

Robust Gaussian Mixture Modelling Based on Spatially Constraints for Image Segmentation

Tai-Song Xiong

College of Applied Mathematics
Chengdu University of Information Technology
Chengdu 610225, Sichuan, P.R. China
xiongtaisong@gmail.com

Yuan-Yuan Huang*

Department of Network Engineering
Chengdu University of Information Technology
Chengdu 610225, Sichuan, P.R. China

*Corresponding author
oyyhuang@gmail.com

Received December, 2014; revised April, 2015

ABSTRACT. *Gaussian mixture model (GMM) has been successfully applied to image segmentation. However, the drawback of GMM is that it lacks of robustness against noise for image segmentation. To effectively reduce negative effects of the noise, in this paper, we propose a variant of GMM which fully considers the spatial relationship between the pixels and the label probability proportions are explicitly modelled as probability vectors. At the same time, the component function of a pixel is also closely relative to its neighboring pixel. In the inference process, gradient descend method is adopted to estimate the parameters of the proposed model. The proposed model compares with some models which are related to mixture models. Several experiments are conducted on both synthetic grayscale images and real-world natural images. The experimental results show the robustness and accuracy of the proposed model outperform some state-of-the-art models.*

Keywords: Spatially varying finite mixture model; Gaussian mixture model; Image segmentation; Gradient descent

1. Introduction. Image segmentation is an important research field in image processing and computer vision. Its target is to learn several parts which have a strong correction with an object [1]. Many models based on various mathematical tools for image segmentation have been proposed in past years and statistical models play an important role in these models.

The finite mixture model (FMM) is one of the most widely applied statistical models because it can model much complicated phenomena. In recent years, the FMM is receiving more and more researchers' attentions because of its simplicities and flexibility. The FMM has been successfully applied to many fields, such as genetics, social science, astronomy, data mining, image segmentation [2]. Its component function may be any probabilistic distribution. We refer to the FMM as Gaussian mixture model (GMM) when Gaussian distribution is chosen by its component function. In image segmentation, a pixel is produced by a GMM [2, 3]. Then the parameters of the GMM can be obtained according to

the image. The K components of the GMM correspond to K clusters when it is applied to image segmentation. According to the Maximum a posteriori (MAP) estimation, a cluster label is assigned to the pixel.

The GMM can obtain good segmentation results for one image without noise, however it usually obtains unsatisfied segmentation results under the noisy conditions. The main reason leads to the unsatisfied segmentation results with noisy images for the GMM is that it assumes that the relationship between the pixels in one image is statistical independent. The spatial relationships between the pixels are not considered in the GMM when it is applied to image segmentation.

To improve the robustness of the GMM against noise, a spatially varying finite mixture model (SVFMM) is proposed in [4] which considers the spatial information of the pixels. The main difference between the GMM and the SVFMM is the representations of the label probability proportions (the probabilities of each pixel belonging to some clusters). In the GMM, the label probability proportions are independent of pixels, but they closely correlate with the pixels in the SVFMM. In general, expectation maximization (EM) algorithm [5] is used to deduce the parameters of the SVFMM. However, the label probability proportions cannot be directly obtained in closed forms. To obtain closed form solutions, a reparatory step should be added in M step. In [4], the reparatory method is Gradient projection. For the SVFMM in [4], convex quadratic programming is used in [6] instead of the Gradient projection in the M step. The experiments in [6] show that the modified model produces better segmentation results than [4]. However, the smooth parameter β in the SVFMM [4, 6] cannot adapt to the image data. It requires a tedious trial-and-error process to obtain an optimal value. To resolve this problem, a novel smoothness priors based on the Gaussian-Markove random fields is proposed [7]. However, the solutions to the label probability proportions cannot satisfy the constraints that they are nonnegative and their sum equals to one. To preserve region boundaries of the segmentation results, two models which consider the MRF priors are presented in [8]. Line process [9] is effectively incorporated with the two models. To estimate the parameters of the two models, variational inference and EM algorithm are applied to these two models, respectively. However, the solutions to the label probability proportions are still not closed forms.

To improve the efficiency and robustness of model, we propose a variant of spatially varying finite mixture model. Firstly, the label probability proportions of the proposed model fully consider the spatial relationships between the pixels, furthermore, the component function of a pixel is closely relative to its neighborhoods. Secondly, the computational cost is reduced because the representation of label probability proportions is a probability vector. Finally, we adopt Gradient descend method to estimate the parameters of the proposed model. The experiments conducted on some images demonstrate the robustness and correctness of the proposed model. The experimental results show the superiority of the proposed model over some exist models.

The rest of this paper is organized as follows. The introduction of GMM and SVFMM are given in Section 2. In Section 3, we describe the proposed model in detail. Experimental results and the relative discussions are given in Section 4. Finally, we present the conclusions in Section 5.

2. The Theoretical Background. In this section, the backgrounds of GMM and SVFMM are discussed in brief. Let x_n denote the n th pixel of an image throughout this paper. We assume that there are N pixels in an image which belong to K clusters.

2.1. Gaussian Mixture Model. The GMM is a superposition of K Gaussian distributions [2, 3] whose form is given by

$$f(x) = \sum_{k=1}^K \pi_k \mathcal{N}(x|\theta_k), \tag{1}$$

where $\mathcal{N}(x|\theta)$ denotes Gaussian distribution and its definition is written as follows

$$\mathcal{N}(x|\theta) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp \left\{ -\frac{1}{2\sigma^2}(x - \mu)^2 \right\}, \tag{2}$$

where $\theta = \{\mu, \sigma^2\}$ and μ is the mean and σ^2 denotes the variance. The coefficient π_k must satisfy the following constraints:

$$0 \leq \pi_k \leq 1, \quad \sum_{k=1}^K \pi_k = 1.$$

The likelihood of the N data can be written as follows

$$F(X) = \prod_{n=1}^N f(x_n) = \prod_{n=1}^N \sum_{k=1}^K \pi_k \mathcal{N}(x_n|\theta_k).$$

In general, the probability of one pixel’s distribution is very little, therefore, the value of their product is even little and it may causes floating-point underflow in computation. To resolve this problem, the logarithm of the likelyhood function is adopted, the log-likelihood function is written in this form.

$$\ln F(X) = \sum_{n=1}^N \ln \sum_{k=1}^K \pi_k \mathcal{N}(x_n|\theta_k).$$

A general technique for determining the parameters of GMM is EM algorithm [5].

2.2. Spatially Varying Finite Mixture Model. The SVFMM [4, 10] which is based on Markov random field (MRF) and MRF prior is imposed on the label probability proportions, is an extension of GMM. The label probability proportion π_{nk} stands for the probability that the n th pixel belongs to the k th cluster. The variable π_{nk} must satisfy the following constraints

$$0 \leq \pi_{nk} \leq 1, \quad \sum_{k=1}^K \pi_{nk} = 1; n = 1, \dots, N; k = 1, \dots, K.$$

The density function of the n th pixel is presented by [4]

$$f(x_n|\Pi, \Theta) = \sum_{k=1}^K \pi_{nk} p(x_n|\theta_k). \tag{3}$$

The component function $p(x_n|\theta_k)$ of (3) can be any probabilistic distribution. In general, Gaussian distribution is adopted in (3). At the same time, $\Theta = \theta_k$, parameter $\theta_k = \{\mu_k, \sigma_k^2\}$ where μ_k is the mean and σ_k^2 stands for the variance of Gaussian distribution, respectively. Comparing (1) with (3), the SVFMM is a special case of GMM when $\pi_{1k} = \pi_{2k} = \dots = \pi_{Nk} = \pi_k$ and the component function of the SVFMM is Gaussian distribution. Let X denote the whole observed data set $\{x_n\}$ in an image, where $n=1, \dots, N$. Let $\Pi = \{\pi^1, \pi^2 \dots \pi^N\}$ denote the set of the label probability proportions

where $\pi^n = (\pi_{n1}, \pi_{n2} \dots \pi_{nK})$ denotes a probability vector of the n th pixel. The probability density function (PDF) of observed data [4] is defined as follows.

$$f(X|\Pi, \Theta) = \prod_{n=1}^N f(x_n|\Pi, \Theta) = \prod_{n=1}^N \sum_{k=1}^K \pi_{nk} p(x_n|\theta_k). \quad (4)$$

The parameter set Π is assumed to follow a Gibbs distribution in [6]. Its probability distribution is defined as follows

$$p(\Pi) = \frac{1}{Z} \exp(-U(\Pi)), \quad \text{with } U(\Pi) = \beta \sum_{n=1}^N V_{\partial_n}(\Pi),$$

where Z represents a normalizing constant, at the same time, β plays a regularization role in the distribution. The n th pixel's clique potential function $V_{\partial_n}(\Pi)$ is closely related to its neighborhood ∂_n . Then the log-density function of data X can be written as follows [4]

$$\log f(\Pi|X; \Theta) = \sum_{n=1}^N \log \sum_{k=1}^K \pi_{nk} p(x_n|\mu_k, \sigma_k^2) + \log p(\Pi). \quad (5)$$

To determine the parameters in (5), the EM algorithm is usually chosen. However, the M-step of the EM algorithm cannot be used to obtain the results of the π_{nk} . The results obtained directly do not satisfy the constraints that the label probability proportions are nonnegative and their sum is equal to 1. To obtain the satisfied results of label probability proportion π_{nk} , an approximate step should be introduced in the M-step. Therefore, the gradient project and convex quadratic programming are adopted in [4] and [6], respectively.

3. The Proposed Model. In this section, a variant of spatially variant finite mixture model which effectively considers the spatial relationships between pixels is presented. At the same time, the representation of the label probability proportions is an explicit probabilistic vector. A second order neighborhood system including itself is used in the proposed model. There are 8 neighbor pixels in a second order neighborhood system, and their positions can be divided into four different directional types which are horizontal, vertical and two diagonal directions [16]. We adopt (6) defined in [16] to denote the weight of the n th pixel belonging to the k th class.

$$\xi_k(x_n) = \sum_{d=1}^D \sum_{x_i \in \partial_n} J_{nd}(x_n, x_i) \exp\left(-\frac{(x_i - c_{kd})^2}{2b_{kd}^2}\right) \quad (6)$$

where the n th pixel's neighborhood is denoted as ∂_n and D is the number of directional types. The value of function exp denotes the distance of x_i and its neighborhood. c_{kd} and b_{kd} stand for the mean and variance of neighborhood belonging to the k th class, respectively. The value of D equals 4 in the proposed model. The definition of identity function $J_{nd}(x_n, x_i)$ is written as follows

$$J_{nd}(x_n, x_i) = \begin{cases} 1, & \text{if } x_i \in \partial_n \text{ and } x_i \text{ is } d\text{th adjacency type of } x_n \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

Then, the definition of the label probability proportion of a pixel [?] can be obtained in the follow form

$$\pi_{nk} = \frac{\xi_k(x_n)}{\sum_{j=1}^K \xi_j(x_n)} = \frac{\sum_{d=1}^D \sum_{x_i \in \partial_n} J_{nd}(x_n, x_i) \exp\left(-\frac{(x_i - c_{kd})^2}{2b_{kd}^2}\right)}{\sum_{j=1}^K \sum_{d=1}^D \sum_{x_i \in \partial_n} J_{nd}(x_n, x_i) \exp\left(-\frac{(x_i - c_{jd})^2}{2b_{jd}^2}\right)}. \quad (8)$$

Clearly, the label probability proportion π_{nk} in (8) naturally satisfy these constraints that $0 \leq \pi_{nk} \leq 1$ and $\sum_{k=1}^K \pi_{nk} = 1$. To consider the neighborhood information of a pixel, its density function of the n th pixel is given by

$$f(x_n | \Theta) = \sum_{k=1}^K \pi_{nk} \frac{1}{N_n} \sum_{x_m \in \partial_n} \mathcal{N}(x_m | \theta_k), \quad (9)$$

where N_n denotes the number of neighborhood of pixel n , its number may be 4, 6, 9. x_m is a neighborhood pixel of pixel n . $\mathcal{N}(x_m | \theta_k)$ is a Gaussian distribution of its neighborhood whose definition is given in (2). Component function $\frac{1}{N_n} \sum_{x_m \in \partial_n} \mathcal{N}(x_m | \theta_k)$ can be normalized one, so it is a probabilistic distribution. Compared with the model in [16], the Gaussian distribution has a simpler form than Student's t-distribution. Therefore, the inference process is also simpler than that of [16]. The component function plays a smooth role because it considers pixel's neighborhood. Furthermore

$$0 \leq \pi_{nk} \leq 1, \sum_{k=1}^K \pi_{nk} = 1; n = 1, \dots, N; k = 1, \dots, K.$$

Given the prior probability $f(\theta_k) = \pi_{nk}$ and the density function (9), we can derive the posteriori probability in the follow form according to Bayes' theorem.

$$f(\theta_k | x_n) = \frac{\pi_{nk} \sum_{x_m \in \partial_n} \mathcal{N}(x_m | \theta_k)}{\sum_{j=1}^K \pi_{nj} \sum_{x_m \in \partial_n} \mathcal{N}(x_m | \theta_k)}. \quad (10)$$

The log-likelihood function of the pixels in an image can be derived [4] according to formulae 9

$$L(\theta) = \sum_{n=1}^N \log f(x_n) = \sum_{n=1}^N \log \left(\sum_{k=1}^K \pi_{nk} \frac{1}{N_n} \sum_{x_m \in \partial_n} \mathcal{N}(x_m | \theta_k) \right). \quad (11)$$

We maximize the log-likelihood function (11) with respect to the parameter set $\Theta = \{\mu_k, \lambda_k, c_{kd}, b_{kd}^2\}$ to obtain the optimization solutions. In [3], the negative logarithm of the log-likelihood function is called an error function. The error function monotonically decreases because of the property of monotonically increasing of the logarithm function. Therefore, we minimize the error function instead of maximizing the log-likelihood function.

$$J(\Theta) = -L(\Theta) = - \sum_{n=1}^N \log \left(\sum_{k=1}^K \pi_{nk} \frac{1}{N_n} \sum_{x_m \in \partial_n} \mathcal{N}(x_m | \theta_k) \right). \quad (12)$$

Given the complete data, we use the new parameters ($(t+1)$ th iteration) to substitute for the old ones (t th iteration) in [11]. Considering the posteriori probability (10), the

change in the error function can be written [12]

$$J(\Theta^{(t+1)}) - J(\Theta^{(t)}) = - \sum_{n=1}^N \log \left(\frac{\sum_{j=1}^K \pi_{nj}^{(t+1)} \sum_{x_m \in \partial_n} \mathcal{N}(x_m | \theta_j)}{\sum_{k=1}^K \pi_{nk}^{(t)} \sum_{x_m \in \partial_n} \mathcal{N}(x_m | \theta_k)} \times \frac{f^{(t)}(\theta_j | x_n)}{f^{(t)}(\theta_k | x_n)} \right). \tag{13}$$

Because the posterior probability $f^{(t)}(\theta_j | x_n)$ is always nonnegative and subjects to $\sum_{k=1}^K f^{(t)}(\theta_k | x_n) = 1$.

1. When the Jensen's inequality is applied, from (13), we have

$$J(\Theta^{(t+1)}) - J(\Theta^{(t)}) \leq - \sum_{n=1}^N \sum_{k=1}^K f^{(t)}(\theta_k | x_n) \times \log \left(\frac{\pi_{nk}^{(t+1)} \sum_{x_m \in \partial_n} \mathcal{N}(x_m | \theta_k)}{f^{(t)}(\theta_k | x_n) \sum_{j=1}^K \pi_{nj}^{(t)} \sum_{x_m \in \partial_n} \mathcal{N}(x_m | \theta_k)} \right). \tag{14}$$

The old parameters (at the t th iteration step) can be dropped according to [12] when we minimize the error function with respect to the new parameters (at the $(t+1)$ th iteration step). The change in error function can be obtained in the following form.

$$E(\Theta^{(t)} | \Theta^{(t+1)}) = - \sum_{n=1}^N \sum_{k=1}^K f^{(t)}(\theta_k | x_n) \times \log \left(\pi_{nk}^{(t+1)} \sum_{x_m \in \partial_n} \mathcal{N}(x_m | \theta_k) \right). \tag{15}$$

We refer to the function E in (15) as an error function. In the proposed model, to obtain the optimization solutions, minimizing E in (15) is equivalent to maximizing the log-likelihood function (11). The derivatives of the function (15) with respect to parameter Θ is used to minimize the value of the error function E . The gradients of the error function $E(\Theta)$ with respect to these parameters $\mu_j, \sigma_j, c_{jd}, b_{jd}^2$ are derived as follows

$$\frac{\partial E}{\partial \mu_j} = - \sum_{n=1}^N f^{(t)}(\theta_j | x_n) \frac{\sum_{x_m \in \partial_n} \mathcal{N}(x_m | \theta_k) \frac{x_m - \mu_j}{\sigma_j^2}}{\sum_{x_m \in \partial_n} \mathcal{N}(x_m | \theta_k)} \tag{16}$$

$$\frac{\partial E}{\partial \sigma_j} = \sum_{n=1}^N f^{(t)}(\theta_j | x_n) \frac{\sum_{x_m \in \partial_n} \mathcal{N}(x_m | \theta_k) \left(\frac{1}{\sigma_j} - \frac{1}{\sigma_j^3} (x_m - \mu_j)^2 \right)}{\sum_{x_m \in \partial_n} \mathcal{N}(x_m | \theta_k)} \tag{17}$$

$$\frac{\partial E}{\partial c_{jd}} = \sum_{n=1}^N \left\{ -f^{(t)}(\theta_j | x_n) \cdot \frac{1}{\xi_j(x_n)} + \sum_{k=1}^K f^{(t)}(\theta_k | x_n) \frac{1}{\sum_{p=1}^K \xi_p(x_n)} \right\} \cdot \rho_{njd}, \tag{18}$$

where

$$\rho_{njd} = \sum_{x_i \in \partial_n} J_{nd}(x_n, x_i) \frac{(x_i - c_{jd})}{b_{jd}^2} \cdot \exp \left(-\frac{(x_i - c_{jd})^2}{2b_{jd}^2} \right). \tag{19}$$

$$\frac{\partial E}{\partial b_{jd}^2} = \sum_{n=1}^N \left\{ -f^{(t)}(\theta_j | x_n) \cdot \frac{1}{\xi_j(x_n)} + \sum_{k=1}^K f^{(t)}(\theta_k | x_n) \frac{1}{\sum_{p=1}^K \xi_p(x_n)} \right\} \cdot \delta_{njd}, \tag{20}$$

where

$$\delta_{njd} = \sum_{x_i \in \partial_n} J_{nd}(x_n, x_i) \frac{(x_i - c_{jd})}{2b_{jd}^4} \cdot \exp\left(-\frac{(x_i - c_{jd})^2}{2b_{jd}^2}\right). \tag{21}$$

We can obtain the value of the posterior probability (10) after finishing the optimization of parameters. We assign a class label to the n th pixel according to the maximum a posteriori by solution of

$$\arg \max_k \{f(\theta_k|x_n)\}, \tag{22}$$

where the $f(\theta_k|x_n)$ is the representation given in (10).

We refer to the proposed model as the spatially smooth relationships-based Gaussian mixture model (SSGMM). The gradient descend method is applied to determine the parameters based on the error function given complete data. The all steps of SSGMM is outlined in algorithm 1.

Algorithm 1 SSGMM.

Initialize:

To determine the mean u_j and the variance σ_j^2 of Gaussian distribution by Using K -means. Then set $c_{jd}=u_j$ and $b_{jd}^2=\sigma_j^2$.

Step 1:

Calculate the Gaussian distribution $\mathcal{N}(x_i|\theta_j)$ in (2).

Compute the weight function $\xi_j(x_i)$ in (6).

The label probability proportion π_{ij} can be updated using (8).

Compute the posterior probability $f(\theta_j|x_i)$ using (10).

Step 2:

The gradient descend method is applied to update the parameters $\Phi = (\mu_j, \sigma_j, c_{jd}, b_{jd}^2)^T$

$$\Phi^{(t+1)} = \Phi^{(t)} - \eta \nabla E(\Phi^{(t)}) \tag{23}$$

where the value of η is enough small and it is the learning rate. $\eta=10^{-5}$ is selected for all experiments in this paper. $\nabla E(\Phi)$ is the derivative of the function E with respect to Φ , where $\nabla E(\Phi)=[\partial E/\partial \mu_j, \partial E/\partial \sigma_j, \partial E/\partial c_{jd}, \partial E/\partial b_{jd}^2]^T$.

Step 3:

If the change of value of log-likelihood function (11) is significant, set $\Phi^{(t)}=\Phi^{(t+1)}$, then return to step 1, otherwise go to step 4.

Step 4:

To get the clusters labels of the pixels, formula (22) is adopted after the posterior probability $f(\theta_k|x_n)$ given in (10) is determined.

4. Experimental Results. In this section, some synthetic images and real-world grayscale images are used to evaluate the performance of the proposed model. The performance evaluation of the proposed model is compared with K -means, the standard GMM [3], Student t-distribution mixture model (StMM) [13, 14], model in [15] based on the maximum of spatial likelihoods (MSL), DCASV [7], CLP, BLP [8], SVFMM [4] and spatially directional information-based Students t-distribution mixture model (SDIStMM) [16]. These models are all implemented in the Matlab environment.

We choose two criteria to quantify the image segmentation results. They are the misclassification ratio (MCR) [17] and the probabilistic rand (PR) index [18], respectively.

The definition of MCR is given by

$$\text{MCR} = \frac{\text{number of mis-classified pixels}}{\text{total number of pixels}}.$$

The value of MCR is lower and the result of image segmentation is better. The PR index is used to quantitatively measure the real-world image segmentation results. The interval of its value ranges from 0 to 1. The segmentation result is better if the value of PR is higher.

4.1. Synthetic Images. In the first synthetic image experiment, we use a three class ($K = 3$) synthetic image used in [17] which is generated by the Gibbs sample [9]. The image has 128×128 pixels and its luminance values are 55, 115 and 225, respectively. The original image is shown in Fig.1(a). The image shown in Fig.1(b) is corrupted by Gaussian noise (0 mean, 0.01 variance). The image segmentation results obtained by K -means, GMM, StMM, MSL, DCASV, CLP, BLP, SVFMM, SDISStMM and the proposed model (SSGMM) are shown in Fig.1(c)-(l), respectively. It is a challengeable work to distinguish the edges and contours of the regions for the corrupted image shown in Fig.1(b). It can be seen from the segmentation results, the MRF-based models obtain better segmentation results than K -means, GMM and StMM. It indicates that the spatial information plays an important role in image segmentation. Furthermore, the proposed model reduces the effect of the noise significantly. This experiment demonstrates that the robustness of the proposed model is superior to the other models except for SDISStMM.

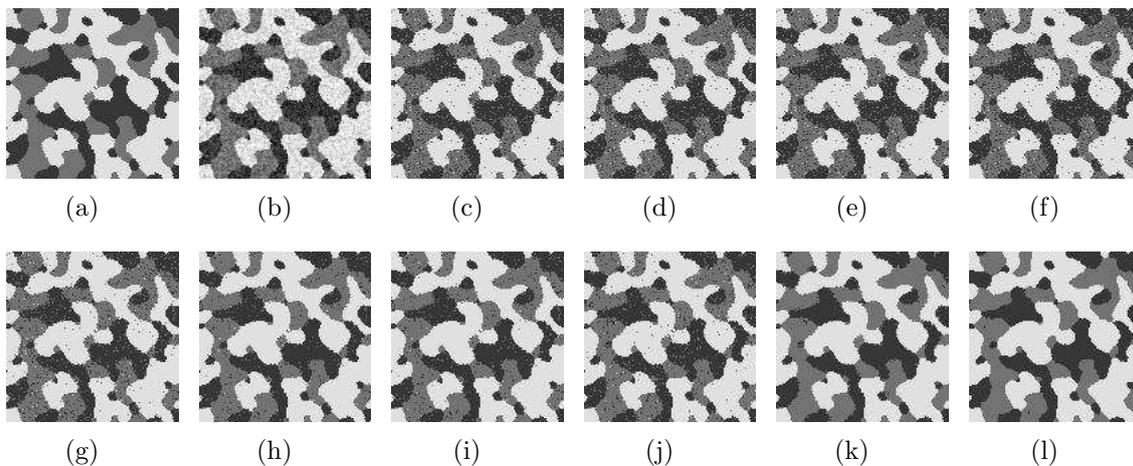


FIGURE 1. First experiment (128×128 image resolution). (a)The original image, (b)Noisy image corrupted with Gaussian noise(0 mean, 0.01 variance), (c) K -mean(MCR=7.68%), (d)GMM(MCR=7.57%), (e)StMM(MCR=7.59%),(f)MSL(MCR=7.63%), (g)DCASV(MCR=5.88%), (h)CLP(MCR=3.37%), (i)BLP(MCR=3.17%), (j)SVFMM(MCR=4.64%), (k)SDISStMM(MCR=1.85%),(l)SSGMM(MCR=2.85%)

To further verify the correctness and robustness of the proposed model, we conduct the experiments on the synthetic image shown in Fig.1(a) corrupted with varying levels of noise. To lessen the effect of the randomness of the noise, we perform experiments ten times at each noise level and the averages of the experimental results are given in Table 1. We can obtain a conclusion from the Table 1 that the models which consider the spatial relationship of pixels obtain better segmentation results than any other models which do

not consider the spatial relationship of pixels. It proves that the spatial relationships of pixels play an important role in image segmentation. Compared with any other state-of-the-art model, the SSGMM produces relatively lower MCR values at each noise level. From the Table 1, we can say that the MCR values of the proposed model increase relatively lower than any other model at every noise level. This experiment shows that the proposed model effectively captures the spatial relationships between the pixels and is more robustness against noise than any other model.

TABLE 1. THE COMPARISON OF THE MCR FOR THE FIRST EXPERIMENT

Methods	Gaussian Noise(0 mean,var)					
	var=0.011	var=0.012	var=0.013	var=0.014	var=0.015	mean
<i>K</i> -mean	9.05%	10.00%	11.06%	11.83%	12.89%	10.97%
GMM	8.94%	9.82%	10.86%	11.69%	12.71%	10.80%
StMM	8.86%	9.71%	10.74%	11.50%	12.56%	10.67%
MSL	9.02%	9.97%	10.99%	11.82%	12.88%	10.94%
DCASV	7.04%	7.79%	8.76%	9.49%	10.45%	8.71%
CLP	4.36%	4.86%	5.62%	6.16%	6.94%	5.59%
BLP	3.91%	4.22%	4.94%	5.42%	6.14%	4.93%
SVFMM	5.48%	6.54%	6.99%	7.67%	8.65%	7.07%
SDIS _t MM	1.81%	2.28%	2.36%	2.44%	2.95%	2.37%
SSGMM	2.95%	3.09%	3.35%	3.65%	4.03%	3.41%

We apply another synthetic image shown in Fig.2(a) to the second experiment. The 128×128 pixel image used in [16] has three classes ($K = 3$) and the luminance values are [85, 170, 255]. The structure of image is more simpler than the structure of image used in the first example. The image shown in Fig.2(b) is obtained by corrupting the image shown in Fig.2(a) with Gaussian noise (0 mean, 0.015 variance). As can be seen from the segment results shown in Fig.2(c)-(1), the proposed model reduces the effect of noise clearly. It produces a lower MCR than the other models except for SDIS_tMM. The averages of the experiments with each noise level are given in Table 2 (ten times experiments are performed on each noise level.). From Table 2, the lower MCR values obtained by the proposed model on each noise level show that the proposed model produces better segmentation results than those of the other models. From the table 2, we can see that the differences of MCR obtained from the image shown in Fig.2(a) between the SSGMM and the other models are larger on each noise level than the image shown in Fig.1(a). It also shows that the proposed model can obtain better segmentation results under simplified image structure. From the synthetic image experimental results, we can see that the proposed model can effectively reduce the Gaussian noise than some models.

4.2. Natural Grayscale Images. It is a challenge work to segment real-world outdoor images because the objects in real-world images are various and complicated. To verify the segment performance of the proposed model for real-world images, the Berkeley's image segmentation database [19] is chosen. There are 300 grayscale images in this database. PR is chosen to quantitatively estimate the segment results. The segment results of image 291000 are shown in Fig.3. From the results, we can see that the result obtained by the proposed model is more smooth than the results of the other models. To further test the effectiveness and accuracy of the proposed model, the PR values of 20 images segment results obtained by some models are given in table 3. We can see that the proposed model obtains relatively better segment results than any other models. It proves that the

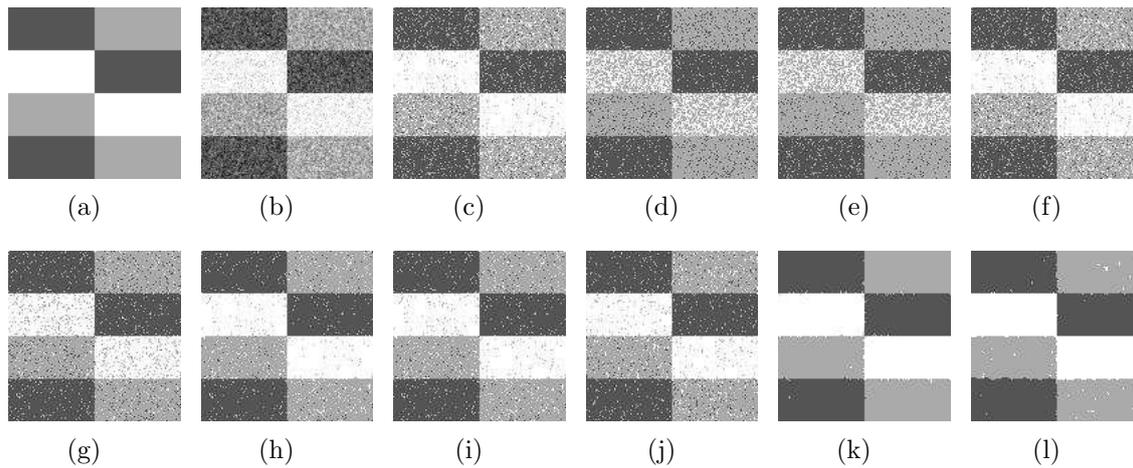


FIGURE 2. Second experiment (128×128 image resolution). (a) The original image, (b) Noise image corrupted with Gaussian noise (0 mean, 0.015 variance), (c) K -mean (MCR=14.25%), (d) GMM (MCR=19.17%), (e) StMM (MCR=19.16%), (f) MSL (MCR=13.68%), (g) DCASV (MCR=9.47%), (h) CLP (MCR=5.94%), (i) BLP (MCR=4.86%), (j) SVFMM (MCR=7.52%), (k) SDISStMM (MCR=0.52%), (l) SSGMM (MCR=0.93%).

TABLE 2. THE COMPARISON OF THE MCR FOR THE SECOND EXPERIMENT

Methods	Gaussian Noise(0 mean, var)					
	var=0.016	var=0.017	var=0.018	var=0.019	var=0.020	mean
K -mean	15.64%	16.99%	17.95%	18.90%	19.87%	17.87%
GMM	13.63%	15.88%	16.02%	16.47%	17.68%	15.94%
StMM	19.55%	20.22%	20.59%	21.23%	21.55%	20.63%
MSL	14.83%	16.20%	17.18%	18.12%	19.07%	17.08%
DCASV	10.19%	11.17%	11.78%	12.61%	13.31%	11.81%
CLP	6.74%	7.80%	8.52%	9.33%	10.19%	8.52%
BLP	5.68%	6.69%	7.44%	8.24%	9.14%	7.44%
SVFMM	7.79%	8.28%	10.01%	10.72%	11.45%	9.65%
SDISStMM	0.93%	1.10%	1.13%	1.50%	1.57%	1.25%
SSGMM	1.29%	1.65%	1.95%	2.27%	2.81%	1.99%

proposed model (SSGMM) is more effective and corrective than some other models for real-world images segmentation.

5. Conclusions. In this paper, we present a GMM which fully incorporates the spatial relationships between the pixels. The component function of a pixel is also closely relative to its neighboring pixels. To effectively inference the unknown parameters of the proposed model, gradient descent method is introduced. Some experiments conducted on synthetic and real-world grayscale images show that the proposed model outperforms some other models for image segmentation. However, the proposed model can only deal with the grayscale image. How to extend the proposed model to segment color images is one of our future's work. At the same time, the proposed model may be applied for medical image segmentation.

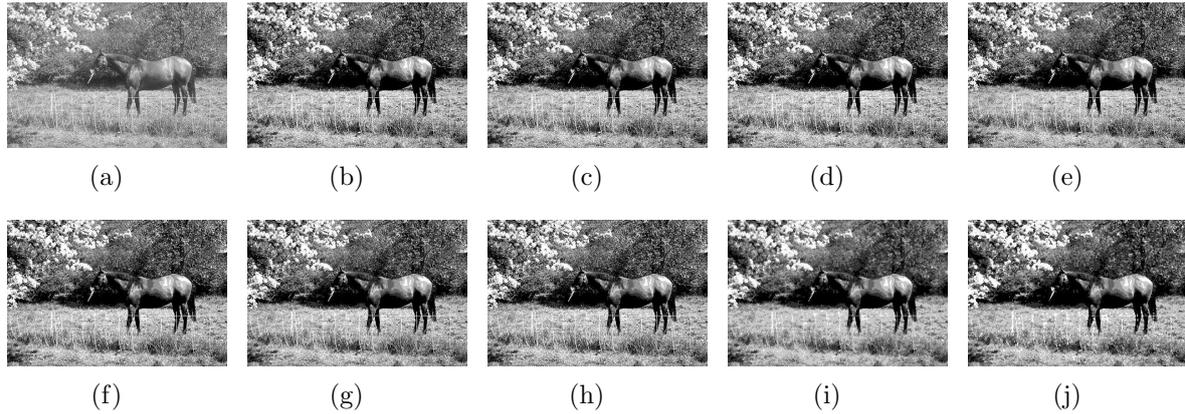


FIGURE 3. Grayscale image segmentation(291000). (a)The original image, (b) GMM(PR=0.676), (c) StMM(PR=0.675), (d)MSL(PR=0.676), (e) DCASV(PR=0.678), (f) CLP(PR=0.679), (g) BLP(PR=0.680),(h) SVFMM(PR=0.678), (i) SDISMM(PR=0.687), (j)SSGMM(PR=0.691)

TABLE 3. COMPARISON OF IMAGE SEGMENTATION RESULTS BASED ON BERKELEY GRAYSCALE IMAGES: PR.

Image	K	GMM	StMM	MSL	DCASV	CLP	BLP	SVFMM	SDISMM	SSGMM
78004	5	0.766	0.765	0.769	0.767	0.769	0.769	0.771	0.771	0.771
78019	4	0.758	0.745	0.760	0.763	0.765	0.765	0.768	0.763	0.777
85048	5	0.752	0.740	0.758	0.755	0.756	0.757	0.761	0.756	0.758
207056	2	0.692	0.672	0.715	0.701	0.712	0.713	0.727	0.731	0.726
181091	4	0.755	0.753	0.758	0.759	0.760	0.760	0.758	0.768	0.761
97017	2	0.782	0.776	0.783	0.784	0.786	0.785	0.790	0.791	0.789
291000	6	0.676	0.675	0.676	0.678	0.679	0.680	0.678	0.687	0.691
147021	2	0.731	0.715	0.785	0.738	0.772	0.776	0.785	0.790	0.789
188005	5	0.744	0.739	0.749	0.747	0.747	0.749	0.750	0.748	0.750
145014	4	0.685	0.682	0.689	0.688	0.688	0.688	0.688	0.689	0.690
189011	2	0.761	0.760	0.779	0.762	0.777	0.779	0.777	0.7719	0.780
239007	5	0.781	0.777	0.777	0.786	0.784	0.784	0.787	0.786	0.785
24063	3	0.858	0.844	0.828	0.860	0.853	0.850	0.688	0.857	0.857
246016	5	0.746	0.745	0.760	0.756	0.755	0.756	0.765	0.761	0.764
23080	5	0.741	0.731	0.752	0.746	0.747	0.748	0.756	0.747	0.752
38092	4	0.788	0.792	0.772	0.797	0.797	0.797	0.799	0.807	0.803
35010	5	0.713	0.714	0.720	0.719	0.720	0.719	0.722	0.722	0.722
55067	3	0.837	0.788	0.822	0.842	0.842	0.842	0.842	0.843	0.843
207056	2	0.692	0.672	0.715	0.701	0.712	0.713	0.727	0.731	0.734
181018	4	0.630	0.613	0.645	0.650	0.645	0.639	0.683	0.667	0.667
average	-	0.744	0.735	0.751	0.750	0.753	0.753	0.751	0.760	0.762

Acknowledgment. This work is partially supported by the National Natural Science Foundation of China (No.61303126), and the Applied Basic Research Program of Sichuan Province (No.2014JY0168) and the Scientific Research Foundation of the Education Department of Sichuan Province (No.14ZA0178) and Foundation of Chengdu University of Information Technology (No.KYTZ201426) and (No.KYTZ201419). The authors also

gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

- [1] M. L. Sonka, V. Hlavac, R. Boyle (eds.), *Image Processing, Analysis, and Machine Vision*, Thomson, 2008.
- [2] G. J. McLachlan, D. Peel (eds.), *Finite Mixture Models*, Wiley and Sons, New York, 2000.
- [3] C. Bishop (eds.), *Pattern Recognition and Machine Learning*, Springer, 2006.
- [4] G. S. Sanjay, T. J. Hebert, Bayesian pixel classification using spatially variant finite mixtures and the generalized EM algorithm, *IEEE Trans. Image Process*, vol.7, no.7, pp.1014-1028, 1998.
- [5] P. Dempster, N. M. Laird, and D. B. Rubin, Maximum likelihood from incomplete data via EM algorithm. *J Roy Statist Soc*, vol.39, pp.1-38, 1977.
- [6] K. Blekas, A. Likas, N. P. Galatsanos, and I. E. Lagaris, A spatially constrained mixture model for image segmentation, *IEEE Trans Neural Netw*, vol.16, no.2, pp.494-498, 2005.
- [7] C. Nikou, N. P. Galatsanos, and A. C. Likas, A class-adaptive spatially variant mixture model for image segmentation, *IEEE Trans Image Process*, vol.16, no.4, pp.1121-1130, 2007.
- [8] G. Sfikas, C. Nikou, N. Galatsanos, and C. Heinrich, Spatially varying mixtures incorporating line processes for image segmentation, *J Math Imag Vis*, vol.36, pp.91-110, 2010.
- [9] S. Geman, D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE TTrans. Pattern Anal. Mach.Intell*, vol.6, no.6, pp.721-741, 1984.
- [10] Y. Y. Zhang, M. Brady, and S. Smith, Segmentation of brain MR images through a hidden Markov random field model and the expectation maximization algorithm, *IEEE Trans Med Imaging*, vol.20, no.1, pp.45-75, 2001.
- [11] R. Unnikrishnan, C. Pantofaru, and M. Hebert, Toward objective evaluation of image segmentation algorithms, *IEEE TTrans. Pattern Anal. Mach.Intell*, vol.29, no.6, pp.929-944, 2007.
- [12] T. Xiong, Y. Zhang, and L. Zhang, Grayscale Image Segmentation by Spatially Variant Mixture Model with Student's t-distribution. *Multimedia Tools and Applications*, vol.72, pp.167-189, 2014.
- [13] D. M. Titterton, A. F. M. Smith, and U. E. Makov(eds.), *Statistical Analysis of Finite Mixture Distributions*, Hoboken, NJ:Wiley, 1985.
- [14] M. N. Thanh, Q. M. J. Wu, and S. Ahuja, An extension of the standard mixture model for image segmentation, *IEEE Trans Neural Netw*, vol.21, no.8, pp.1326-1338, 2010.
- [15] D. Peel, G. McLachlan, Robust mixture modeling using the t distribution. *Stat.Comput*, vol.10, pp.339C348, 2000.
- [16] G. Sfikas, C. Nikou, and N. Galatsanos, Robust Image Segmentation with Mixtures of Student's t-distribution, *IEEE International Conference on Image Processing*, San Antonio, USA, pp.I-273-I-276, 2007.
- [17] R. Mariano, D. Oscar, and M. Washington, Spatial Sampling for Image Segmentation. *The Computer Journal*, vol.55, no.3, pp.313-324, 2012.
- [18] T. Xiong, L. Zhang, Y. Zhang, Robust t-distribution mixture modeling via spatially directional information, *Neural Comput Applic*, vol.24, pp.1269-1283, 2014.
- [19] D. Martin, C. Fowlkes, and D. Tal, A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics, *Proc. of the 8th IEEE Conf Computer Vision*, Vancouver, CA, I-1-I-8, 2001.