Top-down segmentation of ancient graphical drop caps: lettrines

Surapong Uttama, Jean-Marc Ogier and Pierre Loonis

Faculty of Sciences and Technology, L3i Research Laboratory, University of La Rochelle, Avenue Michel Crépeau, 17042 La Rochelle Cédex 1, FRANCE

E-mail: {uttama.surapong, jean-marc.ogier, pierre.loonis }@univ-lr.fr

Abstract

The restauration and preservation of ancient documents is becoming an interesting application in document image analysis. This paper introduces a top-down approach aimed at segmenting the graphical part in historical heritage called lettrine. The research principle was established on the concept of human visual perception and invariant texture analysis (co-occurrence and run-length matrices, autocorrelation function and wold decomposition). The preliminary results concerning segmentation stages were presented by highligthing difficulties related to the nature of strokes and textures in lettrines. Textured background was extracted although there existed some ambiguities. Nonetheless, the segmented areas of interest are informative enough to serve in an indexing method. The prospective bottom-up approach are mentioned and will be added to gain more precise segmentation.

Keywords: Top-down analysis, Lettrine segmentation, Texture analysis, Ancient documents indexing.

1 Introduction

The European cultural and scientific heritage is the public and unique resource that represents the collective memory of our different societies. The international communities (governments, organizations, etc.) have been recognizing an increasing requirement concerning the safeguard of this heritage and the accessibility to this information. This paper deals with a project, called MADONNE [1], aims to emphasize various resources of the international inheritance, and more specifically the books, images collections and any other iconographic documents. In a short term, these numerous documents will constitute a huge amount of data. Consequently, many problems have to be tackled in the digitization process. Among these problems, one of them is to provide set of services allowing to navigate as easy as possible in the document. Thus the aim of our project is to develop a set of tools allowing to extract all the information as automatically as possible from the digitized ancient images as well as to index them in order to navigate through the documents database. This work is performed in collaboration with an historical center (*Centre d'Etudes Supérieures de la Rennaissance*) [2]. In this paper, we illustrate our first contribution concerning the segmentation of ancient graphical drop cap or *lettrine* in French.

2 Definition of lettrine

In view of image analysis, a lettrine is a line-drawing image that is generally handmade in old documents (cf. Figure1). Usually a lettrine is composed of two main parts: a letter (a first character of the first word in the chapter or section) which contains a high semantic level and a drawing painted in background, dealing with a scene with a semantic or only illustrative motif. Mathematically a lettrine (L) can be defines as:

$$L=\{a_j,\, j\in L | a_j \cup a_l = \emptyset,\, \forall\, j, l \in J^2\}$$

where a_j = area of the set of strokes s_i^j so that $d(s_i^j, s_l^j) < \epsilon$

This drawing is specific because it is a line drawing with crosshatch or flat tint used to draw a scene in order to model its volume. Moreover, the objects drawn in the scene have no closed boundary i.e. the objects are represented by groups of parallel lines, reunited in homogeneous texture zone. The segmentation of such image is very specific and the standard segmentation tools are not applicable.

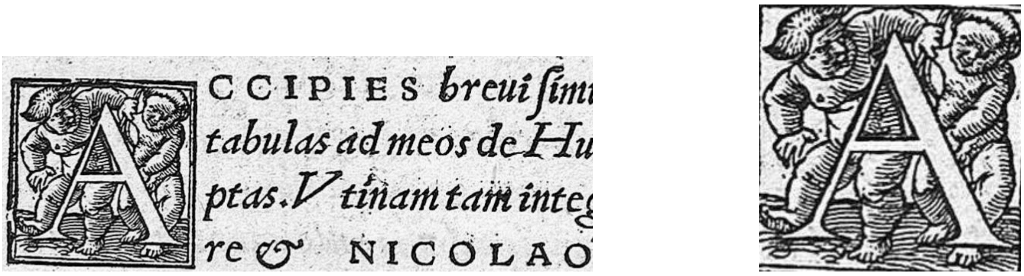


Figure 1: Lettrine in context (left) and close-up of the lettrine (right)

3 Overview of research strategy

The main difficulty concerning this problem is related to the very important density of the graphic objects, and to the lack of research concerning this category of image. Indeed, as one can see on the Figure 1, the lettrine is made principally of well organized strokes in order to represent objects. Considering the density of the lettrine, the high connectivity of the information, it is quite impossible to reuse the techniques issued from the graphic recognition community (GREC). As a consequence, for the implementation of our system, the strategy is to develop specific techniques derived from the computer vision domain, by coupling them with indexing problems.

Our research strategy is strongly inspired from human visual perception principles [3], combining the equivalent of a multi-resolution procedure and the mixing of a top-down and a bottom-up approach: perception cycle [4]. In this paper, we commence our trial on top-down method.

Note that in top-down approach, we aim to label the region with similar pattern to the same class. However, without prior knowledge, there exists the problem in finding number of classes.

At first, our approach consists in trying to distinguish the different information layers (background, textures, drawings, informative zones, etc.) that can be found on the image by separating a lettrine into two regions: the uniform and texture regions. This can be done by measuring the uniformity or entropy of an image. Then the segmentation processes will be applied to each separated region.

To fulfill the top-down segmentation, we combine the classical image analysis tools with invariant texture analysis. In this paper, we present our first results concerning the segmentation stage, highlighting the difficulty related to the notion of strokes image.

4 Lettrine segmentation

The primary step of lettrine segmentation relates to image analysis. In fact, more information from image analysis can guide whether which segmentation methods are the most appropriate and lead to the optimal segmentation results.

4.1 Segmentation principle

The principle of image segmentation composes of two main stages which are global analysis; to distinguish between different image's regions and local analysis; to characterise each previously classified region by using some related attributes. A general overview of segmentation concept is described in Figure 2.

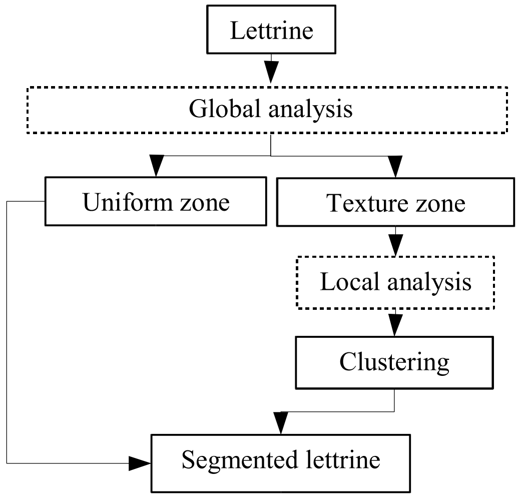


Figure 2: An overview of lettrine segmentation

4.2 Global analysis of lettrine

In global analysis of the lettrine, it is observed that lettrine varies according to its origin. Some contains a lot of shades with multiple parallel and crossing lines, the others have more painted area but fewer strokes. However, most lettrine represent two main areas; the uniform and texture zones.

Thus the primary task becomes to differentiate between uniform and texture regions. One of the effective methods is Gray-level co-occurrence matrix (GLCM) [5] which represents the grey-level transformation of image.

Many features can be derived thanks to the GLCM, for example, the *uniformity*, it describes grey-level transformation through the diagonal directions of the image. The uniformity (*U*) is defined as the average of diagonal members of GLCM with one-unit displacement in four orientations:

$$U = \frac{1}{d} \sum_{j=1}^{d} \sum_{i=1}^{G} P(i,i)$$

where *P(i,i)* is normalized diagonal members of GLCM, *d* is number of directions (which is currently 4) and *G* is number of grey levels.

Thus if calculate the uniformity in four main directions (0 45 90 and 135 degrees), these information could reveal the distinction between high-uniformity and low-uniformity areas which finally present the uniform and texture zones respectively.

The process of global analysis is done by sliding small windows (empirically 4 by 4) throughout the quantized lettrine (4-bit grey level). In each window, we calculate GLCM as well as its uniformity and store it as a value of centre pixel.

4.3 Local analysis of lettrine

After available to partition the homogeneous area and its complement, the local analysis of each zone is considered. The purpose of current process is to prepare necessary information for segmentation process. Concerning the uniform region, it is not necessary to gain supplementary information. The segmentation of this zone is possible by measuring some parameters such as contrast or area. Conversely, for texture region, it is required more information to go further. In other words, it is necessary to define the window's size that presents unique texture characteristics for calculating texture parameters.

4.3.1 Selection of adaptive window's size

For the texture region, the first thing to concern is related to the definition of window's size which represents the size of textured motif. Indeed, the appropriate window's size is a very crucial to the segmentation process. For example, if the textured motif is quite large but the window's size is too small, finally we will lose some significant information. The work in [9] indicates that the *autocorrelation function* (ACF) can be used to determine the minimum window's size in which the structure of texture is homogeneous and well represented. Theoretically, ACF is defined as:

$$ACF(k,l)=\frac{\sum_{i=1}^{M}\sum_{j=1}^{N} I(i,j)I(i+k,j+l)}{\sum_{i=1}^{M}\sum_{j=1}^{N} I^2(i,j)}$$

where *ACF(k,l)* is autocorrelation function of image *I*. *M* and *N* are size of image.

Nonetheless, the above autocorrelation function (*ACF(k,l)*) is in two dimension and not applicable in the process of the determination of window's size. As a result, it is necessary to transform 2-d ACF into one dimension (*ACF(r)*). In this paper, the interpolation is done on a half of a circle. (see Figure 4)

$$ACF(r)=\frac{1}{\pi r}\sum \tilde{ACF}(k,l)$$

where,

$$\tilde{ACF}(k,l) \quad \begin{cases} ACF(k,l) & \text{if} \quad (k,l)\in \mathbb{N} \\ \dfrac{\sum_{i=1}^{4} d_i \cdot ACF(k,l)}{\sum_{i=1}^{4} d_i} & \text{otherwise} \end{cases}$$

d_i is the distance between point *(k,l)* and surrounding points (P_1 to P_4) and r is the radius of circle used to interpolate autocorrelation function

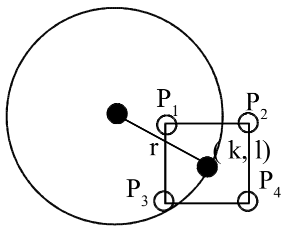


Figure 3. Interpolation of autocorrelation function on half of a circle

In another way, the 1-d autocorrelation function can also be modelled by using the Wold decomposition [7] concept:

$$\tilde{ACF}(r)=e^{-\alpha r}+\gamma\cdot e^{-\beta r}\cos(2\pi fr+\phi)+\delta+\epsilon(r)$$

where,

$\alpha, \gamma, \beta, f, \phi$ and δ are model's parameters and $\epsilon(r)$ is error of modelling.

These parameters signifies whether the interested zone is periodic or random. From our experiment, for evident periodic zone e.g. parallel lines, the window's size is approximately the period of autocorrelation function (*1/f*). In contrast, when the interested zone is not explicitly periodic (because either the window's size is too large and contains different textures or the texture in that window itself is random), the window's size is estimated by considering the size in which $\tilde{F}(r)$ becomes negligible (less than δ).

4.3.2 Computation of texture descriptors

The characteristics of textures of each interested zone are calculated by two texture descriptors: Grey level co-occurrence matrices (GLCM) [5] and run-length matrices [6].

The GLCM reveals certain properties about the spatial distribution of the grey levels in the texture image. For example,if most of the entries in the co-occurrence matrix are concentrated along the diagonals,then the texture is coarse with respect to the direction used in calculation of GLCM.

The run-length matrices displays the length of the primitives (the connected pixels with the same grey value). Larger number of primitives indicates coarser texture.

In the experiment, 8 parameters from GLCM are selected: contrast, entropy, correlation, uniformity, sum average, sum variance, information measure of correlation 1 and 2. In addition, 3 parameters from run-length matrices are calculated i.e. long-run emphasis, grey-level distribution and run-length percentage. Each texture parameter represents distinct signification, for example, contrast shows the change of brightness between pixels, entropy indicates the randomness of pixels, etc.

The processes of local analysis are done by initially divide the entire image into small windows (empirically 20% of image's size). In each divided window, the autocorrelation function is calculated and optimized to identify the sub-window's size. This sub-window is used to compute texture parameters and its size is constant in one window but adaptive for other windows according to the autocorrelation function. This adaptive window's size allows us to obtain more realistic texture descriptors for our segmentation process.

It is necessary to note that the initially fixed window's size (20%) may induce a problem because the first divided window can contain mixed regions between uniform and texture zones or between different texture areas. This occurrence will deviate the results of autocorrelation function.

Figure 4 illustrate mentioned processes of local analysis.

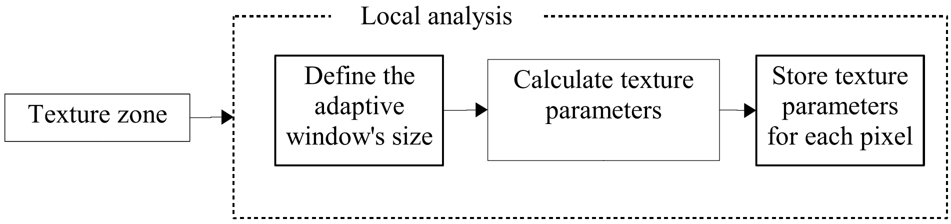


Figure 4: Local analysis of lettrine

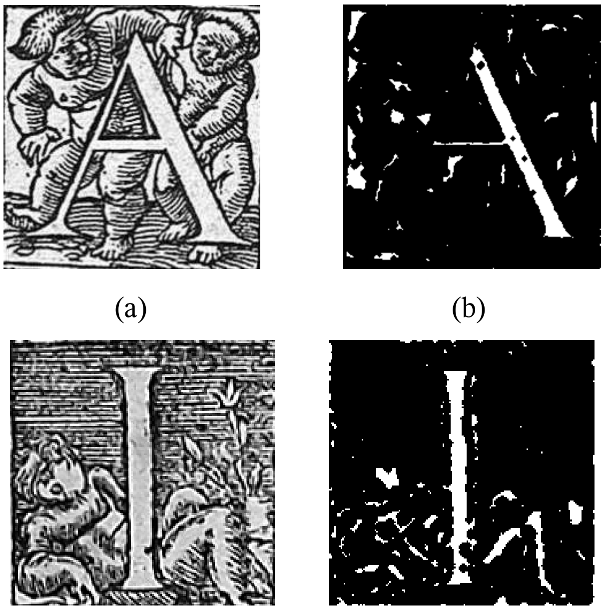
5 Experimental result

The experiment was performed by taking into account the lettrine from *Les Bibliothèques virtuelles humanistes* [2]. According to the processes introduced in previous section, they provide the outcomes as follow:

5.1 Preliminary process results

In global analysis, the original lettrine is partitioned by the property called uniformity. The original lettrine sample and its partitioned result are shown in Figure 5.

At this point there are obvious two zones; the white region refers to the high uniformity or the uniform zone while the black area signifies the low uniformity or the texture zone. This image is intended to use as a mask for applying different segmentation techniques for each region.



(a) (b)

(c) (d)

Figure 5: Original lettrines and their uniform and texture zones

Next in the process of local analysis the lettrine is divided into smaller windows. The one-dimensional autocorrelation function of each image window is calculated and optimized with the Wold decomposition model in order to find the appropriate window's size which is necessary for later process. The criteria used to judge the texture characteristic are experimented on images from Brodatz album [8]. The example of a plot of autocorrelation function and its optimization are presented in Figure 6. Then the sliding neighbourhood operation is implemented for each pixel by using previously calculated window in order to retrieve texture descriptors.

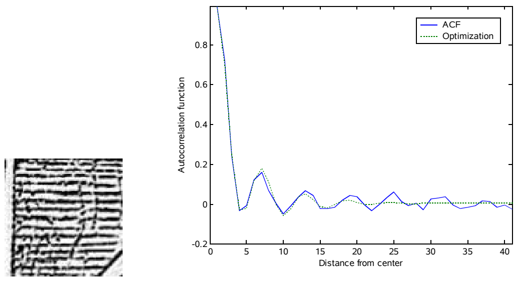


Figure 6: Top-left corner of the lettrine and its autocorrelation function (ACF)

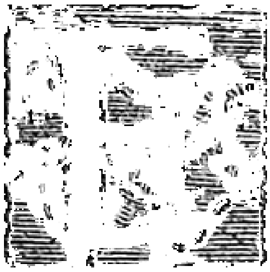
Finally the automatic clustering algorithm (currently k-means) is introduced to cluster the pixels according to their texture properties.

5.2 Textured background extraction

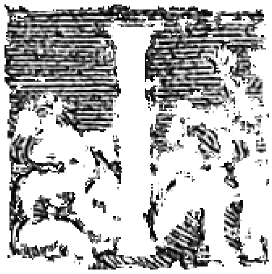
In practice, the background of a lettrine is composed of forms with similar patterns or significations, whether the lettrines are rather simple (cf Figure 7) or more complex (cf Figure 8). It is observed that the parallel lines are across the large uniform areas in Figure 7 (a-b). While in Figures 7 (c-f) the horizontal textures are apparently revealed. Nevertheless, some non-background textures are unexpected to be in the same part of background due to the similarity of texture pattern and invariant texture analysis.

(a) (b)

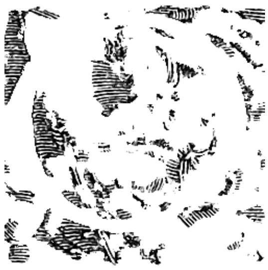
 

(c) (d)

(e) (f)

Figure 7: Textured background extraction of rather simple lettrines

(a) (b)

Figure 8: Textured background extraction of more complex lettrine

This ambiguity gives an appearance of the limitation of the top-down process. As seen in Figure 8 (a-b), it cannot differentiate the horizontal texture belonging to the background and the flag. This problem will be corrected by a bottom-up approach by introducing semantic information in prospective works.

5.3 Lettrine segmentation

The segmentation outcomes vary according to the numbers and types of texture descriptors as well as characteristics of lettrine itself. Figure 9 presents lettrines and some classes of their segmentation results. Four principal groups of segmented classes are formed and revealed:

1. The foreground (Figure 9 (b,h)),
2. The background (Figure 9 (c,i)),
3. The ambiguity between previous two classes (Figure 9 (d,j))
4. The contour (Figure 9 (e,f) et 9 (k,l)).

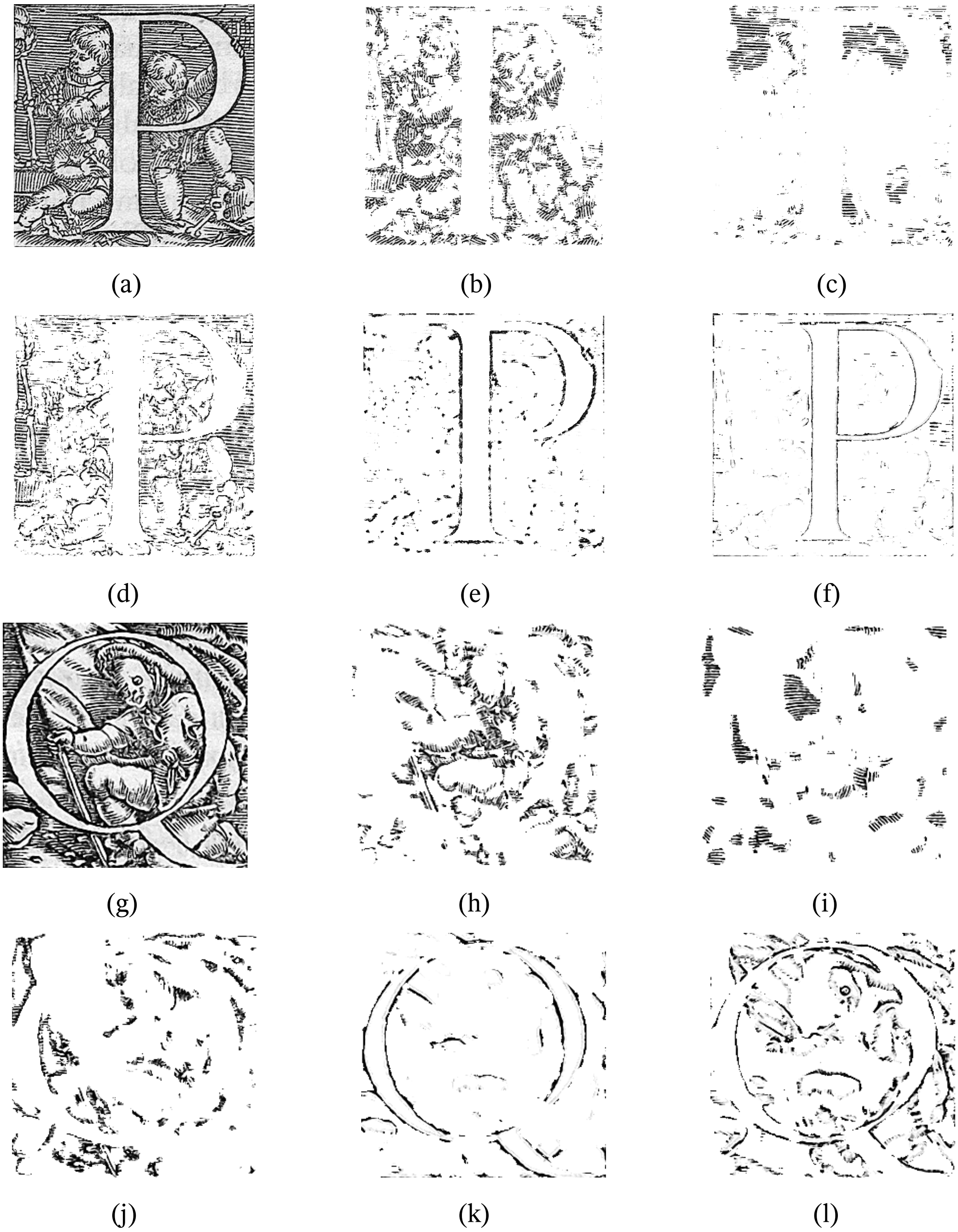


Figure 9. Lettrines with more complexity and their segmented classes

The results show the major limitation of the top-down phase: the unknown number of classes and especially the classes without semantic contents. It is obvious that the labelled images cannot induce a high semantic index, for this first top-down process does not use any expert knowledge.

6 Conclusion and perspectives

Graphics images are specific images for the reason that they are handmade by group of strokes in which their directions and styles denote specific semantic meanings.

Our attempt to analyse the specific case of lettrine, which are lines drawings in ancient document, is based on the identification of homogeneous sets of drawn lines, merged into a coherent texture according to the inspiration of a draftsman.

To segment gray-level lettrines, the top-down approach is implemented in two steps: global and local analyses. The global analysis contributes to distinguish between the interested zones: uniform zone and texture zone. Then the texture parts are characterized by texture parameters in the local analysis, in order to regroup these texture zones according to their similarity. The segmentation processes will help us reduce the insignificant information to the desired and sufficient level for the indexing process.

The explorative work described in this article presents the preliminary segmentation result. The extracted uniform and textured areas of interests are informative enough to serve in an indexing method.

Nevertheless, limitations have been pointed out. First of all, the choice of the initial set of texture areas of interests is important due to it may conduce the local processes to be too far from the optimum. Moreover, the unknown number of classes and number of texture descriptors are existing problems in this unsupervised segmentation.

In conclusion, this initial study toward lettrine segmentation deals with the top-down stage of a whole system. The prospective bottom-up phase will work with notion of fuzzy neighbourhood to extract areas of interest based on the similarity of lines. This will lead us to gain much more sense of segmentation.

References

[1] *Masse de Données issues de la Numérisation du patrimoine*, http://l3iexp.univ-lr.fr/madonne/

[2] *Les Bibliothèques virtuelles humanistes*, http://www.bvh.univ-tours.fr/

[3] U. Neisser, "Cognition and Reality", *Freeman*, New York, 1976.

[4] J. den Hartog, T. ten Kate, and J. Gerbrands, "Knowledge-Based Segmentation for Automatic Map Interpretation", *Lecture Notes in Computer Sciences*, vol. 1072, 159-178, 1996.

[5] R.M. Haralick, K hanmugan, and I. Dinstein, "Textural features for image classification", *IEEE transactions on Systems, Man, and Cybernetics*, vol. 3, 610–621, 1973.

[6] X. TANG, "Texture Information in Run-Length Matrices", *IEEE transactions on image processing*, vol. 7, 1602-1609, 1998.

[7] R. Sriram, J. M. Francos, and W. A. Pearlman, "Texture Coding Using a Wold Decomposition Based Model", *IEEE Trans. Image Proces.*, vol. 5, 1382-1386, September 1996.

[8] Univesity of Oulu, Texture classification (32 Brodatz textures) , http://www.ee.oulu.fi/research/imag/texture/texture.php?page=Brodatz32

[9] C. Rosenberger, "Mise en Oeuvre d'un système Adaptative de Segmentation d'Images", *Ph.D. Thesis.: Rennes*, 172 p, 1999.

[10] D. PuigMiguel, A. García, "Determining optimal window size for texture feature extraction methods", *IX Spanish Symposium on Pattern Recognition and Image Analysis*, Castellon, Spain, vol.2, 237-242, May 2001.

[11] E. Borenstein, E. Sharon, S. Ullman, “Combining Top-down and Bottom-up Segmentation”, *Proceedings IEEE workshop on Perceptual Organization in Computer Vision*, IEEE Conference on Computer Vision and Pattern Recognition, Washington, DC, June 2004.

[12] J. Zhang, T. Tan, “Brief review of invariant texture analysis methods”, *Pattern Recognition*, Vol. 35, Number 3, March 2002.

[13] M. Sharma, M. Markou, S. Singh, “Evaluation of Texture Methods for Image Analysis”, *Pattern Recognition Letters*, 2000.

[14] M. Singh, S. Singh, “Spatial Texture Analysis: A comparative Study”, *16th International Conference on Pattern Recognition (ICPR'02)*, vol. 01, no. 1, 10676, 2002.

[15] Th. Hermes, A. Miene, P. Kreyenhop, “On Textures: A Sketch of a Texture-Based Image Segmentation Approach”, *Classification and Information Processing at the Turn of the Millenium*, 219-226, Bielefeld, Germany,10-12 March 2000.