

# “Inside the Bible”: Segmentation, Annotation and Retrieval for a New Browsing Experience

Costantino Grana  
Università di Modena e  
Reggio Emilia  
Via Vignolese 905  
41100, Modena, Italy  
+39 059 2056265  
costantino.grana@unimore.it

Daniele Borghesani  
Università di Modena e  
Reggio Emilia  
Via Vignolese 905  
41100, Modena, Italy  
+39 059 2056270  
daniele.borghesani@unimore.it

Simone Calderara  
Università di Modena e  
Reggio Emilia  
Via Vignolese 905  
41100, Modena, Italy  
+39 059 2056270  
simone.calderara@unimore.it

Rita Cucchiara  
Università di Modena e  
Reggio Emilia  
Via Vignolese 905  
41100, Modena, Italy  
+39 059 2056136  
rita.cucchiara@unimore.it

## ABSTRACT

In this paper we present a system for automatic segmentation, annotation and image retrieval based on content, focused on illuminated manuscripts and in particular the Borsone D'Este Holy Bible. To enhance the interaction possibilities with this work, full of decorations and illustrations, we exploit some well known document analysis techniques in addition to some new approaches, in order to achieve good segmentation of pages into meaningful visual objects with the relative annotation. We wanted to extend the standard keyword-based retrieval approach in a commentary with a modern visual-based retrieval by appearance similarity: an entire software user interface for exploration and visual search of illuminated manuscripts.

## Categories and Subject Descriptors

I.4.0 [Computing Methodologies]: Document and text processing – Document Capture - *Document analysis, Graphics recognition and interpretation*. I.4.6 [Computing Methodologies]: Image Processing and Computer Vision – Segmentation - *Pixel classification*. I.5.5 [Computing Methodologies]: Pattern Recognition – Implementation - *Interactive systems*.

## General Terms

Algorithms, Design.

## Keywords

Document analysis, Texture analysis, Von Mises distribution.

## 1. INTRODUCTION

Ancient written documents are part and parcel of the entire cultural and historical heritage of mankind. The amount of available printed or painted documents is huge: historical documents, artistic works, technical manuals, journals, newspapers, literary works are just some examples. Their value requires a specific and accurate attention for their preservation, in order to secure them for the generations to come. These

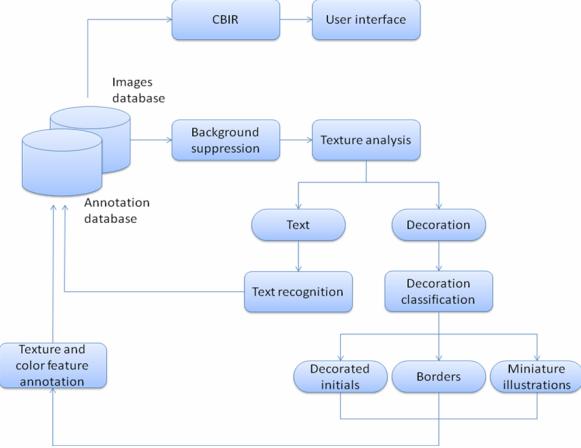
Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MIR '08, October 30–31, 2008, Vancouver, British Columbia, Canada.  
Copyright 2008 ACM 978-1-60558-312-9/08/10...\$5.00.

collections can be safely acquired for backup purposes: tools for digitalization at high resolution are nowadays one of the most easy and effective way to preserve their content. The benefits of this procedure relapse also to the way these data can be managed for analysis and fruition purposes as digital libraries.

In the last years, a great work has been done regarding digital libraries, and document image analysis literature has proposed several techniques for content and layout segmentation to provide the basis for semantic annotation, classification and retrieval. For the implementation over a large collection of digital documents, the accuracy of the analysis and the computational effort required are both significant. Indeed most of the works has been developed for printed documents, while on illuminated manuscripts only a limited amount of works has been carried out. A reference paper is the description of DEBORA [1], which consists of a complete system for the analysis of Renaissance digital libraries. This specific use case, that deals with ancient books or illuminated manuscripts, is of particular interest. These books cannot be flipped through by the public due to their value and delicacy, but computer science can fill the gap between people and all these precious libraries of masterpieces. Digital versions of the artistic works can be publicly accessible, both locally at the museum owning the original version and remotely: in this manner, users – either experts, tourists or people keen on art – can explore a manuscript more comprehensively, choosing their own personal way to browse and enjoy it.

In this paper, we are proposing the on-going results of a project called “Inside the Bible”, since it started with the Holy Bible of Borsone d'Este, duke of Ferrara (Italy) from 1450 d.C. to 1471 d.C. This book is considered one of the best Renaissance illuminated manuscripts of the world, but it is not freely accessible to the public. A lot of valuable illuminated illustrations are depicted: different mythological and real animals, court life illustrations, and some of them even testify the first attempts in exploring prospective for landscapes. Their access in a digital manner is not easy: the amount of data to process is high, a manual annotation is too long to be feasible, moreover the background is sometimes spoiled: segmentation of illustrations and text is not trivial. In the following sections we deal with the segmentation, the annotation and the retrieval based on content tasks. In our proposal, we exploit some texture analysis techniques based on circular statistics, which takes into account angular directions in an efficient way and it is particularly suitable to segment handwritten text and illustrations. Then we address some proposal for user friendly interfaces in order to enhance browsing through visual similarity. Our goal is to embed a masterpiece within a powerful



**Figure 1.** System overview.

and appealing software infrastructure able to catch the more significant details of the work, and present them to the user in an innovative way. We remark that most of the techniques are not tailored for a specific manuscript only, but they are generally applicable in painted documents.

The paper is organized as follows: Section 2 presents an overview of the system; Section 3 focuses on the algorithms tested, starting from the background suppression down to the segmentation and classification stages; Section 4 briefly proposes an idea of annotation standard for the data we are managing; finally Section 5 presents the retrieval stage and the user interface we are proposing for this application.

## 2. OVERVIEW OF THE SYSTEM

The overall scheme of the proposed system is shown in Figure 1. A centralized database stores all the digitalized data and all the necessary annotations, computed automatically by the system or added manually. In particular, the high resolution digitalized replicas of the Bible's pages constitutes the *image database*, while the *annotation database* contains both the automatically extracted knowledge and the historical comments added by experts. A first preprocessing stage of background suppression aims to remove the background from the pages, then a foreground analysis is performed for the remaining objects. Initially

decorations are divided from text, modeling the autocorrelation matrix using mixtures of Von Mises distributions. Then we proceed with the classification of the decorations: we analyze the remaining amount of images separating decorated initials, ornamental sketches and miniature illustrations. As expected, miniature illustrations are the most important parts within the image: we will focus only on them in the content-based image retrieval stage.

The retrieval stage shares the same general structure of any common CBIR system: feature extraction and classification details will be presented in a later section. It is the basis for the user interface module, that integrates the visual (and keyword-based, in the final release) search engine to propose to the user innovative browsing and exploration capabilities.

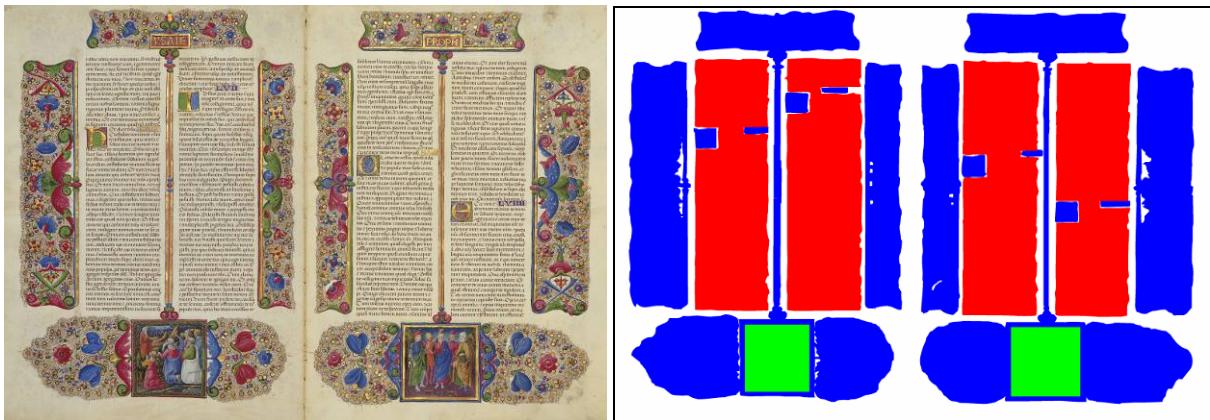
## 3. DOCUMENT SEGMENTATION

Analyzing illuminated manuscripts is a complex task since the quality of the digitalized versions is heavily affected by their preservation status and the acquisition system. Small rotations or scaling can occur, pages can be spoiled, grayscale or low quality acquisition is also possible, generally resulting in a set of noisy textures. Moreover different manuscripts have different contents and layout. For all these reasons, a simple approach based on color, shape or layout analysis may not be effective enough for a large scale implementation. The solution we propose accounts for all these problems to ensure robust segmentation and precise extraction of objects of interest.

The processed documents can be mainly divided in three semantically different parts: background, text and decorations, containing illustrations, decorated initials and ornamental borders (Figure 2).

### 3.1 Background Suppression

A background suppression technique must be exploited to remove the page background focusing on the regions of interest of the digitalized image with the important advantage of decreasing the amount of pixels to analyze in the subsequent stages. We tested several well established techniques for background suppression. In order to balance computational effort and stability of the results, we chose to employ the IGT [2] (Iterative Global Thresholding), an adaptive binarization techniques that selects a global threshold starting from a grayscale image using an iterative



**Figure 2.** Example of illuminated manuscripts and relative ground truth. White identifies background, red identifies text, blue identifies decoration and green identifies decorated illustrations.

procedure, also adopted for printed documents. In this context, IGT has proved to be both effective and easy to implement. The background can be removed successfully with an high precision (Figure 3).

### 3.2 Texture Analysis

We propose a new method for texture analysis to distinguish between textual and pictorial information. Our method is based on an approach proposed in [3], improved with the adoption of circular statistics, in particular the Von Mises distributions. The basic feature we exploited is the autocorrelation matrix, an effective feature in this case, since textual textures have a pronounced orientation that heavily differs both from background and decorations. Formally, the autocorrelation function is the cross correlation of a signal with itself, and it represents a measure of similarity between two signals. Once applied to a grayscale images, it produces a central symmetry matrix, that gives an idea of how regular the texture is.

The image is divided into square blocks whose size  $bs$  must be set according to the scale at which the texture should be analyzed. The definition of the autocorrelation for a block is:

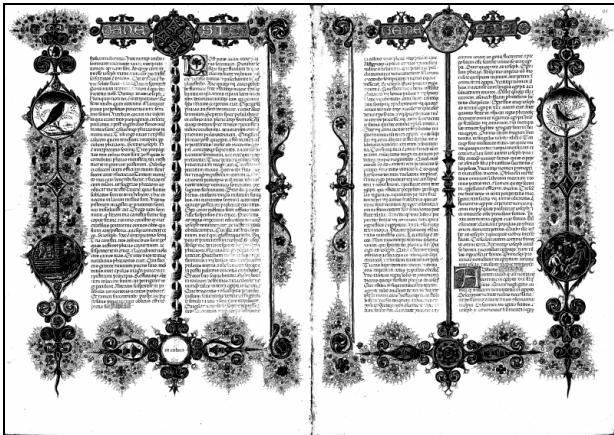
$$C(k,l) = \sum_{y=\max(0,l)}^{bs-1+\min(0,l)} \sum_{x=\max(0,k)}^{bs-1+\min(0,k)} I(x,y) \cdot I(x+k, y+l) \quad (1)$$

where  $l$  and  $k$  are defined in  $[-bs/2, bs/2]$ .

The result of the autocorrelation can be analyzed extracting an estimate of the relevant directions within the texture. Each angle determines a direction, and the sum of all the pixels along each direction is computed to form a polar representation of the autocorrelation matrix, called *direction histogram*. In this way, each direction will be characterized by a weight, indicating its importance within the block.

$$w(\theta) = \sum_{r \in [0, \frac{b}{2}]} C(r \cos \theta, r \sin \theta). \quad (2)$$

Since the autocorrelation matrix has a central symmetry by definition, we consider only the first half of the direction histogram in the range  $[0^\circ, 180^\circ]$ .  $\theta$  and  $r$  are quantized: the step of  $\theta$  is set to  $1^\circ$ ,  $r$  is defined as  $2/b$  of the block size. A text



**Figure 3. Result for the background suppression.**

block will be characterized by peaks around  $0^\circ$  and  $180^\circ$  because of the dominant direction is horizontal, and this behavior is different compared to pictorial textures (described by a generic monomodal or multimodal distribution) and also background textures (described by a nearly uniform flat distribution).

Then the polar distribution obtained by autocorrelation in the previous step is modeled in a statistic framework. The standard Gaussian distributions are inappropriate to model angular datasets: regarding our specific case, there is a discontinuity between  $0^\circ$  and  $180^\circ$ , so these two angles vote for two distinct directions even if they express the same one, producing a bad fitting. Instead, we exploit Von Mises distributions [4] that can correctly represent angular datasets. This statistical formulation has been rarely used for texture analysis, while it has been previously presented for trajectory shape classification [5].

The probability density function is defined as follows:

$$V(\theta | \tilde{\theta}, m) = \frac{1}{2\pi I_0(m)} e^{m \cos(\theta - \tilde{\theta})}. \quad (3)$$

The parameter  $m$  denotes how concentrate the distribution is around the mean angle  $\tilde{\theta}$ . In our context, we used a slightly different formulation (we simply multiply the angles by 2) with a periodicity of  $\pi$  instead of  $2\pi$ , considering only angles in  $[0, \pi)$  representative for valuable and meaningful directions.  $I_0$  is the modified order 0 Bessel function, and is defined as:

$$I_0(m) = \frac{1}{2\pi} \int_0^{2\pi} e^{m \cos \theta} d\theta. \quad (4)$$

To catch the general multimodal behavior of input datasets, we chose a mixture of Von Mises distributions. We used mixtures with 2 components only, because they proved to be sufficient in order to recognize the two most meaningful directions (horizontal and vertical) while keeping an affordable computational cost. An example of fitting for the two types of texture analyzed is shown in Figure 4. The noisy background is also characterized by an almost flat distribution.

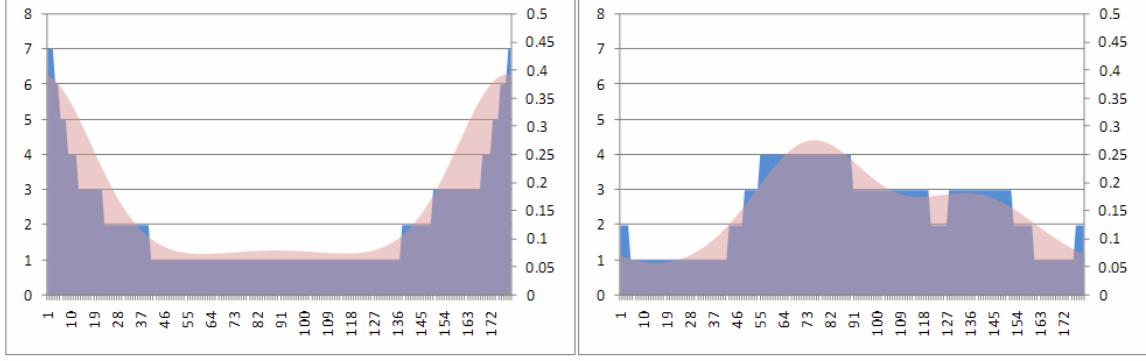
A mixture of  $K$  Von Mises distributions is defined as follows:

$$M(\theta) = \sum_{k=1}^K \pi_k V(\theta | \tilde{\theta}_k, m_k), \quad (5)$$

where  $\pi_k$  represents a weight of the distribution within the mixture. An optimal way to get the maximum likelihood estimates of the mixture parameters is the Expectation-Maximization algorithm [5]. In the E step the expected values for the likelihood are computed, then a set of parameters to maximize such values are obtained, repeating the process until convergence or maximum number of iterations is reached. To maximize the likelihood, a set of responsibilities of the bins for each Von Mises is necessary. Let  $\theta$  be the index of the bin. The responsibilities are computed as follows:

$$\gamma_{\theta k} = \frac{\pi_k V(\theta | \tilde{\theta}_k, m_k)}{\sum_{s=1}^K \pi_s V(\theta | \tilde{\theta}_s, m_s)}. \quad (6)$$

A new set of weights for the Von Mises of the mixture can now be computed:



**Figure 4. Example of directional histograms and the corresponding fitting with Von Mises mixtures for text and decorations.**

$$B(V_1, V_2) = \sqrt{1 - \sqrt{\frac{1}{I_0(m_1)I_0(m_2)}} I_0\left(\frac{\sqrt{m_1^2 + m_2^2 + 2m_1m_2 \cos 2(\tilde{\theta}_1 - \tilde{\theta}_2)}}{2}\right)} \quad (10)$$

$$\pi_k = \frac{\sum_{\theta \in [0, \pi)} w_\theta \gamma_{\theta k}}{\sum_{\theta \in [0, \pi)} w_\theta}. \quad (7)$$

This formulation differs from the one in [5], and the motivation lies on the dataset we used: we do not have a general distribution of angular data to fit, but a sampling of directions and relative weights. For this reason, we consider the weight as a multiplier value for each angle, so formally we have  $w_\theta$  times the angle  $\theta$  in our dataset.

In the M step, we compute the new  $\theta$  and  $m$  values for each Von Mises within the mixture. In particular,  $\theta$  is computed by maximization of the relative likelihood as follows:

$$\tilde{\theta}_k = \arctan\left(\frac{\sum_{\theta \in [0, \pi)} w_\theta \gamma_{\theta k} \sin 2\theta}{\sum_{\theta \in [0, \pi)} w_\theta \gamma_{\theta k} \cos 2\theta}\right). \quad (8)$$

Note the multiplication by 2 in order to relate to a  $\pi$  periodicity. The retrieval of  $m$  by maximization is a bit more complicate, due to the presence of the Bessel functions. Given the derivative of the modified Bessel function  $I_1$ , defined as:

$$I_1(m) = \frac{1}{2\pi} \int_0^{2\pi} e^{m \cos \theta} \cos \theta d\theta, \quad (9)$$

the problem could be mathematically solved using this formulation:

$$A(m_k) = \frac{I_1(m_k)}{I_0(m_k)} = \frac{\sum_{\theta \in [0, \pi)} w_\theta \gamma_{\theta k} \cos(\theta - \tilde{\theta}_k)}{\sum_{\theta \in [0, \pi)} w_\theta \gamma_{\theta k}}. \quad (11)$$

The value of  $m_k$  can be found by the numerical inversion of  $A(m_k)$ . In particular we use the approximation proposed in [6].

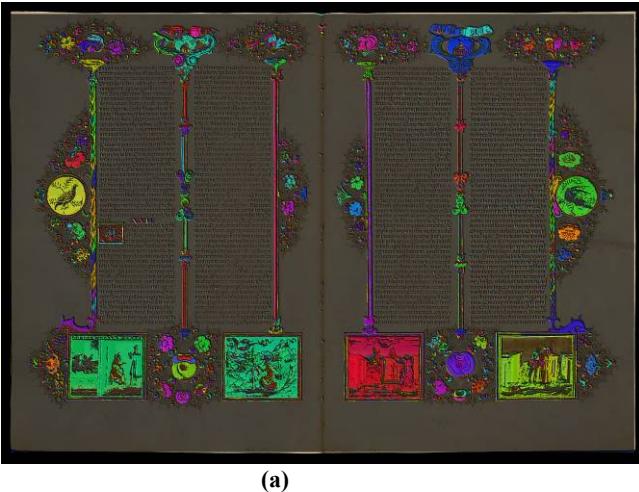
At this point, we have 6 parameters to play with:  $\pi_k$ ,  $\tilde{\theta}_k$ , and  $m_k$  of both Von Mises. This represents a very consistent and compact way to describe a whole distribution, making the retrieval faster and effective.

The similarity between two Von Mises distributions can be defined using the Bhattacharyya distance. Given two Von Mises distributions  $V_1$  and  $V_2$ , the formulation is shown in Equation 10. No explicit form is available for mixtures, so we propose a metric that also takes into account the relative weights of the components of the mixture. Given two mixture distributions  $M^i(\theta) = \sum_{k=1}^2 \pi_k^i V(\theta | \tilde{\theta}_k^i, m_k^i)$ , we computed the Bhattacharyya distance between pairs of distributions and selected the best matching two (calling them  $b$ , while the other two  $o$ ). Then we measure the distance as:

$$d(M^1, M^2) = \frac{WB_b + WB_o}{\pi_b^1 \pi_b^2 + \pi_o^1 \pi_o^2}, \quad (12)$$

**Table 1. SVM classification using radial basis function as kernel. A training set of 20 images with manual ground truth has been used.**

	recall	precision
<b>text</b>	0.931183	0.911579
<b>decoration</b>	0.826138	0.898477



(a)



(b)

**Figure 5. Good (a) and bad (b) results of segmentation using only the morphological approach. The first image highlights the correct segmentation of 6 illustrations (4 rectangular and 2 circular). The second image shows some cases of over-segmentation and under-segmentation**

where

$$WB_x(M^1, M^2) = \pi_x^1 \pi_x^2 B(V_x^1, V_x^2) \quad (13)$$

This metric takes into account the fact that two components can be very similar, but their contribution to the mixtures is quite low.

Some classification results are shown in Table 1. A SVM has been adopted, in particular the best results (in terms of recall and precision) have been obtained exploiting the radial basis kernel.

### 3.3 Text Analysis and Initials Segmentation

Regarding the text obtained by the segmentation process, there are several ways to treat it. A first proposal is the exploitation of OCR system trained with gothic letters. An example is provided in [7] where several OCRs and specialized ICRs (Intelligent Character Recognition) based on convolutional neural networks has been mixed to achieve good performance with ancient fonts unusual in contemporary documents. A word spotting technique is also possible, as shown in [8]. The goal is to retrieve automatically the text by the Bible, in order to produce some potential links between decorations and textual content, and providing a nice and valuable enrichment for the browsing capabilities of this system. This work of automatic extraction of the Bible text is under development.

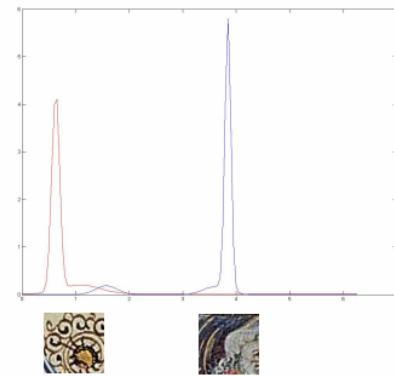
The initials have a very specific characteristic, because it is a decoration texture partially embedded into a wide textual texture. An affordable way to distinguish them is based on the layout segmentation. Text is always arranged into two columns, centered within the page, so a straightforward adaptation of a rectangular template over the two columns can be exploited: small changes on origin, width and height, based on texture features, lead to a near optimal fit of this fixed boundary template, then every decoration texture within or partially within it can be classified as initials.

### 3.4 Decoration Segmentation

The surrounding decorations are the biggest problem. Without any prior knowledge, it is necessary to extract the valuable illustrations within the decoration texture (miniature illustrations

of scenes, symbols, people and animals), rejecting all the border decorations (ornaments). It is often a challenging task to decide whether a given texture represents a meaningful illustration or just an ornament of the page, even by a human being.

The approach we followed in a first instance was based on morphology and labeling of connected components. Exploiting some operations of dilation and erosion over the thresholded image, we can effectively segment the components of the decoration, decreasing the influence of the ornamental parts. Some good segmentation are shown in Figure 5a., but sometimes this operation has the side effect to over-segment a component or to under-segment a region, as shown in Figure 5b. For this reason, we adopted a second approach in order to validate the regions that more probably will contain valuable illustrations. The texture is subdivided in square blocks, and for each block the conversion to HSV color space is computed. We keep only the hue component: it contains information about the color, and it is defined in the range  $[0^\circ, 360^\circ]$ . This angular data is used as input set for a Von Mises modeling, similarly to what we have done in Section 3.2: a mixture of  $k$  components is computed in order to statistically define the color appearance of the block. Figure 6 shows an example of modeling of ornamental texture and illustration. In



**Figure 6. Plot of the mixture of Von Mises modeling for ornamental texture and illustration texture.**



**Figure 7. Simple web interface for the picture navigation, exploiting PicLens, a cross-browser plugin.**

addition, to discriminate patches sharing the same color information, we exploit a Daubechies wavelet transform to compute the energy within the block as square sum of the horizontal (LH), vertical (HL) and diagonal (HH) coefficients. The energy is able to distinguish among different texture patterns because a finely decorated texture generally produces an higher energy response. The parameters of the Von Mises distribution and the image energy values have been used as input feature vector for a SVM, having the appropriate ground truth file. Generally the illustrations have a more heterogeneous color distribution compared to the other decorations, so this property can be used as a discriminative feature to solve our problem.

The joint use of both these techniques improves the performance of the system: we can successfully combine the good boundaries extracted by the previous morphological operations with the good region analysis produced by the color distribution modeling.

## 4. CONTENT ANNOTATION

### 4.1 Visual object

Within a digitalized bible image, only a small amount of visual object can be considered interesting from a browsing and content-based retrieval point of view. The most common figures are animals, scenes from the Bible story, characters, or symbols, and a specific meaning can be associated to them, or a specific part of the Bible story can be linked to them. These connections could help to find all scenes regarding a narrated event, or even understand why a certain object is depicted within a certain part of the Bible.

Each visual object is defined with a set of low level features. The associated annotation is composed by an object identifier and a binary term containing the value of relative features.

### 4.2 Localization schema

Every visual object is defined by an extent within a digitalized image, specifically a structure containing the origin point and values for width and height. The extent will contain both the picture of interest and some additional decoration or background texture: it is not intended to be a perfect boundary of the picture, but only an accurate as possible region where the picture of interest is inserted to. This formulation can be extended considering that the layout of these pages is quite standard: single centered column with text and decorated initials, external border with decorations. Even the structure of decoration is quite

standard, and it can be subdivided into different sections: top left, top, top right, left, right, bottom left, bottom and bottom right.

In order to provide a more detailed definition of the visual object (especially in case of visual object with an heavily decorated boundary), we can provide also a contour, using coordinates relative to the enclosing extent. The detection of the contour could be achieved using several boundary detection algorithms, specifically trained for this kind of images. Actually this problem has not been faced, but it will be included as a future development.

### 4.3 Meaning schema

As previously mentioned, each visual object (either character, animal, scene or object) can be associated with a meaning, part of an ontology. For the masterpiece we are working with, a complete commentary is available by the owner, so a parsing strategy has been implemented in order to match each visual object with a part of the commentary text exploiting keyword expressing locations within the page. Anyway a group of experts can analyze each detail of the Bible, providing a specific annotation and commentary for each visual object extracted. A simple meaning schema is proposed in order to contain this semantic information. A set of keywords and a commentary are part of the XML schema of the content annotation, and several information can be extracted using typical XML-based navigation algorithm through the XML tree.

## 5. CONTENT-BASED IMAGE RETRIEVAL AND BROWSING

An essential goal of this work is the definition and the design of an effective application to present all the details of illuminated manuscripts (and in particular the Borsone d'Este Holy Bible) in an easy and fashionable way. Retrieval of images by content similarity will be improved with a joint use of both pictorial and textual information. An innovative interface will enhance the user experience, providing new methodologies to interact with the work. A set of gestures for browsing, selection, zooming and manipulating contents is under development, in order to provide to the user the possibility to experience a personal way to explore it. We produce some initial application samples using popular web 2.0 applications or browser plugins. An example is shown in Figure 7.

The retrieval of illustrations by similarity exploits the texture features previously described and other color features, in particular: RGB correlogram, spatial chromatic histogram, aHSV histogram and edge histogram. RGB correlogram were firstly introduced in [9] in order to describe the spatial correlation of colors using an arbitrary spatial distance. Spatial chromatic histograms [10] extract the spatial distribution of colors. aHSV histograms [11] consist in an enhanced version of the standard HSV histogram capable to handle low saturated images. Finally we consider the texture distribution using a straightforward edge histogram [12]. Each histogram distance is computed exploiting histogram intersection approach. The fusion of these features has been implemented with a linear combination approach. Results of retrieval are very satisfactory for the browsing goal, even if not directly measured with ground truth yet. Some examples of retrieval results are shown in Figure 8.



Figure 8. Some retrieval results. The green-bordered image is the query image.

## 5.1 Browsing with Sammon Mapping

The presentation of miniature illustrations segmented from digitalized images of Bible's page is performed using a Sammon Mapping. It is a very well known type of data structuring that have the power to visually enhance correlation between data, similarities and clusters. Given  $n$  images to present to the user, the goal of the Sammon mapping in this application is to find  $n$  point in a 2D space in such a way that the corresponding distances between each pair of points in this new two-dimensional space approximate the original ones as close as possible.

The result is shown in Figure 9. The query image is positioned in the center of the map, the other images are placed around it randomly but proportionally with the distance computed using the features previously mentioned. The user can browse through other images simply clicking on them on the map: the mapping will be recalculated in real time and a new disposition will be shown to the user. For each image (both query or not) a set of information will be available to the user, for example explanations about the content of the image from a religious and historical point of view. Also it will be possible to navigate to the relative page in the Bible: this freedom of exploration would let the user discover new interesting links across the work.

## 6. CONCLUSIONS

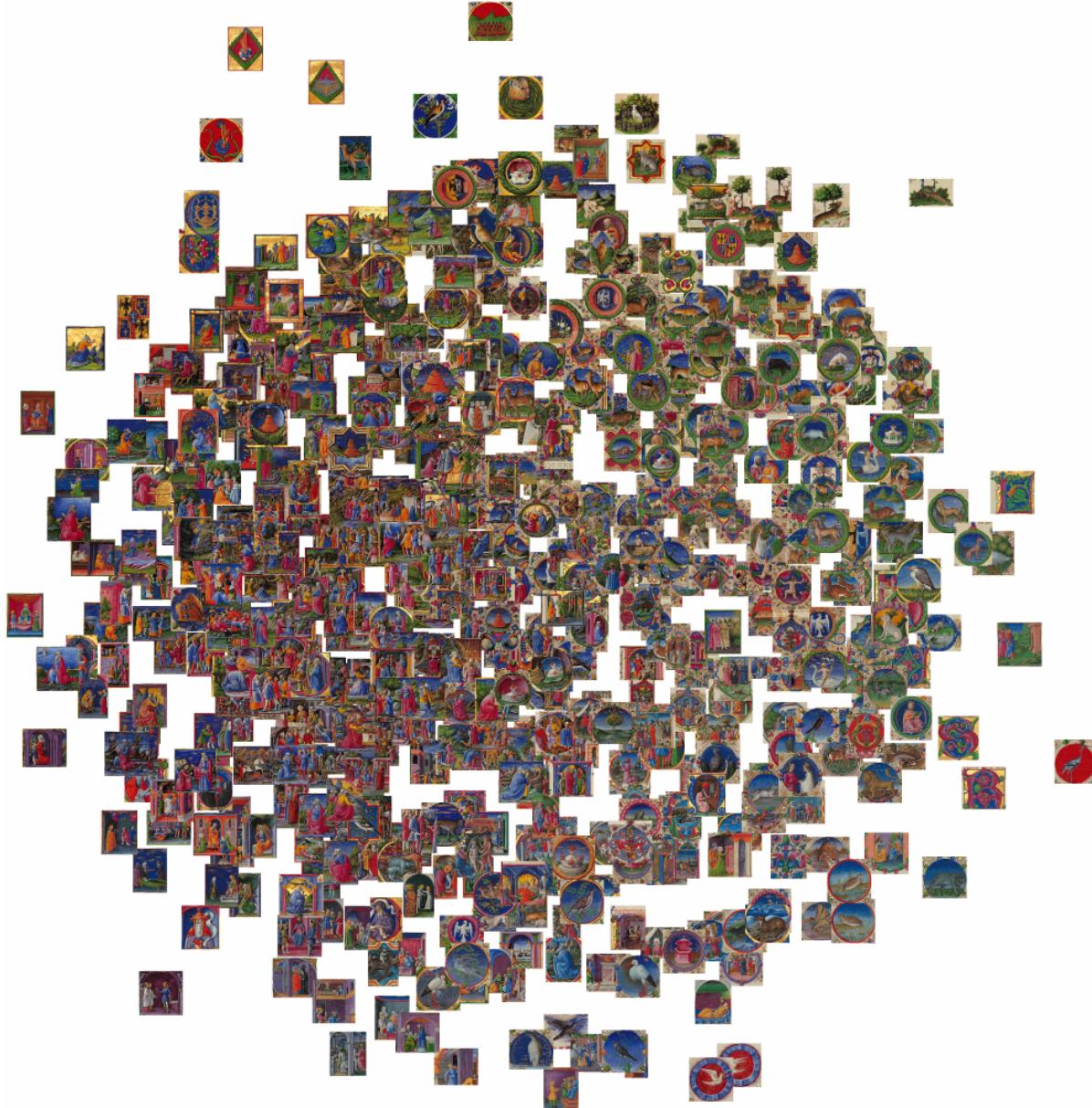
In this work we present a system for the automatic analysis and image retrieval by content, focusing on a specific set of images coming from the digitalization of the Borsone d'Este Holy Bible. A set of techniques has been studied in order to obtain good retrieval results with a limited computational effort. The texture features with Von Mises distribution is very general and could be applied in many other document analysis applications. At the same time we proposed a flexible annotation structure that will be used to store the analysis results and to integrate semantic information (coming from experts) about retrieved images. Retrieval functionalities and commentaries will be proposed as a standalone application, suitable for a large scale implementation in museums, in order to give to the user a new innovative way to interact with masterpieces such the one we exploited in this work.

## 7. ACKNOWLEDGMENTS

Our thanks to Franco Cosimo Panini S.p.a. that S.P.A. that provides as the digital replica of the as The Borsone D'Este Holy.

## 8. REFERENCES

- [1] Le Bourgeois, F., Emptoz, H. 2007. DEBORA: Digital AccEss to Books of the RenAissance. In International Journal on Document Analysis and Recognition, vol. 9, n. 2-4, pp. 193-221.
- [2] Kavallieratou E. 2005. A Binarization Algorithm Specialized on Document Images and Photos. In Proceedings of the 8th Int. Conf. on Document Analysis and Recognition, pp. 463-467.
- [3] Journet, N., Eglin, V., Ramel, J.Y., Mullot, R. 2006. Dedicated texture based tools for characterisation of old books. In Proceedings of the Second International Conference Document Image Analysis for Libraries (DIAL 2006, April 27-28, 2006).
- [4] Bishop C. 2006. Pattern Recognition and Machine Learning", Springer-Verlag.
- [5] Prati A., Calderara S., Cucchiara R. 2008. Using Circular Statistics for Trajectory Analysis. In Proceedings of International Conference on Computer Vision and Pattern Recognition (CVPR 2008, Anchorage, Alaska June 24-26, 2008).
- [6] Hill G.W. 1981. Evaluation and Inversion of the Ratios of Modified Bessel Functions,  $I_0(x)/I_1(x)$  and  $I_{1.5}(x)/I_{0.5}(x)$ . In ACM Transactions on Mathematical Software, vol. 7, n. 2, June 1981, pp. 199-208.
- [7] Cecotti, H., Belaid, A. 2005. Hybrid OCR combination approach complemented by a specialized ICR applied on ancient documents, vol.2, pp. 1045-1049.
- [8] Leydier, Y., Le Bourgeois, F., Emptoz, H. 2005. Omnilingual segmentation-free word spotting for ancient manuscripts indexation. In Proceedings of the Eighth International Conference on Document Analysis and Recognition, vol.1, pp. 533-537.
- [9] Huang J., et. al. 1997. Image indexing using color correlogram. In Proceedings of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Proceedings, pp. 762-768.



**Figure 9. Sammon mapping of the retrieval results. The query image is in the center, while the other images are arranged according to the distances between feature vectors (the near ones are the most similar). This visualization has the power to enhance outdoor scenes (top left), animals (top right), scenes with a lot of people (center left), indoor scenes (bottom left) and birds and aquatic animals (center bottom).**

- [10] Cinque, L., Levialdi, S., Olsen, K.A., Pellicano, A. 1999. Color-based image retrieval using spatial-chromatic histograms. In International Conference on Multimedia Computing and Systems, vol.2, pp 969-973.
- [11] Grana C., Vezzani R., Cucchiara R.. 2007. Enhancing HSV Histograms with Achromatic Points Detection for Video Retrieval. In Proceedings of ACM International Conference on Image and Video Retrieval (CIVR 2007), pp. 302-308.
- [12] Won C.S., Park D.K., Park S.J. 2002. Efficient Use of MPEG-7 Edge Histogram Descriptor, ETRI Journal, vol.24, no.1, pp.23-30
- [13] Sammon, J.W. Jr.. 1969. A nonlinear mapping for data structure analysis. IEEE Transactions on Computers. C-18(5):401-409