Global discrimination of graphics styles

Rudolf Pareti* and Nicole Vincent

Laboratoire Crip5-Sip, Université Paris 5, 45, rue des Saints-Pères, 75270 Paris Cedex 06, France
E-mail: rpareti@free.fr
E-mail: Nicole.Vincent@math-info.univ-paris5.fr

Abstract

Discrimination between graphical drawings is a difficult problem. It can be considered at different levels according to the applications, details can be observed or more globally what could be called the style. Here we are concerned with a global view of initial letters extracted from early renaissance printed documents. We are going to present a new method to index and classify ornamental letters in ancient books. We show how the Zipf law, originally used in mono-dimensional domains can be adapted to the image domain. We use it as a model to characterize the distribution of patterns occurring in these special drawings that are initial letters. Based on this model some new features are extracted and we show their efficiency for style discrimination.

Keywords: Graphics Recognition, Zipf law, indexation, initial letter, illumination, image

1 Introduction

Graphical documents can be considered as images organized with lines and repeated small drawings. An architectural plan or a map contain a lot of symbols and we can determinate from them a style. Indeed they can be indexed and discriminated by the author or the map editor. The style has to be defined in a proper way, either from an objective point of view or form an expert view. Here we consider the second way.

We can consider that initial letters in ancient documents and more especially those that are printed as in the Renaissance period are graphics composed with lines and repeated symbols as illuminations.

Most of the books in the middle age period are reproductions of ancient and religious texts, realised in monasteries in which the copyists work under dictation. The invention of printing led to the multiplication of the books that nevertheless tried to attain the quality and the beauty of the ornamented presentation of previous period manuscripts. Books are rich, well ornamented with ornamental initial letters, illuminations, head bands and various pictures. Here the ornaments are in grey levels because of the technical problems.

In fact the artists working in the editors shops respect a lot of already existing constraints:

- the page setting already existing
- the requirement of the clients
- and all the representative conventions (saints drawn with their attributes, decorative stylization…)

Some artistic schools were created that have specific aspects depending on the location of the shop or on the people working there. Some copies were done by people in order to usurp the origin of a book. Some illuminators copy the others and it is not rare that an illuminator realized almost the same illumination as some already realized in other places. Nevertheless the graphics differ by some details from the original one and a close look at the image is needed to conclude some initial letters come from the same printer or not.

The ornamental initial letter has different functions:

- a religious connotation
- a visual reference to understand quickly the content of the book

- a wealth mark

Since a book contains several media and comprises both a support and content, the study of the text interests many experts in different domains. The text in itself does not contain the same information as can provide the study of the illuminations and particularly the ornamental initial letters. Indexing these pictures will permit us to know whether a book has been made by the same artistic school, or if someone has cribbed from another artist. Even better, the books are often damaged and we cannot know who the author is, when it has been written, if it is an original version of the book or some copy, what edition is being read. To answer these questions and many others, the fonts used in the book can be observed, the figures can be used too. Here we are concerned with the classification of the initial letters. They would teach us a lot of information hidden by the time effect. In the Renaissance period, the period we are interested in this study, the initial letters were some little black and white graphical drawings with illumination. A lot of methods exist to classify them, the most simple relies on the use of the histograms, others rely on pattern matching techniques [1], other rely on indexing methods linked to image indexing and retrieval techniques. We have tried to create a method more robust than the other existing methods taking into account the many aspects contained in these small elements, lines, dashed lines or zones, symbols, letters. Statistical approach seems necessary to solve the problem. Indeed the structural approaches are still not mature enough to handle small differences we cannot model well. Then we have chosen an approach based on elementary windows, which would use a mathematical model to characterise the drawing. Of course, the model is depending on parameters. Zipf law seems to hold in many observed 1D phenomena, then we though to apply it to this type of images.

In a first part we are going to recall Zipf law, as developed in the study of 1D signals, then we will show how to transpose this text law for the purpose of image analysis and more specially for the specific graphical drawings we are studying. Finally results will be presented.

Our process can be divided into two steps, the learning phase and the application phase. In the first step, the training set comprises, for each style already known a small number of ornamental initial letters that experts have classified.

Then in the second step, we are going to take any ornamental initial letter at disposal and we are going to classify it according to the classifier developed. Therefore the output of the system is a classification of the ornamental initial letters according to their style. A verification method can be developed too**.**

2 Zipf law

Zipf law is an empirical law expressed fifty years ago [2]. 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 not organized in a random way. It can be observed that the apparition frequency of the patterns M1, M2, …, Mn, we note N1, N2, .., Nn, are in relation with rank of these symbols, if sorted with respect to their decreasing occurrence frequency. The following relation can be observed:

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

$N_{\sigma(i)}$ represents the occurrence number of 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 logarithmical 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 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, the 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 empirical points of the graph have to be completed by re-scaled points along the horizontal axis.

The validity of this law has been studied and observed first in linguistic study of texts, later it has been observed in other domains but always for mono dimensional signals.

In order to study graphical drawings, 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 Application to images

The ornamental initial letters 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. In spite of the black and white property of the drawings, we have preferred grey level information to achieve the study. The noisy paper and the deformation due to ancient paper need to have this precision, the relative grey levels are more significant than the absolute values. The method we are proposing is invariant under the geometrical transform that leave invariant the shape of the mask. The invariance to change of scale that can occur when images are scanned in different conditions is intrinsically linked to the method itself.

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 in the 2D space the data is imbedded in. We have chosen to use a 3x3 mask because they define the most often considered neighbourhood of a pixel in a 2D space.

Then the principle remains the same, the mask scans the whole picture and the numbers of occurrences of each pattern are computed.

Several problems have to be considered. How to describe the patterns? That is to say how the drawing has to be encoded and what are the properties we want to make more evident?

3.1. Coding of the patterns

We have considered a 3x3 size mask and 256 symbols are used to code pixels. Therefore, there are theoretically 256^9 different patterns. This number is much larger than the number of pixels in an image. Indeed, if patterns are not frequent enough, the model that is deduced from Zipf model cannot be reliable, the statistics loose there significance. For example a 640x480 image can contain only 304964 different patterns. Then, it is necessary to restrict the number of possible patterns to give sense to Zipf model. The coding is decisive in the matter. According to the coding process chosen, 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 phenomenon. If several straight segment appear we can conclude several phenomena are mixed and several models are involved. Besides, different codings allow to make more evident different properties of an image.

Some studies have shown Zipf law was holding in the case of images with different encoding processes [3]. We are looking for coding process that gives models apt at discriminating the graphical drawings we are studying. in our case, this qualifies as effective a coding process. Two drawings that look alike from a style point of view should verify similar power laws. We are going to see different coding methods we have tested.

According to the remarks previously done, the number of different possible must be decreased. This can be done decreasing the number of pixels involved in the mask or decreasing the number of values associated with a pixel.

3.2 The general ranks coding

Here our motivation is to respect the vision of a scene that relies more on differences of grey levels than on the absolute values. So, within the mask, the grey level values are replaced by their rank when they are sorted according to the grey level values. The method affects the same rank when the grey levels have the same value. Then the maximum number of values involved in the mask is 9 and this leads to a very large decrease in the number of different possible patterns.

Image Pattern 1:

2	8	6		
21	31	31		(a)
32	32	32		

Image Pattern 2:

130	136	134		
149	159	159		(b)
160	160	160		

Coded pattern:

0	2	1		
3	4	4		(c)
5	5	5		

Fig 1: patterns (a) and (b) coded in (c) using general rank method

The number of symbols used to represent the grey levels is limited to 9.

It can be noticed that the patterns in figure 1 (a) and (b) are different, yet, the coded pattern (c) that is generated is the same in both cases. It is one of the limitations of this coding. Of course information is lost. Further the method relying on this coding process will be called general rank method.

3.3 Grey level quantisation

The previous method has lead to 9 possible values associated with a pixel in the pattern considered. A simpler way would be to consider only k grey levels to characterise the intensity level of the pixels. Most often it is sufficient to observe an image. More over the images we are dealing with in the graphical drawings are essentialy black and white. A quantisation in k equal classes would lead to unstable results, so we have chosen to use a classsification method of the grey levels in k classes by way of a k-mean algorithm. Further the method relying on this coding process will be called **k-mean**. We have experimented different values of k. In figure 2 we present an example with 9 clusters.

Grey level	[0;20]	[21;76]	[77;96]	[97;120]	[121;146]	[147;174]	[175;203]	[204;229]	[230;255]
center	15	56	86	107	133	159	190	217	242
N° Class	0	1	2	3	4	5	6	7	8

Fig 2 : K-mean example

The example depends on the drawing and considers 9 classes. We can see that some classes may be closer than others.

3.4 Cross Patterns

An other way to decrease the number of possible patterns is to limit the number of pixels in the pattern. To remain coherent with the 2D topology we have chosen to consider the smallest neighborhood of the pixel defining 4-connectivity. It is precised in figure 3.

	3	
4	1	2
	5	

Fig 3: Mask shape indicated as cross mask

In this case we have also 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-mean with k equal to 3 has been processed on the set of pixels of each image. 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. The k-mean classification makes the method independent on the illumination of the scanned image. Further the method relying on this coding process will be called crossmean method.

3.5 Zipf curve construction

Whatever the encoding process used, Zipf graphs can be built. Now in order to study a family of images, these plots have to be compared. A close look at the curves shows they are not always globaly linear, that is to say Zipf law does not hold for the whole patterns. 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. With the second process for example, we observed the left part of the graph is concerned with the regions in the image 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.

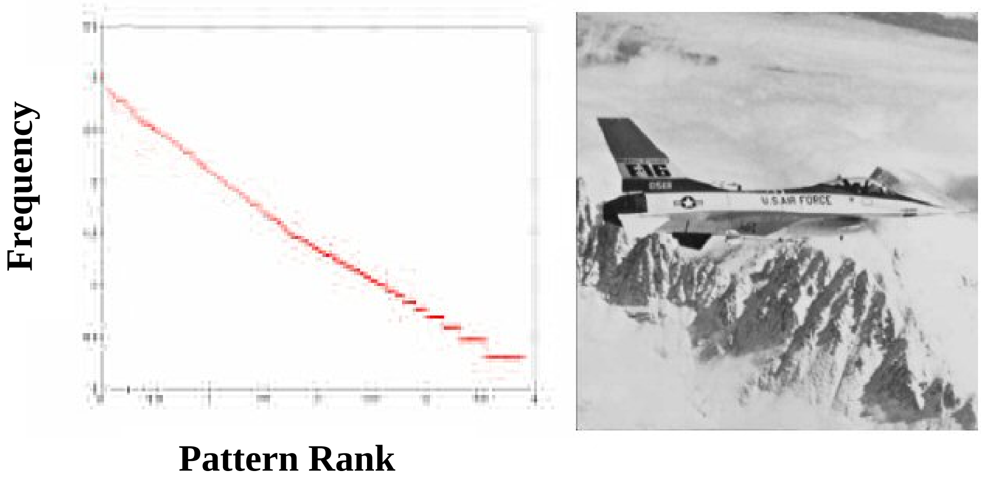

Fig 4: Zipf curve associated with an image with respect to general rank method

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 5 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 in the process and we can model them. Several power laws are involved and then several exponent values can be computed.

The Zipf curve has been drawn with respect to a logarithmic scale, therefore, it is necessary to begin with a resampling of the curve.

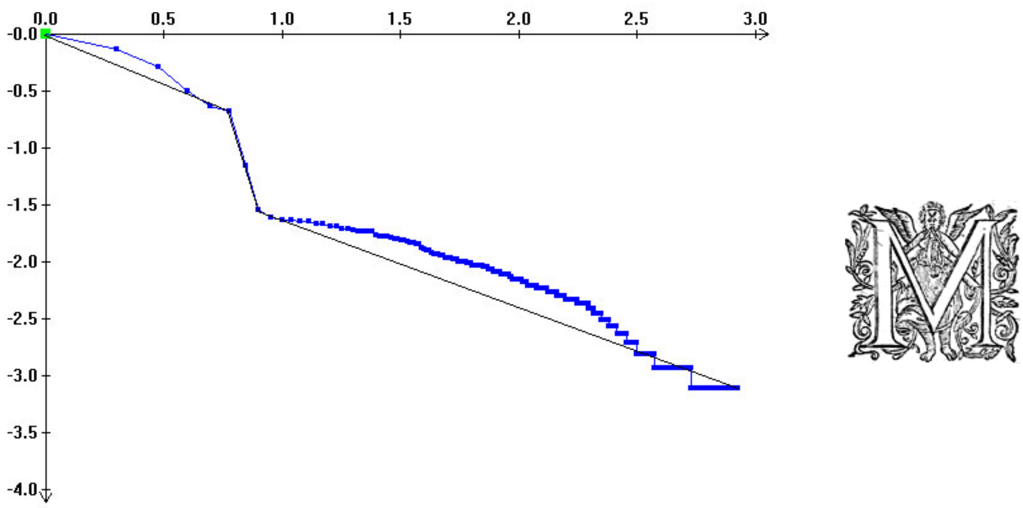

Fig. 5: example of an initial letter and its Zipf plot where are indicated the different straight zones extracted

Then the output of the process is made of 3 meaningful values associated with the picture (the ornamental initial letter figure 5). They correspond to the 3 slopes or leading coefficients.

From this material we are going to learn the characteristics of the different styles.

4. Training

We use a data base made of more than 300 images of initial letters. They have been extracted from different books. They can be considered as belonging to one of three styles, each represented by a number of the ornamental initial letters {S1, S2, S3}. In order to be able to classify the ornamental initial letters we have separated this set in a learning set and a test set according to the ratios indicated in table 1.

Table 1: composition of the learning and test sets

Style	Training set	Test set
S1	28	163
S2	10	22
S3	24	68
other		43

In order to verify the possible efficiency of every exponent associated with one of the models based on Zipf law we have presented, we are going to calculate all the ornamental initial letters Zipf curves of every style and deduce the average and the standard deviation associated with every style. They are indicated in table 2. To have more efficient results we decide to process a normalisation step before the parameter computations. We apply an histogram normalization filter in order to take better advantage of the image spectrum.

Table 2: analysis of the learning test according to extracted features

Methods	style	Average & standard deviation of slope 1		Average & standard deviation of slope 2		Average & standard deviation of slope 3	
General Ranks	S1	-0.15	0.2	-1.19	0.8	-0.66	0.4
	S2	-0.22	0.1	-0.79	0.2	-0.46	0.3
	S3	-0.22	0.1	-1.29	0.2	-0.83	0.1
k-mean k=3	S1	-0.31	0.3	-1.37	0.9	-0.77	0.5
	S2	-1.7	0.7	-0.43	0.2	-0.97	0.3
	S3	-1.38	0.8	-0.65	0.7	-0.82	0.1
Cross mean	S1	-0.99	0.6	-1.06	0.5	-4.49	1.9
	S2	-2.67	0.4	-0.75	0.1	-4.53	0.9
	S3	-1.71	0.2	-0.51	0.1	-4.27	0.3

We can notice that the general rank method is not satisfactory at all. The K-Means method remains a good method for all slopes, but execution time is relatively long. Finally the crossmean method is the most discriminative as far as the first two slopes are concerned.

That is why in the next tests we will use in priority the parameters extracted in the crossmean and k-mean Zipf curves.

5. Results

The first approach we have implemented relies on the modeling of a style by the average values computed in the learning phase. Then distance is calculated using the usual euclidian distance. {µ1, µ2, µ3, …} are the averages learned for every style and Pj are the different parameters considered in the method.

$$\text{distance}(X, S_i) = \sqrt{\sum_j (\mu_i^j - Pj)^2} \qquad (2)$$

The results we obtained in such a way are indicated in table 3.

Table 3: results using Euclidean distance and single prototype style

recognition	S1	S2	S3	other
Style1	**25.15%**	1.23%	26.99%	46.63%
Style2	0.00%	**36.36%**	22.73%	40.91%
Style3	0.00%	0.00%	**94.52%**	5.48%

We can notice that the results are not satisfactory. Indeed, the method does not take into account the classes geometry in the different styles.

In regard to these results we have decided to use the Mahanalobis distance. Besides, some experiments have been done, in order to choose the best value of the grey level number in the k-mean method. They are illustrated in figure 6.

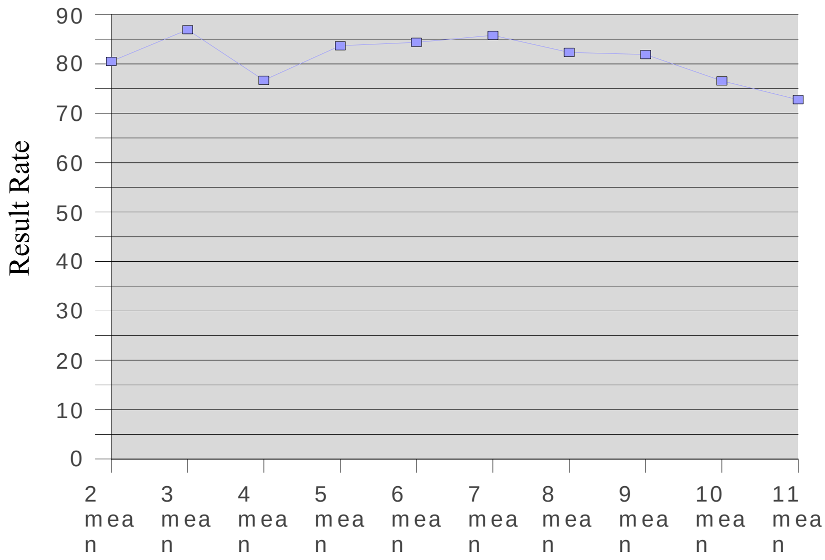

Fig 6: evolution of the recognition rate according to the number of the grey level used in the encoding process

From the previous study we have chosen to use three grey levels, this unables to speed up the process. We notice that the most efficient results have been reached with the k-mean method. Results are indicated in table 4.

Table 4: results using only k-mean parameters and Mahanalobis distance

Style	Images in the test base	Images detected in test base	Recognition rate	False accepted rate
Other	14.77%	39.38%	**64.58%**	75.78%
Style 1	50.46%	41.85%	**75.00%**	9.56%
Style 2	11.69%	0.92%	**7.89%**	0.00%
Style 3	23.08%	17.85%	**69.33%**	10.34%

In table 4 we have indicated the recognition rate as well as the false recognition rate for each class. But the Mahanalobis distance presumes that the slopes of the images to be classified in the class, are distributed according to a normal distribution but we noted that it was not always the case. So we have decided to implement the k nearest neighbours (knn) method where no hypothesis has to be done. Results are indicated in table 5.

Table 5: Results using knn method

Style	Images in the test base	Images detected in test base	Recognition rate	False accepted rate
Other	14.77%	20.00%	**52.08%**	61.54%
Style 1	50.46%	51.08%	**86.59%**	14.46%
Style 2	11.69%	12.00%	**97.37%**	5.13%
Style 3	23.08%	16.92%	**66.67%**	9.09%

The results seem to be more satisfactory. We have experienced the combine use of several encodings but the results have not been significantly better than those presented here.

6. Conclusion

Here we show the use of a model developed in the field of 1D phenomena can give good results in case of images. This law allows to define global parameters based on details. According to the type of encoding used, the nature of information differs. Other encoding processes can be experimented and the method can be applied to other problems involving graphical drawings. The method is invariant under any rotation and chang of scale.

7. 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. Indeed, the initial letters scanned images have been provided as well as the expertise associated with them have been provided by members of the CESR.

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

[3] Y. Caron, H. Charpentier, P. Makris, N. Vincent, Power Law Dependencies to Detect Regions Of Interest, 11th International Conference DGCI 2003, Naples, Italy, November 2003