TEXTURED IMAGE SEGMENTATION BASED ON SPATIAL DEPENDENCE USING A MARKOV RANDOM FIELD MODEL

William Robson Schwartz

University of Maryland
Computer Science Department
College Park, MD, USA, 20742-3275

Hélio Pedrini

Federal University of Paraná
Computer Science Department
Curitiba, PR, Brazil, 81531-990

ABSTRACT

Image segmentation is a primary step in many computer vision tasks. Although many segmentation methods have been proposed in the last decades, there is no generic method that can be applied in a great variety of images. This work presents a new image segmentation method using texture features extracted by wavelet transforms combined with spatial dependence modeled by a Markov random field (MRF). The method initially produces a coarse segmentation, which is refined through a relaxation method based on a new energy function. A set of textured images is used to demonstrate the effectiveness of the proposed method.

Index Terms— Image segmentation, image analysis, minimization methods, image texture analysis.

1. INTRODUCTION

The primary purpose of an image segmentation system is to extract information from the images to allow the discrimination among different objects of interest. Image segmentation is of great interest in a variety of scientific and industrial fields, with applications in medicine, microscopy, remote sensing, control of quality, retrieval of information in graphic databases, among others. The segmentation process is usually based on gray level intensity, color, shape or texture.

Texture can be characterized by local variations of pixel values that repeat in a regular or random pattern on the object or image. It can also be defined as a repetitive arrangement of patterns over a region. Several methods for unsupervised and supervised texture segmentation and classification have been proposed in the literature. However, neither generic methods nor formal approaches exist that are useful for a great variety of images.

Segmentation methods are usually based only on features extracted from pixels located in the image. However, techniques which associate such features with spatial dependence have received increasing attention from scientific community since they present more reliable results than those obtained by using methods based only on features [1, 2]. Besides considering features extracted from image regions, as usually used in region-based segmentation techniques, such methods incorporate information about a region neighborhood through the Bayesian formulation.

Among the main advantages in using segmentation based on random fields are the integration of spatial relationship between adjacent regions of the image [3], the use of several features for image description by means of the Bayesian formulation, the region labeling for generating the final segmentation obtained directly from the random field [4], and the incorporation of constraints into the energy function to be minimized [5].

The energy function commonly presents form $E = E_n + E_d$, where E_d depends on data obtained from the image being segmented or on a training set, and E_n is computed according to the labeling of a neighborhood of each pixel present in the image. The aim of segmentation based on spatial dependence modeled by a MRF is the minimization of such energy function.

This work describes a new image segmentation method using texture features extracted by wavelet transforms combined with spatial dependence modeled by a MRF. This method, which does not require a training set, is composed of two stages. First, the centers of homogeneous regions are located by using the k-means clustering algorithm, resulting in a coarse segmentation. In the second stage, parameters are estimated from the coarse segmentation and a relaxation method is applied to minimize an energy function, resulting in a fine segmentation. Moreover, a new energy function is presented. This function scales both components E_n and E_d to a same range and inserts a penalty when a region should not appear in the final segmentation.

By using this two-step process, the information extracted from the coarse segmentation is used to produce more accurate results to the segmentation. For instance, the initial labeling of MRF is better than usual, which considers only a random labeling. Moreover, parameters are estimated to represent each different region. It also provides clues regarding the range of values of each energy function component and the probability of each region appears in the segmentation.

The paper is organized as follows. Since this work is focused on segmentation based on spatial dependence, Section 2 describes its concepts. In Section 3, the proposed method is presented and discussed. Experimental results obtained by applying the segmentation method and comparisons with other energy functions are shown in Section 4. Finally, Section 5 concludes the paper with final remarks.

2. SEGMENTATION BASED ON SPATIAL DEPENDENCE

Methods based on spatial dependence use the Bayesian formulation to relate features of a region to a certain neighborhood [6]. To divide an image into homogeneous regions by grouping pixels having similar characteristics, such methods consider the existence of an observation (input image with n pixels) and a correctly, but unknown, segmented image.

Each pixel of the image is considered as a random variable that assumes values in $L = \{0, 1, \ldots, G-1\}$, where G denotes the number of regions with distinct characteristics. $Y = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n\}$ denotes the set composed of the feature vectors representing the ob-

served variables. Information about spatial dependence is modeled by a MRF, represented by the random variable set $X = \{X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\}$, where x_i belongs to L.

The Bayes' theorem is used to establish a relationship between features Y and the spatial dependence of variables X, where $P(X)$ is often called *a priori probability*. The correct segmentation for an image is that one that maximizes the *a posteriori probability* $P(X|Y)$, whose estimation is called maximum a posteriori (MAP).

However, the computational cost needed to determine the optimal segmentation with MAP estimation is extremely high, since it is necessary to compute the conditional probability G^n times. Therefore, for practical purpose, segmentation techniques for approximating the optimal solution must be used. Such approximations are obtained by means of relaxation methods such as simulated annealing [7], iterated conditional modes (ICM) [8], maximum marginal probability [9], belief propagation [10], and graph cuts [11], which maximize probability $P(X|Y)$ by iteratively minimizing a energy function [6]. As described before, in general this energy function presents the form shown in Equation 1.

$$E = E_n + E_d \tag{1}$$

Assuming that features extracted to each pixel follow a Gaussian distribution, the components of energy function for each pixel in image are shown in Equations 2 and 3. In these equations, $\boldsymbol{\mu}_v$ and $\boldsymbol{\Sigma}_v$ denote the mean vector and the covariance matrix of vth class in the image, respectively, m represents the number of neighbors for x_i that belong to classes different from v. Finally, β is a weight emphasizing the significance of interaction among adjacent pixels inside a neighborhood [6].

$$E_n = 2m\beta \tag{2}$$

$$E_d = \ln|\boldsymbol{\Sigma}_v| + (\boldsymbol{y}_i - \boldsymbol{\mu}_v)^{\mathrm{T}}\boldsymbol{\Sigma}_v^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}_v) \tag{3}$$

3. PROPOSED METHOD

The aim of the proposed method is to segment a textured image in G regions having similar features. The method is composed of two stages. The first one segments homogeneous regions of the image by using the k-means clustering method with randomly sampled starting points and a bidimensional histogram. In the second stage, the ICM is used to determine the location of boundaries between adjacent regions. The diagram in Figure 1 illustrates the steps of each stage.

During the first stage, the image is divided into a number of square windows, being allowed the overlapping of distinct regions. For each window, a two-level Daubechies wavelet decomposition is applied to compute the energy of six sub-images that present higher frequencies, generating six measures [12], which compose a feature vector. Instead of directly considering the gray level of each pixel to describe each region in the image, such feature vector is used.

Since the result provided by k-means clustering algorithm depends on the cluster centers initially assigned, it is executed a few times with centers randomly sampled each time. After these executions, the clustering which minimizes the mean squared distance from each data point to its nearest center is used to determine the coarse segmentation. This process provides more accuracy to the resulting clustering.

Once the initial clustering is concluded, a bidimensional histogram is calculated, where entry (x, y) contains the occurrence frequency of each G possible labels. Pixel (x, y) is assigned to class i if entry (x, y) has only non zero values in i; otherwise, it will not belong to the coarse segmentation. Therefore, the more overlapping windows, the more precise the result of the first stage will be.

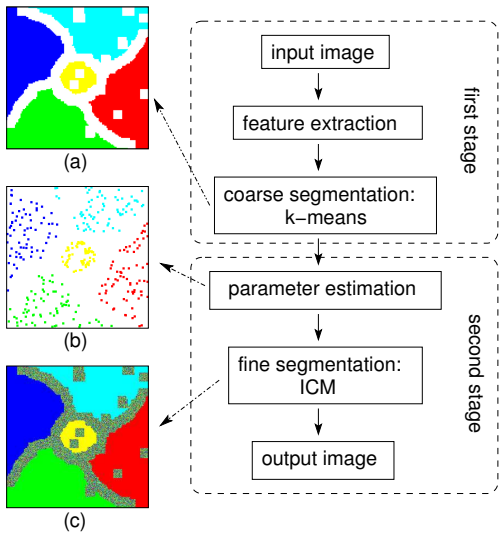

Fig. 1: Diagram illustrating the two stages of the proposed segmentation method. (a) result of coarse segmentation is shown in non-white regions; (b) samples of each class used to parameter estimation; (c) initial labeling of MRF used during the second stage.

Prior execution of ICM, it is performed the parameter estimation and feature extraction to describe each pixel present in the image. According to E_d, shown in Equation 3, parameters $\boldsymbol{\Sigma}_v$ and $\boldsymbol{\mu}_v$ have to be estimated for each class v. After sampling windows within each class from the coarse segmentation, these parameters are estimated from feature vectors composed by six measures resulting of wavelet decomposition. Moreover, feature vectors are created for windows centered in each pixel in the image. Finally, each pixel is associated to a feature vector, which is used by ICM as parameter $\boldsymbol{y}_i$ in Equation 3.

The ICM method is applied to obtain indirectly an approximation of the *a posteriori probability* maximization. This method aims at minimization of energy function shown in Equation 1 applied to each pixel. Although Equation 1 is commonly used, this work proposes a new energy function, in which two important aspects are considered, the probability of a specific class to appear in the segmentation and the scale of the energy function components.

Since the coarse segmentation is available, it is possible to compute an approximation to a priori probability for each class. This may be done through the area occupied for each class in the coarse segmentation. Equation 4 shows this new energy function, which contains an additional component, a penalty term. This term presents high values when the label being considered by ICM violates the prior probability of occurring a class. Since labels change faster in the first iterations, it was experimentally observed that the penalty term in energy function should be used after 20 cycles of ICM.

$$x_i \leftarrow \min_{v=\{0,1,\ldots G-1\}} \{E_n + E_d + Pen(v)\} \tag{4}$$

The penalty term is defined in Equation 5. $P(v)$ is an approximation of percentage that class v occupies in the current cycle of ICM, and $P_{\text{prior}}(v)$ represents the prior probability of class v to appear in the segmentation. If percentage of class v in current segmentation is greater than the prior probability, $Pen(v)$ returns a high value; thus, it is not expected x_i to be labeled as v since ICM performs a minimization in Equation 4.

$$Pen(v) = \log(P(v)/P_{\text{prior}}(v)) \tag{5}$$

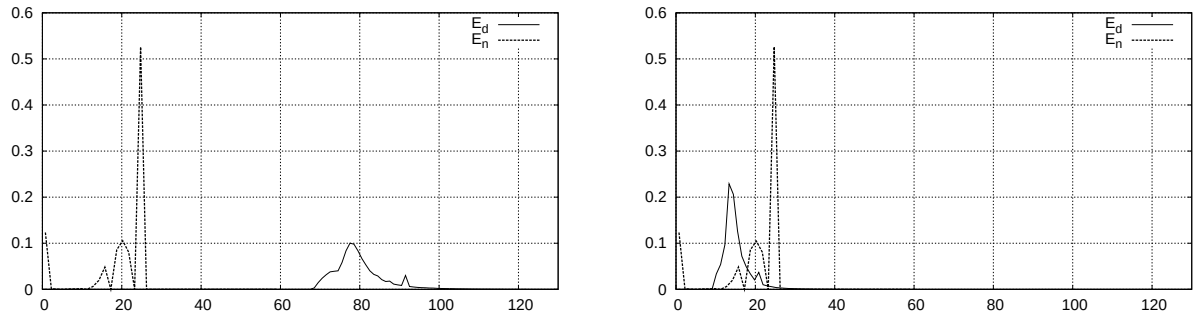

Fig. 2: Histograms showing the interval of energy function components E_d and E_n. (a) interval prior scale transformation of E_d; (b) interval after scale transformation, used in energy function described in this work.

Furthermore, components E_n and E_d are scaled to a same interval of values. This is done because if these two components are in different scale the result of the energy function can be biased. For instance, if E_d is much bigger than E_n, the labeling of current pixel does not depend on its neighbors, yet it depends only on its feature vector and parameters $\mathbf{\Sigma}_v$ and $\boldsymbol{\mu}_v$, for each class v. Such scale transformation in data is performed after computing feature vectors for all pixels, and it is applied according to the range occupied by elements of component E_n, which is determined prior the segmentation and depends on the neighborhood order [6].

4. EXPERIMENTAL RESULTS

To evaluate our methodology, several experiments were conducted using a large image set. However, due to limited space, results for only three texture mosaics are shown in Figures 3(a)-(c). Since an important aspect considered in this work is the described energy function, mosaics are used because they have known ground truth, thus it is possible to perform a quantitative comparison of results obtained by using different energy functions.

In all experiments, images were partitioned into windows with size of 8×8 pixels for feature extraction and parameters estimation. Moreover, in feature extraction executed during the first stage, all 6 measures extracted using wavelet decomposition were normalized so that they follow the Gaussian distribution with mean zero and variance one. K-means clustering method was executed 10 times; starting points were sampled randomly. Furthermore, 100 iterations of ICM, which considered fourth order neighborhood, were performed.

Figure 2 presents the difference in the interval of energy function components E_d and E_n. The first histogram in Figure 2(a) shows these two components in different scales, resulting from original energy function used by ICM. The resulting segmentation is shown in Figure 3(h). On the other hand, Figure 2(b) shows those components after applying a scale transformation. After this transformation, the segmentation obtained is shown in Figure 3(n). It can be clearly observed the improvement by using a uniform scale in components E_d and E_n of Equation 4. The segmentation obtained without scaling is not accurate since the energy function becomes biased by component E_d which presents higher scale and is more spread than E_n.

Results of another experiment are shown in Figure 3. In this figure, the results are compared using the same parameters in first stage but different energy functions in the second one. The first column shows the original mosaic images, while the others show the obtained results. In addition, the label of each figure presents the percentage of correct segmentation obtained according to the ground truth. The second column presents results by using no spatial dependence during the segmentation; this is done by setting parameter β as 0, which is equivalent to apply a simple Gaussian classifier. Results obtained with the original energy function (Equation 1) are shown in the third column.

Furthermore, other two comparisons are performed in Figure 3. In the fourth column, the results obtained with the energy function described in [4] are shown; this function is defined as $E = E_n + \alpha(t)E_d$, where $\alpha(t)$ is a function that depends on the current iteration of the relaxation method. We used $\alpha(t) = 80 \times 0.9^t + 1/6$, as described in that work. Finally, the last column of Figure 3 presents the segmentation obtained by using the proposed energy function shown in Equation 4.

Since all results were obtained using the same parameters during the first stage, different energy functions can be compared. It can be seen that, without considering the neighborhood (without spatial dependence), the results are worse due to the presence of a significant amount of noise. On the other hand, the noise is almost totally removed when spatial dependence is considered. However, if the energy function is not adaptable to the data, the pixel neighborhood does not have enough influence to lead segmentation to accurate results. Finally, when the data considered are in the same scale and a penalty function according to prior probability is used, results usually become more accurate.

5. CONCLUSIONS

This work presented a new image segmentation method using texture features extracted from a wavelet decomposition combined with spatial dependence modeled by a Markov random field. Once the developed method presents two stages, a coarse segmentation can be used for parameter estimation, MRF initialization, and approximation of a prior probability for each class.

Furthermore, the proposed energy function significantly improved segmentation results due to its penalty term and same scale in its components E_d and E_n. The effectiveness of the proposed method was demonstrated by experiments. Finally, although these concepts have been used with ICM method, their application in other energy minimization methods is straightforward.

6. ACKNOWLEDGMENTS

Support of this work was provided by grants from CAPES, Brazil.

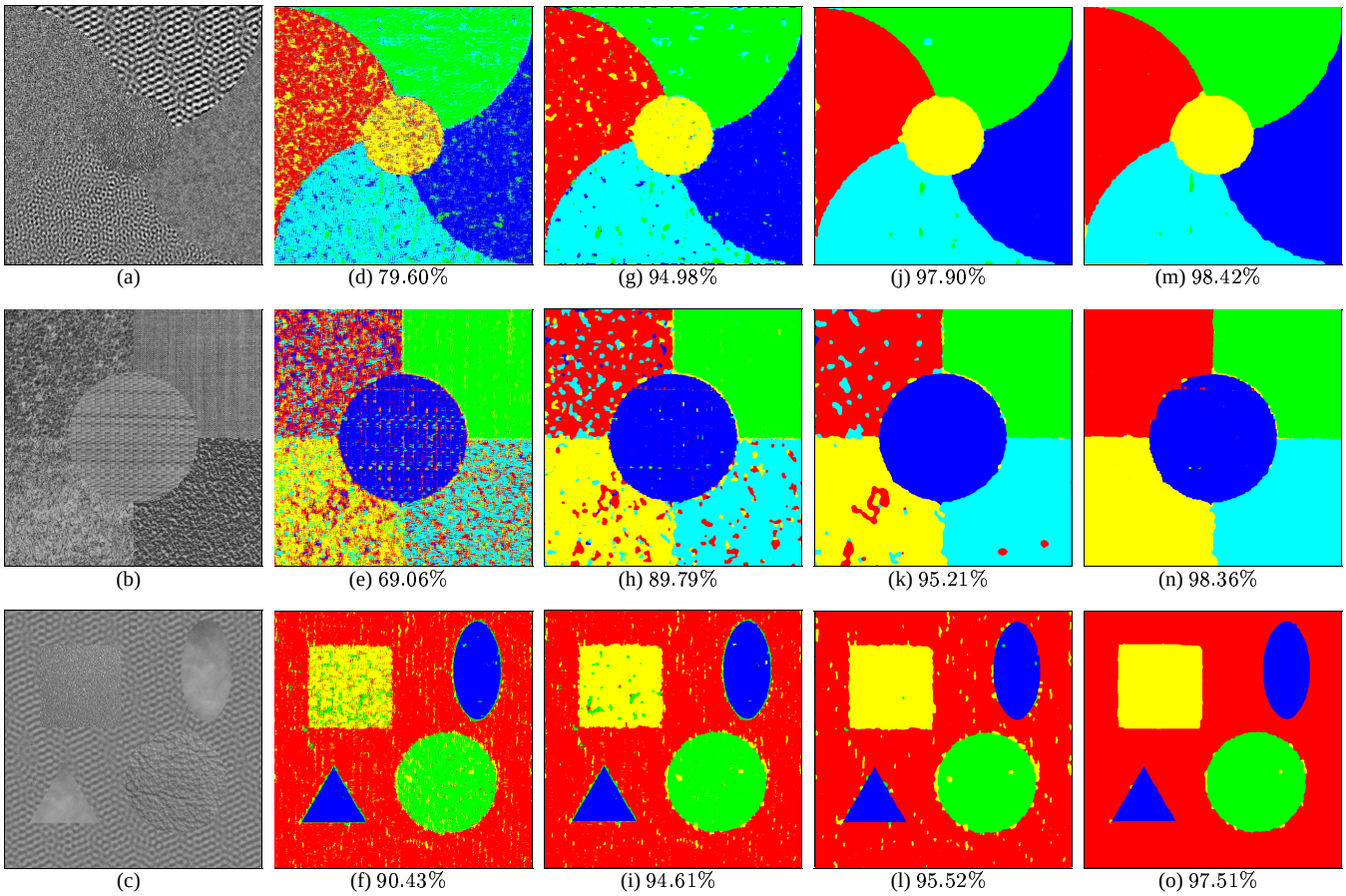

Fig. 3: Results for image segmentation using different energy functions. (a)-(c) original images; (d)-(f) parameter β set as 0; (g)-(i) original energy function shown in Equation 1; (j)-(l) energy function described in [4]; (m)-(o) proposed energy function shown in Equation 4.

7. REFERENCES

[1] C. Palm, "Color Texture Classification by Integrative Co-occurrence Matrices," *Pattern Recognition*, vol. 37, no. 5, pp. 965–976, May 2004.

[2] S. Krishnamachari and R. Chellappa, "Multiresolution Gauss-Markov Random Field Models for Texture Segmentation," *IEEE Transactions on Image Processing*, vol. 6, no. 2, pp. 251–267, Feb. 1997.

[3] R. C. Dubes and A. K. Jain, "Random Field Models in Image Analysis," *Journal of Applied Statistics*, vol. 16, no. 2, pp. 131–164, June 1989.

[4] H. Deng and D. A. Clausi, "Unsupervised Image Segmentation Using a Simple MRF Model with a New Implementation Scheme," *Pattern Recognition*, vol. 37, no. 12, pp. 2323–2335, Dec. 2004.

[5] D. Geman, S. Geman, C. Graffigne, and P. Dong, "Boundary Detection by Constrained Optimization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 609–628, July 1990.

[6] G. Winkler, *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods, A Mathematical Introduction*, Application of Mathematics, Stochastic Modeling and Applied Probability. Springer, 2003.

[7] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, Nov. 1984.

[8] J. E. Besag, "On the Statistical Analysis of Dirty Pictures," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 48, no. 3, pp. 259–302, Sept. 1986.

[9] J. Marroquin, S. Mitter, and T. Poggio, "Probabilistic Solution of Ill-posed Problems in Computacional Vision," *Journal of the American Statistical Association*, vol. 82, no. 397, pp. 76–89, Mar. 1987.

[10] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers, San Mateo, California, 1988.

[11] Y. Boykov, O. Veksler, and R. Zabih, "Fast Approximate Energy Minimization Via Graph Cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, Nov. 2001.

[12] G. van de Wouwer, P. Scheunders, and D. van Dyck, "Statistical Texture Characterization from Discrete Wavelet Representations," *IEEE Transactions on Image Processing*, vol. 8, no. 4, pp. 592–598, Apr. 1999.