Image and Vision Computing 29 (2011) 818–828

Contents lists available at SciVerse ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Dirichlet Gaussian mixture model: Application to image segmentation ☆,☆☆

Thanh Minh Nguyen *, Q.M. Jonathan Wu [1]

Department of Electrical and Computer Engineering, University of Windsor, 401 Sunset Avenue, Windsor, ON, Canada N9B-3P4

ARTICLE INFO

Article history:
Received 3 March 2011
Received in revised form 1 September 2011
Accepted 6 September 2011

Keywords:
Dirichlet Gaussian mixture model
Dirichlet distribution
Spatial constraints
Gradient method
Image segmentation

ABSTRACT

Gaussian mixture model based on the Dirichlet distribution (Dirichlet Gaussian mixture model) has recently received great attention for modeling and processing data. This paper studies the new Dirichlet Gaussian mixture model for image segmentation. First, we propose a new way to incorporate the local spatial information between neighboring pixels based on the Dirichlet distribution. The main advantage is its simplicity, ease of implementation and fast computational speed. Secondly, existing Dirichlet Gaussian model uses complex log-likelihood function and require many parameters that are difficult to estimate. The total parameters in the proposed model lesser and the log-likelihood function have a simpler form. Finally, to estimate the parameters of the proposed Dirichlet Gaussian mixture model, a gradient method is adopted to minimize the negative log-likelihood function. Numerical experiments are conducted using the proposed model on various synthetic, natural and color images. We demonstrate through extensive simulations that the proposed model is superior to other algorithms based on the model-based techniques for image segmentation.

1. Introduction

Image segmentation is one of the most difficult and challenging problem in image processing. Accurately segmented images provide more and useful information for diagnosis and quantitative analysis. However, automated segmentation [1,2] is still a very challenging research topic, due to overlapping intensities and low contrast in images, as well as noise perturbation. In literature, different methodologies proposed for image segmentation include mean shift [3,4], clustering methods [5,6], graph based techniques [7–10], partial differential equations (PDE) based segmentation techniques [11,12], and region competition [13].

During the last decades, much attention has been given to model-based techniques [16–22] to model the uncertainty in a probabilistic manner. In model-based techniques, standard GMM [23,24] is a well-known method used in most applications. An advantage of the standard GMM is that it requires a small amount of parameters for learning. Another advantage is that these parameters can be efficiently estimated by adopting the expectation maximization (EM) algorithm [25,26] to maximize the log-likelihood function. However, a major shortcoming of this method is that it does not take into account the spatial dependencies in the image. Moreover, it does not use the prior knowledge that adjacent pixels most likely belong to the same cluster. In this family of Bayesian segmentation methods, prior probabilities [25] of class membership are considered constant for every pixel of an image. Thus, the performance of Bayesian segmentation methods is too sensitive to noise and image contrast levels.

A possible approach to overcome this problem is to impose spatial smoothness constraints to incorporate the spatial relationships between neighboring pixels [27]. Recently, several mixture models based on Markov random field (MRF) for pixel label are proposed in [28–32]. According to these approaches, prior probabilities capture spatial information by using a MRF. The primary advantage of this family of mixture models is that it incorporates spatial information and reduces complexity and computational cost. Hence, it improves segmentation results, particularly when image is corrupted by high levels of noise.

Another family of mixture models based on MRF for pixel label priors have been successfully applied to image segmentation [33–37]. 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. Their primary disadvantage, however, lies in its additional training complexity. For example, the M-step of the EM algorithm in [33,34] cannot evaluate the prior distribution in a closed form, which therefore corresponds to an increase in the algorithm's complexity. In [33] the gradient projection step was proposed to implement the M-step. Another reparatory projection step based on a closed form update equation was introduced [35,36] to guarantee that the prior probabilities are positive and sum to one.

To eliminate the reparatory projection from the EM algorithm and to guarantee that the prior probabilities are positive and sum to one, a segmentation mixture model with spatial constraints was proposed

☆ This research has been supported in part by the Canada Research Chair Program, AUTO21 NCE, and the NSERC Discovery grant.
☆☆ This paper has been recommended for acceptance by Thomas Brox.
* Corresponding author. Tel.: +1 519 253 3000x4862.
E-mail addresses: nguyen1j@uwindsor.ca (T.M. Nguyen), jwu@uwindsor.ca (Q.M.J. Wu).
[1] Tel.: +1 519 253 3000x4862.

0262-8856/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2011.09.001

in [37]. This model assumes that the prior probabilities follow a Dirichlet distribution [38,39]. The probability to assign a pixel to the class is modeled by a discrete multinomial distribution whose parameters follow a Dirichlet law [40]. The advantage of this model is that that each update to the parameters resulting an E-step followed by an M-step is guaranteed that the prior probability is computed subject to the probability constraints (positive and sum to one) without requiring a reparatory projection step. Moreover, the approach leads to improvements in image segmentation accuracy. However, this model requires many parameters and a complex log-likelihood function. In addition, the cost of this method is still quite high, and is not robust against noise.

In this paper, we propose a new Dirichlet Gaussian mixture model for image segmentation. Our approach differs from those discussed above by the following statements. We propose an alternate way of incorporating local spatial interactions between neighboring pixels based on Dirichlet distribution and Dirichlet law. Secondly, the existing Dirichlet Gaussian model requires a complex log-likelihood function with many parameters. The parameters in the proposed model are less and the log-likelihood function has a simpler form. Thirdly, to estimate the parameters of this Dirichlet Gaussian mixture model, a gradient method is adopted to minimize the negative log-likelihood function.

The remainder of this paper is organized as follows. In Section 2, we present a brief introduction of the Dirichlet Gaussian mixture model, commonly used in the literature for image segmentation. In Section 3, we describe the details of the proposed method, parameter estimation, and the relationship between the proposed model with the existing mixture model. In Section 4, we present the experimental results followed by conclusions in Section 5.

2. Dirichlet Gaussian mixture model

In this section, we start with a brief review of the Dirichlet Gaussian mixture model. Let $\mathbf{x}_i$, $i=(1,2,\ldots,N)$, denote an observation at the i-th pixel of an image with dimension D. The i-th pixel is characterized by the prior probabilities vector $\pi_i=(\pi_{i1},\pi_{i2},\ldots,\pi_{iK})$. Classes are denoted by $(\Omega_1,\Omega_2,\ldots,\Omega_K)$. The discrete probability label at the i-th pixel is denoted by $\mathbf{z}_i=(z_{i1},z_{i2},\ldots,z_{iK})$. According to [37], the discrete probability label z_{ij}, $j=(1,2,\ldots,K)$, is defined as:

$$z_{ij}=\begin{cases}1 & \text{IF : pixel } \mathbf{x}_i \text{ belongs to class } \Omega_j\\ 0 & \text{Otherwise}\end{cases}. \tag{1}$$

In Gaussian mixture model [28–37, 44–51], the density function at an observation $\mathbf{x}_i$ is given by:

$$p(\mathbf{x}_i)=\sum_{j=1}^{K}\pi_{ij}p(\mathbf{x}_i|\Omega_j) \tag{2}$$

where, $p(\mathbf{x}_i|\Omega_j)$ is the Gaussian distribution with its own mean μ_j and covariance Σ_j.

$$p(\mathbf{x}_i|\Omega_j)=\frac{1}{(2\pi)^{D/2}}\frac{1}{\left|\Sigma_j\right|^{1/2}}exp\left\{-\frac{1}{2}\left(x_i-\mu_j\right)^T\Sigma_j^{-1}\left(x_i-\mu_j\right)\right\} \tag{3}$$

Given the density function in Eq. (2), the log-likelihood function is written in the form.

$$L(\Theta)=\sum_{i=1}^{N}log\left\{\sum_{j=1}^{K}\pi_{ij}p\left(\mathbf{x}_i|\Omega_j\right)\right\} \tag{4}$$

As seen from Eq. (4), the pixel $\mathbf{x}_i$ is considered an independent sample. The spatial correlation between the neighboring pixels does not influence the decision process. In order to overcome this problem, the spatially variant finite mixture model (SVFMM) [33,34] was suggested to incorporate the local spatial information by introducing the Gibbs function for the priors. The new log-likelihood function is derived as:

$$L(\Theta)=\sum_{i=1}^{N}\log\left\{\sum_{j=1}^{K}\pi_{ij}p\left(\mathbf{x}_i|\Omega_j\right)\right\}+\log p(\Pi). \tag{5}$$

For more details of $p(\Pi)$, please refer to [33,34,36]. The EM algorithm [25,26] is adopted to maximize the log-likelihood function in Eq. (5) with respect to the parameters $\Theta=(\mu_j,\Sigma_j^{-1},\pi_{ij})$. However, the M-step cannot evaluate the prior distribution π_{ij} in a closed form because of the complexity of the log-likelihood function in Eq. (5). Therefore, for each iteration of the EM algorithm, a reparatory projection step [33,35,36] is added to the M-step to guarantee that the prior probabilities are positive and sum to one, which therefore corresponds to an increase in the algorithm's complexity.

In order to overcome this problem, Dirichlet Gaussian mixture model has been proposed in [37]. This model assumes that the discrete probability label z_{ij} is a random variable following a multinomial distribution [15,40] with M realizations.

$$p(\mathbf{z}_i|\xi_i)=\frac{M!}{\prod_{j=1}^{K}\left(z_{ij}\right)!}\prod_{j=1}^{K}\left(\xi_{ij}\right)^{z_{ij}} \tag{6}$$

where, $\xi_i=(\xi_{i1},\xi_{i2},\ldots,\xi_{iK})$, $i=(1,2,\ldots,N)$, is the parameter. The probability ξ_{ij} satisfies the constraints:

$$\xi_{ij}\geq 0 \text{ and } \sum_{j=1}^{K}\xi_{ij}=1. \tag{7}$$

Besides that, the Dirichlet distribution [39] is defined as:

$$p(\xi_i|\alpha_i)=\frac{\Gamma\left(\sum_{j=1}^{K}\alpha_{ij}\right)}{\prod_{j=1}^{K}\Gamma\left(\alpha_{ij}\right)}\prod_{j=1}^{K}\left(\xi_{ij}\right)^{(\alpha_{ij}-1)} \tag{8}$$

where, $\alpha_i=(\alpha_{i1},\alpha_{i2},\ldots,\alpha_{iK})$, $i=(1,2,\ldots,N)$, is the vector of Dirichlet parameters for ξ_i and α_{ij} is non negative: $\alpha_{ij}\geq 0$. $\Gamma(\cdot)$ in Eq. (8) is the Gamma function. Then, the probability label is given by:

$$p(\mathbf{z}_i|\alpha_i)=\int_0^1 p(\mathbf{z}_i|\xi_i)p(\xi_i|\alpha_i)d\xi_i. \tag{9}$$

After some manipulation, we obtain:

$$p(\mathbf{z}_i|\alpha_i)=\frac{M!}{\prod_{j=1}^{K}\left(z_{ij}\right)!}\frac{\Gamma\left(\sum_{j=1}^{K}\alpha_{ij}\right)}{\Gamma\left(\sum_{j=1}^{K}\left(\alpha_{ij}+z_{ij}\right)\right)}\prod_{j=1}^{K}\frac{\Gamma\left(\alpha_{ij}+z_{ij}\right)}{\Gamma\left(\alpha_{ij}\right)}. \tag{10}$$

We now consider the condition of the discrete probability label z_{ij} in Eq. (1). Substituting Eq. (1) into Eq. (10) and processing with only one realization ($M=1$), the prior probabilities for the i-th pixel corresponding to class Ω_j become:

$$\pi_{ij}=p\left(z_{ij}=1|\alpha_i\right)=\frac{\alpha_{ij}}{\sum_{k=1}^{K}\alpha_{ik}}. \tag{11}$$