Is it possible to automatically identify who has forged my signature? Approaching to the identification of a static signature forger

Miguel A. Ferrer, Aythami Morales
Universidad de Las Palmas de G.C.
Las Palmas de Gran Canaria Spain
mferrer@idetic.eu, amorales@idetic.eu

J.Francisco Vargas
GEPAR, Universidad de Antioquia
Medellín, Colombia
jfvargas@udea.edu.co

Ivan Lemos, Mónica Quintero
LABICI-CTI
Medellín, Colombia
ctilabmed@fiscalia.gov.co

***Abstract*—The automatic handwritten signature verification is an open problem for the scientific community. The most of the published studies examine a generic document trying to locate where the signature has been written, to segment the signature removing complex backgrounds containing lines and letter and to determining whether the signature was made by the owner. However, there are no studies to determine automatically the author of a fake. This paper presents a first approach to the identification of a static signature forger. The underlying hypothesis is the fact that a forger finds difficult to fight against their own free natural way of writing, leading to the second hypothesis that under several conditions it is possible to isolate these features to determine a fake within a population of known forgers. The experiments shown that gray level based features are a good start point to detect who has written the signatures.**

Keywords- Signature identification, forger, pattern recognition, biometrics.

I. Introduction

Handwritten signatures occupy a very special place in the wide set of biometric traits. This is mainly due to the fact that handwritten signatures have long been established as the most widespread means of personal verification. Signatures are generally recognized as a legal means of verifying an individual's identity by government and financial institutions. Moreover, verification by signature analysis requires no invasive measurements and people are familiar with the use of signatures in their daily life.

Unfortunately, a handwritten signature is the result of a complex process depending on the psychophysical state of the signer and the conditions under which the signature process occurs. Therefore, although complex theories have been proposed to model the psychophysical mechanisms underlying handwriting and the ink-depository processes, signature verification still remains an open challenge since a signature is judged to be genuine or a forgery only on the basis of a few reference specimens.

There are three main phases of automatic signature verification: 1) data acquisition and preprocessing, 2) feature extraction and 3) classification. During enrolment phase, the input signatures are processed and their personal features are extracted and stored into the knowledge base. During the classification phase, personal features extracted from an input signature are compared [1]

Leaving apart the problems of locating the signature in a document or segmenting the signature from the document by removing the background, the usual problem is to detect a forgery which use to be classified as: 1) a random forgery which usually is a genuine signature sample belonging to a different writer, 2) simple forgery which occurs when the forger reproduces the signature in his own style and finally 3) the simulated forgery which is a reasonable imitation of the genuine signature model [2]. In this case the verifiers used to be trained with genuine signature and discriminate between other genuine and forgeries. Therefore, the verifier encloses the variability of the training signatures labeling as forgery those signatures that exceed the intra writer variability.

Recently, a forensic signature verification competition called 4NSigComp2010 proposed the detection of simulated and disguised signatures aimed to compare between Forensic Handwritten Experts opinions on authorship of signatures and the systems performances to detect skilled forgeries (simulated and disguised signatures) from genuine signatures of a reference writer [3]. In this case, as the writer change its own way of writing producing a self-forgery, the verifier instead of detecting differences, it should look for similarities between the writer training sequence and the questioned signature. If the similarities are greater than a threshold, the disguise is assigned to the writer.

In this paper we propose an approach to the next step which is to look for who has forged my signature among a given set of forgers. The problem is similar to the disguise detection: the verifier should look for similarities between two signatures. But the disguise detection looks for similarities between two similar signatures (genuine and imitation or self-forged) while to detect the forger looks for similarities between two different signatures (genuine signature of the signer and the signature of other user forged by the same signer).

The outline of this paper is as follows: next section provides information about the imitation procedure, section 3

978-0-7695-4661-2/12 $26.00 © 2012 IEEE
DOI 10.1109/DAS.2012.47

discuss about when is it possible to detect the forged followed by the experiment and results at the section 4. The paper is closed with the conclusion at the section 5.

II. IMITATION PROCEDURE

An imitation is a reproduction by hand of the external features of a pattern. In this modality the aim is to impersonate another person. Spoofing is accomplished through a mimetic work by shadowing the apparent characteristics of the pattern or signature. The forger tries to mimic the movements that he believes led to forms of the model. We can distinguish two groups of imitations [11]:

A. *Slow or servile imitation: specific signs.*

It is a slow and careful imitation of the pattern in sight. In this mode, the subject is forced to simultaneously distribute their attention to two different tasks: constant consultation of the model to capture their peculiarities and careful tracing of the corresponding signs. This division of attention, coupled with inexperience in performing movements that are foreign to own habit, leads to unavoidable anomalies on tracing, to detentions and discontinuous movements, tremors, interruptions, junctions, unnecessary redraws, etc., but above all, to a significant *reduction* of the pen displacements speed. Traces are rigid, uncertain, hesitant and often distensible.

The forger finds it difficult to fight against their own graphic automatism. He must, by other side, to perform in a conscious and voluntary way some signs that the genuine author performs unconsciously and automatic. In any forgery are noticed clues of dedication, shrinkage and the lack of naturalness. The most difficult is not to reproduce shapes, but get the mimic, the "vitality" and the particular expression of manuscript.

B. *Quick imitations*

The forgers are well aware that slow imitations are imperfect and easy to detect. To overcome this they come to quick imitations, which leave no trace of hesitation or psychomotor insecurity. This type of forgery tend to be more spontaneous, loose and natural and therefore more suitable to deceive. Can be classified into: global or comprehensive imitations and free imitation or by assimilation of traces. The first are copies, usually "by memory", of the more relevant and striking features. It captures the global, but neglects the local details. Very often, memory is inaccurate and the writing is almost the forger's. Regarding the latter, are also known as trained imitations. Forger reproduces several times the original signature on a trial basis, to acquire the necessary looseness and fluidity of movements. The training is followed by a self-assessment.

III. DETECTING THE FORGER

Fictitious signatures can be found in fraudulent check, hotel registrations, applications for licenses and services, to purchases of merchandise, etc. It is important to know under what conditions the writer of these signatures can be identified.

There are two requirements which must be satisfied before a positive identification can be approach. First of all, the forged signature must have been written in the natural handwriting of the forger. Signatures for the most part are short, and even a moderate degree of disguise may prevent accurate identification of a single specimen. The second condition is the need for several forgers' known signatures. One specimen of his signature alone is virtually useless for solving the problem.

The writer identification of a signature is established by a combination of a number of personal writing habits in both the questioned and known writer's signatures. Any writing habit, however unusual, helps to individualize his signature. When it appears in the forged signature, it is a link in the chain of evidences which connects the questioned signature with the known one.

In a simulated signature which is not traced there may be somewhat wide divergences between it and the genuine signatures. It is to be expected that these divergences should occur at points where the forger's writing habits differ from those of the model signature. Generally these divergences are of a hybrid nature. They are composed of elements of the forger's writing habits modified to some degree by the handwriting which he is attempting to imitate [4].

As example, figure 1 show the genuine signatures of two writers and their cross forgeries. Notice that the natural way of writer A handwriting produces vertical strokes in his own signature. When signer A forge the signature of write B, the last stroke of the forged signature is more vertical than the genuine one.

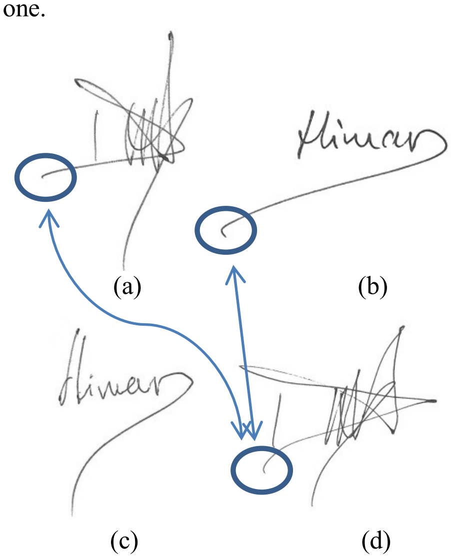

Figure 1. (a) Genuine signature of writer A. (b) Genuine signature of writer B. (c) Signature of writer B forged by signer A. (d) Signature of writer A forged by signer B.

IV. EXPERIMENTS

The aim of this section is to study if some of the already proposed features in the literature for automatic signature verification are useful for forger detection. Concisely, we have tested the features proposed in [5-8]. The authors of [5] characterize the signature by means of geometrical features: width, heights, polar pixel distribution, etc. [6] proposes texture based features. The authors divide the signature image in blocks and calculate statistics measures of the local binary patterns (LBP). The statistic measures form each block are concatenated to obtain the signature feature. The texture measures are interesting because the gray distribution depends on how the writer uses the pencil, which is very personal and difficult to imitate. For this reason we extend the study of this work to the local directional pattern [7] and the local derivative patter [8] which have been recently proposed for biometrics. As classifier we have used a Support Vector Machine (SVM) [10]. With the parameters of [5] and [6] we have used a RBF kernel while with the parameters proposed in [7] and [8], which are a concatenation of histograms, we have used a χ^2 kernel.

A. Database

The database used is the off-line sub corpus of the MCYT database [9]. It includes 75 signers from 4 different Spanish sites. The corpus includes 15 genuine signatures and 15 simulated forgeries for each signer. Therefore we have 1125 genuine signatures and 1125 forgeries available. The name of the files provides information about who is the forger, so it is suitable for our purpose. Genuine signatures were acquired in 2 sessions. Forgers are given the signature images of clients to be forged and, after training with them several times, they are asked to imitate the shape. Therefore, it includes quick and slow imitations. All signature data were acquired with the same inking pen and the same paper templates, over a similar pen tablet. The paper templates were scanned at 600 dpi.

B. Methodology

Let $\text{signature}i\_\text{writer}j\_\text{rep}k.\text{bmp}$ be the signature of user i made by the writer j by k time. In our case $1 \leq i \leq 75$, $1 \leq j \leq 75, 1 \leq k \leq 15$. The model of signature i is trained with $\text{signature}i\_\text{writer}i\_\text{rep}k.\text{bmp}$ $1 \leq k \leq 5$ as positive samples and with the files $\text{signature}j\_\text{writer}j\_\text{rep}k.\text{bmp}$ $j \neq i, 1 \leq k \leq 5$ as negative samples ($74 \cdot 5 = 370$ samples). Highlight that the signature models are only trained with genuine signatures.

Let $x_{i,j,k}$ be the characteristic vector of the $\text{signature}i\_\text{writer}j\_\text{rep}k.\text{bmp}$ and λ_i the model of signer i. The test is done obtaining the scores $s_{i,j,k}^{l} = P\left(x_{i,j,k} | \lambda_l\right)$ with the remainder signatures: $1 \leq i \leq 75$, $1 \leq j \leq 75, 6 \leq k \leq 15$ ($75 \cdot 75 \cdot 10 = 56250$ scores). Our hypothesis is that when working out the scores of a given forgery with the model of all the signers: $s_{i,j,k}^{l}$ $1 \leq l \leq 75$, two of the greater values will be $s_{i,j,k}^{i}$ and $s_{i,j,k}^{j}$. It is expected because the most similar signatures to model i are the forgeries of signature i and, as the forger always left a mark of his natural writing way in the forged signature, it is expected than the score obtained with the forger model $s_{i,j,k}^{j}$ will be also high.

C. Results

Once trained the signature models and made the test, the results are given as probability density functions (pdf). Three pdfs have been worked out:

a. the pdf of $s_{i,j,k}^{i}$, $1 \leq i \leq 75$, $6 \leq k \leq 15$ (750 scores)

b. the pdf of $s_{i,j,k}^{j}$, $1 \leq i \leq 75$, $6 \leq k \leq 15$ (750 scores)

c. the pdf of the remainder scores $s_{i,j,k}^{l}$, $1 \leq i, j \leq 75, l \neq i, j$, $6 \leq k \leq 15$

The results are shown in figure 2. Figure 2.a show the three pdf obtained when charactering the signatures using geometrical parameters [5]. The pdfs obtained using the texture measures known as local binary patterns $\text{LBP}_{8,1}^{\text{riu2}}$ plus $\text{LBP}_{16,2}^{\text{riu2}}$ and statistical measures from gray level co-occurrence matrices (GLCM) [6] are seen in Figure 2.b. Figure 2.c and 2.d show the functions obtained with the Local Directional patterns (LDP) [7] and Local Derivative Patterns (LDerivP) [8] respectively. The pdfs of $s_{i,j,k}^{i}$, $s_{i,j,k}^{j}$ and $s_{i,j,k}^{l}$ are depicted in continuous blue, dotted red and dashed green respectively.

As expected, the $s_{i,j,k}^{i}$ values are greater than $s_{i,j,k}^{j}$ and $s_{i,j,k}^{l}$ in the four systems evaluated. In the case of geometrical features, the pdf of $s_{i,j,k}^{j}$ and $s_{i,j,k}^{l}$ are very similar, that is to say, the geometrical features are not able of detecting the mark of the writer's writing. For this reason we tested with parameters based on texture, where is more difficult to hide the own way of writing. In this case, features of [6] [7] and [8], we find that the scores obtained with the writer (forger) model $s_{i,j,k}^{j}$ are displaced to the right being between the pdf curves of $s_{i,j,k}^{i}$ (model of forged writer) and $s_{i,j,k}^{l}$.

In order to gain a better idea of the difference between scores distribution, the Mahalanobis distance between the score distributions $s_{i,j,k}^{i}$, $s_{i,j,k}^{j}$ and $s_{i,j,k}^{l}$ is given in Table I. The Mahalanobis distance is defined as:

$$\Delta = \frac{|m_1 - m_2|}{\sqrt{0.5\,(\sigma_1^2 + \sigma_2^2)}}$$

where m_1, m_2, σ_1^2, and σ_2^2 are the means and variance of score distributions to be compared, for instance $s_{i,j,k}^{j}$ and $s_{i,j,k}^{l}$. A high Mahalanobis distance means a strong difference between compared scores distribution. As seen in Table I, the best features separating the $s_{i,j,k}^{j}$ and $s_{i,j,k}^{l}$ pdf curves are the proposed in [6]: $\text{LBP}_{8,1}^{\text{riu2}}$ plus $\text{LBP}_{16,2}^{\text{riu2}}$ plus GLCM. . For these parameters the distance between $s_{i,j,k}^{i}$ and $s_{i,j,k}^{j}$ pdf curves is clearly lower than the distance between $s_{i,j,k}^{j}$ and $s_{i,j,k}^{l}$. It means the genuine signature of the forger $s_{i,j,k}^{j}$ contains information potentially useful to identify the forger.

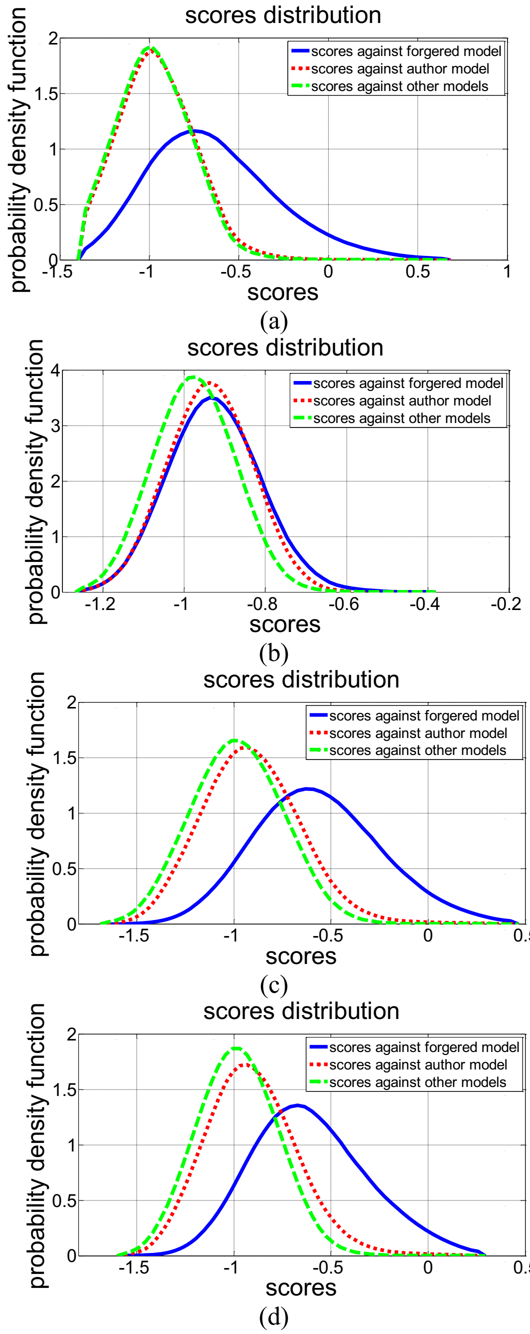

Figure 2. pdfs curves of $s^{i}_{i,j,k}$ (continuous blue), $s^{j}_{i,j,k}$ (dotted red) and $s^{l}_{i,j,k}$ (dashed green) obtained with: (a) geometrical parameters [5]; (b) $LBP^{riu2}_{8,1}$ plus $LBP^{riu2}_{16,2}$ and GLCM [6]; (c) Local Directional patterns (LDP) [7]; and (d) Local Derivative Patterns (LDerivP) [8].

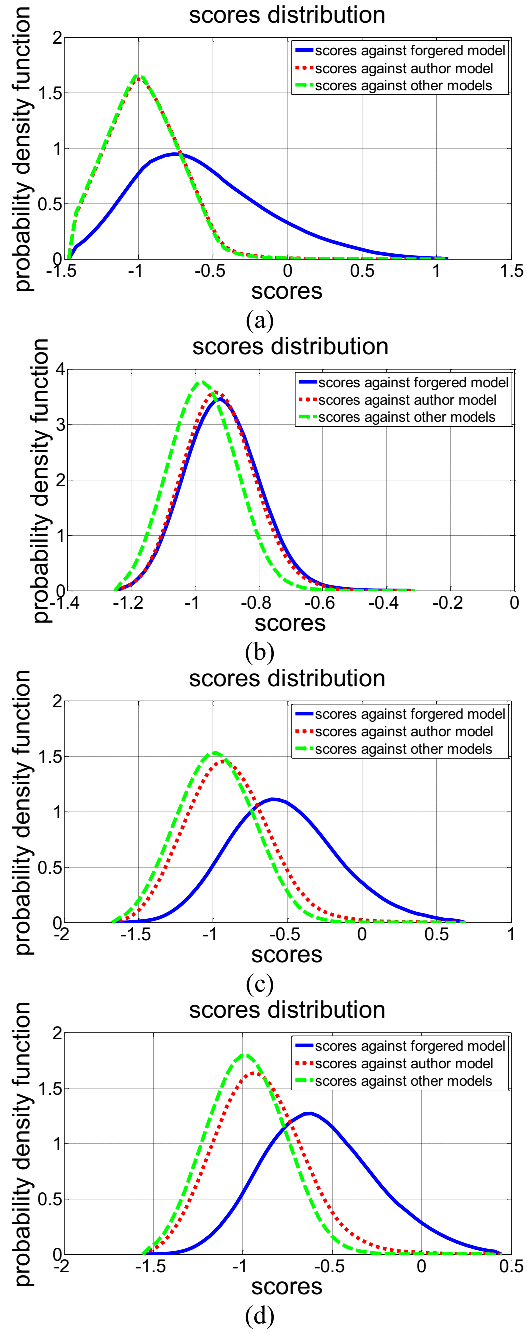

Figure 3. After training the writer models λ_l with 10 samples, pdfs curves of $s^{i}_{i,j,k}$ (continuous blue), $s^{j}_{i,j,k}$ (dotted red) and $s^{l}_{i,j,k}$ (dashed green) obtained with: (a) geometrical parameters [5]; (b) $LBP^{riu2}_{8,1}$ plus $LBP^{riu2}_{16,2}$ and GLCM [6]; (c) Local Directional patterns (LDP) [7]; and (d) Local Derivative Patterns (LDerivP) [8]..

TABLE I MAHALANOBIS DISTANCE BETWEEN THE SCORES DISTRIBUTIONS

Parameters	$s_{i,j,k}^{i}$ and $s_{i,j,k}^{j}$	$s_{i,j,k}^{i}$ and $s_{i,j,k}^{l}$	$s_{i,j,k}^{j}$ and $s_{i,j,k}^{l}$
Geometrical	0.843	0.912	0.113
LBP+GLCM	0.152	0.657	0.562
LDP	1.338	1.688	0.376
LDerivP	1.405	1.886	0.447

Result discussion

So, these last parameters are able to find some of the natural writer's way of writing. We say "some" because the difference between the pdf curves of $s_{i,j,k}^{j}$ and $s_{i,j,k}^{l}$ is not very significant. Therefore, although this parameters are not good for detecting the forger, it can be seen that the forger leave a sort of watermark in the forgery. The intensity of this watermark depends of the natural writing movements done by the forger. More improvements could be done looking for new parameters that increase the difference between both pdfs curves.

In [4] is clearly stated that the ability of identify the forger depends on the amount of handwriting available. As more handwriting is available, more easier is to detect the forger. So we repeat the above experiment training the writer models with 10 samples. Results can be seen in Figure 3. It can be seen that the difference with respect to train with 5 is not significant.

V. CONCLUSIONS

This paper proposes a first approach to the forensic problem of detecting who has forged a signature. The ability of solving this problem by a forensic depends on both the amount of forger handwriting available and whether the forger has written in his natural way. Both limitations also stand for automatic systems.

Obviously, this paper has not solved the question, but has proposed a research line and a measure to start to work: to try to separate the pdf of $s_{i,j,k}^{i}$, $s_{i,j,k}^{j}$ and $s_{i,j,k}^{l}$ The proposed methodology and experiments shown that geometrical parameters are useless for detecting who is the forger but texture based parameters could be a good start point for approaching who has written a signature. More work is being done looking for new parameters, for instance with SIFT descriptors.

ACKNOWLEDGMENTS

This work has been funded by Spanish government MCINN TEC2009-14123-C04 research project and UdeA-CODI project MC11-1-01 E01608.

REFERENCES

[1] Impedovo, D.; Pirlo, G.; , "Automatic Signature Verification: The State of the Art," Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on , vol.38, no.5, pp.609-635, Sept. 2008

[2] D. Bertolini, L.S. Oliveira, E. Justino, R. Sabourin, "Reducing forgeries in writer-independent off-line signature verification through ensemble of classifiers", Pattern Recognition, Volume 43, Issue 1, January 2010, Pages 387-396, ISSN 0031-3203.

[3] M. Liwicki, C. E. an den Heuvel, B. Found, M. Imran Malik, "Forensic Signature Verification Competition 4NSigComp2010: Detection of Simulated and Disguised Signatures", proceedings of 12th International Conference on Frontiers in Handwriting Recognition, pp.715~720, Kolkata, 16-18 November 2010.

[4] Ordway Hilton, "Can the Forger Be Identified from His Handwriting?", The Journal of Criminal Law, Criminology, and police Science, vol. 43, no. 4, pp. 547-555, November-December 1952.

[5] Miguel A. Ferrer, Jesús B. Alonso, Carlos M. Travieso, "Offline geometric parameters for automatic signature verification using fixed-point arithmetic", in IEEE Transactions on pattern analysis and machine intelligence, ISSN 0162-8828, vol. 27, no. 6, pp 993-997, June 2005.

[6] J.F. Vargas, M.A. Ferrer, C.M. Travieso, J.B. Alonso, "Off-line signature verification based on grey level information using texture features", in Pattern Recognition, vol. 44, no. 2, pp. 375-385, February 2011.

[7] M.A. Ferrer, F. Vargas, C.M. Travieso, J.B. Alonso, "Signature verification using local directional pattern (LDP)", IEEE International Carnahan Conference on Security Technology, pp. 336-340, 5-8 Oct. 2010.

[8] B. Zhang, Y. Gao, S. Zhao, J. Liu, "Local Derivative Pattern Versus Local Binary Pattern: Face Recognition With High-Order Local Pattern Descriptor", IEEE Transactions on Image Processing, vol.19, no.2, pp.533-544, February 2010.

[9] J. Fierrez-Aguilar, N. Alonso-Hermira, G. Moreno-Marquez, and J. Ortega-Garcia. "An off-line signature verification system based on fusion of local and global information". In Workshop on Biometric Authentication, Springer LNCS-3087, pages 298–306, May 2004.

[10] J. A. K. Suykens, T. V. Gestel, J. D. Brabanter, B. D. Moor, J. Vandewalle, "Least Squares Support Vector Machines", World Scientific Publishing Co., Pte, Ltd, Singapore, 2002.

[11] L.G. Velásquez, "Falsedad Documental y Laboratorio Forense", Señal Editorial, 2004.