A DETAIL-PRESERVING MIXTURE MODEL FOR IMAGE SEGMENTATION

Thanh Minh Nguyen[1], *Q. M. Jonathan Wu*[1], *Senior Member, IEEE, Dibyendu Mukherjee*[1], *and Hui Zhang*[1,2]

[1]Department of Electrical and Computer Engineering, University of Windsor
[2]School of Computer & Software, Nanjing University of Information Science & Technology
{nguyen1j, jwu, mukherjd, nrzh}@uwindsor.ca

ABSTRACT

This paper presents a new finite mixture model for image segmentation. First, in order to take into account the spatial dependencies in an image, existing mixture models use a constant temperature parameter (β) throughout the image for every label. The constant value of β reduces the impact of noise in homogeneous regions but negatively affects segmentation along the border of two regions. We propose a new way to use a different value of β throughout the image. Secondly, in order to incorporate the correlation between each centre pixel and its neighboring pixels, existing mixture model gives the same importance to all pixels in a neighborhood window. We assign different weights to different pixels appearing in the window, which is based on the fact that the clique strength should be reduced with distance. Thirdly, our model is based on the Student's-t distribution, which is heavily tailed and more robust than Gaussian. We exploit Dirichlet distribution and Dirichlet law to incorporate the spatial relationships between pixels in an image. Finally, expectation maximization (EM) algorithm is adopted to maximize the data log-likelihood and to optimize the parameters. The performance is compared to other existing models based on the model-based techniques, demonstrating superiority of the proposed model for image segmentation.

Index Terms— Mixture model, detail-preserving, spatial constraints, and image segmentation.

1. INTRODUCTION

Image segmentation is one of the heated issues in almost every task of image processing. The objective is to cluster all image pixels into non-overlapped, consistent regions that have common characteristics. Accurate image segmentation provides additional important information for diagnosis and quantitative analysis. However, issues such as poor contrast and the variety and complexity of images complicate accurate segmentation. Many previous works have been carried out on image segmentation. Amongst the various approaches used for segmenting an image, one of the main research directions in the relevant literature is focused on model-based techniques.

In model-based techniques, Gaussian mixture model (GMM) [1–3] is an efficient method used in most applications. However, image segmentation results obtained from GMM are quite poor because the spatial relationship of the neighboring pixels in an image is not taken into its account. In order to overcome this problem, mixture models based on Markov random field (MRF) for pixel label are proposed in [4] to impose spatial smoothness constraints between neighboring pixels. Another family of mixture models based on MRF for pixel label priors have been successfully applied to image segmentation [5, 6]. In these models, instead of imposing the smoothness constraint on the pixel label as in the above category, however, these methods aim to impose the smoothness constraint on the contextual mixing proportions. Although the quality of image segmentation is quite good with noisy images, their primary disadvantage, however, lies in its additional training complexity. The M-step of the EM algorithm cannot evaluate the prior distribution in a closed form.

To overcome this problem, mixture model in [7–9] assumes that the prior probabilities follow a Dirichlet distribution and Dirichlet law. The main advantage is that it guarantees that the prior probabilities are positive and sum to one without requiring a reparatory projection step. However, in order to take into account the spatial dependencies in an image, existing finite mixture models [5, 8, 9] use a constant β throughout the image and for every label. The constant value of β reduces the impact of noise in homogeneous regions but negatively affects segmentation along the border of two regions. Besides that, in order to incorporate the correlation between each centre pixel and its neighboring pixels in a window, the existing finite mixture models give the same importance to all pixels. Clearly, it is usually better to give more weight to the pixels that are closer to the central pixel.

Based on these considerations, in this paper, we propose a new finite mixture model for detail-preserving image segmentation. First, we propose a new way to use a different value of β throughout the image. Secondly, our model assigns differ-

This research has been supported in part by the Canada Research Chair Program, AUTO21 NCE, and the NSERC Discovery grant.

This work was supported in part by the National Natural Science Foundation of China under Grant 61105007.

ent weights to different pixels appearing in a window, based on the fact that the clique strength should be reduced with distance. Thirdly, the spatial information between neighboring pixels of an image is considered during the learning step based on Dirichlet distribution and Dirichlet law. Finally, to estimate the parameters of the proposed model, the EM algorithm is adopted to maximize the data log-likelihood function.

The remainder of the paper is organized as follows. In section 2, we present a brief introduction of the finite mixture model with Dirichlet spatial constraints, commonly used in the literature for image segmentation. In section 3, we describe the details of the proposed algorithm. In section 4, we present the experimental results, and conclusions are given in section 5.

2. FINITE MIXTURE MODEL WITH DIRICHLET SPATIAL CONSTRAINTS

Let x_i, with dimension D, i = (1,2,...,N), denote an observation at the i-th pixel of an image. The i-th pixel is characterized by the prior probabilities vector π_i=$\{\pi_{i1}, \pi_{i2}, ..., \pi_{iK}\}$. The posterior probability is denoted by y_{ij}. In order to partition an image consisting of N pixels into K labels, finite mixture models assumes that each pixel x_i is independent of the label Ω_j. The density function $f(\mathrm{x}_i|\Theta)$ at a pixel x_i is given by:

$$f(\mathrm{x}_i|\Theta) = \sum_{j=1}^{K} \pi_{ij} p(\mathrm{x}_i|\Omega_j) \tag{1}$$

where, the prior probability that pixel x_i is in label Ω_j, which satisfies the constraints $\pi_{ij} \geq 0$ and $\sum_{j=1}^{K} \pi_{ij}$=1. Each distribution $p(\mathrm{x}_i|\Omega_j)$ is called a component of the mixture. Note that, $p(\mathrm{x}_i|\Omega_j)$ can be any kind of distribution. In order to properly account for the neighboring pixels during the learning step, mixture models in [7–9] are based on the Dirichlet distribution and Dirichlet law to incorporate the spatial constraints. According to these models, the probability label is given by:

$$p(\mathbf{z}_i|\alpha_i) = \int_0^1 p(\mathbf{z}_i|\xi_i) p(\xi_i|\alpha_i) d\xi_i \tag{2}$$

where, the discrete label $\mathbf{z}_i$=$(z_{i1}, z_{i2}, ..., z_{iK})$ at the i-th pixel is defined as:

$$z_{ij} = \begin{cases} 1 \text{ IF : pixel } \mathrm{x}_i \text{ belongs to label } \Omega_j \\ 0 \text{ Otherwise} \end{cases} \tag{3}$$

and, ξ_i=$(\xi_{i1}, \xi_{i2}, ..., \xi_{iK})$, i=(1,2,...,N), is the Dirichlet parameter in the K-dimensional probability simplex: $\xi_{ij} \geq 0$ and $\sum_{j=1}^{K} \xi_{ij}$=1. In Eq.(2), α_i=$(\alpha_{i1}, \alpha_{i2}, ..., \alpha_{iK})$ is the vector of Dirichlet parameters ξ_i. The value of α_{ij}, j=(1,2,...,K), is non-negative: $\alpha_{ij} \geq 0$. It is worth mentioning that we can use any non negative value of α_{ij} for the Dirichlet distribution $p(\xi_i|\alpha_i)$. Applying the property of the probability density function [7], after some manipulation, the probability label in Eq.(2) is given by:

$$p(\mathbf{z}_i|\alpha_i) = \frac{M!}{\prod_{j=1}^{K} (z_{ij})!} \frac{\Gamma(\sum_{j=1}^{K} \alpha_{ij})}{\prod_{j=1}^{K} \Gamma(\alpha_{ij})} \frac{\prod_{j=1}^{K} \Gamma(\alpha_{ij} + z_{ij})}{\Gamma(\sum_{j=1}^{K} (\alpha_{ij} + z_{ij}))} \tag{4}$$

We now consider the condition of discrete label $\mathbf{z}_i$ in Eq.(3). Applying this result to the probability label in Eq.(4), with only one realization (M=1) and that $\Gamma(t+1) = t\Gamma(t)$, after some manipulation, the prior probabilities π_{ij} for the pixel x_i corresponding to the label Ω_j is given by:

$$\pi_{ij} = p(z_{ij} = 1|\alpha_i) = \alpha_{ij} \Bigg/ \sum_{k=1}^{K} \alpha_{ik} \tag{5}$$

Given the density function $p(\mathrm{x}_i|\Omega_j)$ and the prior probabilities π_{ij} in Eq.(5), the log-likelihood function is rewritten as:

$$L = \sum_{i=1}^{N} \log \sum_{j=1}^{K} \pi_{ij} p(\mathrm{x}_i|\Omega_j) \tag{6}$$

In order to estimate the model parameters, we need to maximize the log-likelihood function. Once the parameter-learning phase is complete, every pixel x_i is assigned to the label with the largest posterior probability y_{ij}.

3. PROPOSED METHOD

In order to take into account the spatial dependencies in an image, existing finite mixture models [5,8,9] use a constant β throughout the image and for every label. The constant value of β reduces the impact of noise in homogeneous regions but negatively affects segmentation along the border of two regions. Besides that, in order to incorporate the correlation between each centre pixel and its neighboring pixels in the window ∂_i, the existing finite mixture model gives the same importance to all pixels. Clearly, it is usually better to give more weight to the pixels that are closer to the central pixel. Based on these considerations, in this paper, different value of the temperature parameter (β_{ij}) is used throughout the image and for every label.

$$\beta_{ij} = \lambda H_{ij} \tag{7}$$

where,

$$H_{ij} = \exp(-\sum_{m \in \partial_i} (y_{ij}^{(t-1)} - y_{mj}^{(t-1)})^2) \tag{8}$$

where, the neighborhood of the pixel x_i is represented by ∂_i, and (t-1) indicates the iteration of the previous step. By taking a closer look at Eq.(7), it can be visualized that the temperature parameter (β_{ij}) in the proposed method is chosen small along the border of two regions to prevent the image from losing much of its sharpness and details. In other words, in homogeneous regions, the value of β_{ij} has to be chosen large enough to tolerate the noise. In Eq.(7), λ is the parameter.

Next, in order to use more local spatial information, and to control the influence of the neighborhood pixels depending on their distance from the central pixel, a new factor F_{ij} is added in our method.

$$F_{ij} = \frac{\sum\limits_{m \in \partial_i} D_{im} y_{mj}^{(t-1)}}{\sum\limits_{m \in \partial_i} D_{im}} \tag{9}$$

In Eq.(9), D_{im} is the spatial distance between pixel x_i and pixel x_m. The distance D_{im} is given by:

$$D_{im} = \exp(-d_{im}^2/2) \tag{10}$$

where, d_{im} is the Euclidean distance between pixel x_i and pixel x_m. It can be visualized that the distance D_{im} in Eq.(10) makes the influence of the pixels within the local window, to change flexibly according to their distance from the central pixel. The idea to incorporate the spatial information in Eq.(9) is based on a fact that neighboring pixels in an image are similar in some sense. We give the center pixel a greater weight and gradually reduce weight to the pixels that are further to the center pixel.

Next, we propose a novel approach to incorporate the spatial information into the smoothing prior. The Dirichlet parameters α_{ij} is defined as:

$$\alpha_{ij} = \exp(\beta_{ij} F_{ij}) \tag{11}$$

where β_{ij}, and F_{ij} are given in Eq.(7) and Eq.(9), respectively. Given the Dirichlet parameters α_{ij} in Eq.(11), the prior probabilities π_{ij} in Eq.(5) for the pixel x_i corresponding to the label Ω_j is rewriten:

$$\pi_{ij} = \frac{\exp(\beta_{ij} F_{ij})}{\sum\limits_{k=1}^{K} \exp(\beta_{ik} F_{ik})} \tag{12}$$

In our model, $p(\mathrm{x}_i|\Omega_j)$ is the Student's-t distribution $S(\mathrm{x}_i|\mu_j, \Sigma_j, v_j)$ with longer tails and one more parameter compared to the Gaussian distribution $\Phi(\mathrm{x}_i|\mu_j, \Sigma_j)$. Each Student's-t distribution has its own mean μ_j, covariance Σ_j, and degree of freedom v_j. The Student's-t distribution $S(\mathrm{x}_i|\mu_j, \Sigma_j, v_j)$ is given by:

$$\begin{aligned} S(\mathrm{x}_i|\mu_j, \Sigma_j, v_j) &= \frac{\Gamma(v_j/2 + D/2)|\Sigma_j|^{-1/2}}{(v_j\pi)^{D/2}\Gamma(v_j/2)} \\ &\times \frac{1}{\left[1 + v_j^{-1}(\mathrm{x}_i - \mu_j)^{\mathrm{T}} \Sigma_j^{-1} (\mathrm{x}_i - \mu_j)\right]^{(v_j+D)/2}} \end{aligned} \tag{13}$$

Given the prior probability distribution π_{ij} in Eq.(12) and the Student's-t distribution $S(\mathrm{x}_i|\mu_j, \Sigma_j, v_j)$ in Eq.(13), the log-likelihood function in Eq.(6) is written in the form.

$$L = \sum_{i=1}^{N} \log \sum_{j=1}^{K} \pi_{ij} S(\mathrm{x}_i|\mu_j, \Sigma_j, v_j) \tag{14}$$

The next objective is to optimize the parameter set $\Theta = \{\mu_j, \Sigma_j, v_j, \lambda\}$ in order to maximize the log-likelihood function in Eq.(14). To maximize this function, the EM algorithm [11] is applied. Note that, there is no closed form solution for maximizing the log-likelihood under a Student's-t distribution. To overcome this problem, the Student's-t distribution in previous models [10] is represented as a Gaussian distribution with scaled precision u_{ij}:

$$S(\mathrm{x}_i|\mu_j, \Sigma_j, v_j) \sim \Phi(\mathrm{x}_i|\mu_j, \Sigma_j/u_{ij}) \mathcal{G}(u_{ij}|v_j/2, v_j/2) \tag{15}$$

Given the Student's-t distribution in Eq.(15), application of the complete data condition in [1, 2]. After some manipulation, we have the estimates of μ_j and Σ_j at the (t+1) step:

$$\begin{aligned} \mu_j^{(t+1)} &= \frac{\sum\limits_{i=1}^{N} y_{ij}^{(t)} u_{ij}^{(t)} \mathrm{x}_i}{\sum\limits_{i=1}^{N} y_{ij}^{(t)} u_{ij}^{(t)}} \\ \Sigma_j^{(t+1)} &= \frac{\sum\limits_{i=1}^{N} y_{ij}^{(t)} u_{ij}^{(t)} (\mathrm{x}_i - \mu_j^{(t+1)})(\mathrm{x}_i - \mu_j^{(t+1)})^T}{\sum\limits_{i=1}^{N} y_{ij}^{(t)}} \end{aligned} \tag{16}$$

where, the posterior probability $y_{ij}^{(t)}$ is:

$$y_{ij}^{(t)} = \frac{\pi_{ij}^{(t)} S(\mathrm{x}_i|\mu_j, \Lambda_j, v_j)}{\sum\limits_{k=1}^{K} \pi_{ik}^{(t)} S(\mathrm{x}_i|\mu_j, \Lambda_j, v_j)} \tag{17}$$

and,

$$u_{ij}^{(t)} = \frac{v_j^{(t)} + D}{v_j^{(t)} + (\mathrm{x}_i - \mu_j^{(t)})^T \Sigma_j^{-1(t)} (\mathrm{x}_i - \mu_j^{(t)})} \tag{18}$$

The estimates of the degrees of freedom v_j are given by the solution of the equation

$$\begin{aligned} -\Psi\left(\frac{v_j}{2}\right) + \log\left(\frac{v_j}{2}\right) + 1 + \frac{\sum\limits_{i=1}^{N} y_{ij}^{(t)} (\log u_{ij}^{(t)} - u_{ij}^{(t)})}{\sum\limits_{i=1}^{N} y_{ij}^{(t)}} \\ + \psi\left(\frac{v_j^{(t)} + D}{2}\right) - \log\left(\frac{v_j^{(t)} + D}{2}\right) = 0 \end{aligned} \tag{19}$$

where $\Psi(\cdot)$ in Eq.(18) is the digamma function. The next step is to update the estimate of the parameter λ by using Newton-Raphson method

$$\lambda^{(t+1)} = \lambda^{(t)} - \left(\frac{\partial^2 L}{\partial \lambda^2}\right)^{-1} \frac{\partial L}{\partial \lambda} \tag{20}$$

In the next section, we will demonstrate the robustness, accuracy and effectiveness of the proposed model, as compared with other approaches.

4. EXPERIMENTS

In this section, two experiments are conducted to evaluate and compare the effectiveness of the proposed technique with others. In order to evaluate the segmentation performance quantitatively, we employ the misclassification ratio (MCR) [13] and probabilistic rand index (PRI) [14] in our experiments, which is the number of misclassified pixels divided by the total number of pixels. Note that, for MCR, the lower the value, the better the quality of the segmentation, while the higher value of PRI indicates better segmentation results.

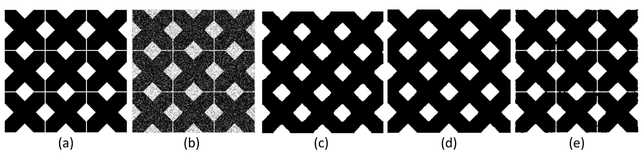

Fig. 1. Segmentation results of the synthetic image experiment, (a): original image, (b): Gaussian noise (0 mean, 0.15 variance), (c): FLICM (MCR=4.29%), (d): SMM-SC (MCR=2.68%), (e): Proposed Method (MCR=0.69%).

In the first experiment, we generated an image (271x271 image resolution) that contains two labels with luminance values [0, 1] as shown in Fig. 1(a). The image shown in Fig. 1(b) is obtained by corrupting the original image with Gaussian noise (0 mean, 0.15 variance). In Fig. 1(c)(e), we present the segmentation results of FLICM [12], SMM-SC [8], and the proposed method, respectively. As shown in Fig. 1(d), the MCR obtained by employing FLICM (MCR=4.29%) is quite high compared with the SMM-SC method (MCR=2.68%). As shown in Fig. 1(d), SMM-SC method reduce the impact of noise in homogeneous tissues but negatively affects segmentation along the border of two tissues. The over-smoothing behavior can be seen in Fig. 1(d). We can see that some details are lost in the segmented images. The accuracy of the proposed method, as shown in Fig. 1(e), is higher than other methods.

In the second experiment, we show the segmentation results of real-world color images from the Berkeley's image segmentation dataset [15]. Fig. 2 shows the real-world images from used for segmentation by employing FLICM, SMM-SC, and the proposed method, respectively. the segmentation accuracy for FLICM method is quite poor. The

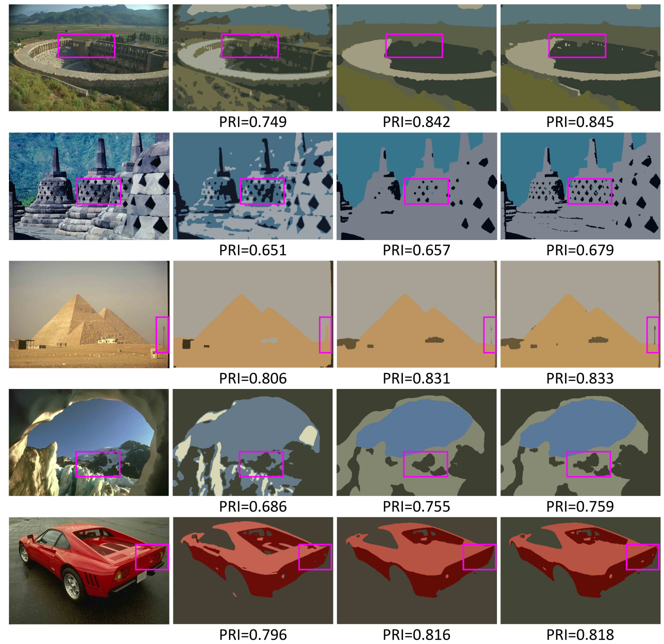

Fig. 2. Color natural image segmentation (166081, 217090, 161062, 176051, 29030), (1^{st} column): original image, (2^{nd} column): FLICM, (3^{rd} column): SMM-SC, (4^{th} column): Proposed Method.

effect of noise on the final segmented images are still quite highly noticeable in the marked boxes. SMM-SC can produce a better segmentation. However, looking closely in the marked red box, we can see that a small portion of pixels have been misclassified. Compared with these methods, the details and edges of the marked regions are better preserved by the proposed method.

5. RELATION TO PRIOR WORK AND CONCLUSIONS

We have presented a new mixture model for detail-preserving segmentation in this paper. In order to take into account the spatial dependencies in an image, existing mixture models gives the same importance to all pixels in a neighborhood window and use a constant temperature parameter (β) throughout the image for every label. We use a different value of β throughout the image and assign gradually reducing weights to pixels appearing in a window neighborhood system in accordance with the distance from the center pixel of the window. We exploit Dirichlet distribution and Dirichlet law to incorporate the spatial relationships between pixels in an image. Our method is based on the Student's-t distribution, which is heavily tailed and more robust than Gaussian. The proposed method has demonstrated an excellent performance as compared to other existing models based on the model-based techniques.

6. REFERENCES

[1] Bishop C. M., "*Pattern Recognition and Machine Learning*", Springer, 2006.

[2] McLachlan G., "*Finite Mixture Models*", Wiley, 2000.

[3] Thanh M. N., Wu Q. M. J., and Ahuja S., "An Extension of the Standard Mixture Model for Image Segmentation," *IEEE Transactions on Neural Networks*, vol. 21, no. 8, pp. 1326–1338, 2010.

[4] Celeux G., Forbes F., and Peyrard N., "EM procedures using mean field-like approximations for Markov model-based image segmentation," *Pattern Recognition*, vol. 36, no. 1, pp. 131–144, 2003.

[5] Blekas K., Likas A., Galatsanos N. P., and I. E. Lagaris, "A spatially constrained mixture model for image segmentation," *IEEE Transactions on Neural Networks*, vol. 16, no. 2, pp. 494–498, 2005.

[6] Nikou C., Galatsanos N. P., and Likas A., "A Class-Adaptive Spatially Variant Mixture Model for Image Segmentation," *IEEE Transactions on Image Processing*, vol. 16, no. 4, pp. 1121–1130, 2007.

[7] Nikou C., Likas A., and Galatsanos N. P., "A Bayesian Framework for Image Segmentation with Spatially Varying Mixtures," *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2278–2289, 2010.

[8] Thanh M. N., and Wu Q. M. J., "Robust Student's-t Mixture Model with Spatial Constraints and Its Application in Medical Image Segmentation," *IEEE Transactions on Medical Imaging*, vol. 31, no. 1, pp. 103–116, 2012.

[9] Thanh M. N., and Wu Q. M. J., "Dirichlet Gaussian Mixture Model: Application to Image Segmentation," *Image and Vision Computing*, vol. 29, no. 12, pp. 818–828, 2011.

[10] Peel D., and McLachlan G., "Robust Mixture Modeling Using the t Distribution," *Statistics and Computing*, vol. 10, pp. 339–348, 2000.

[11] Dempster P., Laird N. M., and Rubin D. B., "Maximum likelihood from incomplete data via EM algorithm," *Journal of the Royal Statistical Society*, vol. 39, no. 1, pp. 1–38, 1977.

[12] Krinidis S., and Chatzis V., "A Robust Fuzzy Local Information C-means Clustering Algorithm," *IEEE Transactions on Image Processing*, no. 19, vol. 5, pp. 1328–1337, 2010.

[13] Zhang Y., Brady M., and Smith S., "Segmentation of brain MR images through a hidden Markov random field model and the expectation maximization algorithm," *IEEE Transactions on Medical Imaging*, vol. 20, no. 1, pp. 45–57, 2001.

[14] Unnikrishnan R., Pantofaru C., and Hebert M., "A measure for objective evaluation of image segmentation algorithms," *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 3, pp. 34–41, 2005.

[15] Martin D., Fowlkes C., Tal D., Malik J., "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," *International Conference on Computer Vision*, vol. 2, pp. 416–423, 2001.