Writer Identification based on the fractal construction of a reference base

A. SEROPIAN[1], M. GRIMALDI[1], N. VINCENT[2]
[1]Laboratoire d'Optique Appliquée - Université de Toulon et du Var, Bâtiment R
BP 132 83957 La Garde Cedex – France
[2]Laboratoire d'Informatique Université de Tours , 64, avenue Jean Portalis,
37200 Tours - France
E-mails : seropian@univ-tln.fr, grimaldi@univ-tln.fr, vincent@univ-tours.fr

Abstract

Our aim is to achieve writer identification process thanks to a fractal analysis of handwriting style. For each writer, a set of characteristics is extracted. They are specific to the writer. Advantage is taken from the autosimilarity properties that are present in one's handwriting. In order to do that, some invariant patterns characterizing the writing are extracted. During the training step these invariant patterns appear along a fractal compression process, then they are organized in a reference base that can be associated with the writer. This base allows to analyze an unknown writing the writer of which has to be identified. A Pattern Matching process is performed using all the reference bases successively. The results of this analyze are estimated through the signal to noise ratio. Thus, the signal to noise ratio according to a set of bases identifies the unknown text's writer.

Keywords : identification, fractal compression, extraction of invariant patterns, pattern matching.

1. Introduction

The writer identification problem is a very old problem that is actual from the birth of writing. In our environment, the need to authentify a handwritten text is a recurrent problem in the world of law as well as in the medical field where the prescription has to come from an authorized personality.

The writer identification problem has to become an important element in the domain of handwriting recognition. Actually, the important variability of writings makes specially difficult the problem of recognition. In the case of a mono-writer approach some solutions exist. A process that would identify the concerned writer, or his style, could lead to a multi-writer approach that could automatically adapted to each individual.

Several studies are on to solve the problem. The styles can be defined by characteristic elements of local patterns that frequently appear in the concerned writing [1]. J.-C. Simon refers to characteristics of the writing as invariants.

These invariant elements have got miscellaneous natures [4], geometrical (loops, straight vertical lines, etc.) and/or topological (crossing points, extreme points, etc.). L. Heutte in [3] has already handled the specific properties of each writing using some attributes that can be extracted such as line slope or the number of connected components of the text outline. N. Vincent in [7] and [8] has developed more global approaches, proving writing images have a fractal behavior. The fractal dimension comes out to be a robust parameter, constant for each writer

The extraction of invariant patterns that has been realized in [5] from the writing image itself is a major progress in the identification process. The author is assuming that invariant elements are characteristics that are proper to each writer. This technique is quite similar to the way the graphology expert works.

In this study, we are interested in the writer identification process. It will lead to obtain some information on each individual. From a handwritten text, we are to extract some characteristics, i.e. some invariant patterns. The method implemented is derived from fractal compression in order to take into account the autosimilarities within an image and also from pattern matching to relate these inner similarities to other types of handwriting.

In the first part, a recall about fractal compression and decompression processes is done. In the second part, the invariant patterns in the writing are extracted. These invariants are used to characterize the writing. In the third part, the similarities are used to identify a new text's writer. The last part presents the results.

2. Fractal Compression

2.1. Theoretical recall

Fractal compression is a technique that has been developed by Y. Fisher [2]. Its basic principle is to try and consider a given image I as the fixed point of a geometrical transform T. Most often, the transformation T is complex and the image is defined as the attractor of an iterative function system (IFS). The fixed point is obtained as the limit of an image sequence $I = \lim_{n \to \infty} T(I_n)$ that is iteratively defined by $I_{n+1} = T(I_n)$. Here, the problem of compression is to get back from a known image, a system of transforms that would precisely admit this image as its fixed point.

Before presenting the compression process in itself, the definition of an IFS has to be described with more details.

2.1.1. IFS (Iterative Function System) and PIFS (Partitioned Iterative Function System). An IFS is a set of geometrical elementary linear or affine contractive transformations that allows to generate fractal images.

These transformations make possible the definition of a function T. Its action on an image I is calculated according to the equation :

$$T(I) = \bigcup_{i=1}^{n} T_i(I)$$

The image that is obtained has some specific properties. In particular it is autosimilar: it is made of copies of itself entirely (cf. Figure 1) but modified by the transforms. In the following example, the system is made of a reduction, followed by a repositioning in a triangle shape.

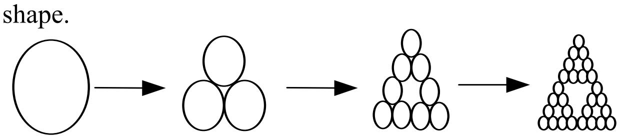


Figure 1. Generation of fractal image

In order to adapt this method to natural images, the transformations T_i are limited to applications highlighting similarities between parts of the image. Then, the chosen model is a partitioned iterative function system (PIFS). A PIFS defines a transformation T that is the union of affine contractive transformations of the PIFS :

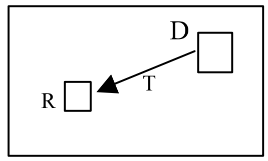
$$T(I) = T_1(I_1) \bigcup T_2(I_2) \ldots \bigcup T_n(I_n)$$

where the set of all images obtained from all the transforms of sub-images I_i enables to partition the spatial domain of I.

So, The fractal image is the attractor of the IFS and can be derived from any image. And as far as real images are concerned, PIFS are used.

2.1.2. Fractal compression. The aim of compression step is to determine the transformations that are part of a PIFS having the initial image *I* as fixed point. To construct the PIFS, the image is partitioned into sub-images R called Ranges. These Ranges have to be interpreted as a result of a geometrical affine contractive transformation of Domains D with a T_i Transformation. These Domains have to be themselves sub-images of the initial image. In a usual way, the Domains D are twice the size of the corresponding Ranges to be sure to define contractive T_i. We have : $R = T_i(D)$.

From a practical point of view, the parts R of the image with the transforms are approximated by minimizing the distance between R and T_i(D). There the usual metric is used.



R: a Range
D: a Domain
T: the transform by which the image of D is R

Figure 2. Principle of the fractal compression

The search of the best transformation is limited to affine transforms with the following analytical definition :

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & s \end{bmatrix} * \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} e \\ f \\ o \end{bmatrix} \qquad (1)$$

The parameters (x,y) indicate the coordinates of a image pixel and z its gray level. The coefficients *a, b, c* and *d* determine the geometrical spatial transformation, *e* and *f* determine the translation, *o* the contrast and *s* the luminosity. It has been proven that in order to obtain a set of contractive transformations, the parameter s has to be strictly less than 1.

Fractal compression process replaces the image by the system of transformations that are each defined by 8 parameters and the position of associated R and D.

2.1.3. Fractal decompression. In the decompression step the transformations are iteratively applied to all sub-images of any image till the fixed point is obtained. It is assumed to be obtained when the difference between two successive images of the sequence is small enough. To quantify the quality of the fractal compression of an image, beside the compression ratio, the peak signal/noise ratio is generally used. For an image that is coded using 256 gray levels, that comprises n pixels and if we note z_i and z'_i the gray levels of pixel i respectively in the initial image and in the decompressed image, then:

$$PSNR = 10 * \log\left(\frac{255^2}{rms}\right) \text{ with } rms = \frac{1}{n}\sum_{i=1}^{n}\left(z_i - z'_i\right)^2$$

(2)

This principle is the starting point from which the concept of our method of writer identification is derived. We will show how a base constituted from invariant elements can be built for each writer. The transformations defined in the compression process are to be used.

2.2. Writing application

The handwriting samples are scanned and images are defined in gray levels so that all possible information is retained. More than the compression process itself, we are going to use the properties of the inner similarities brought up in the construction of the PIFS. The images to be processed are all in 256 gray levels (from 0 to 255) to visualize the details of the styles more precisely. Some details can appear in the gray level image, where as it is not the case in binary images.

The partition of the image for the fractal compression process could be done in different ways (Quadtree, Delaunay triangulation …). Indeed, the occidental writing admits privileged directions. To adapt the fractal compression to the handwriting, the vertical and horizontal directions are privileged. So, the Ranges as well as the Domains are chosen as square areas the sizes of which are to be determined according to the desired precision and to the resolution of the image.

It can be noted, a decompressed image is considered quite good when the estimated PSNR is near 30dB and excellent with 35dB [2]. Such a quality measurement is a criterion that can make us confident about the hypothesis we are to consider.

3. Writer style learning

In this section, the possibility to extract the invariant elements form handwriting [7] is demonstrated. We consider a handwritten text image. The method is relying on the inner similarities of the handwriting. These elements are extracted during the fractal compression process. The search for the transformations of the PIFS is in fact a search for some inner similarities. Then they are revealed in the handwritten text.

During the compression step, the objective is that all the Ranges can be obtained from Domains of the image as the result of a transformation. The best transform corresponds to the transform that minimizes the RMS (Root Mean Square) between the two R and $T_i(D)$. For each Range, the memorized parameters are the position of the associated Domain and the transformation that minimizes the criterion. The Domains are chosen among all sub-images of appropriate size contained in the image.

Thus, for a writer, a reference base B_s is made of the set of Domains that have just been extracted. They contain some characteristics of the writer. They have been selected during the compression phase in a text used as a learning reference. This text has to be long enough to be significant of the writer style.

In a compression process, the number of selected different Domains increases when the length of the text is increasing, but the increase is not linear and, for each writer, the number of different Domains in the reference base tends to stabilize. We are to suppose the set of selected Domains contains the whole information concerning the patterns present in the writing of a writer. Actually these Domains are sufficient to generate the image of the learning text writing. Indeed they are the only information in the decompression process with the associated transform.

It must be noticed the reference base doesn't depend on the content of the learning text. These texts could be different for each writer. It is essential that each base includes all the invariant aspects of the writer.

This reference base represents all the inner similarities contained in the writing. In order to characterize the writer as best as possible, we have chosen to retain in the reference base only the most representative of the inner similarities of the writing. That is to say the autosimilarities the most exact between ranges and the transformed Domains. The precision is measured by a quadratic error. It is not reasonable to keep all the Domains (or the Ranges) when they are not well connected. So, we keep only 80% of Domains to build the reference base of a writer. Besides, for each writer, a threshold corresponding to the rate of acceptable similarities is defined. Afterwards, it is noted ε_s.

When looking for the possible transforms to best associate Ranges and Domains, we have limited the search to only one transformation: the enlargement. Besides, the choice of a partition of the image induces the geometrical shape of the sub-images that constitute the Ranges and the Domains. At this stage we have chosen some square windows and a contraction ratio of ½.

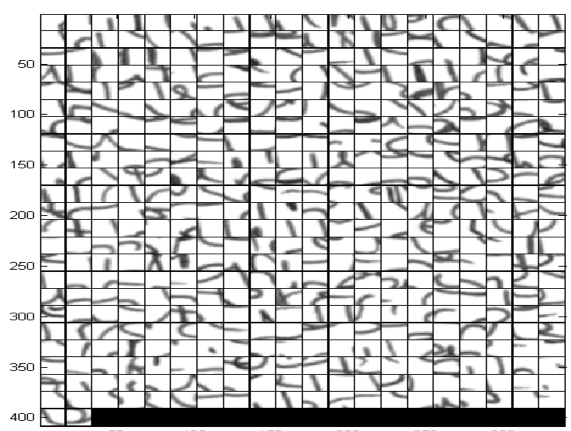


Figure 3. Example of reference base

4. Writer Identification

Of course, in order to identify the writer from a new handwriting text, the learning phase concerning the writer handwriting must have occurred and the corresponding reference base must have been stored. This reference base has been established form a learning text. *A priori* we have got N reference bases.

It is obvious that the comparison between handwritings is possible only if the sizes of the writings are identical. Then, it is necessary to apply a normalization step on the writings. The height is linked to the choice of the dimensions for the fixed size of the range sub-images and have to be adapted to the usual details of the writing. The writings are normalized with respect to the height of the letter's body. More precisely, an enlargement is computed according to the vertical direction and applied according to both directions on the image.

4.1. Similarities research

The new text the author of which has to be identified is not necessarily quite as long as the text used during learning step. A solution could be to compare the quality of the images of the new text after fractal compression and decompression steps using each of the reference bases associated with all the known writers [6]. Here, we propose an other approach. Likewise in the fractal compression, the image is partitioned in the same way in Ranges. We use a pattern matching process between the text to be identified and the elements of the reference base. The text will be assigned with respect to the reference that allows the greatest number of correspondences. The quality of the correspondence is measured by comparing the initial text and the reconstruction of the image that is achieved by a pattern matching process.

The ranges have the same shape as the sub-images contained in the reference bases. So, it is not any more the inner similarities of the writing that are searched for but we are looking for the similarities that can exist between parts of the image and the elements of the reference base B. The Domains are chosen only in the writer base B_k that it is tested. Thus, with each Range R_i of the non identified image, a sub-image of the base B_k is associated when minimizing the square error criterion $[R_i - T(Dj)]^2$. Here again, the only transformation considered is the enlargement with ratio equal to ½.

For each writer, and for any reference base, only the reference sub-images that frequently appear in the learning phase are considered because they are those that most characterize the writer. So, in each base, only 40% of sub-images, are retained. They are the most frequent ones. Moreover the matches are not considered to occur when the similarity between the Range and the Domain in the base is not sufficient. The threshold was fixed at ε_s during the learning step according to the tested writer. That allows to limit the comparisons to the elements the most representative of the writer and therefore those contained in the base.

The principle of the writer identification step is showed in figure 4.

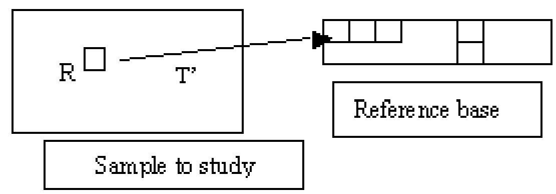


Figure 4. Similarities search with respect to the reference base

4.2. Reconstruction of the image

At the same time, while the similarities between the writing sample ranges and the reference bases are searched for, starting from a white image, the reference base sub-images that are most similar to the initial Ranges are copied at the appropriated location. If neither of the base Domains are found to be similar enough, the image will remain partially white.

The quality of the similarity between the initial image and the reconstructed image by the process is quantified using the PSNR parameter.

A change within the writing style results in the image reconstruction with a lower quality.

5. Application

We have worked with twenty different writers who provided samples of handwriting texts. Texts of long size were used for building the reference base and texts of smaller size to perform the writer identification.

Figure 5 shows two long texts handwritten by two different writers.

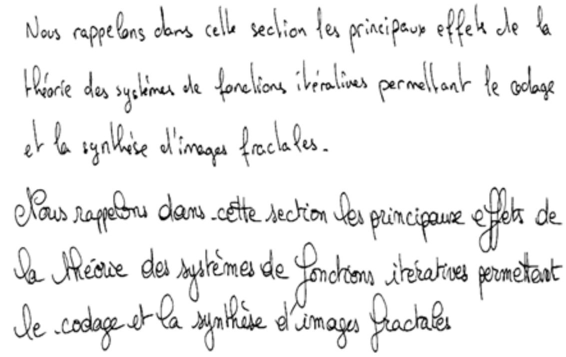


Figure 5. Two examples of texts used to build reference bases

The results that we present have been obtained with the use of Ranges of size 8x8. The Domains are size 16x16 pixels for the base construction. In this step, only 80% of base sub-images are taken into account for the most representative writer's elements. This percentage was chosen because it is a quite good compromise between the elimination of too common elements and the non elimination of elements that characterize the writer.

In the experiment, the calculated PSNR between the initial images and decompressed images takes values around 30dB.

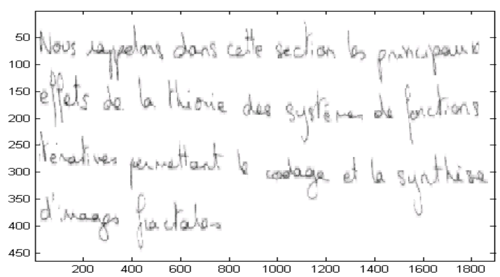


Figure 6. Example of decompressed image

The first test that we made allows to verify only that texts used in the learning step can be identified with no confusion as handwritten by his author. The identification rate is 100%.

With the new texts that have been handwritten. Some samples of texts used in the second part of the test are represented in figure 7.

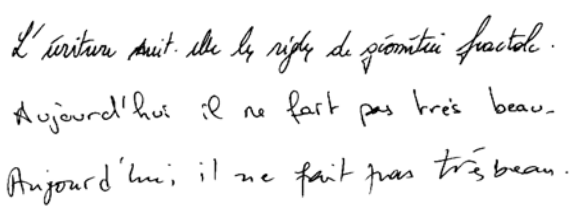


Figure 7. Examples of writing used in the identification process.

In the same way, the texts are normalized and using different reference bases, images are reconstructed with the help of similarities found in the reference bases. Then our identification process is applied with a threshold of acceptable similarity at 40%. The associated PSNR with each reconstructed image is computed and the best reference base provides the name supposed to be the writer of the handwritten text. We have got an identification rate superior to 85%.

6. Conclusion

Most of the problems are solved. Here a representation of writing in a finite dimension space is considered. The problem has received a response with an original approach. The image is treated as a whole taking into account inner characteristics of each writing, without any extraction nor details comparison.

The results are convincing, of course they have to be confirmed with a larger number of writers.

Besides, the application of the method can be extended out of the field of occidental texts.

At present the handwriting recognition process is the next study. This process will have to use the reference bases extracted during the learning step and then to adapt itself automatically to each writer.

7. References

[1] J.-P Crettez., "A set of handwriting families : style recognition", International Conference on Document Analysis and Recognition (ICDAR'95), Montreal (Canada), 1995, pp. 489-494.

[2] Y. Fisher, "Fractal Image Compression : Theory and Application", Springer-Verlag, New York, 1995.

[3] L. Heutte, T. Paquet, A. Nosary, C. Hernoux, "Handwritten text recognition using a multiple-agent architecture to adapt the recognition task", 7th International Workshop on Frontiers in Handwriting Recognition, IWFHR VII, Amsterdam, LRB Schomaker & LG Vuurpijl (Eds.), pp. 413-422 (2000).

[4] K. Khan, I.K. Sethi, "Handwritten signature retrieval and identification", Pattern Recognition Letters, 1996, vol.17, pp. 83-90.

[5] A. Nosary, L. Heutte, T. Paquet and Y. Lecourtier, "Defining writer's invariants to adapt the recognition task", International Conference on Document Analysis and Recognition (ICDAR'99), Bangalore (India), 1999, pp. 765-768.

[6] A. Seropian, N. Vincent, "Writers Authentication and fractal Compression", 8th International Workshop on Frontiers in Handwriting Recognition, IWFHR VIII, Niagara-on-the-Lake (Canada), 2002, pp. 434-439.

[7] N. Vincent, S. Barbezieux, "Compression of handwriting images : a way to define a writing style", Vision Interface, 1998, pp. 347-354.

[8] N. Vincent, V. Bouletreau, R. Sabourin and H. Emptoz, "How to use fractal dimensions to qualify writings and writers", Fractals, World Scientific, 2000, Vol 8, n°1, pp. 85-97.