J. Vis. Commun. Image R. 34 (2016) 135–145

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Double Gaussian mixture model for image segmentation with spatial relationships ☆

Taisong Xiong [a], Lei Zhang [b,*], Zhang Yi [b]

[a] College of Applied Mathematics, Chengdu University of Information Technology, Chengdu 610225, Sichuan, PR China
[b] Machine Intelligence Laboratory, College of Computer Science, Sichuan University, Chengdu 610065, PR China

ARTICLE INFO

Article history:
Received 26 June 2015
Accepted 23 October 2015
Available online 10 November 2015

Keywords:
Markov random model
Gaussian mixture model
Image segmentation
Expectation maximization (EM) algorithm
Gradient descent
Spatial relationships
Synthetic noisy grayscale images
Real-world color images

ABSTRACT

In this paper, we present a finite mixture model based on a Gaussian distribution for image segmentation. There are four advantages to the proposed model. First, compared with the standard Gaussian mixture model (GMM), the proposed model effectively incorporates spatially relationships between the pixels using a Markov random field (MRF). Second, the proposed model is similar to GMM, but has a simple representation and is easier to implement than some existing models based on MRF. Third, the contextual mixing proportion of the proposed model is explicitly modelled as a probabilistic vector and can be obtained directly during the inference process. Finally, the expectation maximization algorithm and gradient descent approach are used to maximize the log-likelihood function and infer the unknown parameters of the proposed model. The performance of the proposed model at image segmentation is compared with some state-of-the-art models on various synthetic noisy grayscale images and real-world color images.

1. Introduction

The goal of image segmentation [1] is to label image pixels and divide them into clusters according to similarity of attributes. Image segmentation has been widely applied to image processing, video surveillance systems and medical image analysis. Because of its widespread applications, image segmentation is receiving greater attention. In recent years, many different algorithms for image segmentation have been developed. Among these algorithms, representative techniques include region splitting and merging [2,3], normalized cut [4,5], mean shift [6,7], and level sets [8–10]. Statistical algorithms based on clustering have also been successfully applied to image segmentation.

Among the statistical models, the finite mixture model (FMM) [11] is receiving increasing attention because of its simple form and ease of implementation. The FMM comes in several different forms. Two representative FMMs are the Gaussian mixture model (GMM) [12,13] and Student's-t mixture model (SMM) [14,15]. The component functions for these models are the Gaussian distribution and Student's t-distribution, respectively. The expectation maximization (EM) [16,17] algorithm is generally used to infer the parameters of these models. The GMM and SMM obtain good segmentation results when applied to image segmentation [13,15].

However, experiments show that FMM is not robust against noise, with unsatisfactory segmentation results obtained when images are degraded by noise. The experiments show that FMM is not robust against noise. The main cause of this is that FMM supposes that the pixels are statistically independent and so the spatial information of pixels is not taken into account.

To overcome the aforementioned drawback and improve the quality of image segmentation for FMM, the Markov random field (MRF) [18,19] model which incorporates the spatial relationships between image pixels has been proposed. MRF models have been widely applied in image processing [20,21] and image segmentation [22]. The MRF models are divided into two types according to the methods used to incorporate spatial information. One approach imposes spatial information on pixel labels [22,23]. These MRF models obtain better segmentation results than FMM because of the spatial information, but this also increases the burden of computation cost for the MRF model. To improve the computational efficiency of MRF model, a Bethe approximation [24] is used to approximate the Gibbs free energy function. However, the computational cost in [24] still remains very high because of the complex object function.

Another type of MRF, the spatially variant FMM (SVFMM) was proposed in [25]. In this model, the contextual mixing proportion

☆ This paper has been recommended for acceptance by M.T. Sun.

* Corresponding author.
E-mail address: leizhang@scu.edu.cn (L. Zhang).

http://dx.doi.org/10.1016/j.jvcir.2015.10.018
1047-3203/

π_{nk} is assumed to be a random variable upon which a spatial smoothness prior is imposed. The EM algorithm is used to infer the parameters of the model. The model obtains better segmentation results than FMM. However, the contextual mixing proportion π_{nk} cannot be obtained in closed form in the M-step of the EM algorithm. In other words, the contextual mixing proportion π_{nk} is a probabilistic vector and must be nonnegative and satisfy the constraints $\sum_{k=1}^{K}\pi_{nk} = 1$. To obtain a closed form solution for π_{nk}, the gradient projection algorithm is added in the M-step [25]. Blekas et al. [26] improved the SVFMM proposed in [25] using convex quadratic programming instead of gradient projection. The primary drawback of the model in [25,26] is that the contextual mixing proportion π_{nk} cannot be obtained directly from the given data. To resolve this problem, a prior distribution based on the Gauss–Markov random field is imposed on the contextual mixing proportion [27]. The advantage of this model is that a closed form equation can be obtained by the EM algorithm. To preserve the region boundaries of the image segmentation results, two models were proposed in [28]. The first of these integrates a binary-Bernoulli-distributed line process (BLP) with MRF, and the other incorporates a continuous-Gamma-distributed line process (CLP). However, one drawback of the models in [27,28] is that the contextual mixing proportion π_{nk} still cannot be obtained directly as a probabilistic vector.

To improve the computational efficiency of the model, a new representation of the contextual mixing proportion was given in [29]. In this model, the contextual mixing proportion is explicitly modelled as a probabilistic vector. Therefore, a closed form for the contextual mixing proportion can be obtained directly. However, this model can only be applied to grayscale image segmentation. A detail-preserving mixture model for image segmentation was proposed in [30]. Different weight values are assigned to different pixels in each pixel's neighborhood. However, the computational complexity of the model in [30] is still very high because of the complex representation of the log-likelihood function.

To simplify the representation of the model and improve the computational efficiency, we propose a mixture model for image segmentation in this paper. The proposed model is very different from most of the aforementioned models. First, the proposed model is based on GMM. Therefore, it has a simple form and is easy to implement. Second, the proposed model incorporates spatial relationships between the pixels. Therefore, it is more robust against noise than FMM. Third, fewer parameters need to be estimated in the proposed model than in other models based on MRF, which improves the computational efficiency of the proposed model. Finally, the EM algorithm and gradient descent method are used to estimate the parameters of the proposed model directly. Therefore, the inference process is much simpler than in some other models based on MRF.

The remainder of this paper is organized as follows. In Section 2, the theoretical background related to our proposed model is introduced briefly. In Section 3, a detailed description of the proposed model is presented. Experimental results obtained by our model for various synthetic noisy grayscale images and natural color images and some discussions are presented in Section 4. Our conclusions are given in Section 5.

2. Theoretical background

In this section, the GMM and the mixture model based on MRF, both of which are closely related to our proposed model, are described briefly.

2.1. Gaussian mixture model

A GMM is a linear combination of more than one Gaussian distribution [11]. Its definition is given in the following equation:

$$f(x_n) = \sum_{k=1}^{K}\pi_k N(x_n|\Theta_k), \tag{1}$$

where each component function $N(x_n|\Theta_k)$ is a Gaussian distribution. The multivariate Gaussian distribution of a D-dimensional vector x has the following form [12]:

$$N(x|\Theta) = \frac{1}{(2\pi)^{D/2}}\frac{1}{|\Sigma|^{1/2}}\exp\left\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\}, \tag{2}$$

where μ represents a D-dimensional mean vector, Σ denotes a $D \times D$ covariance matrix, and $|\Sigma|$ is the determinant of the matrix Σ. The prior distribution π_k denotes the probability that observation x_n belongs to the kth class Ω_k. It can be seen that π_k is independent of observation x_n. Furthermore, π_k should satisfy the following constraints:

$$0 \leqslant \pi_k \leqslant 1, \quad \sum_{k=1}^{K}\pi_k = 1; \quad k = 1,\ldots,K. \tag{3}$$

When the density function for an observation has been determined, the log-likelihood function of N observations is given by [12]

$$L(\Theta) = \sum_{n=1}^{N}\log\left(\sum_{k=1}^{K}\pi_k N(x_n|\Theta_k)\right). \tag{4}$$

According to (1) and (4), the main advantage of the GMM is that its form is very simple and it requires a small number of parameters. However, when GMM is applied to image segmentation, these observations are considered to be independent of each other. To determine the parameters (π_k, μ_k, Σ_k), the EM algorithm [16,17] is usually applied to maximize the log-likelihood function in (4). In the E-Step of EM, the posterior probability can be obtained by [12]

$$p^{(t)}(\Theta_k|x_n) = \frac{\pi_k N(x_n|\Theta_k)}{\sum_{j=1}^{K}\pi_j N(x_n|\Theta_j)}. \tag{5}$$

In the M-Step of EM, the parameters (π_k, μ_k, Σ_k) are updated iteratively according to the following formulas [12]:

$$\mu_k^{(t+1)} = \frac{\sum_{n=1}^{N}p^{(t)}(\Theta_k|x_n)x_n}{\sum_{n=1}^{N}p^{(t)}(\Theta_k|x_n)}, \tag{6}$$

$$\Sigma_k^{(t+1)} = \frac{\sum_{n=1}^{N}p^{(t)}(\Theta_k|x_n)(x_n-\mu_k)(x_n-\mu_k)^T}{\sum_{n=1}^{N}p^{(t)}(\Theta_k|x_n)}, \tag{7}$$

$$\pi_k^{(t+1)} = \frac{\sum_{n=1}^{N}p^{(t)}(\Theta_k|x_n)}{N}, \tag{8}$$

where t denotes the iteration number. The loop is terminated when the convergence condition is satisfied. Using the maximum posterior criterion and the value from (5), we can obtain the class label for each one pixel.

2.2. Mixture model based on MRFs

To improve the accuracy and robustness of GMM for image segmentation, MRF models [18,19] that consider spatial dependent relationships between pixels have been introduced. These MRF models [20] have been applied successfully to image segmentation, restoration and other processing tasks. In MRF models, the parameter π_{nk} represents the probability that pixel x_n belongs to the kth class Ω_k. It is quite obvious that the prior distribution π_{nk} of the nth pixel is closely related to itself. The parameter π_{nk} is referred to as the contextual mixing proportion. The density function of the nth pixel x_n is defined as follows: