2013 International Conference on Biometrics (ICB). IEEE, 2013, 1-6

DOI: http://dx.doi.org/10.1109/ICB.2013.6612974

Variations of Handwritten Signatures with Time: A Sigma-Lognormal Analysis

Marta Gomez-Barrero, Javier Galbally, Julian Fierrez, Javier Ortega-Garcia
Biometric Recognition Group - ATVS, EPS, Universidad Autonoma de Madrid
C/ Francisco Tomas y Valiente 11, 28049 Madrid, Spain
{marta.barrero,javier.galbally,julian.fierrez,javier.ortega}@uam.es

Réjean Plamondon
Laboratoire Scribens, Département de Génie Électrique, École Polytechnique de Montreal
C.P. 6079, Succursale Centre-Ville, Montréal, QC, Canada H3C 3A7
rejean.plamondon@polymtl.ca

Abstract

The variation of dynamic signatures with time is analysed for the first time using the Kinematic Theory, following a general, consistent and fully reproducible protocol. Experiments are carried out on a new long-term database captured in 6 sessions uniformly distributed over a 15 month time span, under almost identical conditions. Signatures are represented with the Sigma Log-Normal model, which takes into account the effects of body ageing closely related to handwriting, such as neuromuscular response times. After studying the evolution of signatures with time, an analysis on age groups based on the model parameters is carried out.

1. Introduction

As any novel field of research, most efforts in the biometric community (which includes researchers, vendors, evaluators, etc.) have been focused on improving the performance of the recognition systems, so that lower error rates are achieved. As a consequence, other important related aspects such as the performance degradation effect known as *ageing* have been frequently overlooked. In order to make biometric recognition systems more reliable and advance in the development of this rapidly emerging technology, some efforts need to be directed towards this topic.

The term *ageing* is generally used to refer to the gradual decrease in a system performance caused by the changes suffered by the users' trait in the long-term (which cannot be avoided as is inherent to human nature) [8]. This accounts for part of the intra-user variability (i.e., variability within the samples of the same user): a subject's signature may considerably differ from the enrolled one after a sufficiently long period. This results in lower similarity scores and as a consequence an increase in the system error rates.

The present work addresses the problem of the variation of signatures with time from a novel perspective and studies the potential of this new approach for the study of signatures belonging to different age groups.

The first task of the design of a biometric recognition system is choosing features that will be used to model the biometric trait, in the present case, signatures. Most signature parametrizations are based on either global features, such as average speed, total duration or number of pen-ups [9, 11], or time signals in general derived from the coordinate x y or pressure functions given by most current on-line acquisition devices [5, 7, 10]. Another recent and novel approach is based on the Kinematic Theory of rapid human movements: on the one hand, single handwriting strokes can be characterized with a Delta-Lognormal model [1]; on the other hand, on-line signatures can be represented as a sequence of strokes (i.e., a summation of log-normal curves) [12]. The main advantage of this model is that it takes into account physical body features such as the neuromuscular system responsible for the production of a signature, as it has been proved on previous works linking this theory to neuromuscular disorders diagnosis [13] or stroke risk factors prediction [14]. Since *ageing* in signatures ultimately comes from *ageing* of the neuromuscular system, this parametrization is specially suited for the problems addressed in the present study.

Some related works on signature variability and the study of *ageing* in handwritten signatures have been already published [3, 6]. Furthermore, the Kinematic Theory of rapid human movements has already been used to study different handwritten tasks. In [2], the variability observed in handwriting patterns for a single word with a fixed number of strokes is analysed based on the Sigma-Lognormal model. This model integrates in its parameters some control motor knowledge. Since *ageing* affects the neuromuscular system of individuals, it is reasonable to assure that its effect may produce changes on the signatures Log-Normal parameters. Furthermore, the problem posed by *ageing* was studied for single strokes characterized by the Delta-Lognormal Model in [16]. Taking into account these works and combining the methodologies followed in each of them, in the present article we carry out the first study on how the *ageing* of the neuromuscular system affects signatures (multiple stroke tasks).

In addition, using this modelling approach, the differences between two age groups are also analysed using the Sigma Log-Normal parameters. After studying the evolution of signatures across time, Sigma Log-Normal parameters between two different age groups (i.e., elders, youngsters) are compared, in order to explore the potential of these features as a way to estimate the user's age.

All the experiments are carried out on a new dynamic signature database: the Signature 15M-Term dataset. It comprises signatures of the 29 common users of two multimodal and publicly available databases: the BiosecurID DB [4] and the Biosecure DB [15], which were captured in 6 uniformly distributed acquisition sessions over 15 months.

The rest of the article is structured as follows. In Sect. 2, a brief introduction to the Sigma Log-Normal model for online signature representation is presented. Then, the experimental protocol followed in the experiments and the database used are presented in Sect. 3. Results are given in Sect. 4 and final conclusions are drawn in Sect. 5.

2. Overview of the Sigma-Lognormal Model

The Sigma-Lognormal model was first applied to online signatures in [12]. As a high level representation of the models supported by the Kinematic Theory, it considers single strokes as primitives from which complex patterns are built. Each primitive has a lognormal velocity profile and a complex pattern is produced by summing up strokes, which results in:

$$\vec{v}(t) = \sum_{i=1}^{N} \vec{v}_i(t) \quad \text{with } L \geq 2$$

where N represents the number of strokes involved in the generation of a given pattern and $\vec{v}_i(t)$ is the velocity profile of the i-th stroke.

Each curved stroke is modelled by a Sigma-Lognormal, which reflects both the motor control process and the neuromuscular response. It is represented by a feature vector representing the parameters of the model:

$$P_i = (t_{0i}, D_i, \theta_{di}, \theta_{fi}, \mu_i, \sigma_i)$$

where t_{0i} is the starting time of the stroke, D_i its length, θ_{id} the starting direction angle, θ_{fi} the ending direction angle, μ_i the logtime delay and σ_i the logresponse time. These last two parameters characterize the lognormal impulse response of the neuromuscular system.

The Sigma Log-Normal model establishes the theoretical *ideal* representation of the signature. The variations between this ideal model ($\vec{v}(t)$ and the actual behaviour of a signature ($\vec{v_r}(t)$) may be measured in terms of the SNR, defined over the velocity signals as

$$SNR = 20 \log \left(\frac{\int_{t_s}^{t_e} \left[v_x^2(t) - v_y^2(t) \right] \mathrm{d}t}{\int_{t_s}^{t_e} [(v_x(t) - v_{x\_r}(t))^2 + (v_y(t) - v_{y\_r}(t))^2] \mathrm{d}t} \right)$$

A low SNR denotes problems in the motor control system and can have a potential use in the early detection of certain decreases related to ageing [16].

3. Database and Experimental Protocol

3.1. The On-Line Signature 15M-Term Database

The dataset used in the experiments comprises the online signature data of the 29 users shared by the BiosecurID and the Biosecure databases, which were acquired in a 15 month time span. Both databases are fully compatible in terms of the acquisition scenario, protocol and capturing device used (Wacom Intuos 3 pen tablet). This way, we may discard acquisition-related external factors as the cause of possible changes in a user's signature.

- The BiosecurID Signature Subset [4]. It comprises 16 original signatures and 12 skilled forgeries per user, captured in 4 separate acquisition sessions (named here BID1, BID2, BID3 and BID4). The sessions were captured leaving a two month interval between them, in a controlled and supervised office-like scenario.

- The Biosecure Signature Subset [15]. This dataset was captured 6 months after the BiosecurID acquisition campaign had finished. It comprises 30 original signatures per user, and 20 skilled forgeries, distributed in two acquisition sessions separated three months (named here Bure1 and Bure2). The 15 original samples corresponding to each session were captured in three groups of 5 consecutive signatures with an interval of around 15 minutes between groups.

For the final dataset used in the present work only the original signatures were considered, this way it comprises 1,334 signatures coming from the 29 common users of the two databases with 46 samples per user (16 from BiosecurID, and the remaining 30 from Biosecure) which are distributed in 6 sessions (BID1-2-3-4 and Bure1-2). It constitutes the first signature dataset where we can track over a 15 month time span the signature of a given user (as there are 6 almost uniformly distributed acquisition sessions in this interval). Given the limited number of users and the short acquisition time span of the database, the results here presented show a preliminary general trend. However, the Longterm database is the most adequate to the task at hand to the best of our knowledge.

3.2. Signature Evolution with Time: Protocol

In order to analyse the evolution of the signature parameters across time, several steps are taken.

Feature extraction. The first task of the experiments is computing the Sigma Log-Normal model parameters of each signatures of the database.

Normalization. As stated in [2], the parameters that describe each stroke can be broadly divided into two categories: control $(t_{0i}, D_i, \theta_{di}, \theta_{fi})$, related to the motor control system, and peripheral (μ_i, σ_i), related to the impulse response of the neuromuscular system reacting to the commands generated by the controller. Both sets of parameters might be affected in different ways by the ageing of the human body, and therefore require a different treatment.

Since each signature of each individual has a varying number of strokes, N, a given signature will be represented by a $N \times 6$ matrix. Therefore, an additional parameter will be studied: the number of strokes, N.

In order to study the variability of each of the 6 parameters of the Sigma Log-Normal model, we normalise the parameters according to the transformations proposed in [2]:

$$
\begin{aligned}
D_i &\rightarrow \frac{D_i}{D_{max}} \\
\theta_{di} &\rightarrow \theta_{di} - \theta_{d1} \\
\theta_{fi} &\rightarrow \theta_{fi} - \theta_{d1} \\
t_{0i} &\rightarrow t_{0i} - t_{01} \\
\mu_i &\rightarrow \bar{\mu} = \sum_{i=1}^{N_{jk}} \mu_i \\
\sigma_i &\rightarrow \bar{\sigma} = \sum_{i=1}^{N_{jk}} \sigma_i
\end{aligned}
$$

where $D_{max} = \max\{D_i\}$ and $i = 1, \ldots, N_{jk}$. This way, control parameters result in sequences normalised by the initial or maximum value of the original sequence, so that different scales or orientations of signatures do not affect the parameters. On the other hand, peripheral parameters, related to neuromuscular responses, are reduced to their mean values: as we want to analyse the variations of the neuromuscular responses in the whole signature, not in each single stroke.

For the experimental analysis we will consider Δt_{0i} instead of t_{0i}, defined as the difference between the starting time of two consecutive strokes, i.e., $\Delta t_{0i} = t_{0i+1} - t_{0i}$. This derived feature is more consistent for each signer than t_{0i} and gives very valuable information about how much time in advance the signer plans the execution of each of the strokes.

Statistical measures. In order to show the evolution of the signatures with time, we will analyse statistical differences between acquisition sessions. These variations will be measured in terms of their medians and 25th and 75th percentiles, so that two different questions may be answered: i) do the parameters (their medians) change with time?, and ii) does the range of variability of each parameter (its percentiles) remain constant, or, on the contrary, increase or decrease?

In order to compute those statistical differences, we take the average of each parameter of a given subject's signatures belonging to one session. Then, the differences across sessions are computed subtracting the averaged values. That is, when computing the difference between sessions j and k, we average all the signatures belonging to session j, leading to a vector of the averaged parameters P_j, then average the signatures belonging to session k, leading to a second vector P_k, and finally compute $P_{jk} = P_j - P_k$.

3.3. Age Group Analysis: Protocol

In this case, the goal of the experiments is to determine if there exists a correlation between the value of the Sigma Log-Normal parameters and the age of the signer. For this purpose, results based on the same parametrized features as well as the Signal to Noise Ratio (SNR) are represented by their distributions, so that differences between the two age groups compared are more clearly shown.

4. Experiments and Results

So far, only one signature-related study focused on *ageing* has been carried out in the literature [16]. However, this very valuable work was performed on a single stroke task and the analysis was based on the Delta-Lognormal model. The study compared single strokes from two groups: young (mean:27.5 years old) and aged (mean: 66.9 years old). It was shown that for the elder group t_0 were longer, D were smaller, μ and σ_1 were larger, and σ_2 smaller. This was observed for a 40 years age difference under a very controlled experiment. In other words, previous results indicate that, as we get older, we have a tendency to protect ourselves by planning slower and smaller movements, minimizing abrupt changes.

In the present work, we try to extend those results to handwritten signatures; that is, multiple stroke tasks modelled with Sigma Log-Normals as presented in Sects. 3.2 and 3.3. The experiments have two main objectives: i) analyse the effects of time on the Sigma Log-Normal parameters of the handwritten signature, and ii) study the differences between signatures of different age groups in terms of their Sigma Log-Normal parametrizations.

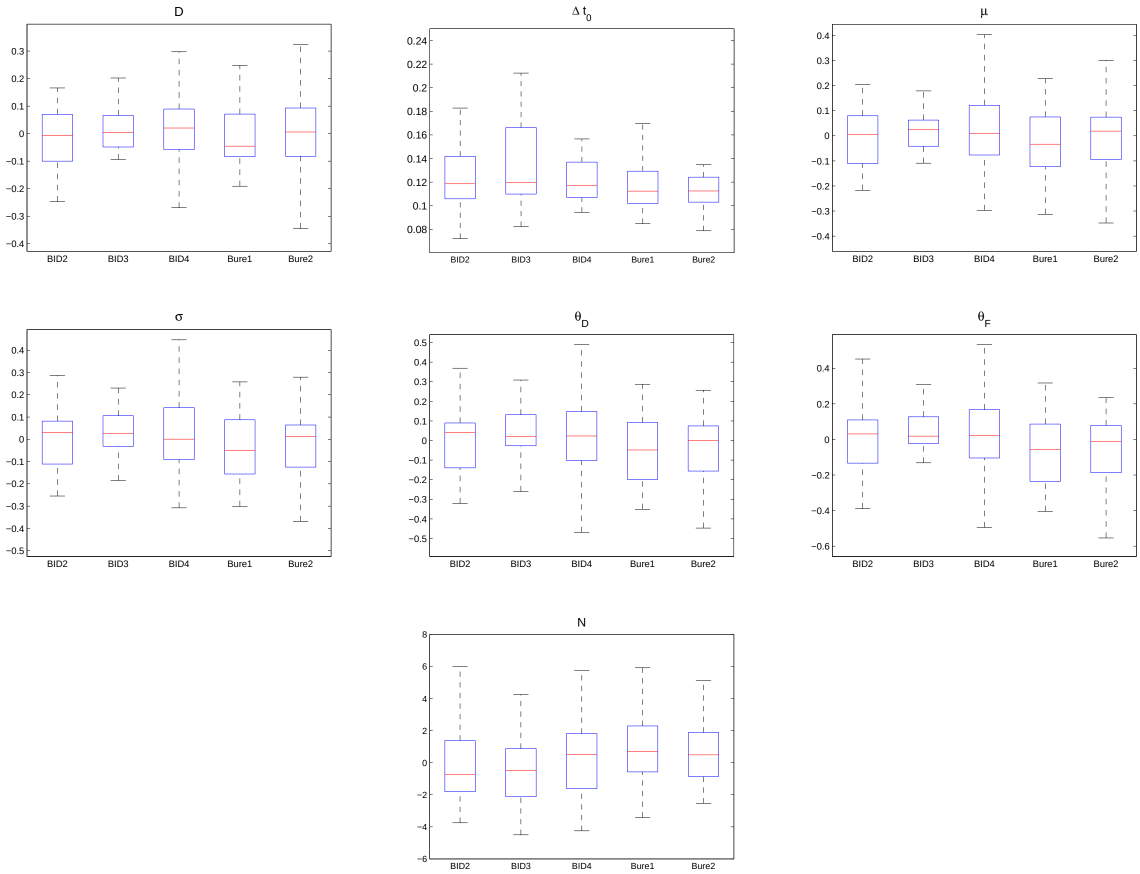

Figure 1. Boxplot of the differences between session 1 (BID1) and sessions 2 (BID2) to 6 (Bure2) for each parameter (D, t_0, μ, σ, θ_d, θ_f) and for the number of strokes (N).

4.1. Signature Evolution with Time

In the first set of experiments, we use the On-Line Signature 15M-Term database. As explained in Sect. 3.1, it comprises 29 users with 6 sessions per user. Therefore, we will have five different sets of results for each parameter, each of them with 29 values per parameter:

1. BID2 - BID1: time difference of two months.
2. BID3 - BID1: time difference of four months.
3. BID4 - BID1: time difference of six months.
4. Bure1 - BID1: time difference of twelve months.
5. Bure2 - BID1: time difference of fifteen months.

In Fig. 1, boxplots of the differences between sessions are depicted. On each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles and the whiskers extend to the most extreme data points considered.

As it can be observed, the number of strokes N shows an upward trend and the difference between starting times Δt_0 a downward trend, as we should expect: as we get older, we tend to plan smaller movements (here, shorter strokes). However, we try to preserve the overall signature, therefore needing a higher number of movements or strokes.

On the other hand, on several parameters, like μ, values tend to increase within each of the databases: BID and Bure. That variability may be therefore attributed more to boredness or loss of interest from the user than to ageing: in the first sessions, subjects show more interest on the task at hand (signing several times) than on the last ones.

4.2. Age Group Analysis

Our aim now is to compare the signatures of subjects of considerably different ages and to study the potential use of the Sigma Log-Normal parameters to estimate the signers' age. For this purpose, in the BioSecure database [15], we isolate two different groups of similar sizes out of its 210 users, with a significant age difference between them:

- Subjects between 18 and 21 years: 28 subjects. Mean: 20.0 years old.

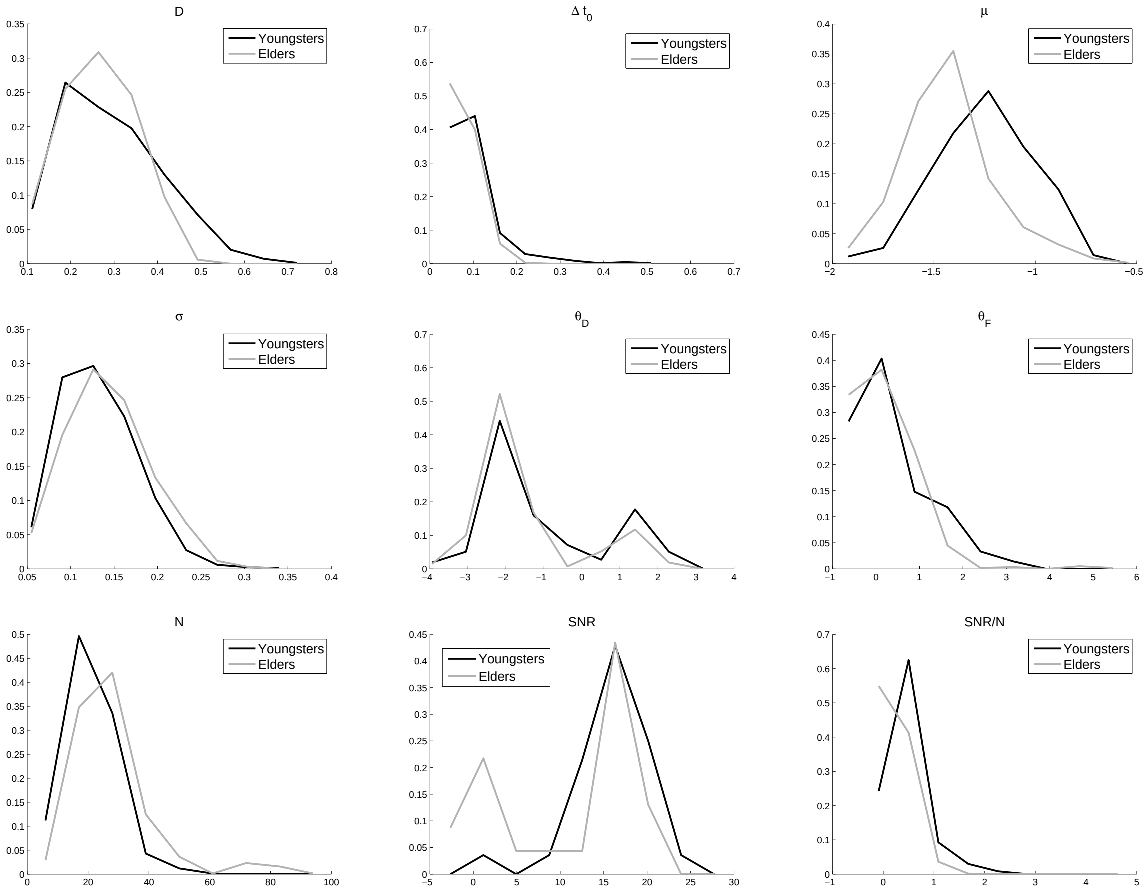

Figure 2. Distributions of the age groups (young and elder) for each parameter (D, t_0, μ, σ, θ_d, θ_f), for the number of strokes (N) and for the Signal to Noise Ratio (SNR).

- Subjects between 59 and 72: 23 subjects. Mean: 63.7 years old.

Now, we have only two sets to compare, with $28 \times 30 = 840$ observations per Sigma Log-Normal parameter for the younger group and $23 \times 30 = 690$ for the elder group.

In Fig. 2, we can see the distributions of each Sigma Log-Normal parameter (D, Δt_0, μ, σ, θ_d, θ_f) as well as the number of strokes (N), the Signal to Noise Ratio (SNR) and the ratio between the SNR and the number of strokes (SNR$/N$) for both age groups: youngsters in black and elders in light grey. Several observations can be made:

- As expected, θ_d and θ_f distributions do not differ greatly: they represent the starting and ending direction angles of each stroke. Therefore, if the signature's shape remains more or less the same, so will these angles.

- Similarly to the results in [16] with the Delta Log-Normal model, longer strokes (that is, larger values of D) seem to be less frequent for the elder group.

- Reinforcing the results on Sect. 4.1, the number of strokes N is in general bigger for the elder group: the mean distribution is higher and low values of N are less frequent for the aged subjects. The starting time difference between strokes Δt_0 distribution shows as well that small values appear to be more frequent for the elder group, resulting the combination of trends in N and Δt_0 in signatures with more and shorter strokes as we get older.

- The μ distributions show a displacement of the peak towards the left for the elder group. This may be due to a tendency noted above: as we get older, we tend to use a higher number of strokes for the signatures. Those strokes are therefore shorter (see D figure, with a bigger peak for small values and a lower tail for the bigger values), requiring a lower logtime delay (μ) but a similar logresponse time (see σ figure).

- The SNR distribution has a significantly bigger peak on lower values (around 0 dBs) for the aged group. This means that, in general, elder subjects tend to present a smaller SNR: the degradation associated with ageing makes us move away from lognormality (i.e., ideal case).

- The ratio SNR$/N$ accentuates the trend shown by SNR and N: age results in more trembling, which means more small lognormals (higher N), and a bigger deviation from lognormality, which means a smaller SNR. Globally, that should lead to a higher probability for small values in elders for the ratio. As expected, in Fig. 2 we can observe that the distribution of SNR$/N$ has a peak in -0.1 for the elder group and a peak in 0.5 for the younger group.

5. Conclusions

The Kinematic Theory and its associated Sigma Log-Normal model provides a solid framework for the study of the production of rapid human movements that takes into account different psychophysical features closely related to the ageing in the human body, such as the neuromuscular response time.

In this context, in the present work we have carried out the first study on the handwritten signature variation with time based on the Sigma Log-Normal model. The experiments, carried out on a medium size database captured on 6 uniformly distributed sessions over 15 months, have shown certain tendencies such as:

- The number of strokes, N, has an upward trend: as we get older, we tend to plan more movements.
- In order to preserve the shape of the signature, those strokes must be shorter, leading to smaller values in Δt_0.

Furthermore, the analysis carried out on the signatures of two groups of subjects with a 40 year age difference between them has confirmed certain observations made on a previous similar study carried out on a controlled single stroke task [16]:

- Strokes tend to be shorter for the elder group, and there are more strokes per signature.
- As we get older, the ratio SNR/N presented decreases: we move away from lognormality (i.e., the ideal case) and the number of lognormals per signatures shows an upward trend.

Given the limited amount of data available, these results are only preliminary: further experiments should be conducted where signatures are tracked over longer periods of time for a higher number of users. However, the protocol here presented may be extended to larger databases and applied to different tasks: single strokes, handwriting, Arabic or Chinese signatures, etc. We thus believe that studies such as the one presented here can help to bring some insight into the difficult problem of biometric ageing in order to put this rapidly emerging technology into practical use.

6. Acknowledgements

This work has been partially supported by projects Contexts (S2009/TIC-1485) from CAM, Bio-Challenge (TEC2009-11186) and Bio-Shield (TEC2012-34881) from Spanish MINECO, Guardia Civil, *Cátedra UAM-Telefónica* and grant RGPIN-915 from NSERC Canada.

References

[1] M. Djioua and R. Plamondon. A new algorithm and system for characterization of handwriting strokes with delta-lognormal parameters. *IEEE TPAMI*, 31:2060–2072, 2009.

[2] M. Djioua and R. Plamondon. Studying the variability of handwriting patterns using the kinematic theory. *HMS*, 28:588–601, 2009.

[3] M. Erbilek and M. Fairhurst. Framework for managing ageing effects in signature biometrics. *IET Biometrics*, 1(2):136–147, 2012.

[4] J. Fierrez, J. Galbally, et al. BiosecurID: a multimodal biometric database. *PAA*, 13:235–246, 2009.

[5] J. Fierrez, L. Nanni, et al. An on-line signature verification system based on fusion of local and global information. In *Proc. IAPR Intl. Conf. AVBPA*, pages 523–532, 2005.

[6] R. Guest. Age dependency in handwritten dynamic signature verification systems. *PRL*, 27:1098–1104, 2006.

[7] A. Jain, F. Griess, and S. Connell. On-line signature verification. *PR*, 35:2963–2972, 2002.

[8] A. Lanitis. A survey of the effects of ageing on biometric identity verification. *IJB*, 2:34–52, 2010.

[9] L. Lee, T. Berger, and E. Aviczer. Reliable on-line human signature verification systems. *IEEE TPAMI*, 18:643–647, 1996.

[10] M. Martinez-Diaz, J. Fierrez, J. Galbally, and J. Ortega-Garcia. Towards mobile authentication using dynamic signature verification: useful features and performance evaluation. In *Proc. ICPR*, 2008.

[11] W. Nelson, W. Turin, and T. Hastie. Statistical methods for on-line signature verification. *IJPRAI*, 8:749–770, 1994.

[12] C. O'Reilly and R. Plamondon. Development of a sigma-lognormal representation for on-line signatures. *PR*, 42:3324–3337, 2009.

[13] C. OReilly and R. Plamondon. Impact of the principal stroke risk factors on human movements. *HMS*, 30:792806, 2011.

[14] C. O'Reilly and R. Plamondon. Design of a neuromuscular disorders diagnostic system using human movement analysis. In *Proc. ISSPA*, pages 787–792, 2012.

[15] J. Ortega-Garcia, J.Fierrez, et al. The multi-scenario multi-environment BioSecure multimodal database (BMDB). *IEEE TPAMI*, 32:1097–1111, 2010.

[16] A. Woch, R. Plamondon, and C. O'Reilly. Kinematic characteristics of bidirectionl delta-lognormal primitives in young and older subjects. *HMS*, 30:1–17, 2011.