A Markov Random Field Image Segmentation Model Using Combined Color and Texture Features

Zoltan Kato[1] and Ting-Chuen Pong[2]

[1] National University of Singapore, School of Computing, 3 Science Drive 2, Singapore 117543, Tel: +65 874 8923, Fax: +65 779 4580, kato@comp.nus.edu.sg
[2] Hong Kong University of Science and Technology, Computer Science Dept., Clear Water Bay, Kowloon, Hong Kong, China. Tel: +852 2358 7000, Fax: +852 2358 1477, tcpong@cs.ust.hk

Abstract. In this paper, we propose a Markov random field (MRF) image segmentation model which aims at combining color and texture features. The theoretical framework relies on Bayesian estimation associated with combinatorial optimization (Simulated Annealing). The segmentation is obtained by classifying the pixels into different pixel classes. These classes are represented by multi-variate Gaussian distributions. Thus, the only hypothesis about the nature of the features is that an additive white noise model is suitable to describe the feature values belonging to a given class. Herein, we use the perceptually uniform CIE-L*u*v* color values as color features and a set of Gabor filters as texture features. We provide experimental results that illustrate the performance of our method on both synthetic and natural color images. Due to the local nature of our MRF model, the algorithm can be highly parallelized.

Keywords: image segmentation, Markov random field model

1 Introduction

Image segmentation is an important early vision task where pixels with similar features are grouped into homogeneous regions. There are many features that one can take into account during the segmentation process: gray-level, color, motion, different texture features, etc. However, most of the segmentation algorithms presented in the literature are based on only one of the above features. Recently, the segmentation of color images (textured or not) received more attention [3,6,7]. In this paper, we are interested in the segmentation of color textured images. This problem has been addressed in [7], where an unsupervised segmentation algorithm is proposed which uses Gaussian MRF models for *color textures*. These models are defined in each color plane with interactions between different color planes. The segmentation algorithm is based on agglomerative hierarchical clustering. Our approach is different in two major points. First, we use a stochastic model based segmentation framework instead of clustering. Second, we use a combination of classical, gray-level based, texture features and

W. Skarbek (Ed.): CAIP 2001, LNCS 2124, pp. 547–554, 2001.

color instead of a direct modelization of color textures. The segmentation model consists of a MRF defined over a nearest neighborhood system. The image features are represented by multi-variate Gaussian distributions (basically a noise model). Since the different texture-types are described by a set of Gaussian parameters, it is possible to classify or recognize textures based on a prior learning of the possible parameters. Of course, parameter estimation can be a difficult task, if we do not have training data. Herein, we do not address this problem but we note that the estimation task can be solved using an adaptive segmentation technique [5,10].

We use the perceptually uniform CIE-L*u*v* color values and texture features derived from the Gabor filtered gray-level image. Of course, the nature of the texture features is not crucial to the algorithm from the segmentation point of view. The only requirement in the current model is that an additive white noise model should be suitable to describe the texture features. Most of the filtering approaches (see [8] for a comparative study of different filtering techniques) fall into this category but stochastic texture models (such as Gaussian Markov random fields [7,10]) are also suitable for our segmentation model. Herein, we use a real-valued even-symmetric Gabor filter bank. Segmentation requires simultaneous measurements in both spatial and frequency domain. However, spatial localization of boundaries requires larger bandwidths whereas smaller bandwidths give better texture measurements. Fortunately, Gabor filters have optimal joint localization in both domains [4]. In addition, when we are combining texture features with color, the spatial resolution is considerably increased. We have tested the algorithm on a set of both natural and synthetic color textured images.

2 MRF Segmentation Model

We assume that images are defined over a finite lattice $\mathcal{S} = \{s_1, s_2, \ldots, s_N\}$, where $s = (i, j)$ denotes lattice sites (or pixels). For each pixel s, the region-type (or pixel class) that the pixel belongs to is specified by a class label, ω_s, which is modeled as a discrete random variable taking values in $\Lambda = \{1, 2, \ldots, L\}$. The set of these labels $\omega = \{\omega_s, s \in \mathcal{S}\}$ is a random field, called the *label process*. Furthermore, the observed image features (color and texture) are supposed to be a realization $\mathcal{F} = \{\boldsymbol{f}_s | s \in \mathcal{S}\}$ from another random field, which is a function of the label process ω. Basically, the *image process* $\mathcal{F}$ represents the deviation from the underlying label process. Thus, the overall segmentation model is composed of the hidden label process ω and the observable noisy image process $\mathcal{F}$. Our goal is to find an optimal labeling $\hat{\omega}$ which maximizes the a posteriori probability $P(\omega \mid \mathcal{F})$, that is the *maximum a posteriori* (MAP) estimate [2]: $\arg\max_{\omega \in \Omega} \prod_{s \in \mathcal{S}} P(\boldsymbol{f}_s \mid \omega_s))P(\omega)$, where Ω denotes the set of all possible labelings.

Herein, the image process is modeled by an additive white noise. Thus, we suppose that $P(\boldsymbol{f}_s \mid \omega_s)$ follows a Gaussian distribution and pixel classes $\lambda \in \Lambda = \{1, 2, \ldots, L\}$ are represented by the mean vectors $\boldsymbol{\mu}_\lambda$ and the covariance matrices $\boldsymbol{\Sigma}_\lambda$. Furthermore, ω is supposed to be a MRF with respect to a first order neighborhood system. Thus, according to the *Hammersley-Clifford theorem* [2], $P(\omega)$ follows a Gibbs distribution:

$$P(\omega) = \frac{\exp(-U(\omega))}{Z(\beta)} = \frac{\prod_{C \in \mathcal{C}} \exp(-V_C(\omega_C))}{Z(\beta)} , \quad (1)$$

where $U(\omega)$ is called an *energy function*, $Z(\beta) = \sum_{\omega \in \Omega} \exp(-U(\omega))$ is the normalizing constant (or *partition function*) and V_C denotes the *clique potential* of clique $C \in \mathcal{C}$ having the label configuration ω_C. $\mathcal{C}$ is the set of spatial second order cliques (ie. *doubletons*). Note that the energies of *singletons* (ie. pixel sites $s \in \mathcal{S}$) directly reflect the probabilistic modeling of labels without context, while doubleton clique potentials express relationship between neighboring pixel labels. In our model, these potentials favor similar classes in neighboring pixels. Thus the energy function of the so defined MRF image segmentation model has the following form:

$$\begin{aligned} U(\omega, \mathcal{F}) = & \sum_{s \in \mathcal{S}} \left(\ln(\sqrt{(2\pi)^3 \mid \boldsymbol{\Sigma}_{\omega_s} \mid}) + \frac{1}{2}(\boldsymbol{f}_s - \boldsymbol{\mu}_{\omega_s}) \boldsymbol{\Sigma}_{\omega_s}^{-1} (\boldsymbol{f}_s - \boldsymbol{\mu}_{\omega_s})^T \right) \\ & + \beta \sum_{\{s,r\} \in \mathcal{C}} \delta(\omega_s, \omega_r), \end{aligned} \quad (2)$$

where $\delta(\omega_s, \omega_r) = 1$ if ω_s and ω_r are different and 0 otherwise. $\beta > 0$ is a parameter controlling the homogeneity of the regions. As β increases, the resulting regions become more homogeneous. Now, the segmentation problem is reduced to the minimization of the above energy function. Since it is a nonconvex function, some combinatorial optimization technique is needed to tackle the problem. Our experiments used Simulated Annealing (Gibbs sampler [2]) and Iterated Conditional Modes (ICM) [1].

2.1 Color Features

The first question, when dealing with color images, is how to measure quantitatively color difference between any two arbitrary colors. Experimental evidence suggests that the RGB tristimulus color space may be considered as a Riemannian space [9]. Due to the complexity of determining color distance in such spaces, several simple formulas have been proposed. These formulas approximate the Riemannian space by a Euclidean color space yielding a perceptually uniform spacing of colors. One of these formulas is the $L^*u^*v^*$ [9] color space that we use herein.

2.2 Texture Features

Many different techniques have been proposed for analyzing image texture. Herein, we will focus on the multi-channel filtering approach where the channels are represented by a bank of real-valued, even-symmetric Gabor filters. The

Fourier domain representation of the basic even-symmetric Gabor filter oriented at 0^o is given by [4]:

$$H(u,v) = A\left(\exp\left(-\frac{1}{2}\left(\frac{(u-u_0)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right)\right)+\exp\left(-\frac{1}{2}\left(\frac{(u+u_0)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right)\right)\right), \tag{3}$$

where $\sigma_u = 1/2\pi\sigma_x$, $\sigma_v = 1/2\pi\sigma_y$, $A = 2\sigma_x\sigma_y$, u_0 is the frequency of a sinusoidal plane wave along the x-axis, and σ_x and σ_y are the deviations of the Gaussian envelope along the x and y axes. Filters with other orientations can be obtained by rotating the coordinate system. In our tests, we used four orientations: $0^o, 45^o, 90^o, 135^o$ and the radial frequencies u_0 are 1 octave apart: $\sqrt{2}, \sqrt{2}/2, \sqrt{2}/4, \sqrt{2}/8, \ldots$ For an image with a width of 2^W pixels, the highest radial frequency falling inside the image array is $\sqrt{2}/2^{W-2}$. From each filtered image g, we compute a *feature image* using the nonlinear transformation $\mid \tanh(\alpha g_s) \mid, s \in \mathcal{S}$; followed by a Gaussian blurring with a deviation proportional to the center frequency of the Gabor filter: $\sigma = k/u_0$. In our experiments, the Gabor filtered images are scaled to the interval $[-1, 1]$ and we set $\alpha = 40$ and $k = 1$.

Table 1. Computing times and segmentation error on the synthetic images.

Image:	Fig. 2				Fig. 3			
feature	Gibbs	error	ICM	error	Gibbs	error	ICM	error
texture	185 sec.	13.5%	3.4 sec.	16.6%	1349 sec.	19.0%	13 sec.	20.6%
color	105 sec.	2.5%	5.5 sec.	8.8%	732 sec.	19.7%	20 sec.	23.1%
combined	319 sec.	1.2%	16 sec.	2.7%	2581 sec.	8.9%	73 sec.	11.4%

3 Experimental Results

The proposed algorithm has been tested on a variety of color images including synthetic images (Fig. 1, Fig. 2, Fig. 3), outdoor (Fig. 4) and indoor (Fig. 5) scenes. We have used MIT's VisTex database to compose the synthetic color textured images. The test program has been implemented in C and run on a UltraSparc1 workstation. Herein, we present a few examples of these results and compare segmentation results using color only, texture only and combined features. Furthermore, we also compare the results obtained via a deterministic (ICM [1]) and stochastic (Simulated Annealing using the Gibbs sampler [2]) relaxation. The mean vectors and covariance matrices were computed over representative regions selected by the user and we set $\beta = 2.5$ in all cases. This value has been found to provide satisfactory results in all cases. An optimal Gabor filter set containing 2–4 filters has been picked manually for each image. We remark, however, that it is also possible to automatically select these filters (see [8] for a collection of methods). We found in all cases that segmentation based purely on texture gives fuzzy boundaries but usually homogeneous regions. Whereas segmentation based on color is more sensitive to local variations in color but provides sharp boundaries. As for the combined features, the advantages of both color and texture based segmentation are quite well preserved: we obtained sharp boundaries and homogeneous regions (see Fig. 1 as a good example). For example, the ball on the *tennis* image (Fig. 5) is much better detected than on either the texture or color segmentations. In terms of sharpness and homogeneity, the combined segmentation clearly outperforms the others. The power of combined features is well demonstrated by Fig. 3. Three regions contain a wooden texture with nearly matching colors and a small difference in the direction (right and lower part) or scale (middle part) in texture. The two other regions have similar texture but completely different color. Comparing color and texture only segmentations, the latter two regions are perfectly separated in the color segmentation but they are mixed in the texture one. On the other hand, the former three regions are much better separated in the texture segmentation than in the color one. As for the combined segmentation, the five regions are well separated, the error rate is half of the separate segmentation's rates. Regarding the different optimization techniques, we can see that the suboptimal ICM provides somewhat lower quality compared to the optimal Gibbs sampler but it converges much faster (see Table 1). However, this difference is less important in the case of combined features and for the real images it is nearly invisible. We also run a test on one of the images published in [7] (see Fig. 4) and we found that our results are close to the ones in [7].

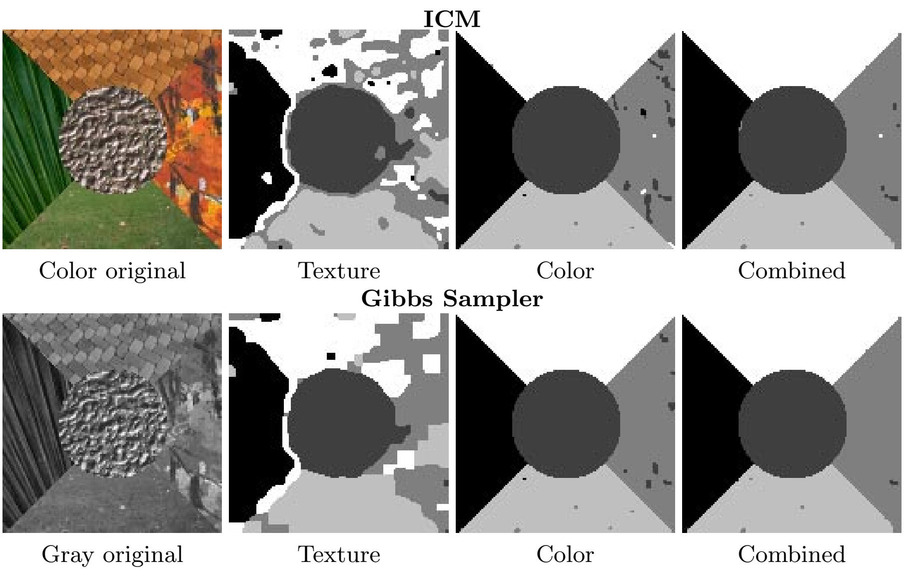

Fig. 1. Segmentation of a 128 × 128 synthetic image with 5 classes. We have used 2 Gabor filters to extract texture features.

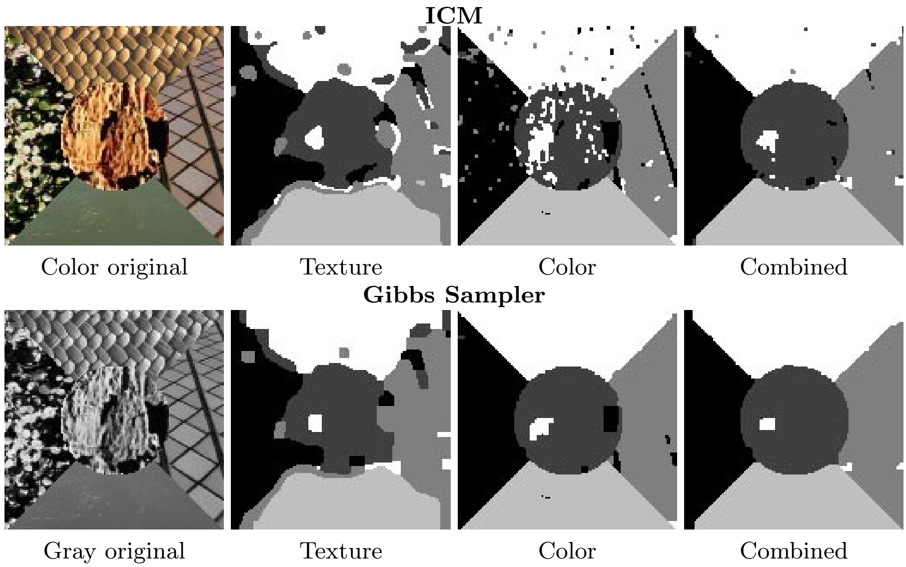

Fig. 2. Segmentation of a 128 × 128 synthetic image with 5 classes. We have used 3 Gabor filters to extract texture features.

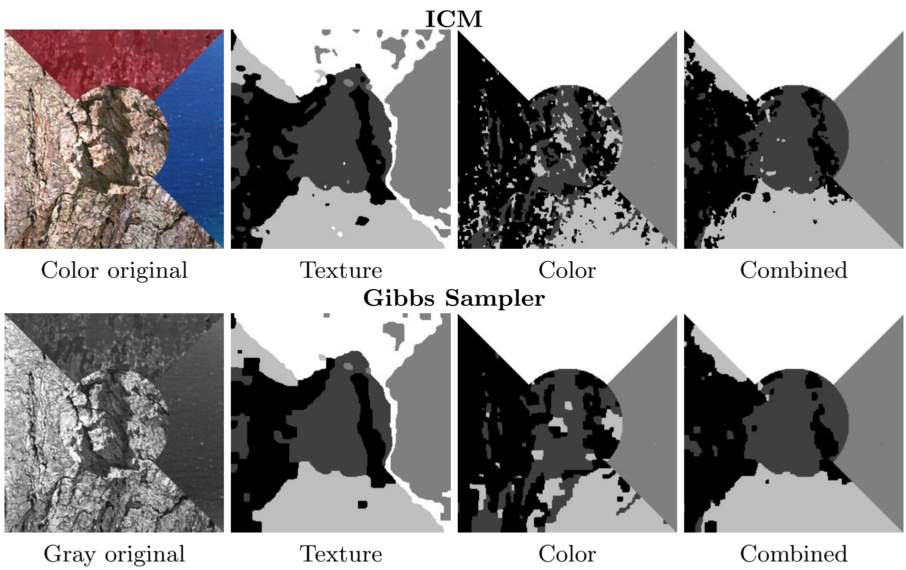

Fig. 3. Segmentation of a 256 × 256 synthetic image with 5 classes. We have used 4 Gabor filters to extract texture features.

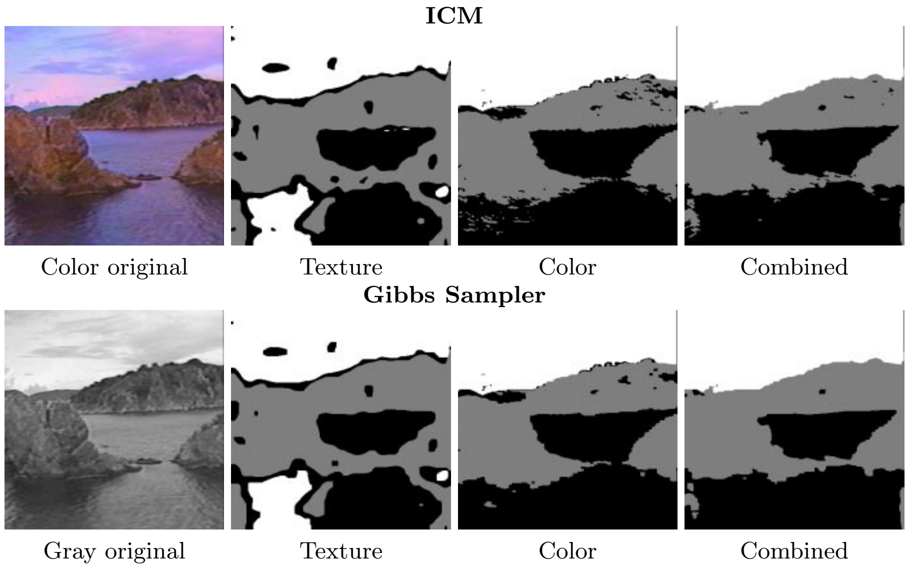

Fig. 4. Segmentation of a 384 × 384 real image with 3 classes. We have used 3 Gabor filters to extract texture features.

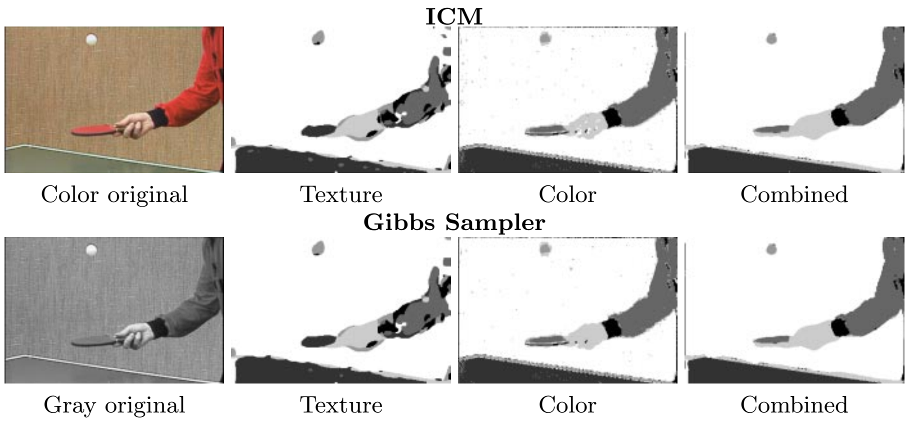

Fig. 5. Segmentation of a 360 × 240 real image with 6 classes. We have used 3 Gabor filters to extract texture features.

4 Conclusions

We have proposed a MRF image segmentation model which is able to combine color and texture features. The model itself is not restricted to a specific texture feature. In fact, any feature is suitable as far as feature values belonging to a pixel class can be modeled by a random white noise. Due to our model based approach, it is also possible to classify different kind of textures based on a prior training of the corresponding parameters. The quality of the segmentation is improved with respect to color only and texture only segmentations. We also tested different optimization methods and found that the suboptimal but fast ICM is a good tradeoff between quality and computing time when using combined features. Although our implementation is sequential, the algorithm is highly parallel due to the local nature of the MRF model. Thus, a parallel implementation can further improve the computing speed.

References

1. Besag, J.: On the statistical analysis of dirty pictures. J. Roy. Statist. Soc., ser. B, 1986
2. Geman, S., Geman, D.: Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Trans. on Pattern Analysis and Machine Intelligence **6** (1984) 721–741
3. Huang, C.L., Cheng, T.Y., Chen, C.C.: Color images segmentation using scale space filter and Markov random field. Pattern Recognition **25(10)** (1992) 1217–1229
4. Jain, A.K., Farrokhnia, F.: Unsupervised texture segmentation using Gabor filters. Pattern Recognition **24(12)** (1991) 1167–1186
5. Kato, Z., Zerubia, J., Berthod, M.: Unsupervised parallel image classification using Markovian models. Pattern Recognition **32(4)** (1999) 591–604
6. Liu, J., Yang, Y.H.: Multiresolution color image segmentation. IEEE Trans. on Pattern Analysis and Machine Intelligence **16(7)** (July 1994) 689–700
7. Panjwani, D.K., Healey, G.: Markov random field models for unsupervised segmentation of textured color images. IEEE Trans. on Pattern Analysis and Machine Intelligence **17(10)** (October 1995) 939–954
8. Randen, T., Husoy, J.H.: Filtering for texture classification: A comparative study. IEEE Trans. on Pattern Analysis and Machine Intelligence **21(4)** (April 1999) 291–310
9. Sangwine, S.J., Horne, R.E.N. (eds): The Colour Image Processing Handbook. Chapman & Hall (1998)
10. Won, C.S., Derin, H.: Unsupervised segmentation of noisy and textured images using Markov random fields. Computer Graphics and Image Processing: Graphical Models and Image Processing **54(4)** (July 1992) 208–328