Quantitative Evaluation of Features for Forensic Handwriting Examination

Angelo Marcelli, Antonio Parziale and Claudio De Stefano

***Abstract*— We propose a quantitative approach to both feature evaluation and comparison that combines Forensic Handwriting Examination best practices with Pattern Recognition methodologies. The former provide a set of features that are meant to capture the distinctive aspects of handwriting, the latter the computational tools for the quantitative evaluation of the features values as well as for their comparison. We will show that such a combined approach leads to a procedure that is theoretically sounds and can be expressed in terms the document examiners are familiar with. Eventually, we will suggest possible ways of using the results of the proposed approach in forensic handwriting examiners casework.**

***Index Terms*— writer identification, forensic handwriting examination, handwriting generation**

I. Introduction

One of the main obstacles for the adoption of automatic tools in forensic handwriting examination lies in the difference between the way the human expert and the machine look at the problem.

Forensic Handwriting Examiners (FHEs) assume that the distinctive features of handwriting are visually observable in the document and therefore their evaluation is generally achieved by visual inspection, sometime supported by physical measurements. The features refer to visual clues at *document* level, such as margins and base line inclination, at *line* level, such as interline and word spacing, at *word* level, such as word layout, and eventually at *character* level, such as slant or inter-character spacing. Additionally, their comparison is rather qualitative than quantitative and it is usually expressed in a non-standard textual form or with reference to some scale [1,2].

On the other hand, automatic tools assume that the distinctive features of a subject handwriting are captured by the way some basic elements are arranged in the digital image of the document and their values can be computed from those of the basic elements. The basic elements the features refer to can be the *pixels* of the image, *group of pixels* configuration, or *piece of the digital ink*. Eventually, their comparison is achieved by measuring some form of "difference" between their values, sometimes combined into a confidence score [3, 4].

This difference makes FHEs uncomfortable when using automatic tools, as they do not usually understand what the tool is doing and to which extent the values of features computed by the tool support the proposed final decision and, even more important, how trustable such a decision is.

To partially overcome this inconvenient and provide empirical evidence of the confidence of the conclusion proposed by automatic tools, competitions are regularly organized to evaluate the performance of automatic tools on forensic application, such as signature verification and writer verification/identification [5-7]. The results of the competitions have shown that automatic tools have much improved in terms of both error rate and the capability to provide a confidence measure of the final decision. Nonetheless, FHEs sustain that the results of the competitions show that the systems were able to model the handwriting of a given population of subjects so as to be able to effectively perform the task. Thus, it is not ensured that the performance of the systems will remain the same when dealing with the handwriting of a subject outside of the initial population. On the other hand, they also remark, in most of their caseworks the available material does not contain enough samples to include the new subjects into the population and retrain the automatic system to perform the task.

To fill this gap, we propose a method for linking the features FHEs are familiar with to measurements on the digital image of the document. As the measurements are expressed in terms of the values of some mathematical entities, such a link will allow to use the mathematical properties of these entities also for evaluating the difference between measurements performed on different documents, and thus on the variations between features. In a nutshell, the method should provide a quantitative estimation of the variability of the features when estimated on different documents.

In the following section we introduce the two main stages of our method, namely the set of measurements we use to represent each feature and the mathematical procedure we adopt to compare their values. In Section 3 we report the data gathered from some preliminary experiments, and in the last section we discuss what the quantitative dimension of the proposed measurements may suggest to FHEs.

Angelo Marcelli is with the Department of Information and Electrical Engineering and Applied Mathematics, Via Giovanni Paolo II, 132, University of Salerno, Italy (e-mail: amarcelli@unisa.it).

Antonio Parziale is with the Department of Information and Electrical Engineering and Applied Mathematics, Via Giovanni Paolo II, 132, University of Salerno, Italy (e-mail: anparziale@unisa.it).

Claudio De Stefano is with the Department of Electrical and Information Engineering, Via Si Biasio, 32, University of Cassino and Southern Lazio, Italy (e-mail: destefano@unicas.it).

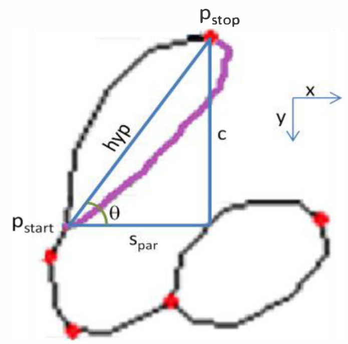

Fig. 1. In magenta is highlighted the stroke under investigation. The triangle used for computing the slant of the stroke is reported. p_{start} and p_{stop} are respectively the first and the last point of the stroke.

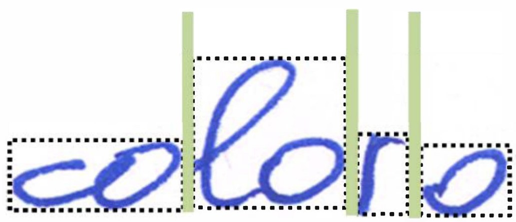

Fig. 3. The inter-fragment spacing is highlighted in green. The bounding-boxes surrounding each fragment are depicted with a dotted line.

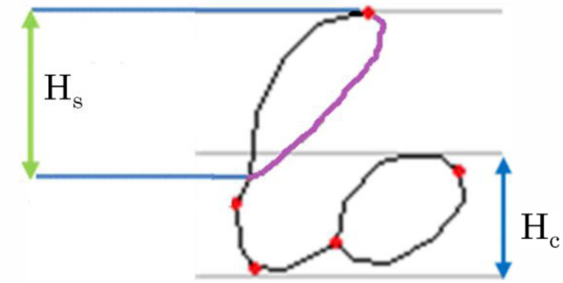

Fig. 2. The first stroke of the fragment is highlighted in magenta, the boundaries of the core and the upper zone are depicted in gray. The stroke is classified as an ascender up, H_c is the height of the core region while H_s is the height of the ascender up.

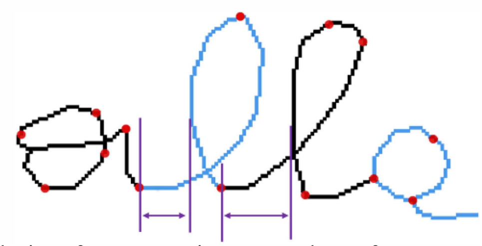

Fig. 4. The intra-fragment spacing computed on a fragment containing the text "alla". The spacing is highlighted in magenta. Only two ligatures are detected.

II. From Features to Measurements

In order to automatically compute the set of features that will be used in this paper we need to identify the basic elements the features are based upon. For the purpose, we use a procedure that was presented in [8]. The software executes the following processes on the image of the document:

A. *Words Extraction*

All words in a document are extracted using a procedure based on the algorithm presented in [9].

B. *Fragments Extraction and Classification*

Usually words are not produced by keeping the pentip in constant contact with the paper, so as to have a continuous ink, therefore sub-images that correspond to pieces of ink produced without lifting the pen are extracted from each word image. Each sub-image is then classified as cursive, isolated character, vertical line, horizontal line, dot, noise, or reject, as described in [10].

Each cursive component and isolated character is manually associated with its transcription.

C. *Writing Order recovery*

The static handwriting of fragments labeled as isolated character and cursive is represented by its skeleton, which is then converted into a graph, whose arcs correspond to the skeleton branches, and nodes to either end point or branch point of the skeleton. Criteria derived by handwriting generation [11] are then applied to reconstruct the writing order of the ink of the fragment.

D. *Stroke Segmentation and Classification*

The sequence of points belonging to the ink is then segmented in strokes by using an algorithm based on the concept of saliency introduced for modeling visual attention shift [12]. Each stroke is classified on the basis of its position across the upper, the core and the lower regions of the word. In particular:

- a stroke starting in the core region and ending in the upper region is labeled as *ascender up*;
- a stroke starting in the upper region and ending in the core region is labeled as *ascender down*;
- a stroke starting in the core region and ending in the lower region is labeled as *descender down*;
- a stroke starting in the lower region and ending in the core region is labeled as *descender up*;
- a stroke starting and ending in the core region is labeled as *center*;
- a stroke starting in the lower/upper and ending in the upper/lower region is labeled as *pipe*;

E. *Character Segmentation*

Each elementary movement of a fragment is associated with the label corresponding to the character it belongs to, as described in [13].

Once the handwritten words have been extracted and the elementary movements have been identified and classified a set of 7 features are measured for each word and they are divided in to two groups depending on the basic unit that they characterize:

- *stroke level* measurements (up-slant, down-slant, up-ratio, down-ratio, curvature) ;
- *word level* measurements (inter-fragment spacing, intra-fragment spacing);

The **slant** of a stroke is evaluated by building the triangle that has the segment between the first (p_{start}) and the last (p_{stop}) point of the stroke as hypotenuse (hyp), and one of the sides (s_{par}) parallel to the baseline, as shown in Fig. 1. The slant is computed as:

$$slant = \begin{cases} \cos^{-1}\dfrac{s_{par}}{hyp} \; if \; p_{start.x} < p_{stop.x} \\ 180° - \cos^{-1}\dfrac{s_{par}}{hyp} \; if \; p_{start.x} > p_{stop.x} \\ 90° \; otherwise \end{cases}$$

The **up-slant** is defined as the slat of an ascender up, whereas the **down-slant** is the slant of a descender down.

The **up-ratio** and the **down-ratio** are computed, respectively, as the height of an ascender and a descender divided by the height of the core region. The height of a stroke is defined as ($|p_{stop.y} - p_{start.y}|$), as shown in Fig. 2.

The **curvature** of a stroke is computed as the ratio between the length of the segment connecting p_{start} and p_{stop} and the number of pixels belonging to the stoke. Because the width of a stroke is one pixel, the curvature of a straight stroke is one.

The **inter-fragment spacing** is computed as the distance between the bounding boxes of two consecutive fragments belonging to the same word, as shown in Fig. 3.

The **intra-fragment spacing** measures the length of ligatures between characters within a fragment. A ligature is defined as the piece of ink that doesn't contain a segmentation point, starts from the segmentation point between two consecutive characters and ends when the histogram of the vertical projections shows a transition from 1 to another value, as shown in Fig. 4.

III. Comparing Measurement

In order to compare two or more documents, each of them is represented by seven statistical distributions, one for each of the measurements presented before. We chose to use a Gaussian distribution for each sample distribution, therefore a document is described as:

$$D_j = \{N_1^j(\mu_1, \sigma_1), \ldots, N_i^j(\mu_i, \sigma_i), \ldots, N_7^j(\mu_7, \sigma_7)\}$$

The difference between two documents is computed by summing up the Kullback-Leibler divergences (*KL)* [14] between the Gaussian distributions of each feature. The Kullback-Leibler divergence is a measure of the difference between two probability distributions but it is not a metric because it doesn't satisfy the symmetry condition. So, given two documents D_x and D_y, the distance between them is:

$$d(D_x, D_y) = \frac{1}{7} * \sum_{i=1}^{7} KL(N_i^x, N_i^y)$$

and

$$d(D_x, D_y) \neq d(D_y, D_x)$$

IV. Experiments

As we are interested in establishing to which extent difference in the measurements relate to difference in the features, we have performed some experiments on a small set of documents prepared for the purpose.

In the experiments, 4 subjects were requested to produce 2 documents each, in different sessions. In the first session, the subjects were required to write down a text whose reading (including punctuation mark and new line) was previously recorded by a professional speaker. The pace of reading was determined in a preliminary section on different texts, so as to be sure that all the subjects were comfortable with, and, at the same time, feel the pressure to write a little faster than they would usually do. In the second session, the subjects were simply requested to copy the same text of the reading, at their own pace. This procedure was adopted to investigate the role of the proposed approach in estimating both the variability within the same subject in the different experimental conditions and the dissimilarity between different subjects.

The measurements were obtained by using Masquerade [15], a software tool that provides a large set of measurements that can be easily customized to the user need. The values provided by the tool for each features were then interpreted as values of a random variables whose probability distribution function was assumed to be Gaussian, as mentioned before. The values of the mean and variance of each feature where eventually used to compute the Kullback-Liebler divergence for that features between each pair of documents.

In the first experiments, to investigate the ability of the features to characterize the *intra-writer* variability, we have computed the divergences between the features extracted from each pair of document produced by each subject. Table 1 reports, the values of the Kullback-Leibler divergences for each feature for each pair of documents: in the table, Gij denotes the document produced by the i-th subject in the j-th session.

In the second experiments, to investigate *inter-writer* variability, for each subject we computed the divergences between each pair of features measured when compared the document of that subject with the documents produced by any other subjects. In Table 2 we report the values of the Kullback-Leibler divergences when comparing the documents with respect to the ones produced in the first session, while Table 3 reports the values when comparing the documents with those produced in the second session.

V. Discussion and Concluding Remarks

The data in the tables can be interpreted by keeping in mind that the Kullback-Leibler divergence estimates the dissimilarity between the ways the features values vary within each documents. In other words, small values of divergences correspond to features whose values exhibit similar variations.

Thus, we would expect small values of the divergence when a feature is computed on documents produced by the same subject. This means that, the values of the divergences can tell us something about the intra-writer variability.

Let us consider, for instance the first two entries in Table 1. We can note that, as expected, the values of the divergences are generally small, but with a few exceptions: in the first entry, the third, sixth and seventh features show a divergence that is larger than those of the other ones, in the second entry the same happens for the third and fourth feature. Those values seem to suggest that those features vary much more than the others between documents produced by the same subjects. In turn, this may imply that the subject handwriting aspects captured by those features vary more than others and therefore that those features are less specific of that subject. We argue that this should be taken into account when comparing a genuine document of the subject with a questioned one. Eventually, if we consider both the entries together, they show that while the third features exhibits larger-than-other values in both comparisons, this is not true for the other ones. This suggests that changes in the writing modalities affect only few of the subject handwriting characteristics, namely those corresponding to the fourth, sixt and seventh features, while the subject by itself and rather independently of the writing modalities introduces large variation in his handwriting characteristics captured by the third features. We argue that this observation also seems to suggest that, for this subject, the most distinctive features are the first, the second and the fifth, as they correspond to the lowest values of the divergences across the two documents. Let us consider now the entries in the table that refers to the third subject. By following the same line of thoughts as previously, we can conclude that all the features have the same discriminative power, as they all exhibit comparable values, suggesting that this subject handwriting characteristics were not affected by the writing modalities and that all the features should be taken into account when comparing the documents of this subject against the questioned one. The data referring to the other subjects involved in the experiments show a behavior that is somehow between the one of first and third subject.

As noticed before, as we expect small values of divergences between documents produced by a subject, we would also expect larger values of the divergences when comparing documents produced by different subjects. Or, more precisely, we expect that at least some of the divergences are larger than those computed on genuine documents. This follows from the previous observations on the data reported in Table I, according to which not all the features exhibit a small intra-writer variability. Let us now consider what the data in Table 2 and Table 3 suggest about the inter-writer variability. As we are comparing documents produced by different subjects we would expect to see larger values than those reported in Table I. In statistics, the standard deviation is a measure that is used to quantify the amount of variation or dispersion of a set of data values. A standard deviation close to 0 indicates that the data points tend to be very close to the mean (also called the expected value) of the set, while a high standard deviation indicates that the data points are spread out over a wider range of values. As we are interested in investigating just the variation of the features across the subjects, Table 4 reports the values of the standard deviation of the values of each features divergence across the data in Table 2 (first row) and Table 3 (second row). They tell us that not all the features exhibit the expected behavior. In particular, we see that the first, the fifth and the sixth features have a small standard deviation, meaning that the values of the features across the documents tends to exhibit the same amount of variations for each pair of documents. Accordingly, we argue that these features have much less discriminative power with respect to the other ones, and that also should be taken into account when comparing the genuine with the questioned.

Let us now addressing the last issue, i.e. the *comparison between the intra-writer and the inter-writer variability*. Equipped with the observations we have made on both from the data reported above, we argue that we should give more importance to the features that exhibits both properties of being stable and discriminative. Stable means that they exhibit similar variations among genuine documents, while discriminative means that they exhibit different variations across a population of writer. Thus, the ideal set of features should include the ones exhibiting both small intra-writer variations, i.e. small divergence in Table I and large inter-writer variations, i.e. large values in Table 2 and 3.

To implement this concept, let us take a hard-decision approach based on the following criteria:

1. the set of features that is considered for the comparison includes only the features that enjoy both the above mentioned properties;
2. each feature is a decision-maker that provides a binary output, specifying whether the documents are produced by the same subject or not;
3. the final decision is achieved by majority among the features; in case of a tie no decision is proposed.

As it can be seen, the criteria described above are kind of Draconian, as each feature acts independently of the others and the final response is achieved by counting the votes, i.e. again without evaluating the relations between the features. We believe, however, that testing the proposed approach under such conditions will provide a lower bound on the performance.

When applied to the data reported in the Tables I and II above, they give the following results:

- the first, fifth and sixth features are removed from the set used during the comparison because of their small inter-writer variations;
- when assuming as genuine the documents produced by the first subject, the third and fourth features are removed because of their large intra-writer variations;

TABLE I
DIVERGENCE MEASUREMENTS WITHIN SUBJECTS. GIJ DENOTES THE DOCUMENT PRODUCED BY THE I-TH SUBJECT DURING THE J-TH SESSION.

	up-slant	down-slant	up-ratio	down-ratio	curvature	intra-fragment Spacing	inter-fragment Spacing
G12 vs G11	0,065	0,149	1,050	0,008	0,009	1,240	1,240
G11 vs G12	0,054	0,260	5,077	12,280	0,007	0,009	0,410
G22 vs G21	0,022	0,650	0,347	1,190	0,007	0,065	0,022
G21 vs G22	0,019	2,910	0,810	6,660	0,008	0,070	0,018
G32 vs G31	0,030	0,130	0,480	0,620	0,002	0,007	0,025
G31 vs G32	0,025	0,084	0,609	0,524	0,003	0,007	0,020
G42 vs G41	0,100	0,140	42,060	0,780	0,005	0,069	0,110
G41 vs G42	0,070	0,139	1,750	2,950	0,006	0,090	0,070

TABLE II
DIVERGENCE MEASURES BETWEEN THE DOCUMENTS WITH RESPECT TO THOSE PRODUCED IN THE FIRST SESSION.

	up-slant	down-slant	up-ratio	down-ratio	curvature	intra-fragment Spacing	inter-fragment Spacing
G32 vs G11	0,103	0,327	1,780	1,750	0,001	0,008	7,330
G31 vs G11	0,056	0,390	1,310	1,430	0,008	0,001	6,400
G22 vs G11	0,050	2,150	2,060	2,470	0,005	0,034	0,017
G21 vs G11	0,060	1,060	1,340	1,240	0,014	0,009	0,030
G42 vs G11	0,054	0,018	4,850	1,850	0,024	0,150	2,120
G41 vs G11	0,140	0,180	1,290	0,630	0,039	0,024	0,950
G32 vs G21	0,199	4,160	0,171	2,830	0,011	0,011	11,010
G31 vs G21	0,170	1,680	0,080	0,650	0,007	0,005	9,750
G12 vs G21	0,140	3,120	0,100	1,610	0,002	0,005	2,120
G11 vs G21	0,047	6,390	8,420	11,250	0,012	0,010	0,040
G42 vs G21	0,097	7,630	51,830	2,490	0,052	0,230	3,460
G41 vs G21	0,280	4,210	0,015	1,170	0,057	0,068	1,640
G22 vs G31	0,179	1,507	0,664	0,630	0,000	0,039	1,010
G21 vs G31	0,250	0,490	0,100	0,530	0,007	0,004	1,120
G12 vs G31	0,028	0,379	0,095	0,170	0,001	0,007	0,277
G11 vs G31	0,064	0,546	9,810	12,530	0,007	0,001	0,950
G42 vs G31	0,066	0,630	67,600	0,490	0,025	0,170	0,140
G41 vs G31	0,020	0,140	0,200	1,110	0,025	0,034	0,360
G32 vs G41	0,025	0,015	0,160	0,710	0,027	0,036	0,980
G31 vs G41	0,023	0,096	0,130	0,390	0,029	0,030	0,870
G22 vs G41	0,280	1,840	0,340	1,360	0,028	0,020	0,378
G21 vs G41	0,400	0,780	0,013	0,400	0,070	0,050	0,480
G12 vs G41	0,045	0,272	0,103	0,480	0,045	0,056	0,014
G11 vs G41	0,150	0,140	6,960	1,800	0,039	0,022	0,350

– implementing the criteria 2 and 3 mentioned above, out of the 6 comparisons, 2 of them (namely G31 vs G11 and G32 vs G11) are correctly identified as not produced by the first subject, while for the remaining 4 no conclusions are reached;

Applying the same procedure to each subjects, of the 24 comparisons, 3 of them (namely G31 vs. G41, G32 vs G41, and G12 vs. G41) erroneously conclude that the documents under comparison are produced by the same subject, 6 of them do not reach a conclusion, and the remaining 15 lead to correct conclusions.

The small number of subjects involved in the experiment, the many constraints on the experimental settings, the use of a subset of the features currently adopted in forensic examination of handwriting, do not allow to draw any conclusion. The experiments need to be performed on large data set, as the ones used for the most recent international competitions on writer identification, more features need to be implemented in terms of quantitative measurements, and FHEs will have to be involved in the measurement and evaluation processes. Still, we believe that the work reported here has shown that pursuing a method for quantitative evaluation of the features currently used by FHEs is feasible, at least to some extent, and that looking at the quantitative dimension of the features can offer many solid methodologies

TABLE III
DIVERGENCE MEASURES BETWEEN THE DOCUMENTS WITH RESPECT TO THOSE PRODUCED IN THE SECOND SESSION.

	up-slant	down-slant	up-ratio	down-ratio	curvature	intra-fragment Spacing	inter-fragment Spacing
G32 vs G12	0,004	0,900	0,419	0,117	0,004	0,023	0,680
G31 vs G12	0,022	0,610	0,060	0,140	0,001	0,006	0,580
G22 vs G12	0,081	2,030	0,630	0,850	0,001	0,075	0,450
G21 vs G12	0,150	0,940	0,080	1,100	0,002	0,005	0,550
G42 vs G12	0,009	0,480	37,640	0,099	0,035	0,230	0,040
G41 vs G12	0,035	0,610	0,094	1,330	0,038	0,057	0,012
G32 vs G22	0,098	49,000	0,088	8,750	0,001	0,027	8,210
G31 vs G22	0,110	24,710	1,360	2,740	0,000	0,040	7,300
G12 vs G22	0,065	38,790	1,990	4,950	0,001	0,098	1,430
G11 vs G22	0,040	71,070	31,560	121,080	0,005	0,042	0,018
G42 vs G22	0,030	82,840	152,840	6,410	0,022	0,060	2,460
G41 vs G22	0,170	50,090	0,890	17,670	0,025	0,025	1,060
G22 vs G32	0,125	1,820	0,061	1,220	0,001	0,026	1,040
G21 vs G32	0,220	0,770	0,270	1,950	0,012	0,012	1,160
G12 vs G32	0,004	0,360	0,860	0,110	0,004	0,028	0,280
G11 vs G32	0,087	0,240	17,670	14,640	0,001	0,010	0,990
G42 vs G32	0,025	0,220	89,240	0,030	0,018	0,160	0,160
G41 vs G32	0,020	0,014	0,280	2,100	0,026	0,045	0,360
G32 vs G42	0,022	0,240	2,110	0,039	0,021	0,178	0,317
G31 vs G42	0,048	0,360	1,970	0,580	0,031	0,214	0,240
G22 vs G42	0,030	2,180	2,370	1,140	0,027	0,065	0,605
G21 vs G42	0,090	1,090	1,850	2,410	0,069	0,270	0,710
G12 vs G42	0,008	0,227	1,700	0,130	0,044	0,308	0,030
G11 vs G42	0,040	0,010	0,830	20,260	0,027	0,190	0,550

TABLE IV
STANDARD DEVIATION OF THE DIVERGENCES ACROSS THE SUBJECTS AND DOCUMENTS.

	up-slant	down-slant	up-ratio	down-ratio	curvature	intra-fragment Spacing	inter-fragment Spacing
G12 vs G11	0,065	0,149	1,050	0,008	0,009	1,240	1,240
G11 vs G12	0,054	0,260	5,077	12,280	0,007	0,009	0,410

developed in the realm of Pattern Recognition to gather data and derive the kind of evidences needed in Forensic Handwriting Examination. Eventually, we will investigate to apply the same method to the analysis of offline signatures by taking also into account the dependence of the lexical and morphological features studied in [16].

REFERENCES

[1] R. A. Huber and A. M. Headrick, “Handwriting Identification: Facts and Fundamentals”, CRC Press, 1999

[2] Ron N. Morris, “Forensic Handwriting Identification”, Academic Press, 2000

[3] R. Plamondon and S. N. Srihari, “On-line and off-line handwriting recognition: a comprehensive survey,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 63–84, 2000.

[4] 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, Sep. 2008.

[5] G. Louloudis; B. Gatos.; N. Stamatopoulos; A. Papandreou., "ICDAR 2013 Competition on Writer Identification," Proc. ICDAR 2013, August 2013, pp.1397-1401.

[6] F. Slimane; S. Awaida; A. Mezghani; M.T. Parvez; S. Kanoun; S.A. Mahmoud; V. Margner, "ICFHR2014 Competition on Arabic Writer Identification Using AHTID/MW and KHATT Databases," Proc. ICFHR 2014, September 2014, pp.797-802.

[7] M.I. Malik; M. Liwicki; L. Alewijnse; W.Ohyama; M.Blumenstein; B. Found, "ICDAR 2013 Competitions on Signature Verification and Writer Identification for On- and Offline Skilled Forgeries (SigWiComp 2013)," Proc. ICDAR 2013, August 2013, pp.1477-1483.

[8] A. Parziale, A. Santoro, A. Marcelli et al., “An Interactive Tool for Forensic Handwriting Examination,” Proc. ICFHR 2014, pp.440-445, September 2014

[9] V. Papavassiliou, T. Stafylakis, V. Katsouros and G. Carayannis, “Handwritten document image segmentation into text lines and word,” Pattern Recognition, vol. 43, Jan. 2010, pp. 369-377.

[10] A. Parziale, C. De Stefano and A. Marcelli, “Segmenting Isolated Characters within Cursive Words,” Proc. IGS 2011, June 2011, pp. 156-159.

[11] L.P. Cordella, C. De Stefano, A. Marcelli and A. Santoro, “Writing Order Recovery from Off-Line Handwriting by Graph Traversal,” Proc. ICPR 2010, August 2010, pp. 1896-1900.

[12] C. De Stefano, G. Guadagno and A. Marcelli, “A saliency-based segmentation method for on-line cursive handwriting,” Int. J. Patt. Recogn. Artif. Intell., vol. 18, November 2004, pp. 1139-1156

[13] R. Senatore and A. Marcelli, “Where are the Characters? Characters Segmentation in Annotated Cursive Handwriting”, Proc. IGS 2013, June 2013, pp. 171-174.

[14] Kullback, S.; Leibler, R. A. On Information and Sufficiency. Ann. Math. Statist. 22 (1951), no. 1, 79-86. doi:10.1214/aoms/1177729694.

[15] Masquerade, Natural Intelligent Tecnologies S.r.l., www.nitesrl.com

[16] M Diaz-Cabrera, MA Ferrer, A Morales (2015) Modeling the Lexical Morphology of Western Handwritten Signatures. PLoS ONE 10(4): e0123254.