On defining signatures for the retrieval and the classification of graphical drop caps

Rudolf Pareti[1], Surapong Uttama[2], Jean-Pierre Salmon[3]
Jean-Marc Ogier[2], Salvatore Tabbone[3], Laurent Wendling[3], Sebastien.Adam[4], Nicole Vincent[1]
1 Laboratoire CRIP5, Université de Paris 5,
{rudolf.pareti},{nicole.vincent} @math-info.univ-paris5.fr
2 Laboratoire L3i, Université de la Rochelle, Avenue Michel Crépeau
17042 La Rochelle CEDEX, France
{uttama.surapong},{jean-marc.ogier}@univ-lr.fr
3 LORIA, INRIA, B.P.239
54506 Vandoeuvre-lès-Nancy CEDEX, France
{jean-pierre.salmon}{antoine.tabbone}{laurent.wendling}@loria.fr
Laboratoire PSI, Université de Rouen, UFR Sciences
76821 Mont Saint Aignan CEDEX, France

Abstract

Ancient graphical documents are invaluable heritages which have been handed down since generations. They possess both intellectual and spiritual worth for humanity. In this context, many digitization processes have been started, producing very large warehouse of images. These huge amount of data raise the problem of indexing the information in order to make easier navigation in the databases. In the context of a French research program, called MADONNE, this paper proposes a set of complementary contributions concerning ancient graphic images indexing.

1. Introduction

Ancient graphical documents are invaluable heritages which have been handed down since generations. They possess both intellectual and spiritual worth for humanity. Unfortunately, due to their ages and publishing technology, they decay with time. As a result, they are very rare and difficult for public access. The governments and international organizations recognized this problem and attempted to preserve, restore and make these ancient documents available for interested communities.

In domain of document image analysis, ancient documents are digitized and studied in various disciplines. This paper, however, focuses on the difficult problems of classification and indexing of graphical parts especially the graphical drop cap namely lettrine in French (see Figure 1). The results that are proposed in this paper are issuing from a large collaborative research project, called MADONNE[1], that involves several French laboratories (L3i laboratory- University of La Rochelle, LORIA in Nancy, PSI lab - University of Rouen, the LI lab - University of Tours, the LIRIS lab - Lyon, the IRISA - Rennes, and the CRIP5 - University of Paris 5)

MADONNE project aims at designing methods for going beyond plain digitization projects for historical documents, as such, all the projects tend to yield large databases of images with very little structure for navigating and indexing these databases. The project investigates the use of document image analysis methodology for providing useful browsing and indexing features in these large collections.

In this context, the works presented in this paper focus on the results of a sub-group in MADONNE's project, the aim of which is to provide services on graphic objects.

The graphic objects on which our consortium started to work on correspond to drop caps, because of a set of requirements proposed by a historical center from the University of Tours called CESR[2]. However

[1] MAsse de DONNEes : http://l3iexp.univ-lr.fr/madonne/
[2] Centre d'Etudes Supérieures de la Renaissance :

all the concepts and results presented in this paper appear to be completely generic and may be re-used in another application context.

This paper aims at proposing a set of operational tools to classify, and index ancient graphics especially drop caps by using different image modelling techniques, signature computations, and selection of relevant primitives

This paper is organized as follows: section 2 refers to the general problem of drop caps image characterization by using two complementary descriptions -signature-, while section 3 tackles the problem of selecting relevant primitives. Section 4 demonstrates the experimental results. In section 5 we discuss concerning the complementarity of presented systems. The conclusion and perspective works are presented in section 6.

2. Images, specificities and features

2.1 Observation

Generally, a drop cap is printed by a handmade wood block carved with different strokes and patterns. Thus each drop cap is unique in quality and property. The patterns found in drop caps are normally parallel lines or curves and hatches without definite boundaries. This leads to main difficulties for segmentation and classification processes. There are different styles of drop caps, depending on the period it has been produced at, the author … We illustrate in Figure 1 a set of drop caps, showing the variability of representation as well as the variety of the information shared in the drawing.

The research problems dealing with navigation into graphic image databases are very numerous. Indeed, one problem may correspond to develop slight signatures allowing to perform drop caps spotting, like hyper-text navigation (Figure 2). In this case, the problem is solved by developing text-graphic separation and implementing drop cap detection algorithm.

Another problem suggested by historians is to categorize drop caps as a function of their graphic style, allowing thus to group them by period, or by author for instance. Of course, some research questions are relative to the Uppercase recognition problem. But one of the most difficult problem is to peform content based image retrieval, allowing to find a drop cap or a part of the drop cap among a large number of images (graphic image query) (Figure 3). An other problem would be to detect failure in the drawing and to track such a characteristic in the related future works.

At present we are working on initial letters that have been extracted form full digitized documents after a text-graphic separation step, followed by an initial letter recognition process based on spatial rules between text and images.

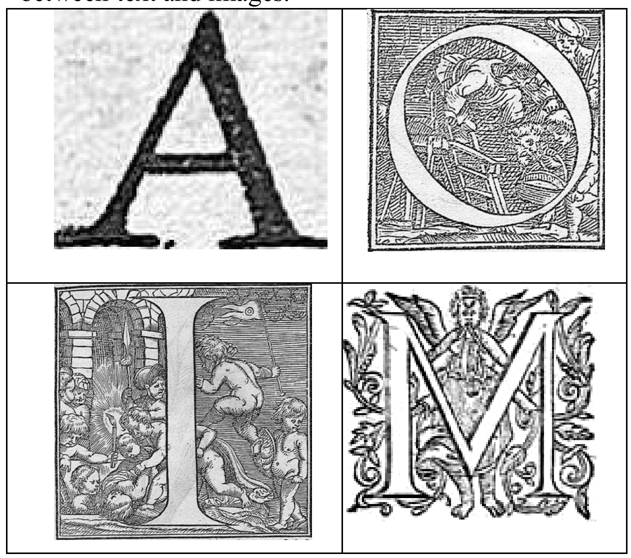


Figure 1 : Illustration of the variability of representation of drop caps.

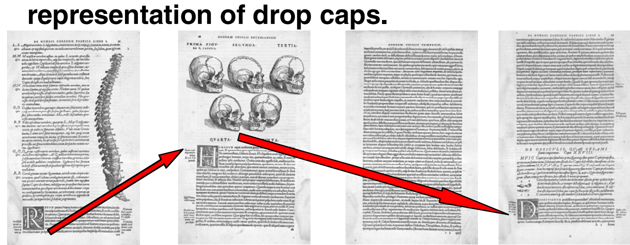


Figure 2 : Drop cap spotting.

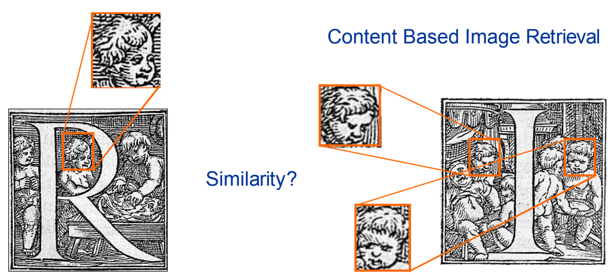


Figure 3 : Content Based Image Retrieval.

2.2 images characterization

The research problems deal with integrating graphical features into the general indexing and browsing schemes. For this, we have to characterize the different kinds of graphical components which are present in such documents, to design adequate segmentation tools to localize them, and to work on

http://www.cesr.univ-tours.fr/

relevant features and signatures which can be used to perform indexing on these features.

In our context, the MADONNE cooperation aspects deal with the definition of relevant signature in order to classify/index drop caps databases. It also concerns the problem of selecting relevant features when dealing with a specific problem, on the basis of the results of the classification process.

The litterature is very rich in the domain concerning image based signature computation, but most of the classical techniques used in Content Based Image Retrieval (CIBR) are not usable in our context, because of the features of the image that are so specific that classical computer vision approaches are not usable (stroke images).

In this paper, we focus on different and complementary contributions issuing from the MADONNE consortium. These works are regularly "synchronized" through meetings between our research teams, allowing to merge our works.

In the following part, we will focus on the problem of image characterization through different approaches (CRIP5 and L3i works) allowing to compute relevant signatures for indexing. Then we will present one manner to combine these signatures in order to optimize the classification/indexing process (LORIA's work).

3. Characterization of the image

Our first approach is a statistical and global approach relying on no hypothesis. We are not looking for a specific shape that would be too problem dependent but we are studying the inner structuration of the grey level pixels within the image. The structure deals with patterns that are not predefined but that appears thanks to their frequency.

3.1. Zipf law based strategy

Among the numerous laws that have been exhibited in nature, gaussian and linear laws are those that have been most studied in a theoretical ways. Nevertheless modeling nature needs more powerfull tools. Here we have made use of power laws.

3.1.1 The Zipf Law

Zipf law is an empirical law expressed fifty years ago [1]. It relies on a power law. The law states that in phenomena figured by a set of topologically organized symbols, the distribution of the occurrence numbers of n-tuples named patterns is organized in such a way that the apparition frequency of the patterns M_1, M_2, …, M_n, noted N_1, N_2, .., N_n, are in relation with rank of these symbols when sorted with respect to their occurrence frequency. The following relation holds:

(1) $$N_{\sigma(i)} = k \times i^a$$

$N_{\sigma(i)}$ represents the occurrence number of the pattern with rank i. k and a are constants. This power law is characterized by the value of the exponent a. k is more linked to the length of the symbol sequence studied. The relation is not linear but a simple transform leads to a linear relation between the logarithm of N and the logarithm of the rank. Then, the value of exponent a can be easily estimated by the leading coefficient of the regression line approximating the experimental points of the 2D graph ($\log_{10}(i)$, $\log_{10}(N_{\sigma(i)})$) with $i=1$ *to* n. Further on, the 2D graph is called Zipf graph. One way to achieve the approximation is to use the least square method. As points are not regularly spaced, the points of the 2D graph are re-scaled along the horizontal axis.

The validity of this law has been observed in many domains but rather for mono dimensional signals [2].

In order to study images, we are going to try and apply Zipf law to images and therefore to adapt the concepts introduced in the statement of Zipf law to two dimensional data.

3.1.2. Application of Zipf law to images

In this section, we are to point out some problems that may occur when images are concerned.

In the case of the mono dimensional data, the mask was limited to successive characters. When images are concerned, the mask has to respect the topology of the 2D space the data is imbedded in. We have chosen to use 3x3 masks as a neighbourhood of a pixel in a 2D space.

Then the principle remains the same, the number of occurrences of each pattern is computed. Nevertheless as 256 symbols are used to code pixels, there are theoretically 256^9 different patterns. This number is much larger than the number of pixels in an image. Indeed, if all patterns are represented only once, the model that is deduced from Zipf curve would not be reliable, the statistics would lose there significance. For example a 640x480 image contains 304964 patterns. Then it is necessary to restrict the number of perceived patterns to give sense to the model. The coding of the image is decisive in the matter.

Then several problems have to be considered. What are the properties we want to make more evident? How many classes of patterns are to be considered? This will be solved through a new encoding of the image.

According to the coding process chosen and to the image content, the Zipf curve general shape can vary a lot. When it differs from a straight line, the model of a power law is not suited for the global image modelling. Nevertheless, if several straight segments appear we can conclude several phenomena are mixed. Different codings allow to make more obvious different properties of an image. As a consequence, our idea is to apply this principle for drop cap classification and indexing.

3.1.3. In the context of drop caps

Some studies have shown a Zipf law was holding in the case of images with different encoding processes [3]. We are looking for some coding process that gives models apt at discriminating the images we are studying. This qualifies as effective a coding process. We are not concerned by details in the image but by a global view. Two images that look alike from a style point of view should give similar power laws. Indeed the nominal letter is not important and should not have any effect on the conclusion of the study.

The initial drop caps we are studying have been scanned as grey level images where each pixel is encoded with 8 bits (256 different levels). The intensity is the information encoded.

According to the remarks previously done, the number of different symbols must not be as high as the number of 3x3 patterns build from 256 grey levels. Two ways to achieve the decrease are possible, either the number of symbols used to characterize a pixel is decreased or the number of the pixels involved in the pattern is diminished.

Here our motivation is to respect the vision of a scene that relies more on differences of grey levels than on the absolute values.

A simple way to reduce the number of distinguished patterns is to consider only k grey levels to characterise the intensity level of the pixels. Most often it is sufficient to observe an image. A quantisation in k equal classes would led to unstable results, so we have chosen to use a classsification method of the grey levels in k classes by way of a k-means algorithm based on the grey levels. We have experimented different values of k. A study of the method efficiency with respect to the k value has been performed. If no hypothesis is done on the classes different grey levels give good results

An other way to decrease the number of different patterns is to limit the number of pixels in the pattern. To remain coherent with the 2D topology we can consider the smallest neighborhood of the pixel defining 4-connectivity. It is precised in Figure 4. Of course we omit some information when taking this one rather than the 3x3 window but confidence is increased then.

	3	
4	1	2
	5	

Figure 4: Mask shape indicated as cross mask.

In this case we can further and lower the number of grey levels involved. Taking into account the nature of the drawings, we have achieved a drastic decrease in the number of grey levels considered as we have considered only 3 of them. This number is in fact issued from the nature of the images we are working on. They are rather black and white images. A k-means with k equal to 3 has been used on each image. We talk here about crossmean method to refer this process. The number of possible patterns is therefore equal to 3^5=243, that is about the same as the initial number of grey levels but the information contained in the value is not the same. We are going to show the efficiency of the methods compared to the only use of an histogram. Besides, the k-means classification makes the method independent of the drop cap of the scanned image and is well adapted to a mainly black and white image. The next Figure 5 gives an example of built Zipf curve following our crossmean process.

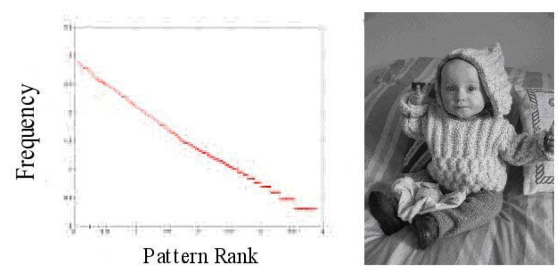


Figure 5 : Zipf curve associated with an image with respect to method.

A close look at the Zipf plots shows they are not always globaly linear, that is to say Zipf law does not hold for all the codings. It depends on the coding process. Nevertheless some straight line segments can be observed. According to the coding process used these zones can be interpreted. Here we are only concerned by the k-mean coding and the use of a neighborhood of each pixel. We have observed, with a k-means process, the left part of the graph is concerned with the regions in the image, they occupy more space in the image and then the frequencies are higher, whereas the right part gives information on the contours present in the image. Then, we can extract on the one hand some structure indication of the regions and on the other hand the structure of the contours within the images. If several objects are important in the image, the number of straight line segments increases in the Zipf plot.

Then we have chosen to consider in each curve up to three different linear zones. They are automatically extracted into three zones as shown in Figure 6 using a recursive process. The splitting point in a curve segment is defined as the furthest point from the straight line linking the two extreme points of the curve. The fact is that the image carries a mixt of several phenomena that are highlighted. Several power laws are involved and then several exponent values can be computed.

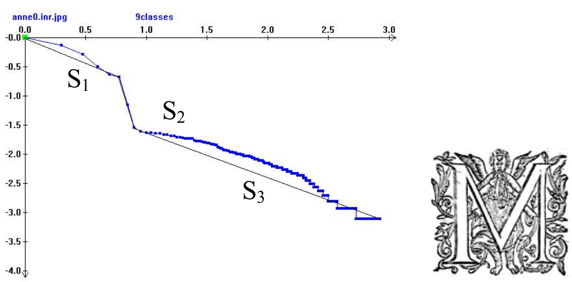


Figure 6: Example of an initial drop cap and its Zipf plot where are indicated the different straight zones extracted.

Then we have chosen to define the output of the process as 3 meaningful values associated with the picture (the drop cap of Figure 6 gives a representative exemple). The three values correspond to the 3 slopes, leading coefficients and are associated as the indexes with the image. The representation space is then a 3D space in which some distance computation can be achieved. As the domain of the 3 components are equivalent, no feature weighting is needed. The experiments and results of our approach dedicated to MADONNE project will be given in the section (4).

3.2. MST and PGA strategy

3.2.1. General presentation

Concerning the characterization of the drop caps, our other idea, that is complementary to the previous Zipf law's one, is strongly inspired from keypoints or salient zones detection (Schmidt, Harris detectors [4]).

The idea is to try to characterize the content of an image through the spatial organization of specific information. Actually, we consider the spatial organization of the different layers of information extracted thanks to an original segmentation stage as a good basis for the classification and the indexing approach.

In our context, at first, we experimented graphs for the description of the topology of the image and for the spatial organization description of the different layers. From this point, our current studies rely on the works presented in the last GREC conference [5], for which the topology of the image is based on the Minimum Spanning Tree Length (MST). Actually, the MST is computed on each layer containing relevant regions issuing from a consistent segmentation process.

Secondly, in order to compare this MST based approach with a relevant signature classically used in CBIR, and in order to reduce the complexity of the algorithm, we proposed to implement a signature representing the spatial relations of the regions issuing from the same segmentation process.

3.2.2. Segmentation Process

Our processing strategy to retrieve drop caps is based on a several stages, the first one consists in segmenting the drop cap images in different layers of information : textures layers, uniform areas, outline layers…. This operation is based on a strategy inspired from visual perception principles, that is summarized in Figure 7..

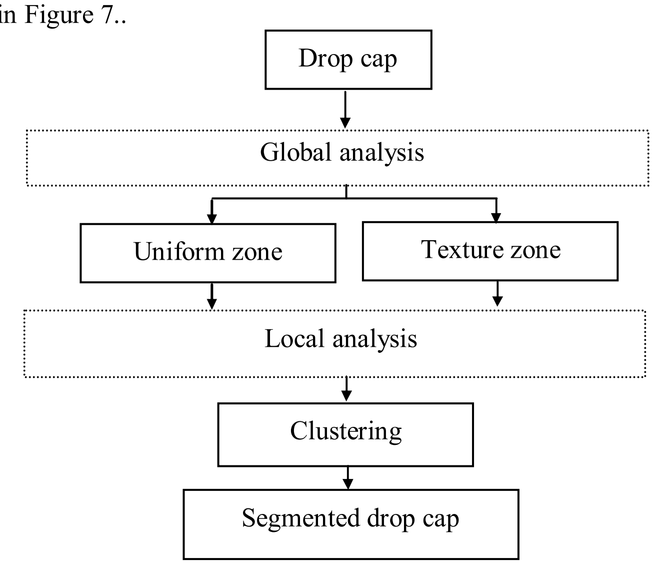


Figure 7 : General Scheme of a drop cap segmentation process

Actually, this segmentation stage is a first top-down technique, that will be followed by a bottom-up approach in our future works, in order to improve the quality of segmentation.

This segmentation process provides a set of layers characterizing respectively a category of information in the original image. As a consequence, for instance, this stage can provide a layer corresponding to textures,, a layer corresponding to homogeneous areas, a layer corresponding to outlines. Figures 8 and 9 illustrate an example of the homogeneous and textured areas computed from a drop cap image.

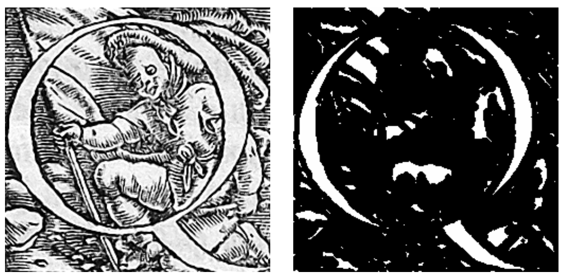


Figure 8. Original drop cap and its homogeneous areas

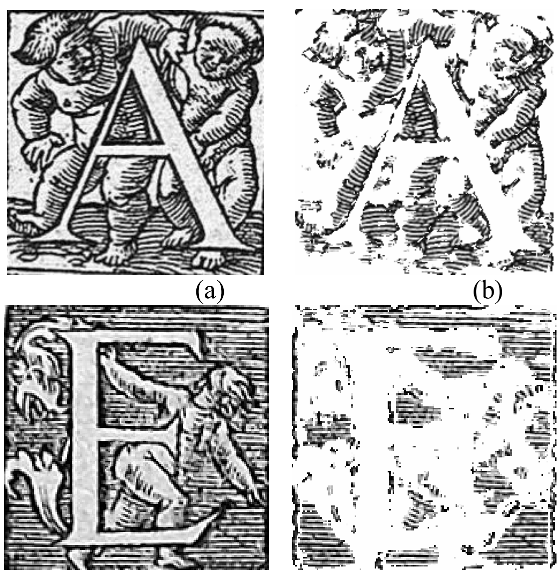


Figure 9. Textured background extraction of rather simple drop cap

3.2.3. Drop caps characterization

We focus on this part on the characterization of drop cap in order to permit to find a query drop cap among a wide set of images. For that purpose we have experimented to types of signatures: based on MST and PGA.

- **MST Based Signature**

The first idea for finding a signature of the drop cap is to consider the MST [6].as a relevant structure for representing the spatial organization of the drop cap's layers.

In order to reduce the complexity of the MST computation, and in order to apply it on a relevant set of information, the MST is applied on each of the layers issuing from the segmentation process. More precisely, it is applied on the set of regions issuing from the segmentation process.

Doing like this, a features vectors integrating the MST length computed on each layer can be computed, and this vector is considered as a relevant signature of the drop cap.

In order to reduce the complexity of the MST computation, that is a quite costly operation, we applied the MST on each connected component of each layer, considering the gravity center as a relevant information (Figure 10). As a consequence, we computed the MST on each of these connected components, as shown in Figure 11

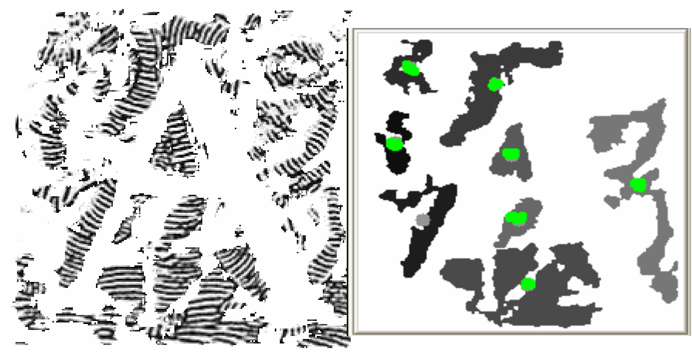


Figure 10 : Strategy of drop cap description

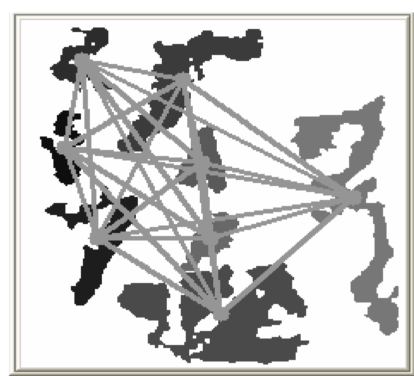


Figure 11 : Computation of the MST

The principle of this work is summarized in the following Figure 12. Based on the computation of the MST for each layer of information, we compute a features vector describing the spatial organization of the drop cap. This technique allows to compute a feature vector, the size of which depends on the number of layers extracted from the drop cap.

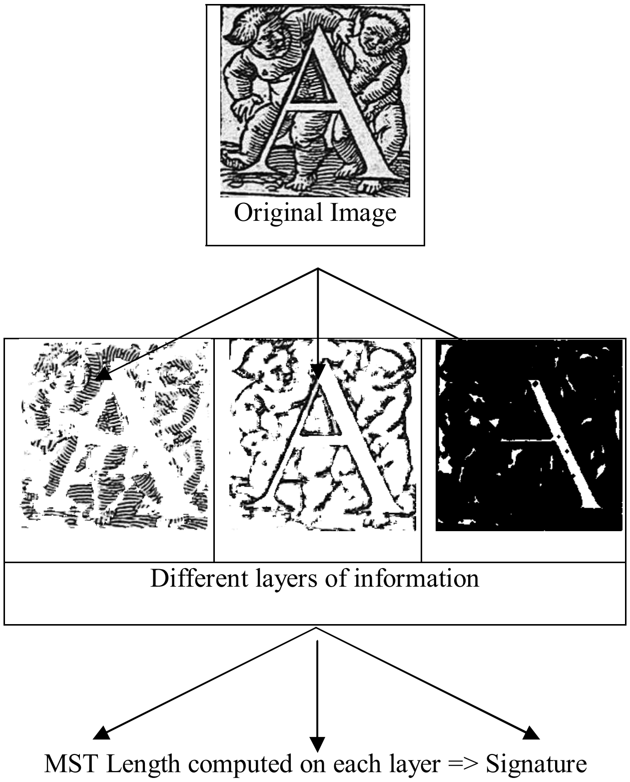


Figure 12 : Computation of the MST Based signature

Even if this kind of approach is very interesting, the computation of the MST is a quite expensive operation, and the MST Length computation is not a bijective operation that do not guarantee that it is the unique manner to describe an image. As a consequence, another approach, based on a relevant Pairwise Geometric Attributes signature issuing from CBIR systems, has been experimented. This complementary signature is computed from the different layers issuing from the segmentation process. The application of this technique in the context of MADONNE's project and the corresponding results are detailed in the experimentation part.

- **PGA Signature**

We have experimented a technique coming from CBIR system, proposed in 2003 by [7]. This approach takes as input the one as the MST Length strategy presented in the previous part.

The principle of this approach is to consider the relationships of each pair of connected components of the image, in terms of distance and orientation. In the original paper proposed by the authors, the input that were considered were vectors issuing from a vectorization process computed on graphic images. In our case, the strategy that is applied is based on the connected components of each layer issuing from the segmentation process, and more specifically, on the principal inertia axis computed on each of them.

The different stages of the algorithm are the following ones :

1/ Segmentation of the different layers (Figure 13 a)

2/ Computation of the principal inertia axis on each connected component of each layer (Figure 13 b)

(a) (b)

Figure 13 : Computation of the principal inertia axis on a segmented layer

3/ Computation of the relations between two principal inertia axis : Pairwise Geometric Attribute. This stage is the most "complicated" one of the process, even if it is quite simple. The next Figure 14 explains this computation. For each pair of the principal inertia axis, it consists in computing a relation between them, in terms of angle and distance. More precisely, as shown in Figure 14, a computation of the angle and distance is operated between each principal inertia axis.

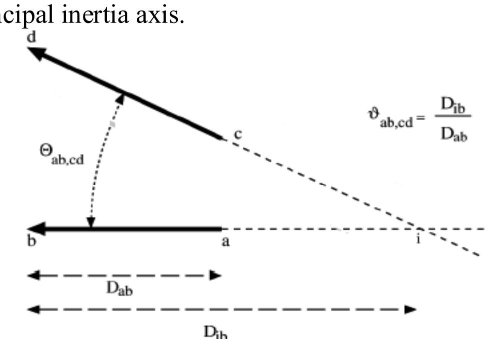


Figure 14 : Pairwise Geometric Attribute Construction

For each pair of object, we compute The angle and the distance between these two axis are respectively computed:

$$\vartheta_{ab,cd} = \frac{1}{\frac{1}{2} + \frac{D_{ib}}{D_{ab}}}. \quad \textbf{(2)}$$

$$\alpha_{ab,cd} = \arccos\left[\frac{\mathbf{x}_{ab}\cdot\mathbf{x}_{cd}}{|\mathbf{x}_{ab}||\mathbf{x}_{cd}|}\right];$$

(3)

4/ On the basis of the previous stage, computation of the histogram of relations: two-dimensional pairwise geometric histogram. Actually, this histogram represents the frequency of occurrence of relations between two similar situations:

(4)

$$H(I,J) = \begin{cases} H(I,J)+1 & \text{if } (i,j)\in E \text{ and } \alpha_{i,j}\in A_I \text{ and } \vartheta_{i,j}\in R_J \\ H(I,J) & \text{otherwise,} \end{cases}$$

Where

E = {couples ($a_{i,j}$, $V_{i,j}$) i et j are two principal inertia axis }

A_I = {angles computed between two principal inertia axis }

R_J ={ratio computed between two principal inertia axis}

As a consequence, each original image is characterized by a histogram describing these relations between the principal inertia axis. In terms of similarity measurement, the technique that is applied for the comparison between two images consists in applying a distance between these histograms. The application of this technique in the context of MADONNE's project is detailed in the experimentation part in section (4).

3.3. One way to combine descriptors

After having presented different approaches for signing drop cap images, we propose here an original technique allowing to combine different signatures or to select relevant features within a vector of primitives. We deal in this section with an approach allowing to improve the recognition by studying and integrating the behavior of basic descriptors.

Let us consider a basic confusion matrix to assess the power of recognition of descriptors. By definition if symbols are correctly recognized they should be assigned to the corresponding bin cluster. Bad recognition remains to a disparate distribution of the symbols on the lines of the confusion matrix. We can conclude that it is important to take into account not only the recognition rate but also the distribution of the errors in order to improve the recognition process.

Then, weight maps are defined from studies of distribution of distances in order to take into account the symbols maybe far away from a model. The idea consists in increasing the weight corresponding to such symbols to bring them closer to the model. Moreover the influence of other clusters is set by integrating negative weights in the map. The score is given by:

$$Score = \sum_{i=1,d} F_{dc}(1-(\vec{oc}-\vec{ox}))$$

(5)

Where d is the number of descriptors, (oc) and (ox) are respectively the vector of descriptor for an object of the cluster c and of the target cluster. Fdc is a weighted map of descriptor d for the cluster c calculated as follows: For each selected distance the ratio is set by the number of symbols belonging to the cluster divided by all the symbols found at this distance. Obviously the maps are generally not similar because the influence of the descriptors differs and the clusters to be recognized are different. The followinf Figure 15 gives an example of computed weight map.

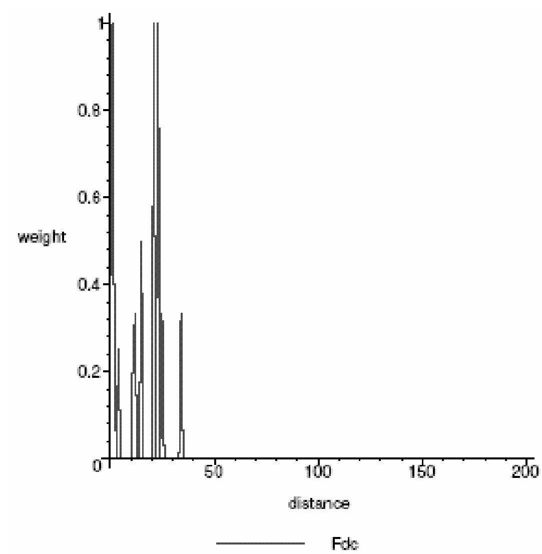


Figure 15 : Example of computed weight map

We also combine such maps in an additive way to improve the recognition as follows: Let us consider a learning database. Let $D=\{d_1,\ldots,d_n\}$ be the set of descriptors and $\Omega=\{\omega_1,\ldots,\omega_n\}$ be the associated normalized recognition rates reached using such database. Let x be a sample. The normalized weighed sum W_s calculated from x is given by Equation (6). The following figure 16 gives some examples of maps combination.

$$W_s(x) = \sum_{i=1,n} d_i(x) \cdot w_i \quad \textbf{(6)}$$

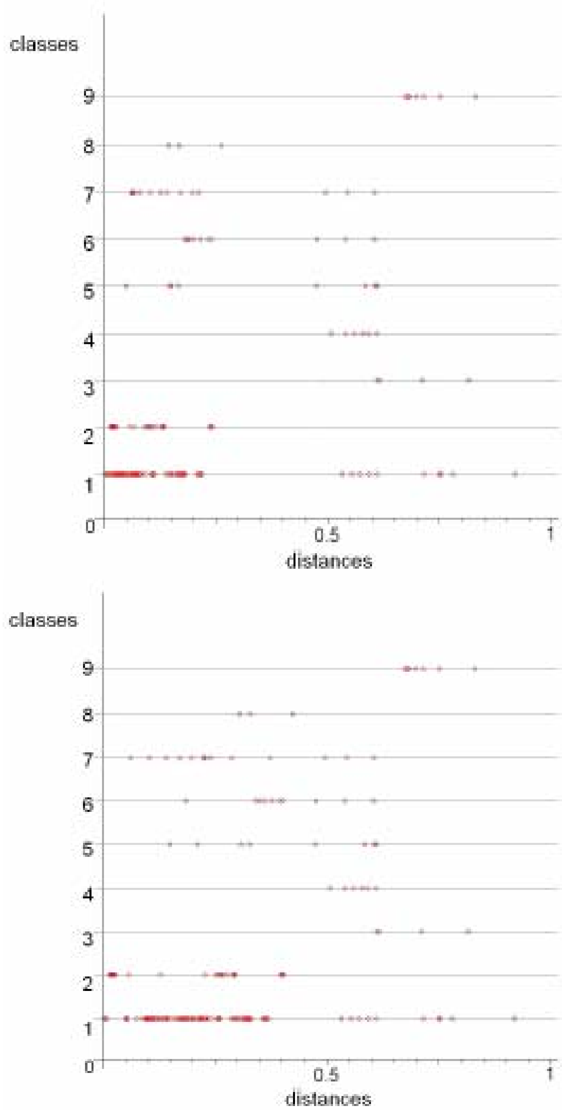


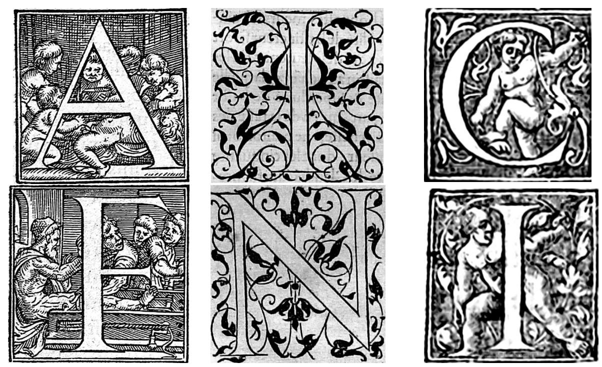
Figure 16 : Examples of combination weight map

4. Application and experiments

In order to measure the relevance of these different approaches, we have experimented these different signatures and selection techniques on a set of images provided by our CESR partner. The following parts illustrate the application of these different algorithms.

4.1. Zipf Law Results

We use a data base made of more than 200 images. We decide to index all of them using a k-means quantization with k=3, here our pattern is a 3x3 one the most usually used neighborhood in a 2D space. In order to verify the efficiency of all the exponents associated with the Zipf models we have presented, we compute all the drop cap Zipf curves. Each initial drop cap is represented by 3 numbers corresponding to the 3 slopes S1, S2 and S3 defined in the previous section (3.1.3). Besides, each initial drop cap is labelled by the experts who indicates let us say the style of the drop cap. In our database we count three different styles and some that cannot be classified because their number is not significant.



Style 1 Style 2 Style 3

Figure 17 : style samples

The calculation of a distance between two initial drop caps I and I' is given by the Hamming distance in the parameter space :

$$\textbf{(7)} \quad \text{distance}\,(I, I') = \sum_{i=1}^{3} \left| s_i - s'_i \right|$$

The information extracted from the images can be used in different ways. It can be used either for a recognition or a classification purpose.

- First Application

With a new initial drop cap of an unknown style the system can indicate the corresponding style depending on the number N of different style occuring in the nearest neighbors. We do this with all the initial drop cap of our database, based on the N-1 other images of the base. That is a recognition process.

The results obtained are presented in table 1. They depend on the number of k nearest neighbors involved in the decision process.

N	Style 1	Style 2	Style 3
1	92%	70%	84%

3	100%	94%	91%
5	100%	100%	97%

Table 1 : Results with 3-means quantization and a 3×3 window

- Second Application

The system can also be used for image retrieval from an image sample. From a request initial drop cap the user can ask for the n most similar drop caps contended in the data base. An illustration of the application is presented in Figure 18. Here six images are presented. Either the number of images can be indicated or a threshold similarity beyond which the image is not displayed. We can notice our method indexes the initial drop cap in a global manner and does not take in account the nominal letter.

Figure 18 : Zip law based system

4.2 MST signature and PGA Algorithm

The development of our drop cap retrieval system is currently under way. The retrieval results by using this signature appear to be encouraging (Figure 19). However, with larger database of diverse drop caps, the MST becomes a weak point in cost of computation as well as the ambiguity in non-bijective problem. The results, analyzed in terms of Recall/precision, are very encouraging. The database on which our system is tested contains close to 500 drop caps. Current research is working on the PGA and other signatures in order to compare their stability and rate of retrieval. We illustrate in the Figure 19 a result of our system applied to drop caps. In the case that is shown in this figure, it is interesting to see that the first returned drop cap is the good one, but it is also interesting to see that the second drop cap contains the same uppercase letter, and that the background of the drop cap "looks like" the query : two persons in the background.

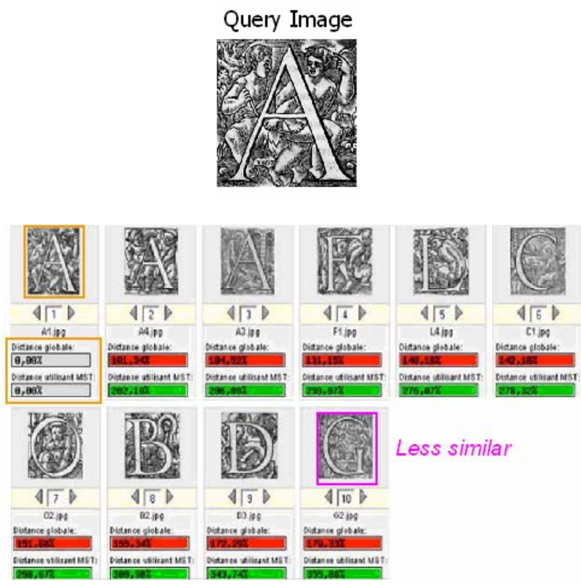


Figure 19 : Content Based Image Retrieval

4.3 Feature Selection Results

A test database consisting of graphical dropped initials extracted from archival documents is used here. Such images require a first processing in order to be usable within the framework of our application.

First a classical algorithm of binarization, based on an entropy criterion calculated on the grey level histogram, is carried out to extract both objects and background. Obviously the useful information can be contained in any cluster of the binary partition. Then after applying dilation and erosion steps to clean the region, a set of rules and measures (size, compactness...) is applied to choose the "good" related component. That is to extract the target letter.

An example of decomposition is provided in Figure 20 below. The first line shows a set of dropped initial coming from noisy ancient documents. Binary images with removed artifacts are given in the second line. Then the main connected components are presented. We can remark that the letters are generally included in component 3 to 5. The last line presents the extracted letters. A database of drop caps has been defined using such method. We consider here 11 shapes per cluster. The scale factor should be taken into account to compare graphical symbols because the size of the dropped initial differs following the archival documents. Moreover few letters could be rotated.

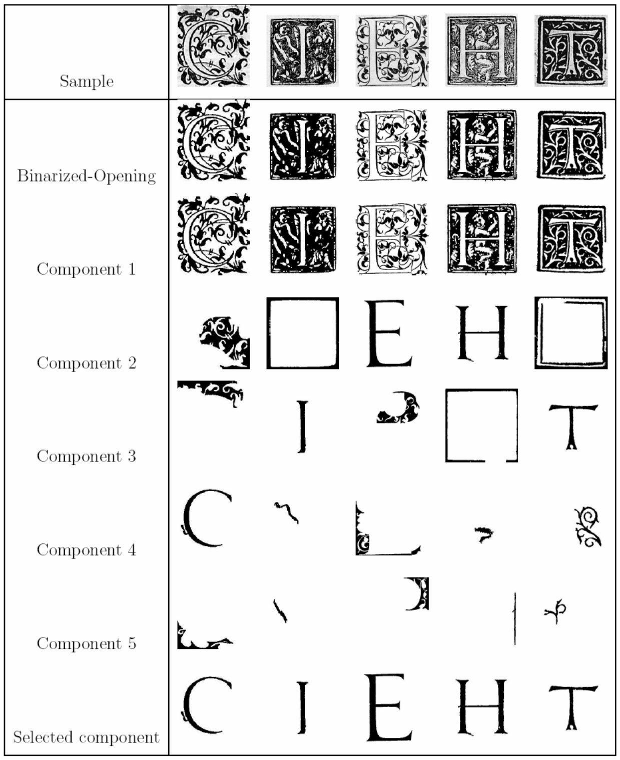


Figure 20 Example of drop cap decomposition

A set of descriptors able to manage with such constraints has also been implemented to check our approach that is: C for compactness, E for elliptic feature [8], S_A angular signature [6], T_R^f for signature based on Radon transform [9], G_{FD} for generic Fourier descriptor, FM for Fourier-Mellin [9] and MZ for moments of Zernike [10]. The recognition rates achieved using such methods are provided before and after integrating the proposed approach, we present these ones in the next Table 2:

Descriptors	C	E	S_A	G_{FD}	T_{Rf}	FM	MZ	W_S
Before	49	41	70	59	50	81	51	79
After	52	39	75	69	50	84	55	94

Table 2 : Recognition results with and without feature selection

We can remark a real improvement using our method. Around 94% of well classified symbols are reached using the global combination. We can also denote that classical weighted sum gives rise to a score lower than the best descriptor (*FM*).

5. Discussions

As one can observe when analyzing these different systems, these approaches proposed by different laboratories are very complementary and should be combined in order to improve the global quality of the system.

In terms of results, the signatures based on Zipf law may be used for the content based retrieval system, while PGA and MST based signatures may be used for the classification of the drop caps. The third combining approach proposes an interesting approach for an optimal combination of the relevant primitives. One manner to integrate this approach could be, for instance, to perform to this selection of relevant primitives when performing a relevance feed-back, during the interactions with the end-user.

6. Conclusions and Perspectives

This paper proposes a set of contributions issuing from different labs cooperating in the context of a French National research program, called MADONNE. The project investigates the use of document image analysis methodology for providing useful browsing and indexing features in these large collections. The MADONNE project is thus focused on indexing, organization and incremental enrichment of heritage data warehouses, in order to provide generic services to the users, including researchers in human and social sciences and to work towards interoperability of the data and of the browsing tools.

This paper focuses on complementary approaches concerning graphic classification/indexing problems. Different contributions concerning the computation of original signatures are proposed, while a relevant primitives selection strategy is proposed. These signatures are respectively based on Zipf law, Minimum Spanning Tree, and Pair wise Geometric Attribute. We have also proposed a new approach to automatically extract an appropriate subset of shape descriptors dedicated to a given application.A model based on the Choquet integral and on Shapley values has been designed to extract a subset of descriptors associated to each cluster and related to the classes of shapes.

The perspectives of these works deal with the integration of these different signatures within the same drop cap indexing system, while the selection primitives process should permit to integrate the specific requirement of a end-user, inserted in a relevance feed-back process.

Acknowledgements

Here we want to thank all the persons in the CESR (Centre d'Etudes Supérieures de la Renaissance) located in Tours University, France. They have made this study possible. The initial drop caps scanned images have been provided as well as the expertise associated with them.

We want to thank too Mathieu Delalandre (L3i laboratory, La Rochelle University, France) for his corrections and comments concerning this paper.

References

[1] Eakins, J P Content base image retrieval – can we make it deliver ? 2nd UK Conference on image retrieval, Newcastle upon tyne, February 99

[2] K.Melessanaki, V.Papadakis, C.Balas, D.Anglos, Laser induced breakdown spectroscopy and hyper-spectral imaging analysis of pigments on an illuminated manuscript, Spectrochimica Acta Part B 56 (2001) 2337-2346

[3] G.K. Zipf, Human Behavior and the Principle of Least Effort, Addison-Wesley, 1949

[4] E Dellandrea, P Makris, N Vincent (2004), "Zipf analysis of audio signals" Fractals, 12(1):73-85.

[5] S. Uttama, J.-M. Ogier, and P. Loonis. Top-down segmentation of ancient graphical drop caps: lettrines. In Proceedings of 6th IAPR International Workshop on Graphics Recognition, Hong Kong, pages 87 96, August 2005.

[6] Asymptotic theory of greedy approximations to minimal K-point random graphs A.Hero and O.Michel, IEEE Transactions on Information Theory, IT 45, pp 1921-1939, 1999

[7] Line Pattern Retrieval Using Relational Histograms, Benoit Huet, and Edwin R. Hancock, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 21, NO. 12, DECEMBER 1999

[8] R. Teag. “Image Analysis via the General Theory of Moments”. Journal of the Optical Society of America, 70(8):920-921, 1980.

[9] S. Tabbone and L. Wendling. “Binary shape normalization using the Radon transform”. Discrete Geometry for Computer Imagery (DGCI), LNCS 2886, Naples, Italy, vol. 2886, pp. 184-193, 2003.

[10] A. Khotanzad and Y. H. Hong. “Invariant Image Recognition by Zernike Moments”. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(5):489-497, 1990.