

Handwriting and Signature : One or Two Personality Identifiers ?

*V. Boulétreau, **N. Vincent, ***R. Sabourin & *H. Emptoz

* Laboratoire RFV
INSA de Lyon
vboul@rfv.insa-lyon.fr

** Laboratoire LI - E3i
Univ. F. Rabelais
vincent@balzac.univ-tours.fr

*** Laboratoire LIVIA
Ecole de Technologie Supérieure
sabourin@gpa.etsmtl.ca

Abstract

Handwriting and signature are often studied without any connection. In this paper, we present a method applied both to handwriting and signature classification that is based on their fractal behavior. First is presented the method we have developed for the computation of the fractal dimension and the secondary dimension of writing. We describe how these parameters allow us to define a pertinent representation space. We also show how this approach has permitted to extract classes related to writing and signature styles. Lastly, this method has allowed us to give evidence of the independence between the behaviors of the writer when he signs and when he writes. Such an independence will be a source of very enriching information within the context of signature authentication.

Introduction

For both automatic recognition and signature authentication, the elaboration of a universal processing system comes up against a major obstacle : the variability of the writing and of the signature styles. As any human production, handwriting is subject to many variations from very diverse origins : historic, geographic, ethnic, social, psychological,... For a few years, researchers have agreed that it is necessary to reach a better knowledge of writing, in its patterns, of course, but also in the psychological and motor mechanisms it involves.

Within the context of automatic recognition, one of the approaches described consists in establishing a handwriting classification. The constitution of adapted learning sets or the choice of recognition tools for each of the extracted classes should allow to increase significantly the recognition rates [1-2].

Besides, if it were possible to establish a first rough classification permitting to automatically distinguish North American from French, Arabic or Asiatic signatures, adaptation the authentication process to each of the defined classes should be interesting.

An approach based on the study of the fractal behavior of handwriting has been proposed by N. Vincent [3]. We will show that under certain conditions of observation, the

fractal parameters we define are stable and discriminant enough to establish a handwriting classification according to styles. As well, we will show that it is possible to apply these parameters to the study of signatures. Finally, we will make evident the independence of the writers behavior when they write and when they sign. This independence will permit to describe more completely the writer himself.

In some kinds of documents as bank checks, we have both a written text and a signature. The combined study of these two kinds of writing will be a source of complementary information.

1. The fractal dimension

1.1. Definition

The fractal dimension, as defined by B. Mandelbrot [4], is a « number which measures the degree of irregularity or of fragmentation of a set », or the measure of the complexity of the studied set. The application of the tools provided by the fractal geometry to handwriting images permits to provide an interesting qualitative and quantitative description of them.

The fractal dimension calculus we apply lays on the measure of the Minkowski-Bouligand dimension. In the plan, the Minkowski-Bouligand dimension of a X set is :

$$D(X) = \lim_{\eta \rightarrow 0} \left(1 - \frac{\log[A(X_\eta)/\eta]}{\log \eta} \right)$$

where $A(X_\eta)$ is the area of the optimal covering of X by balls of radius η .

For a fractal curve, the behavior of $\log[A(X_\eta)/\eta]$ versus $\log(\eta)$ is **linear**. Then, the limit we search takes the value $D(X) = 1 - p(X)$ where $p(X)$ is the slope of the graph we have called **evolution graph** :

$$(x = \log \eta, y = \log[A(X_\eta)/\eta])$$

The construction of such a graph requires the calculus of the area of the covering of X by balls with radius η . We have chosen to dilate X η times by a ball with radius 1. Figure 1 shows the look the of evolution graph associated to a handwriting image.

It is possible to distinguish 3 zones with different slopes in this graph, characterizing a particular behavior and corresponding to a particular scale of observation.

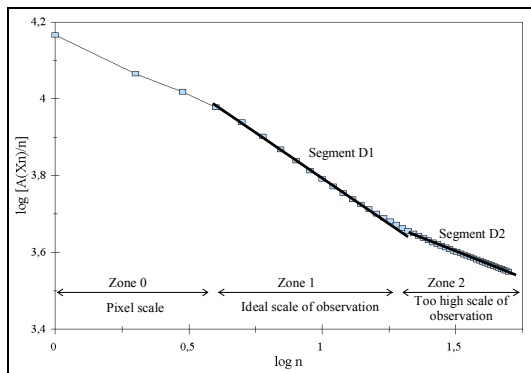


Figure 1 : Appearance of an evolution graph - Segmentation.

The slope of these different zones will allow us to define various parameters among which :

- **The fractal dimension of handwriting (FD)** computed from the slope of zone 1. This part of the graph matches with the dilations for which the text semantic content is apparent. The FD is thus a parameter describing the perception one can have of a handwriting from an adapted distance of observation.
- **The secondary dimension (D2)** is calculated from the slope of zone 2. This zone of the graph corresponds to the values of η from which the text is hidden by dilations. D2 describes the perception one has of the same writing observed from a more remote point of view.

The analysis of handwriting which we propose to achieve and the classification we establish are relying on these two parameters. Studies have been carried out in order to define which phenomena are susceptible to affect the values of the parameters and how we can manage to get free of such variations.

1.2. Stability of the fractal dimension of handwriting

We have first tested the stability towards physical constraints linked to both writing and acquisition [5]. We have verified that the use of various writing tools did have an influence only on the very first points of the evolution graph (zone 0) that we do not consider for the computation of the FD. We have demonstrated [6] that resolution changings induce a translation of the different zones of the graph but do not bring any modification in the slopes of zone 1 and 2 from which are computed our parameters. Besides we have shown that this result did make sense only for binarisation resolution upper than 200 dpi. Under this resolution, the image of handwriting is sensible to binarisation thresholds ; even if the semantic content is perceptible, the measures we compute from images binarised at a resolution lower than 200 dpi are not representative of the writing.

We have shown that to provide reliable results, the computation of the FD of handwriting has to be achieved from a text of at least 4 or 5 words. Actually, the FD computed from an only word would be more representative

of the grapheme than of the writing ; basing our computation on a sufficient quantity of text, we have managed to reach a measure independent from the analyzed text.

We have also looked into the stability of the FD in the time. This analysis is based on the study of writings taken from 48 post graduate students at the rhythm of once a week (about 1500 images). The one way analysis of variance and repeatability tests did show that our parameters did constitute a measure stable in time, and that they did have a satisfactory discriminant power.

2. A classification of writing

2.1. Handwriting classification

FD and D2 have allowed us the definition of the space in which we can establish our classification : if we represent handwritings in the plan, FD vs. D2, we notice their distribution is linked to legibility. We call this representation the **legibility graph**. Using methods of classification by aggregation, we did extract different classes characterized by particular writing styles [7] as schematically shown in figure 2.

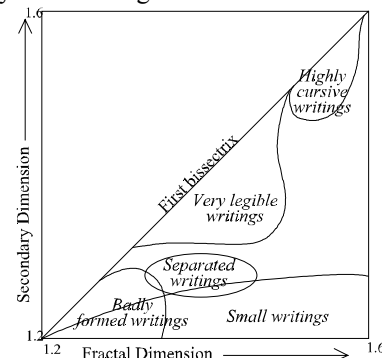


Figure 2 : schematic representation of the legibility graph.

We have chosen a quite low number of classes thus each of the extracted classes corresponds to a particular writing style for which it is reasonable to look for specified process and/or learning sets. A finer classification is possible, but the stability of the different classes would be clearly lower. Lastly, a multiplication of the classes would bring more processes or learning sets which is not necessarily efficient.

2.2. Application to the study of signatures

The study of signatures is a very particular field of pattern recognition. First because the objective is different, also because signatures themselves constitute very particular writing. A signature is characterized by two aspects : one, conscious, is the pattern of the signature, the other, unconscious, leads to the series of spontaneous movements constituting the plotting. A study achieved in collaboration with neuropsychologists has shown [6] that the fractal parameters were strongly correlated to parameters used for more cognitive and perceptive

approaches of writing. The fractal parameters we have developed for the study of handwriting could thus bring a great richness to traduce this double aspect conscious / unconscious of the signatures. Let us note that in the case of the study of signatures, the quantity of text considered for the computation of the FD is no more a problem since the analyzed pattern is always the same. Next figure shows the different classes of signatures we have extracted.

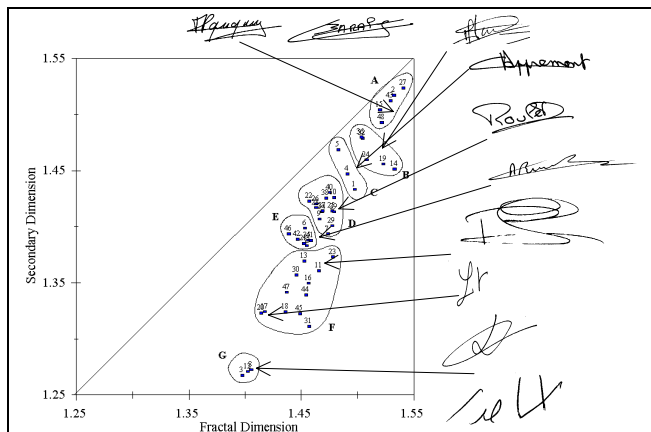


Figure 3 : Families of signatures.

These classes, defined in the legibility graph, really correspond to what may be qualified of families of signatures. To low values of FD and D2 (class G) correspond very simple signatures such as paraphs. The more complex signatures become, the more the dimensions increase. The class in the center of the graph matches with signatures not too overelaborated, which FD and D2 are quite similar to the average values for handwriting. Classes A, B and C group typical French signatures.

3. Combined study of handwriting and signature

This study is based on the comparison of 20 signatures and 20 writing of each of our 48 writers. Next table summarizes this comparison.

We compare three measures :

- the mean value obtained from all the images ;
- the between-class or between-writer variance ;
- the within-class or within-writer variance.

	FD Sig	FD HandWriting
Mean value	1.46994	1.4725613
Between-class variance	0.02481	0.00824
Within-class variance	0.00097	0.00035

First we see that the mean values taken by each of our parameters are only slightly different. Nevertheless, the difference of behavior of the fractal parameters for writing and signature is striking if we compare the variances :

- the between variance is clearly higher for signatures. This is the result of the great variety of signatures ; it is not surprising that the variation of the FD is important between a simple paraph and a very embellished.

- The within variance is also higher for signatures. In other words, despite the uniformity of the semantic content or more generally of the signature wording, the values taken by our parameters for signatures are more variable than that computed from writing. This large variability is actually compensated by the variability between-writers. We have shown, with statistical studies (repeatability test, analysis of variance), that the FD did constitute a suitable tool for classification of signatures.

Besides, we have calculated the linear correlation coefficient between the FD of writing and signature : its value is $r \approx 0.33$. The D2 are even less correlated since we obtain $r \approx 0.24$. The very low values of these two coefficients show that the fractal parameters for handwriting and signatures are not linearly correlated. This conclusion is indubitably only valid for French signatures. North American signatures should not present a so important independence towards handwriting, since they are usually constituted simply by the name of the writer, without any particular personalization.

We have finally compared the values taken by our parameters on samples of writing and signatures of the same writers. The independence of behavior of our parameters regarding writing and signature shows that, when it is possible, the compared study of both writing and signature of a same writer may be the source of interesting *complementary information*

Conclusions

In the field of authentication of documents, *e.g.* deeds, using these two sources of information shall indubitably be helpful. Last, the quantity of text represented by the literal amount of bank checks is enough to obtain a stable FD, the signature authentication should then also be enriched by a knowledge of the writer's writing, under the condition he has written the amount himself, and it is not obvious.

Bibliography

- [1] Hull J.J., Ho T.K., Favata J., Govindaraju V. & Srihari S.N. Combination of segmentation-based and wholistic handwritten word recognition algorithms. In: *From pixels to features III, Frontiers in handwriting recognition*, S. Impedovo an J.C. Simon. North Holland: Elsevier Sciences Publisher B.V., 1992. p. 261-272.
- [2] Crettez J.-P., Gilloux, M. & Leroux, M. A set of handwriting families : style recognition. *ICDAR'95, Montréal, August 16-19 95*. Vol. 1, p 489-494.
- [3] Vincent N. & H. Emptoz H. A classification of writings based on fractals. In: *Fractal Reviews in the Natural and Applied Sciences*. M. M. Novak. London : Chapman & Hall. pp 320-331. 1995.
- [4] Mandelbrot B. *Les objets fractals*. Flammarion 1975, 203p.
- [5] Boulétreau V., Vincent N. & Emptoz H A writing qualification invariant towards line thickness and resolution changings. *Proc. of ACCV'95, Singapore, December 5-8 1995*. Vol. I, p. 325-329.
- [6] Boulétreau V. *Vers un classement de l'écrit par des méthodes fractales*. Thèse de doctorat 1997. 196 p.
- [7] Boulétreau V., Vincent N., Sabourin R. & Emptoz H. Synthetic parameters for handwriting classification. *Proc. of ICDAR'97, Ulm, August 18-20 1997*. Vol.1, p.102-106.