generally accepted that it constitutes the most important aspect to be taken into account by a quality metric [6] (i.e., samples assigned to a higher quality should lead to a more accurate identification of individuals).

These standardization efforts have led to the proposal of specific quality measures for certain traits such as the fingerprint (where quality may be computed in terms of the ridge strength, continuity, clarity, uniformity or integrity [8]) or the iris (where quality is measured in terms of the occlusion, the focus, the blurring or the angular deformation [9]). However in behavioral biometric traits such as the signature [10], the proposal of objective and practical metrics for quality estimation is not an easy task. Although some work has been carried out to try to propose indicators of the quality of both on-line [11], [12] and off-line signature [13], [14], there is still no general consensus on how the quality of a signature should be measured.

In the present work, the term quality is considered from the *utility* point of view in order to investigate the cause that makes some signatures more suitable for automatic recognition than others. For this purpose, the Kinematic Theory of rapid human movements [15], [16] and its associated Sigma-Lognormal model are used to analyze the humanly-produced features that differentiate well-performing on-line signatures from those giving higher error rates. Furthermore, the work also studies the possibility to use the intrinsic information that lies behind the production of well-performing samples to propose a set of measures that can help to predict how a certain signature will perform in a given automatic recognition system.

Reported results show that there is a direct connection between the human-based kinematic information present in the single strokes of handwritten signatures and their behaviour in automatic recognition systems, and that this information may be used to estimate the expected performance of a given signature in a biometric application.

The rest of the paper is structured as follows. The Sigma-Lognormal model is reviewed in Sect. II. The database, system used in the tests, and the experimental protocol are presented in Sect. III. Results are given in Sect. IV, while the conclusions are finally drawn in Sect. V.

1520-5363/11 $26.00 © 2011 IEEE
DOI 10.1109/ICDAR.2011.133

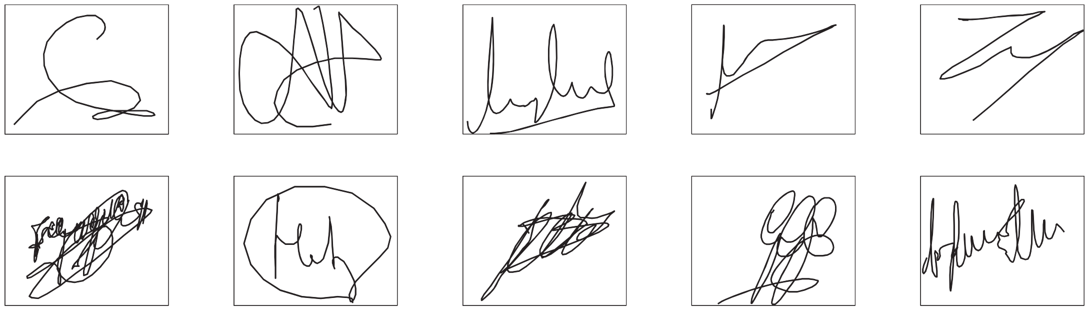

Figure 1. One sample of the best 5 performing users in MCYT (top row), and of the 5 worst performing users (bottom row).

II. The Sigma-Lognormal Model

The Kinematic Theory of rapid human movements, which was first introduced in [15], [16], relies on the Sigma-Lognormal model to represent the information of both the motor commands and the timing properties of the neuromuscular system involved in the production of complex movements like signatures. Being a theory based on the human writing behavior, its application to the analysis of signature quality can bring some insight into the difficult issue of what makes some signatures perform better than others in automatic recognition systems.

The Sigma-Lognormal model considers the resulting speed of a single stroke j as having a lognormal shape Λ scaled by a command parameter (D) and time-shifted by the time occurrence of the command (t_0) [17].

$$|\vec{v}_j(t;P_j)| = D_j\Lambda(t-t_{0j};\mu_j,\sigma_j^2) = \\ = \frac{D_j}{\sigma(t-t_{0j})\sqrt{2\pi}}\exp\{\frac{[\ln(t-t_{0j})-\mu_j]^2}{-2\sigma_j^2}\},$$

where $P_j = [D_j, t_{0j}, \mu_j, \sigma_j, \theta_{sj}, \theta_{ej}]$ represents the set of Sigma-Lognormal parameters:

- D_j: the amplitude of the input commands.
- t_{0j}: the time occurrence of the input commands, a time-shift parameter.
- μ_j: the log-time delays, the time delays of the neuromuscular system expressed on a logarithmic time scale.
- σ_j: the log-response times, which are the response times of the neuromuscular system expressed on a logarithmic time scale.
- θ_{sj}: starting angles of the circular trajectories described by the lognormal model along a pivot.
- θ_{ej}: ending angles of the circular trajectories described by the lognormal model along a pivot.

In this context, a signature can be seen as the output of a generator that produces a set of individual strokes superimposed in time. The resulting complex trajectory can be modeled as a vectorial summation of lognormals (being N_{LN} the total number of lognormal curves in which the signature is decomposed):

$$\vec{v}(t) = \Sigma\Lambda(t) = \sum_{j=1}^{N_{LN}} \vec{v}_j(t;P_j).$$

The reconstruction error of a velocity profile using the Sigma-Lognormal parameters $\vec{v}(t)$ can be evaluated by computing the SNR between the reconstructed specimen and the original one:

$$\mathrm{SNR} = 10\log\Big(\frac{\int_{t_s}^{t_e}[v_{xo}^2(t)+v_{yo}^2(t)]dt}{\int_{t_s}^{t_e}[(v_{xo}(t)-v_{xa}(t))^2+(v_{yo}(t)-v_{ya}(t))^2]dt}\Big), \tag{1}$$

where t_s and t_e are respectively the starting and ending times of the signature, and the subindex o refers to the original velocity profile (x or y) while a corresponds to the reconstructed functions.

This fitness evaluation metric (SNR) will be used in the experiments, together with the number lognormal curves (N_{LN}) and the parameters defining each of those strokes ($P_j = [D_j, t_{0j}, \mu_j, \sigma_j, \theta_{sj}, \theta_{ej}]$), to analyze the quality of well- and bad-performing signatures.

III. Experimental Protocol

In the experiments, the dynamic signature data of the MCYT database is used (the whole multimodal corpus comprises signature and fingerprint information of 330 users) [18]. The signature dataset comprises 25 original samples and 25 skilled forgeries per user (captured in five different acquisition sets). These data are used to estimate the performance under a random forgeries scenario of a state of the art HMM-based signature recognition system (with 12 states and 4 mixtures per state) using as feature set 23 time sequences derived from the coordinate (x and y) and pressure (p) functions of each signature [19], [20]. The performance evaluation is carried out in a realistic working scenario where a reduced number of samples (five) of each user are available to train its model.

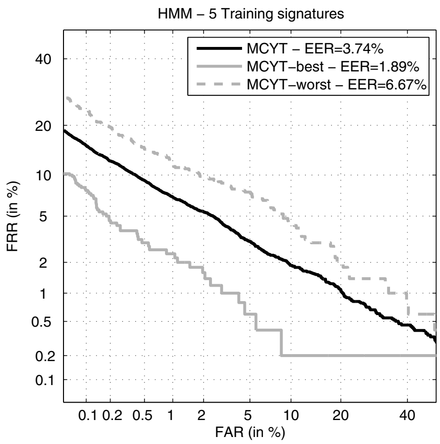

Figure 2. DET curves of the HMM system evaluated on MCYT, and on the best and worst performing 25 users of MCYT (MCYT-best and MCYT-worst, respectively).

The set of genuine scores is computed matching the enrollment data (five samples) with the last 20 original signatures of the user resulting in $330 \times 20 = 6,600$ similarity scores. For the impostor scores each user's model is compared with one signature of the remaining clients (i.e., $330 \times 329 = 108,529$ impostor scores).

Skilled forgeries are discarded in the experiments as the performance of a system under this scenario does not only depend on the quality of the signature (utility), but also on other factors such as the skill of the forger or the intrinsic difficulty of the signature to be copied. These issues are out of the scope of this work and have been addressed elsewhere [21], [22], [12].

IV. Results

The objective of the experiments is twofold, i) on the one hand to look for the humanly-produced kinematic information which differentiates well-performing from bad-performing signatures, and ii) on the other hand to determine if that information may be used to predict a signature's performance on a given automatic recognition system.

In order to reach these two goals, two pools of clients are chosen from the MCYT database comprising the best (MCYT-best) and worst (MCYT-worst) performing users in the database (in our case 25). In order to carry out the selection, both the intraclass variability (reflected by the genuine scores) and the interclass variability (responsible for the impostor scores) of the clients should be taken into account. Thus, the metric used to select the two sets of users is computed as $Q = \overline{S_g}/\overline{S_i}$, where $\overline{S_g}$ and $\overline{S_i}$ are the average of the genuine and impostor scores of each user, respectively. In Fig. 1 one sample of the five best and worst performing users in MCYT is shown (top and bottom row, respectively).

In Fig. 2 we show the Detection Error Trade-off (DET) curves for the system evaluated on the whole MCYT database and on the two sets of 25 best and worst performing users. It can be observed the significant difference in performance between the two sets of clients with the Equal Error Rate (EER) increasing almost by a factor three from MCYT-best to MCYT-worst.

A. Experiment 1: Feature Analysis

In this first experiment we analyze the different kinematic information comprised in well and bad performing signatures. For this purpose, the Sigma-Lognormal 8-feature set defined in Sect. II, FS $= [D, t_0, \mu, \sigma, \theta_s, \theta_e, N_{LN}, \text{SNR}]$, is extracted from each signature in MCYT-best and MCYT-worst taking into account that the first six features are stroke-based, while the last two (N_{LN} and SNR) are related to the whole signature. The individual distributions for each of the parameters in the best (solid) and worst (dashed) sets of users are shown in Fig. 3.

Several observations may be extracted from the results shown in Fig. 3:

- Most of the kinematic information present in well- and bad- performing signatures is almost the same (see distributions for D, μ, σ, θ_s, and θ_e).
- The starting point of the strokes of signatures with a good performance are in general closer to the beginning of the signing process than those belonging to signatures with a worse performance (see distribution for t_0). This characteristic (smaller t_0) is typical of shorter signatures and of signatures composed of better learned movements (the commands for the start of all the strokes are given very close to the start of the signature).
- The distribution of the number of lognormals clearly shows that the fewer the number of strokes the better the performance of the signature, which is consistent with having shorter samples as expressed in the previous observation, as well as a better fine motor control, with no shaking.
- Finally, we may say that signatures that perform well are better represented by the Sigma-Lognormal model (their SNR is higher) than those with worse error rates.

B. Experiment 2: Performance Prediction

In this experiment we use the information extracted from the previous test to predict the expected performance (good or bad) of a given signature. With this objective the samples in MCYT-best and MCYT-worst are parameterized according to the three features that shown some discriminant power between well- and bad-performing signatures in experiment 1 (i.e., t_0, N_{LN}, and SNR). The 25 samples corresponding to the first 10 users from each set are then used to train a Gaussian Mixture Model (GMM) of four mixtures, which is later employed to classify the signatures of the remaining 15 users from each set. In order to avoid biased results, two

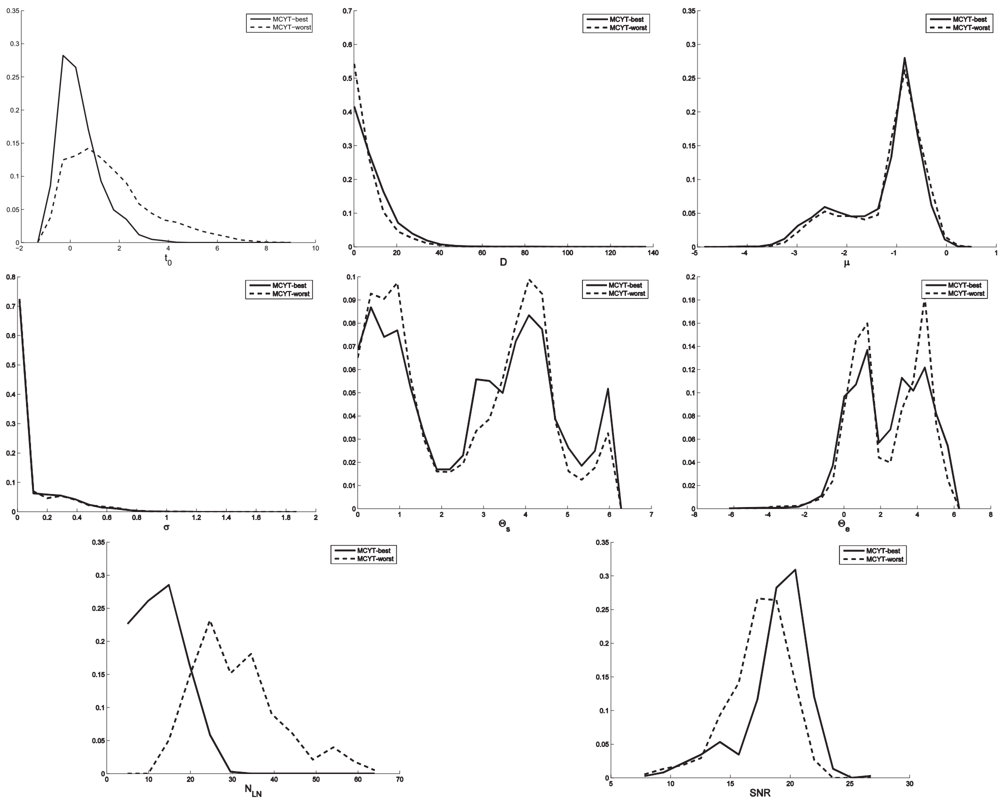

Figure 3. Distributions of the Sigma LogNormal features for the best (solid) and worst (dashed) performing users in MCYT.

Table I
CLASSIFICATION ERROR RATES FOR WELL AND BAD PERFORMING SIGNATURES OF THE FEATURE SUBSET $[N_{LN}, \text{SNR}]$. FWR STANDS FOR FALSE WORST RATE, FBR FOR FALSE BEST RATE AND ACE FOR AVERAGE CLASSIFICATION ERROR.

Classification Error Rates (%)		
FWR	FBR	ACE
6.5	18.3	**12.4**

fold cross-validation is carried out exchanging training and test sets in two successive steps of the classification process.

It is possible that the best classifying results are not obtained using the set of three proposed features ($t0$, N_{LN}, and SNR), but a subset of them. As we are dealing with a three dimensional problem there are just seven possible feature subsets, which allows to apply exhaustive search as feature selection technique. This process showed that the optimal subset for classification was the one formed by the N_{LN} and SNR parameters, while t_0 was discarded. The classification results obtained with this scheme are shown in Table I where FWR stands for False Worst Rate (a signature of the best performing set is classified as not suitable for recognition), FBR stands for False Best Rate (a signature with a low performance is assigned to MCYT-best), and ACE is the Average Classification Error (average of FWR and FBR).

The results given in Table I show the feasibility of using the proposed features as possible quality measures in order to predict the expected performance of a certain signature in a given automatic recognition system, and their high discriminant potential which may be improved in combination with other complimentary quality metrics such as the complexity [12] or the stability [11] of the signature.

V. CONCLUSIONS

A study of the quality of on-line handwritten signatures has been carried out based on the Kinematic Theory of rapid human movements and its associated Sigma-Lognormal model. Two main issues have been addressed in the work: i) what humanly-produced dynamic information differentiates well-performing from bad-performing signatures, and ii) if that information can be used to predict the performance of a signature on a given automatic recognition system.

The experimental results, carried out on the MCYT database comprising over 16,000 signatures, have shown that shorter signatures with better learned movements and being very accurately modeled by the Sigma-Lognormal features are more suited for personal recognition (i.e., they present lower error rates). Furthermore, two of the Sigma-Lognormal parameters have proven a very high potential for the *a priori* estimation of signature performance and have been proposed as possible quality indicators of its *utility* (from a recognition error rate perspective).

This type of quality assessment study has numerous applications in the context of biometric systems [6], for instance: i) quality algorithms may be used as a monitoring tool [23]; ii) quality of enrolment templates and/or samples acquired during an access transaction can be controlled for acquiring-until-satisfaction purposes (recapture); or iii) some of the steps of the recognition system can be adjusted based on the estimated quality (quality-based adaptation [24], [25]).

ACKNOWLEDGMENT

This work has been partially supported by projects Contexts (S2009/TIC-1485) from CAM, Bio-Challenge (TEC2009-11186) from Spanish MICINN, *Dirección General de la Guardia Civil*, and *Cátedra UAM-Telefónica*.

J. G. would also like to thank R. P. for hosting him at the École Polytechnique de Montréal during the development of part of this research work.

REFERENCES

[1] A. K. Jain, A. Ross, and S. Pankanti, "Biometrics: a tool for information security," *IEEE Trans. on Information Forensics and Security*, vol. 1, no. 2, pp. 125–143, 2006.

[2] A. Mansfield and J. Wayman, "Best practices in testing and reporting performance of biometric devices," CESG Biometrics Working Group, Tech. Rep., 2002.

[3] R. Cappelli, D. Maio *et al.*, "Performance evaluation of fingerprint verification systems," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 3–18, 2006.

[4] NIST, "Biometric quality workshop," 2011, (http://www.nist.gov/itl/iad/ig/bio-quality.cfm).

[5] D. Maio, D. Maltoni *et al.*, "FVC2004: Third Fingerprint Verification Competition," in *Proc. of ICBA*, D. Zhang and A. K. Jain, Eds. Springer LNCS-3072, 2004, pp. 1–7.

[6] P. Grother and E. Tabassi, "Performance of biometric quality measures," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, pp. 531–543, 2007.

[7] ISO/IEC 29794-1, "ISO/IEC 29794-1:2006 biometric quality framework standard," 2006.

[8] F. Alonso-Fernandez, J. Fierrez *et al.*, "A comparative study of fingerprint image quality estimation methods," *IEEE Trans. on Information Forensics and Security*, vol. 2, no. 4, pp. 734–743, 2008.

[9] A. Abhyankar and S. Schuckers, "Iris quality assessment and bi-orthogonal wavelet based encoding for recognition," *Pattern Recognition*, vol. 42, pp. 1878–1894, 2009.

[10] R. Plamondon and S. Srihari, "On-line and off-line handwriting recognition: a comprehensive survey," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 63–84, 2000.

[11] S. Muller and O. Henniger, "Evaluating the biometric sample quality of handwritten signatures," in *Proc. of ICB*. Springer LNCS-4642, 2007, pp. 407–414.

[12] J. Brault and R. Plamondon, "A complexity measure of handwritten curves: modeling of dynamic signature forgery," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 23, pp. 400–413, 1993.

[13] F. Alonso-Fernandez, M. Fairhurst *et al.*, "Impact of signature legibility and signature type in off-line signature verification." in *Proc. of BSym*, 2007, pp. 1–6.

[14] ——, "Automatic measures for predicting performance in off-line signature," in *Proc. of ICIP*, 2007, pp. 369–372.

[15] R. Plamondon, "A kinematic theory of rapid human movements. Part I. Movement representation and generation," *Biological Cybernetics*, vol. 72, pp. 295–307, 1995.

[16] ——, "A kinematic theory of rapid human movements. Part II. Movement time and control," *Biological Cybernetics*, vol. 72, pp. pp. 309–320, 1995.

[17] C. O'Reilly and R. Plamondon, "Development of a Sigma-Lognormal representation for on-line signatures," *Pattern Recognition*, vol. 42, pp. 3324–3327, 2009.

[18] J. Ortega-Garcia, J. Fierrez *et al.*, "MCYT baseline corpus: a bimodal biometric database," *IEE Proc. Vision, Image and Signal Processing*, vol. 150, no. 6, pp. 391–401, 2003.

[19] J. Fierrez, J. Ortega-Garcia *et al.*, "HMM-based on-line signature verification: feature extraction and signature modeling," *Pattern Recognition Letters*, vol. 8, pp. 2325–2334, 2007.

[20] M. Martinez-Diaz, J. Fierrez *et al.*, "Towards mobile authentication using dynamic signature verification: useful features and performance evaluation," in *Proc. of ICPR*, 2008.

[21] F. Alonso-Fernandez, J. Fierrez *et al.*, "Robustness of signature verification systems to imitators with increasing skills," in *Proc. of ICDAR*, 2009.

[22] J. Hennebert, R. Loeffel *et al.*, "A new forgery scenario based on regaining dynamics of signature," in *Proc. of ICB*. Springer LNCS-4642, 2007, pp. 366–375.

[23] T. Ko and R. Krishnan, "Monitoring and reporting of fingerprint image quality and match accuracy for a large user application," in *Proc. of AIPR*, 2004, pp. 159–164.

[24] J. Fierrez-Aguilar, J. Ortega-Garcia *et al.*, "Discriminative multimodal biometric authentication based on quality measures," *Pattern Recognition*, vol. 38, no. 5, pp. 777–779, 2005.

[25] F. Alonso-Fernandez, J. Fierez *et al.*, "Quality-based conditional processing in multi-biometrics: application to sensor interoperability," *IEEE Trans. on Systems, Man and Cybernetics. Part A.*, vol. 40, pp. 1168–1179, 2010.