

Describing Texture Directions with Von Mises Distributions

Costantino Grana, Daniele Borghesani, Rita Cucchiara
Dipartimento di Ingegneria dell'Informazione
Università degli Studi di Modena e Reggio Emilia
name.surname@unimore.it

Abstract

In this work we describe a new approach for texture characterization. Starting from the autocorrelation matrix an elegant description through a mixture of Von Mises distributions is proposed. A compact 6 valued descriptor is produced for each block and served as input to an SVM classifier. Tests are carried out on high resolution illuminated manuscripts images.

1. Introduction

Document image analysis has a quite long story in pattern recognition. Several techniques have been proposed 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. A specific use case are ancient books or illuminated manuscripts, that cannot be flipped through by the public due to their value and delicacy. Computer science has the power to fill the gap between people and all these precious libraries of masterpieces: in fact digital versions of the artistic works can be publicly accessible, either locally or remotely, giving the user the freedom to choose his personal way to navigate and enjoy.

The quality of images of illuminated manuscripts heavily depends on the way they have been acquired or the preservation status of the work. 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 would not be effective enough for a large scale implementation. In this paper, a flexible approach to the problem is proposed based on texture analysis.

As shown in Fig. 1, images are analyzed by blocks through autocorrelation, and a texture direction histogram is computed for each block. Then we describe it with a mixture of Von Mises distributions (MoVM), a

statistical formulation that is more suitable for angular data than a mixture of Gaussians. We implemented an EM algorithm for parameters extraction and finally we exploited Support Vector Machine (SVM) for block classification. The goal is to provide a very fast and compact representation for each class, in order to make the retrieval as fast and effective as possible. Experiments have been carried out on illuminated manuscripts of the Borso D'Este Holy Bible.

2. Related work

Document segmentation is normally based on partitioning of image in blocks and then texture analysis. Several works for text segmentation have been proposed: a clustering approach is presented in [1], while in [2] a classification using Gabor filters has been used. A comprehensive survey is proposed by Busch et.al. [3] exploring and comparing several techniques and situations. More general approaches dealing also with background and pictures segmentation have been proposed. Some works exploit geometric constraints over the layout: a literature survey is provided in [4]. Many others compute specific descriptors followed by classification: an example is provided by [5] with hidden tree Markov models.

The majority of works have been developed for printed documents, while only limited work has been carried out on illuminated manuscripts. A reference paper is the description of the DEBORA system [6], which consists of a complete system for analysis of Renaissance. In [7] an effective technique for texture

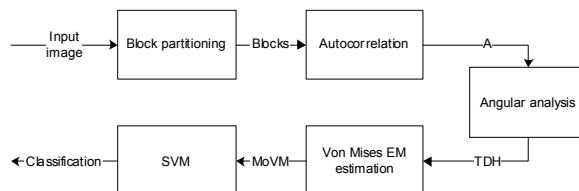


Figure 1. Block diagram of the proposed approach.

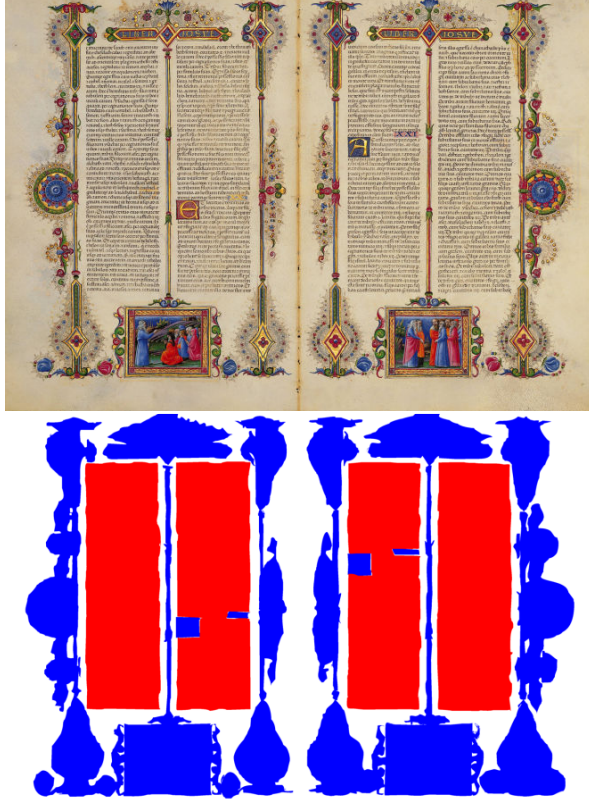


Figure 2. Example of illuminated manuscripts and relative ground truth. White identifies background, red identifies text and blue identifies pictures or decorations.

characterization in old books has been proposed, exploiting the autocorrelation matrix in order to extract the relevant directions within the texture (called directional rose). In our work, we extended this approach formulating an elegant description of the different directions within the block with the use of mixtures of Von Mises distributions.

3. The proposed approach

The image (Fig. 2) is divided into square blocks of size bs , set according to the scale, on which the autocorrelation is computed. The autocorrelation function is the cross correlation of a signal with itself and, applied to a grayscale image, it produces a central symmetry matrix, that gives an idea of how regular the texture is. It is defined as:

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

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. The sum of all the pixels along each direction angle is computed to form a polar representation of the autocorrelation matrix, that we call Texture Direction Histogram. In this way, each direction will be characterized by a value, indicating its importance within the block.

$$h(\theta) = \sum_{r \in (0, bs/2]} C(r \cos \theta, r \sin \theta). \quad (2)$$

Since the autocorrelation matrix has a central symmetry by definition, we consider only the first half of the direction histogram, from 0° to 179° . θ is quantized with a 1 degree step and r is divided in steps of $2/bs$.

The TDH can be modeled using Von Mises distributions. Gaussian distributions are inappropriate to model periodic datasets because of the discontinuity at 0 and π . Instead, a Von Mises distribution is circularly defined, so it can correctly represent angular datasets. Its probability density function is:

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

The parameter m denotes how concentrate the distribution is around the mean angle $\tilde{\theta}$. In our context, we used a slightly different formulation setting a periodicity of π instead of 2π . I_0 is the modified zero order Bessel function.

To catch the general multimodal behavior of input datasets, we chose a mixture of K Von Mises distributions defined as:

$$M(\theta) = \sum_{k=1}^K \pi_k V(\theta | \tilde{\theta}_k, m_k), \quad (4)$$

where π_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 EM algorithm [8]. To maximize the likelihood in the E step, a set of responsibilities of the bins for each Von Mises is computed (θ is the index of the bin):

$$\gamma_{\theta k} = \frac{\pi_k V(\theta | \tilde{\theta}_k, m_k)}{\sum_{s=1}^K \pi_s V(\theta | \tilde{\theta}_s, m_s)}. \quad (5)$$

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

$$\pi_k = \frac{\sum_{\theta \in (0, \pi)} h_\theta \gamma_{\theta k}}{\sum_{\theta \in (0, \pi)} h_\theta}. \quad (6)$$

This formulation differs from the one in [8], and the motivation lies on the dataset we used: we do not have a general distribution of angular data to fit, but a sam-

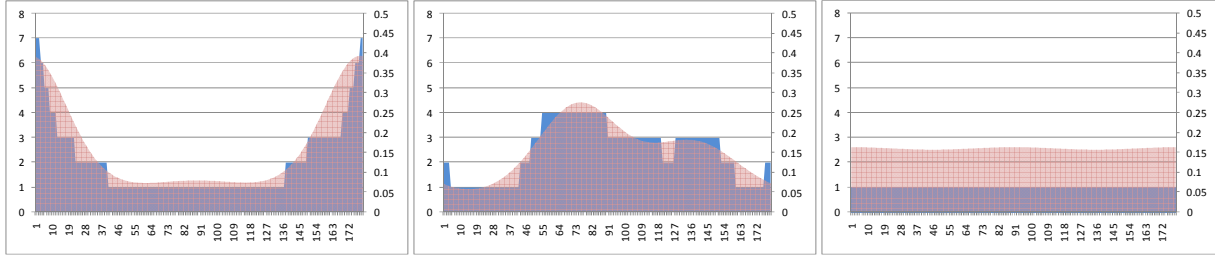


Figure 3. Directional histograms and corresponding fitting with Von Mises mixtures. A text block is characterized by peaks around 0° and 180° because of the dominant direction is horizontal. An image has a generic monomodal or multimodal distribution, while a background texture is described by a nearly uniform flat distribution.

$$B(V_1, V_2) = \sqrt{1 - \frac{1}{\sqrt{I_0(m_1)I_0(m_2)}} I_0\left(\frac{\sqrt{m_1^2 + m_2^2 + 2m_1m_2 \cos 2(\tilde{\theta}_1 - \tilde{\theta}_2)}}{2}\right)} \quad (9)$$

pling of directions and relative weights. For this reason, we consider the histogram value as a multiplier for each angle.

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

$$\tilde{\theta}_k = \arctan \left(\frac{\sum_{\theta \in [0, \pi]} h_\theta \gamma_{\theta k} \sin 2\theta}{\sum_{\theta \in [0, \pi]} h_\theta \gamma_{\theta k} \cos 2\theta} \right). \quad (7)$$

Note the multiplication by 2 in order to relate to a π periodicity. The retrieval of m by maximization is a more complicate, due to the presence of the Bessel functions. The value of m_k can be obtained by the numerical inversion [9] of:

$$A(m_k) = \frac{I_1(m_k)}{I_0(m_k)} = \frac{\sum_{\theta \in [0, \pi]} h_\theta \gamma_{\theta k} \cos(\theta - \tilde{\theta}_k)}{\sum_{\theta \in [0, \pi]} h_\theta \gamma_{\theta k}}. \quad (8)$$

where I_1 is the derivative of I_0 .

At the end of the EM optimization we have 3 parameters π_k , $\tilde{\theta}_k$, and m_k for every Von Mises. This is a

	text	background	image	recall
text	431	5	54	0.879592
background	7	446	172	0.7136
image	27	63	631	0.875173

Table 1. Average confusion matrix relative to a 1-nearest neighbor classification.

consistent and compact way to describe a whole distribution, making the retrieval faster and effective.

In this work we used mixtures with two components only, because they proved to be sufficient to recognize the three textures within the manuscript images. An example of fitting for the three types of analyzed textures is shown in Fig. 3.

The similarity between two Von Mises distributions can be measured using the Bhattacharyya distance, as in Eq. 9. No explicit form is available for mixtures, so we propose a new 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 test the Bhattacharyya distance between the two components of one distribution and the two of the other, select the best matching two (call them b , and the other two o) and then measure the distance as:

$$d(M^1, M^2) = \frac{WB_b + WB_o}{\pi_b^1 \pi_b^2 + \pi_o^1 \pi_o^2}, \quad (10)$$

where

$$WB_x(M^1, M^2) = \pi_x^1 \pi_x^2 B(V_x^1, V_x^2) \quad (11)$$

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

4. Experimental results

Tests have been performed on about 36000 blocks taken from 20 uncompressed 8373x6039 24bit pictures of biblical illustrated manuscripts. The data set has been manually annotated in three classes: text, images

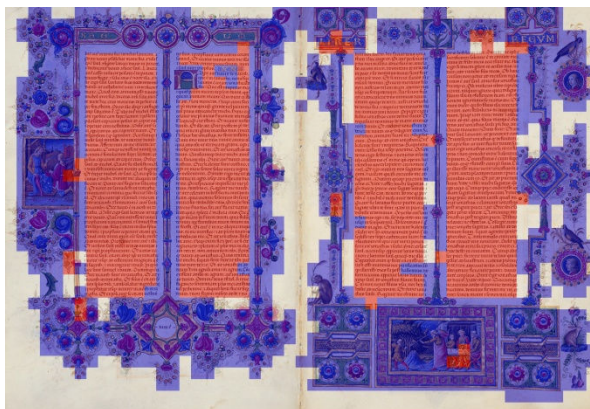


Figure 4. Visualization of the proposed classification.

and background. For each block the TDH has been computed.

In order to represent the data distribution coherently with the real shape distribution, we chose to subtract to each bin the minimum of the entire block. In this way, a noisy block like the background will show a nearly flat distribution with a very small variance. Instead a text block, with a dominant direction over the horizontal axis, will continue to show a coherent distribution, with peaks near 0° and 180° . The histogram bins have been quantized to a fixed step. The normalized distribution has been used to fit a mixture of Von Mises distributions.

To perform a first evaluation, a confusion matrix has been computed using a 1-nearest neighbor approach. The training set was analyzed, then a test set was classified. The results (Table 1) were quite promising: this feature has a good discriminative power with all three kinds of texture used.

In order to produce a more efficient classifier, we employ Support Vector Machines. The best results (in terms of recall and precision) have been obtained using the Radial Basis Function kernel. An example output is shown in Fig. 4.

The results of Table 2 confirm that the mixture of two Von Mises is particularly suitable for manuscript text classification. We would like to underline that no post processing stage has been carried out, like morphological operators which have been tested to be very effective in removing small and localized outliers. The approach is still very general and can be safely applied to different contexts.

5. Conclusions

In this paper a new technique for document analysis is presented, characterizing each texture with a 6-vector feature. The autocorrelation computation and the fitting of the direction histogram with the mixture of Von Mises distributions has a reasonable processing

	recall	precision
text	0.931183	0.911579
background	0.854086	0.87976
image	0.826138	0.898477

Table 2. SVM classification using radial basis function as kernel.

time for this application: a high resolution page with more than 1800 blocks can be processed in about 100 seconds on a standard PC. The result of the feature extraction provides a very compact description of blocks. For this reason, these features can be easily exploited by a content-based image retrieval system: the computational time needed for comparison and searching will be extremely low.

We thank Franco Cosimo Panini S.p.A. for providing access to their digital library.

References

- [1] S. Bres, W. Eglin and A. Gagneux. Unsupervised clustering of text entities in heterogeneous grey level documents. In *Int. Conf. on Pattern Recognition*, pages III-224-227, 2002.
- [2] A. K. Jain and S. Bhattacharjee. Text segmentation using Gabor filters for automatic document processing. *Machine Vision and Applications*, 5(3):169-184, 1992.
- [3] A. Busch, W.W. Boles and S. Sridharan. Texture for script identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(11):1720-1732, November 2005.
- [4] S. Mao, A. Rosenfeld and T. Kanungo. Document structure analysis algorithms: a literature survey. In *Proc. SPIE Electronic Imaging*, pages 197-207, 2003.
- [5] M. Diligenti, P. Frasconi and M. Gori. Hidden tree Markov models for document image classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(4):519-523, April 2003.
- [6] F. Le Bourgeois and H. Emptoz. DEBORA: Digital AccEss to BOoks of the RenAissance. *International Journal on Document Analysis and Recognition*, 9(2-4):193-221, 2007
- [7] N. Journet, V. Eglin, J.Y. Ramel and R. Mullot. Dedicated texture based tools for characterisation of old books. In *2nd Second International Conference on Document Image Analysis for Libraries*, April 2006.
- [8] A. Prati, S. Calderara and R. Cucchiara. Using Circular Statistics for Trajectory Analysis. In *Proceedings of International Conference on Computer Vision and Pattern Recognition*, Anchorage, Alaska (USA), June 2008.
- [9] G.W. Hill. Evaluation and Inversion of the Ratios of Modified Bessel Functions, $I_0(x)/I_1(x)$ and $I_{1.5}(x)/I_{0.5}(x)$. *ACM Transactions on Mathematical Software*, 7(2):199-208, June 1981.