

# Fast and Robust Spatially Constrained Gaussian Mixture Model for Image Segmentation

Thanh Minh Nguyen and Q. M. Jonathan Wu, *Senior Member, IEEE*

**Abstract**—In this paper, a new mixture model for image segmentation is presented. We propose a new way to incorporate spatial information between neighboring pixels into the Gaussian mixture model based on Markov random field (MRF). In comparison to other mixture models that are complex and computationally expensive, the proposed method is fast and easy to implement. In mixture models based on MRF, the M-step of the expectation-maximization (EM) algorithm cannot be directly applied to the prior distribution  $\pi_{ij}$  for maximization of the log-likelihood with respect to the corresponding parameters. Compared with these models, our proposed method directly applies the EM algorithm to optimize the parameters, which makes it much simpler. Experimental results obtained by employing the proposed method on many synthetic and real-world grayscale and colored images demonstrate its robustness, accuracy, and effectiveness, compared with other mixture models.

**Index Terms**—Expectation-maximization (EM) algorithm, Gaussian mixture model, image segmentation, Markov random field, spatial information.

## I. INTRODUCTION

IN ORDER to analyze the contents of a given image easily, it is often useful to construct a simpler representation. The process employed to partition an image into nonoverlapping regions is called image segmentation. An accurately segmented image provides detailed information about the objects present in an image and their respective boundaries. The process involves grouping or clustering pixels based on their intensity and spatial locations. However, images corrupted with high levels of noise make it harder to achieve accurate image segmentation. Various image segmentation algorithms, such as graph-based methods [1], [2], mean shift-based algorithms [3], binary partition tree [4], [5], and clustering approaches [6], [7] have been proposed.

Over the past decades, a number of algorithms based on the model-based techniques [8]–[10] have been proposed. The main advantage of these models is their capability to use prior knowledge to model the uncertainty in a probabilistic manner.

Manuscript received November 15, 2011; revised April 16, 2012 and June 7, 2012; accepted June 20, 2012. Date of publication August 2, 2012; date of current version April 1, 2013. This work was supported in part by the Canada Research Chair Program, AUTO21 Networks of Centers of Excellence, and in part by the Natural Sciences and Engineering Research Council of Canada. This paper was recommended by Associate Editor F. Lavagetto.

The authors are with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON N9B-3P4, Canada (e-mail: nguyen1j@uwindsor.ca; jwu@uwindsor.ca).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCSVT.2012.2211176

This prior knowledge may concern the spatial relationships between pixels in an image [11]–[13], or the whole structure of an image [14], [15]. Another advantage of the model-based techniques is that they provide a natural way to cluster data based on the components of the mixture that generated them. The methods based on the model-based techniques are successfully applied in visual surveillance [16] and statistical analysis [17]. In the model-based techniques, standard Gaussian mixture model (GMM) [8]–[10], [18], [19], is a well-known method that has been widely used due to its simplicity and ease of implementation. One of the main drawbacks of this model is that the prior distribution  $\pi_j$  does not depend on the pixel index  $i$  and the spatial relationships between the labels of neighboring pixels. Therefore, segmentation is extremely sensitive to noise and illumination.

Recently, mixture models based on Markov random field (MRF) has received great attention in most applications in the areas of image analysis, such as image sequence analysis [20], image compressions [21], and still image segmentation [22]–[24]. In order to reduce the segmentation sensitivity to noise in a still image, mixture models with MRF have been employed for pixel label [22]–[24]. The most important distinction is that in standard GMM, a common prior distribution  $\pi_j$  for all pixels  $x_i$  is evaluated, whereas, in these approaches, the prior distribution  $\pi_{ij}$  varies for every pixel  $x_i$  corresponding to each label  $\Omega_j$  and depends on the neighboring pixels and their corresponding parameters. Although the effect of noise on the final segmentation result is reduced, it lacks enough robustness with respect to noise. In addition, the computational cost of the MRF-based methods remains quite high.

Another group of mixture models based on MRF are proposed in [25]–[32] to impose spatial smoothness constraints between neighboring pixels. In this group, an MRF models the joint distribution of the priors of each pixel label  $\pi_{ij}$ . These models work well for noisy image segmentation. However, in [29]–[31], due to the introduction of the prior distribution to accurately evaluate the influence of the neighboring pixel labels during the learning step, the M-step of the expectation-maximization (EM) algorithm [33], [34], [36] cannot be directly applied to estimate the model parameters from the observations. Various approximations have been introduced in order to tackle this problem. For example, the MAP algorithm in [29] cannot evaluate the prior distribution  $\pi_{ij}$  in a closed form, and thus the gradient projection algorithm was proposed to implement the M-step. In [30] and [31], another method based on a closed-form update equation was used to implement

the M-step and estimate the parameters. For this reason, compared to the standard GMM, the cost of these methods are still quite high.

Based on these considerations, we propose a new mixture model for image segmentation, which differs from the above methods in the following manner. First, our proposed method incorporates spatial relationships among neighboring pixels in a simpler metric. Therefore, the proposed method is fast and easy to implement, compared with other mixture models that are complex and computationally expensive. Generally, in these models, the M-step of the EM algorithm cannot be directly applied for the maximization of the log-likelihood with respect to the parameters. In our proposed method, we can directly apply the EM algorithm to optimize the parameters, which makes it simpler. Finally, the proposed model is quite robust with respect to noise, accurate, and effective as compared to other GMM-based methods.

This paper is organized into five sections. In Section II, a brief background on mixture models based on MRF for image segmentation is presented. In Section III, we describe the details of the proposed algorithm. Section IV contains the experimental results. Conclusions are drawn in Section V.

## II. MIXTURE MODELS BASED ON MRF

Let  $x_i$ ,  $i = (1, 2, \dots, N)$ , denote an observation at the  $i$ th pixel of an image with dimension  $D$ . The neighborhood of the  $i$ th pixel is presented by  $\partial_i$ . Labels are denoted by  $(\Omega_1, \Omega_2, \dots, \Omega_K)$ . In order to partition an image consisting of  $N$  pixels into  $K$  labels, GMM [10], [29] assumes that each observation  $x_i$  is considered independent of the label  $\Omega_j$ . The density function  $f(x_i|\Pi, \Theta)$  at an observation  $x_i$  is given by

$$f(x_i|\Pi, \Theta) = \sum_{j=1}^K \pi_{ij} \Phi(x_i|\Theta_j) \quad (1)$$

where  $\Pi = \{\pi_{ij}\}$ ,  $i = (1, 2, \dots, N)$ ,  $j = (1, 2, \dots, K)$  is the set of prior distributions modeling the probability that pixel  $x_i$  is in label  $\Omega_j$ , which satisfies the constraints

$$0 \leq \pi_{ij} \leq 1 \text{ and } \sum_{j=1}^K \pi_{ij} = 1 \quad (2)$$

and  $\Phi(x_i|\Theta_j)$  is the Gaussian distribution, called a component of the mixture. Each Gaussian distribution can be written in the form

$$\Phi(x_i|\Theta_j) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma_j|^{1/2}} \exp \left\{ -\frac{1}{2} (x_i - \mu_j)^T \Sigma_j^{-1} (x_i - \mu_j) \right\} \quad (3)$$

where  $\Theta_j = \{\mu_j, \Sigma_j\}$ ,  $j = (1, 2, \dots, K)$ . The  $D$ -dimensional vector  $\mu_j$  is the mean, the  $D \times D$  matrix  $\Sigma_j$  is the covariance, and  $|\Sigma_j|$  denotes the determinant of  $\Sigma_j$ . Note that the observation  $x_i$  in (1) is modeled as statistically independent, the joint conditional density of the data set  $X = (x_1, x_2, \dots, x_N)$  can be modeled as

$$p(X|\Pi, \Theta) = \prod_{i=1}^N f(x_i|\Pi, \Theta) = \prod_{i=1}^N \sum_{j=1}^K \pi_{ij} \Phi(x_i|\Theta_j). \quad (4)$$

Since the observation  $x_i$  is considered to be independent given the pixel label, the spatial correlation between the neighboring pixels is not taken into account. As a result, the segmented image is sensitive to noise and illumination [29]. To overcome this problem, MRF distribution [11] is applied to incorporate the spatial correlation among label values

$$p(\Pi) = Z^{-1} \exp \left\{ -\frac{1}{T} U(\Pi) \right\} \quad (5)$$

where  $Z$  is a normalizing constant,  $T$  is a temperature constant, and  $U(\Pi)$  is the smoothing prior. The posterior probability density function given by Bayes' rules can be written as

$$p(\Pi, \Theta|X) \propto p(X|\Pi, \Theta)p(\Pi). \quad (6)$$

By incorporating (6), the log-likelihood function can be derived as

$$\begin{aligned} L(\Pi, \Theta|X) &= \log(p(\Pi, \Theta|X)) \\ &= \sum_{i=1}^N \log \left\{ \sum_{j=1}^K \pi_{ij} \Phi(x_i|\Theta_j) \right\} + \log p(\Pi) \\ &= \sum_{i=1}^N \log \left\{ \sum_{j=1}^K \pi_{ij} \Phi(x_i|\Theta_j) \right\} - \log Z - \frac{1}{T} U(\Pi). \end{aligned} \quad (7)$$

Depending on the type of energy  $U(\Pi)$  selected in (7), we can have different kinds of models. In the Bayesian autologistic model [11], [35], the function  $U(\Pi)$  is chosen to incorporate the spatial correlation as

$$U(\Pi) = \sum_{i=1}^N \sum_{j=1}^K \alpha_{ij} \pi_{ij} + \sum_{i=1}^N \sum_{j=1}^K \sum_{m \in \partial_i} \beta_{ijm} \pi_{ij} \pi_{mj} \quad (8)$$

where  $\alpha_{ij}$  and  $\beta_{ijm}$  form the parameter set. In this model, in order to maximize the log-likelihood function, we need to optimize many parameters. Another disadvantage is that its segmentation is not sufficiently robust to noise.

Other mixture models based on MRF have been successfully applied to image segmentation [29]–[31] and different ways are adopted to select the energy  $U(\Pi)$ . In [29],  $Z$  and  $T$  in (7) are set to one ( $Z=1$ ,  $T=1$ ), and  $U(\Pi)$  is given by

$$U(\Pi) = \beta \sum_{i=1}^N \sum_{m \in \partial_i} \sum_{j=1}^K (\pi_{ij} - \pi_{mj})^2 \quad (9)$$

while in another MRF model-based method [30],  $U(\Pi)$  is given by

$$U(\Pi) = \beta \sum_{i=1}^N \sum_{m \in \partial_i} \left[ 1 + \left( \sum_{j=1}^K (\pi_{ij} - \pi_{mj})^2 \right)^{-1} \right]^{-1} \quad (10)$$

where  $\beta$  in (9) and (10) represents a constant value.

In [31], spatial information is taken into account and  $U(\Pi)$  is given as

$$U(\Pi) = \sum_{i=1}^N \sum_{j=1}^K \sum_{s=1}^S \left[ \frac{1}{2} \log \beta_{js}^2 - \frac{1}{2} \frac{\left( \sum_{m \in \partial_i} (\pi_{ij} - \pi_{mj}) \right)^2}{\beta_{js}^2} \right] \quad (11)$$

where  $S$  is the total number of the considered directions. In the general case,  $S$  is equal to four ( $S=4$ : horizontal, vertical, and two diagonal directions).  $\beta_{js}$  in (11) is a variable parameter.

As shown in (9)–(11), the incorporation of local information adds complexity. In order to maximize the likelihood in (7) with respect to the parameters  $\Pi$  and  $\Theta$ , an iterative EM algorithm can be applied. However, due to the complexity of the log-likelihood function, the M-step of EM algorithm cannot be applied directly to the prior distribution  $\pi_{ij}$ . Note that the prior distribution  $\pi_{ij}$  should satisfy the constraints in (2). Thus, the resulting algorithms are computationally complex and utilize large amounts of computational power to solve the constrained optimization problem of the prior distribution  $\pi_{ij}$ . For details regarding the maximization of the log-likelihood function in (7), we refer the readers to [29]–[31].

### III. PROPOSED METHOD

Various mixture models differ based on the way they derive the strength of the smoothing prior  $U(\Pi)$ . In [11], given in (8), the smoothing prior  $U(\Pi)$  has a simple form, thus, it is easy to optimize the parameter set  $\{\Pi, \Theta\}$  to maximize the log-likelihood function. However, one of its main drawbacks is that the segmentation result is not robust to noise. Models in [29]–[31], represented by (9)–(11), make use of a complex smoothing prior. Their primary disadvantage lies in its additional training complexity. The M-step of the EM algorithm cannot be applied directly to the prior distribution, which, therefore, corresponds to an increase in the algorithms complexity. In order to overcome these disadvantages, we introduce a novel factor  $G_{ij}$  defined as

$$G_{ij}^{(t)} = \exp \left[ \frac{\beta}{2N_i} \sum_{m \in \partial_i} (z_{mj}^{(t)} + \pi_{mj}^{(t)}) \right] \quad (12)$$

where  $z_{mj}$  is the posterior probability and  $\beta$  is the temperature value that controls the smoothing prior. In this paper, it has been set to 12 ( $\beta=12$ ).  $\partial_i$  is the neighborhood of the  $i$ th pixel, including itself. A square window of size  $5 \times 5$  is used in this paper.  $N_i$  is the number of neighboring pixels around the pixel  $x_i$  in this window ( $N_i=25$ ). By taking a closer look at (12), it can be visualized that the factor  $G_{ij}$  is defined as a multiplication of both posterior probabilities and prior distributions. Based on a fact that neighboring pixels in an image are similar in some sense, we can use this kind of relationship by replacing each posterior probability  $z_{ij}$  and posterior probability  $\pi_{ij}$  in an image with the average value of their neighbors, including themselves. Note that the factor  $G_{ij}$  is only dependent on the value of the priors and posteriors

at the previous step (at the  $t$  step). It plays a role as a linear filter for smoothing and restoring images corrupted by noise. For this reason, the main advantage of  $G_{ij}$  is the ease of implementation, and incorporation of the spatial relationships among neighboring pixels in a simpler metric.

Next, we propose a novel approach to incorporate the spatial information into the smoothing prior. The new smoothing prior  $U(\Pi)$  is given by

$$U(\Pi) = - \sum_{i=1}^N \sum_{j=1}^K G_{ij}^{(t)} \log \pi_{ij}^{(t+1)} \quad (13)$$

where  $t$  indicates the iteration step. The intuition of (13) that the derivative of the smoothing prior  $U(\Pi)$  with respect to prior distribution  $\pi_{ij}$  at the current step (at the  $t+1$  step) is only dependent on the term  $\pi_{ij}^{(t+1)}$ . For this reason, the M-step of the EM algorithm in our method is simple and computationally efficient. The MRF distribution  $p(\Pi)$  in (5) is given by

$$p(\Pi) = Z^{-1} \exp \left\{ \frac{1}{T} \sum_{i=1}^N \sum_{j=1}^K G_{ij}^{(t)} \log \pi_{ij}^{(t+1)} \right\}. \quad (14)$$

Given the MRF distribution  $p(\Pi)$ , the log-likelihood function in (7) is written in the form

$$\begin{aligned} L(\Pi, \Theta | X) = & \sum_{i=1}^N \log \left\{ \sum_{j=1}^K \pi_{ij}^{(t+1)} \Phi(x_i | \Theta_j^{(t+1)}) \right\} \\ & - \log Z + \frac{1}{T} \sum_{i=1}^N \sum_{j=1}^K G_{ij}^{(t)} \log \pi_{ij}^{(t+1)}. \end{aligned} \quad (15)$$

Application of the complete data condition in [29], maximizing the log-likelihood function  $L(\Pi, \Theta | X)$  in (15) will lead to an increase in the value of the objective function  $J(\Pi, \Theta | X)$

$$\begin{aligned} J(\Pi, \Theta | X) = & \sum_{i=1}^N \sum_{j=1}^K z_{ij}^{(t)} \left\{ \log \pi_{ij}^{(t+1)} + \log \Phi(x_i | \Theta_j^{(t+1)}) \right\} \\ & - \log Z + \frac{1}{T} \sum_{i=1}^N \sum_{j=1}^K G_{ij}^{(t)} \log \pi_{ij}^{(t+1)}. \end{aligned} \quad (16)$$

The conditional expectation values  $z_{ij}$  of the hidden variables can be computed as follows:

$$z_{ij}^{(t)} = \frac{\pi_{ij}^{(t)} \Phi(x_i | \Theta_j^{(t)})}{\sum_{k=1}^K \pi_{ik}^{(t)} \Phi(x_i | \Theta_k^{(t)})}. \quad (17)$$

The next objective is to optimize the parameter set  $\{\Pi, \Theta\}$  in order to maximize the objective function  $J(\Pi, \Theta | X)$  in (16). Similar to the MRF-based methods [29]–[31],  $Z$  and  $T$  in (16) are set proportional to one ( $Z=1$ ,  $T=1$ ). From (16), the new objective function is given by

$$\begin{aligned} J(\Pi, \Theta | X) = & \sum_{i=1}^N \sum_{j=1}^K z_{ij}^{(t)} \left\{ \log \pi_{ij}^{(t+1)} + \log \Phi(x_i | \Theta_j^{(t+1)}) \right\} \\ & + \sum_{i=1}^N \sum_{j=1}^K G_{ij}^{(t)} \log \pi_{ij}^{(t+1)}. \end{aligned} \quad (18)$$

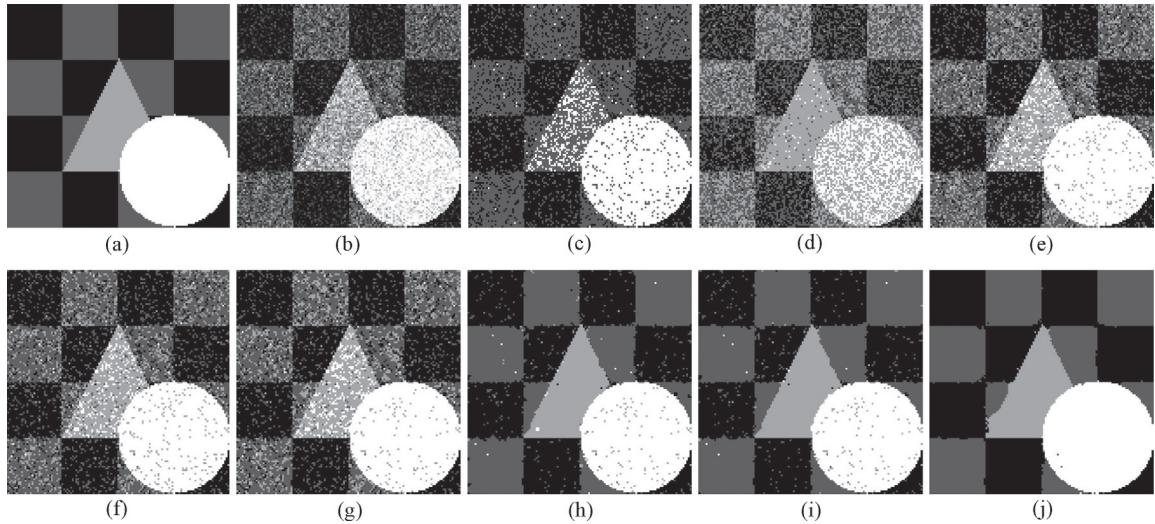


Fig. 1. First experiment (128 × 128 image resolution). (a) Original image. (b) Corrupted original image with Gaussian noise (0 mean, 0.03 variance). (c)  $K$ -means (MCR = 28.74%). (d) Standard GMM (MCR = 41.67%). (e) SVFMM (MCR = 23.28%). (f) CA-SVFMM (MCR = 20.29%). (g) ICM (MCR = 20.23%). (h) SIMF (MCR = 3.83%). (i) MEANF (MCR = 3.55%). (j) Proposed method (MCR = 1.13%).

TABLE I

COMPUTATIONAL COST (IN SECONDS) COMPARISON FOR THE SYNTHETIC IMAGE IN THE FIRST EXPERIMENT

Methods	Standard GMM	SVFMM	CA-SVFMM	ICM	SIMF	MEANF	Proposed Method
Time (s)	3.2	80.3	218.1	85.4	163.2	121.8	4.9
PSNR (dB)	14.15	16.89	17.21	16.60	24.41	24.58	27.99
RMSE	0.20	0.14	0.13	0.15	0.08	0.08	0.04
MCR (%)	41.67	23.28	20.29	20.23	3.83	3.55	1.13

From (3), the function in (18) can be rewritten as

$$\begin{aligned} J(\Pi, \Theta | X) = & \sum_{i=1}^N \sum_{j=1}^K z_{ij}^{(t)} \left\{ \log \pi_{ij}^{(t+1)} - \frac{D}{2} \log(2\pi) - \frac{1}{2} \log |\Sigma_j^{(t+1)}| \right\} \\ & + \sum_{i=1}^N \sum_{j=1}^K z_{ij}^{(t)} \left\{ -\frac{1}{2} (x_i - \mu_j^{(t+1)})^T \Sigma_j^{-1(t+1)} (x_i - \mu_j^{(t+1)}) \right\} \\ & + \sum_{i=1}^N \sum_{j=1}^K G_{ij}^{(t)} \log \pi_{ij}^{(t+1)}. \end{aligned} \quad (19)$$

To maximize this function, the EM algorithm [29], [30], [33], [34] is applied. Let us now consider the derivation of the function  $J(\Pi, \Theta | X)$  with the means  $\mu_j$  at the  $(t+1)$  iteration step. We have

$$\frac{\partial J}{\partial \mu_j^{(t+1)}} = \sum_{i=1}^N z_{ij}^{(t)} \left[ -\frac{1}{2} (2 \Sigma_j^{-1(t+1)} \mu_j^{(t+1)} - 2 \Sigma_j^{-1(t+1)} x_i) \right]. \quad (20)$$

The solution of  $\partial J / \partial \mu_j = 0$  yields the minimizer of  $\mu_j$  at the  $(t+1)$  step

$$\mu_j^{(t+1)} = \frac{\sum_{i=1}^N z_{ij}^{(t)} x_i}{\sum_{i=1}^N z_{ij}^{(t)}}. \quad (21)$$

Thus, setting the derivative of the function in (18) with respect

to  $\Sigma_j^{-1}$  at the  $(t+1)$  iteration step, we have

$$\frac{\partial J}{\partial \Sigma_j^{-1(t+1)}} = \sum_{i=1}^N z_{ij}^{(t)} \left[ \frac{1}{2} \Sigma_j^{(t+1)} - \frac{1}{2} (x_i - \mu_j^{(t+1)}) (x_i - \mu_j^{(t+1)})^T \right] \quad (22)$$

and equating it to zero yields

$$\Sigma_j^{(t+1)} = \frac{\sum_{i=1}^N z_{ij}^{(t)} (x_i - \mu_j^{(t+1)}) (x_i - \mu_j^{(t+1)})^T}{\sum_{i=1}^N z_{ij}^{(t)}}. \quad (23)$$

An important consideration is that the prior distribution should satisfy the constraints in (2). In order to enforce these constraints, we use the Lagrange's multiplier  $\eta_i$  for each data point

$$\frac{\partial}{\partial \pi_{ij}^{(t+1)}} \left[ J - \sum_{i=1}^N \eta_i \left( \sum_{j=1}^K \pi_{ij}^{(t+1)} - 1 \right) \right] = 0. \quad (24)$$

Equation (24) can be rewritten in the following form:

$$\frac{z_{ij}^{(t)}}{\pi_{ij}^{(t+1)}} + \frac{G_{ij}^{(t)}}{\pi_{ij}^{(t+1)}} - \eta_i = 0. \quad (25)$$

The constraint  $\sum_{j=1}^K \pi_{ij} = 1$  enables the Lagrange multiplier  $\eta_i$  to satisfy the following condition:

$$\eta_i = 1 + \sum_{j=1}^K G_{ij}^{(t)}. \quad (26)$$

TABLE II  
COMPUTATIONAL COST (IN SECONDS) COMPARISON FOR THE SYNTHETIC IMAGE IN THE SECOND EXPERIMENT

Methods	Standard GMM	SVFMM	CA-SVFMM	ICM	SIMF	MEANF	Proposed Method
Time (s)	1.2	117.1	147.1	78.3	108.2	97.7	3.1
PSNR (dB)	10.45	12.93	12.72	16.46	17.13	17.08	23.68
RMSE	0.31	0.23	0.23	0.16	0.13	0.14	0.06
MCR (%)	35.02	12.01	11.16	7.65	5.65	5.94	1.05

TABLE III  
COMPARISON OF THE PROPOSED METHOD WITH OTHER METHODS, FOR THE THIRD EXPERIMENT, IN THE PRESENCE OF VARYING LEVELS OF NOISE

Methods	Gaussian Noise (0 Mean, Var)			Mixed Noise: Salt & Pepper Noise (sp) + Gaussian Noise (0 Mean, Var)		
	Var = 0.01	Var = 0.03	Var = 0.05	sp = 0.005 Var = 0.01	sp = 0.01 Var = 0.03	sp = 0.02 Var = 0.05
Standard GMM	27.06%	30.31%	37.20%	29.74%	31.23%	38.22%
SVFMM	4.85%	18.11%	25.56%	5.13%	18.88%	28.40%
CA-SVFMM	4.80%	17.73%	25.49%	5.02%	18.25%	28.58%
ICM	0.62%	5.90%	27.32%	0.93%	11.65%	29.88%
SIMF	0.56%	2.86%	19.35%	0.68%	8.16%	22.03%
MEANF	0.59%	2.70%	17.65%	0.64%	7.14%	21.94%
Proposed method	0.12%	0.21%	0.36%	0.29%	0.35%	0.42%

The necessary condition for determining the prior distribution  $\pi_{ij}$  at the  $(t+1)$  iteration step becomes

$$\pi_{ij}^{(t+1)} = \frac{z_{ij}^{(t)} + G_{ij}^{(t)}}{\sum_{k=1}^K (z_{ik}^{(t)} + G_{ik}^{(t)})}. \quad (27)$$

So far, the discussion has focused on estimating  $\{\Pi, \Theta\}$  of the model in order to assign a label  $\Omega_j$  to the pixel  $x_i$ . The various steps of the proposed mixture model incorporating spatial information based on MRF can be summarized as follows.

Step 1) *Initialize the parameters  $\{\Pi, \Theta\}$ :* the means  $\mu_j$ , covariance values  $\Sigma_j$ , and prior distributions  $\pi_{ij}$ .

Step 2) *E step.*

- a) Evaluate the values  $z_{ij}$  in (17) using the current parameter values.
- b) Update the factor  $G_{ij}$  by using (12).

Step 3) *M step:* Reestimate the parameters  $\{\Pi, \Theta\}$ .

- a) Update the means  $\mu_j$  by using (21).
- b) Update covariance values  $\Sigma_j$  by using (23).
- c) Update prior distributions  $\pi_{ij}$  by using (27).

Step 4) Evaluate the log-likelihood in (15) and check the convergence of either the log-likelihood function or the parameter values. If the convergence criterion is not satisfied, then go to step 2.

Once the parameter-learning phase is complete, every pixel  $x_i$  is assigned to the label with the largest posterior probability  $z_{ij}$

$$x_i \in \Omega_j : \text{IF } z_{ij} \geq z_{ik}; \quad j, k = (1, 2, \dots, K). \quad (28)$$

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

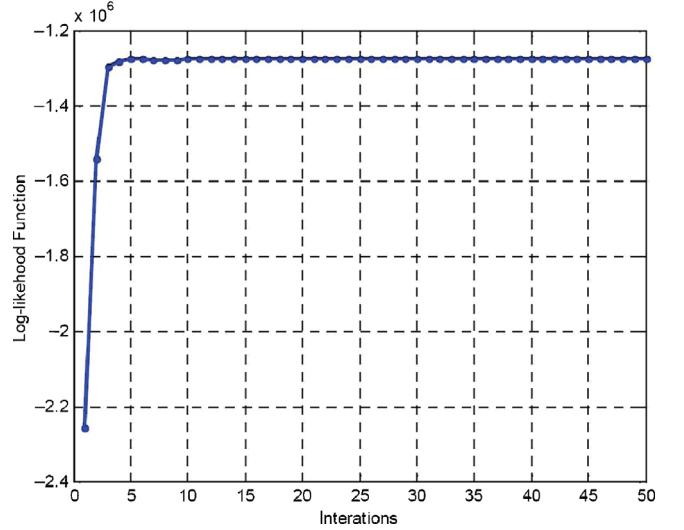


Fig. 2. Maximization progress of the log-likelihood of the proposed method of the first experiment.

#### IV. EXPERIMENTS

In this section, the performance of the proposed algorithm is compared with algorithms based on the  $K$ -means algorithm [6], standard GMM [8], and mixture models based on MRF, such as iterated conditional modes (ICM) [37], simulated field algorithm (SIMF), and mean field algorithm (MEANF) [22], [24]. We also compare the results with the spatially variant finite mixture model (SVFMM) [30] and the class-adaptive spatially variant finite mixture model (CA-SVFMM) [31]. For ICM, SIMF, and MEANF methods, a standard isotropic Potts model with a second-order (8-neighbor) neighborhood system is utilized, where the temperature value  $\beta$  is automatically optimized. Parameter  $\beta$  in SVFMM algorithms is manually

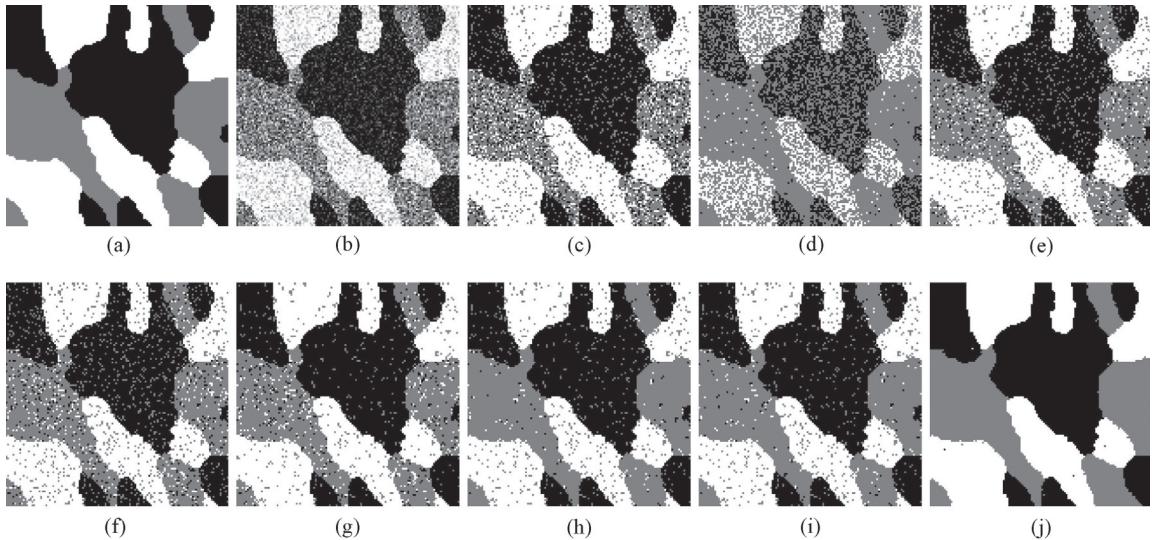


Fig. 3. Second experiment (128 × 128 image resolution). (a) Original image. (b) Corrupted original image with Gaussian noise (0 mean, 0.05 variance). (c)  $K$ -means (MCR=17.23%). (d) Standard GMM (MCR=35.02%). (e) SVFMM (MCR=12.01%). (f) CA–SVFMM (MCR=11.16%). (g) ICM (MCR=7.65%). (h) SIMF (MCR=5.65%). (i) MEANF (MCR=5.94%). (j) Proposed method (MCR=1.05%).

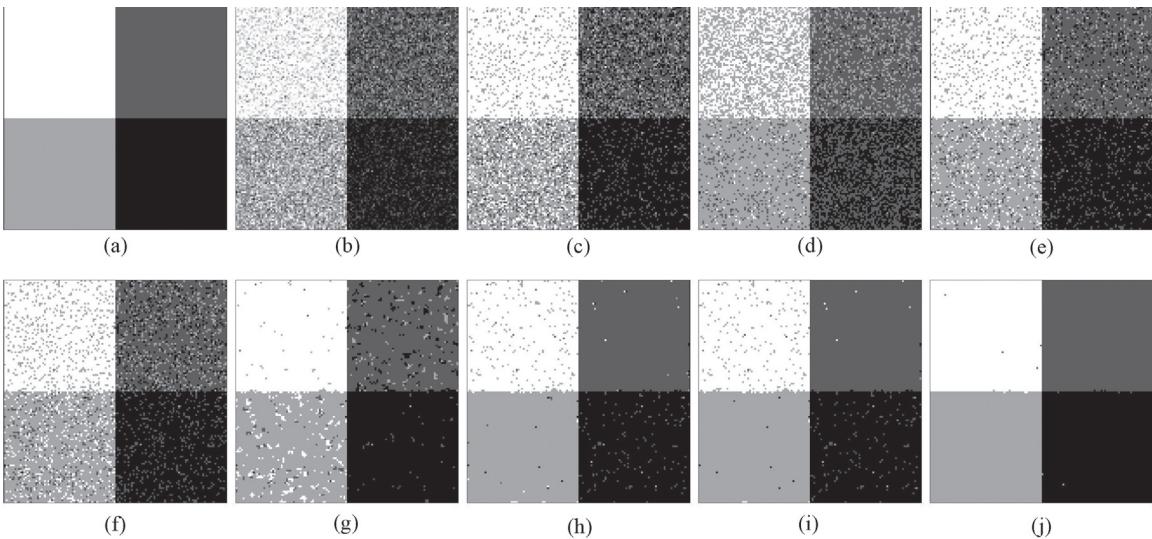


Fig. 4. Third experiment (128 × 128 image resolution). (a) Original image. (b) Corrupted original image with Gaussian noise (0 mean, 0.03 variance). (c)  $K$ -means (MCR = 22.51%). (d) Standard GMM (MCR = 30.31%). (e) SVFMM (MCR = 18.11%). (f) CA–SVFMM (MCR = 17.73%). (g) ICM (MCR = 5.90%). (h) SIMF (MCR = 2.86%). (i) MEANF (MCR = 2.70%). (j) Proposed method (MCR = 0.21%).

set a value of 0.1 based on the user experience. SVFMM and CA–SVFMM methods were implemented in the MATLAB environment. All these methods were run until convergence. The proposed method was implemented and tested on both synthetic and real-world images. For synthetic images in Section IV-A, in order to compare the results obtained, misclassification ratio (MCR) [38], peak signal-to-noise ratio (PSNR) [39], and average root mean squared error (RMSE) have been used. For real images in Sections IV-B and IV-C, in order to evaluate the segmentation performance quantitatively, variation of information (VI) [40], global consistency error (GCE) [41], and normalized probabilistic rand (NPR) index [42] are employed. Note that, for MCR, RMSE, VI, and GCE, the lower the value, the better the quality of the segmentation, while the higher values of PSNR, NPR indicate better segmentation results. All these methods

were implemented and tested on a personal computer (PC) (Pentium 4, running at 3 GHz with 1 GB RAM).

#### A. Segmentation of Synthetic Images

In the first experiment, a synthetic image (128 × 128 image resolution), as shown in Fig. 1(a), was used to compare the performance of the proposed algorithm with others. The image has four labels with luminance values [0, 1/3, 2/3, 1]. The image shown in Fig. 1(b) was obtained by corrupting the original image with Gaussian noise (0 mean, 0.03 variance). Fig. 1(c) presents the segmentation result obtained by employing the  $K$ -means algorithm, which is used during the initialization step for all the remaining methods. In Fig. 1(d)–(i), we present the segmentation results obtained by employing standard GMM, SVFMM, CA–SVFMM, ICM, SIMF, and MEANF methods, respectively. Under given

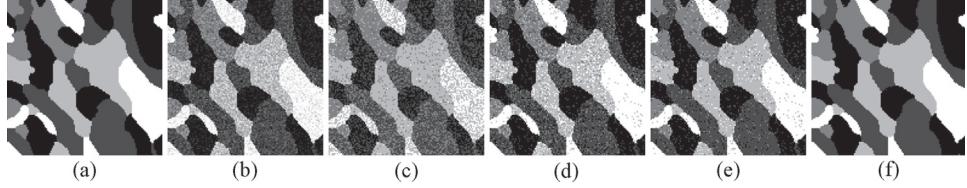


Fig. 5. Fourth experiment ( $128 \times 128$  image resolution). (a) Original image. (b) Corrupted original image with Gaussian noise (0 mean, 0.01 variance). (c) Standard GMM (MCR = 33.92%). (d) Auto-logistic (MCR = 17.89%). (e) SVFMM (MCR = 10.13%). (f) Proposed method (MCR = 0.73%).

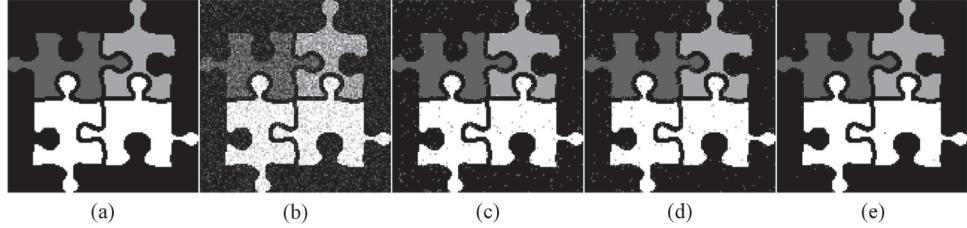


Fig. 6. Fifth experiment ( $128 \times 128$  image resolution). (a) Original image. (b) Corrupted original image with Gaussian noise (0 mean, 0.02 variance). (c) SIMF with 8-neighbor (MCR = 2.53%, time = 81.3 s). (d) SIMF with 24-neighbor (MCR = 2.19%, time = 109.6 s). (e) Proposed method (MCR = 0.35%, time = 0.81 s).

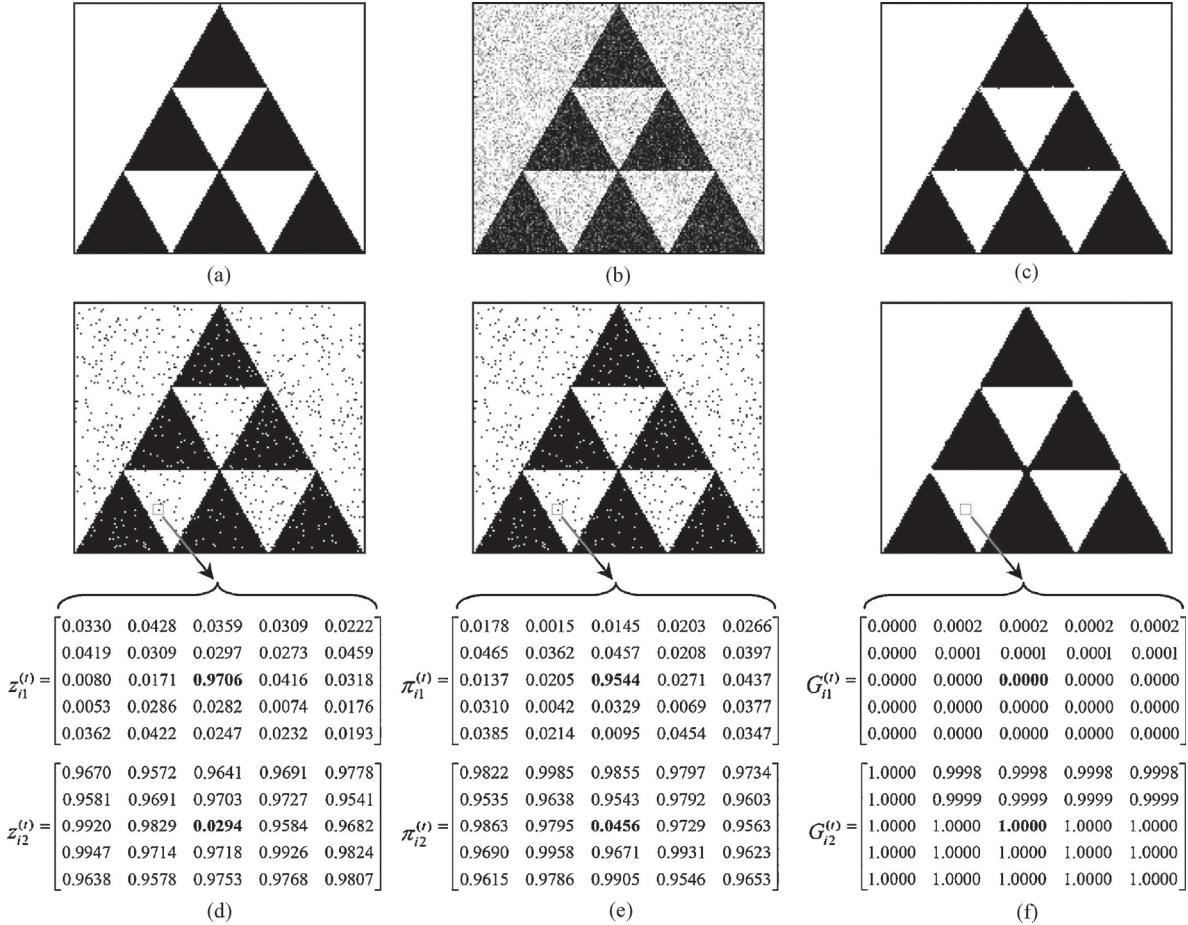


Fig. 7. Physical intuition of the factor  $G_{ij}^{(t)}$ . (a) Original image. (b) Corrupted original image with Gaussian noise (0 mean, 0.07 variance). (c) Result of our method. (d) Posterior probability  $z_{ij}^{(t)}$  at the first iteration. (e) Prior distribution  $\pi_{ij}^{(t)}$  at the first iteration. (f) Factor  $G_{ij}^{(t)}$  at the first iteration.

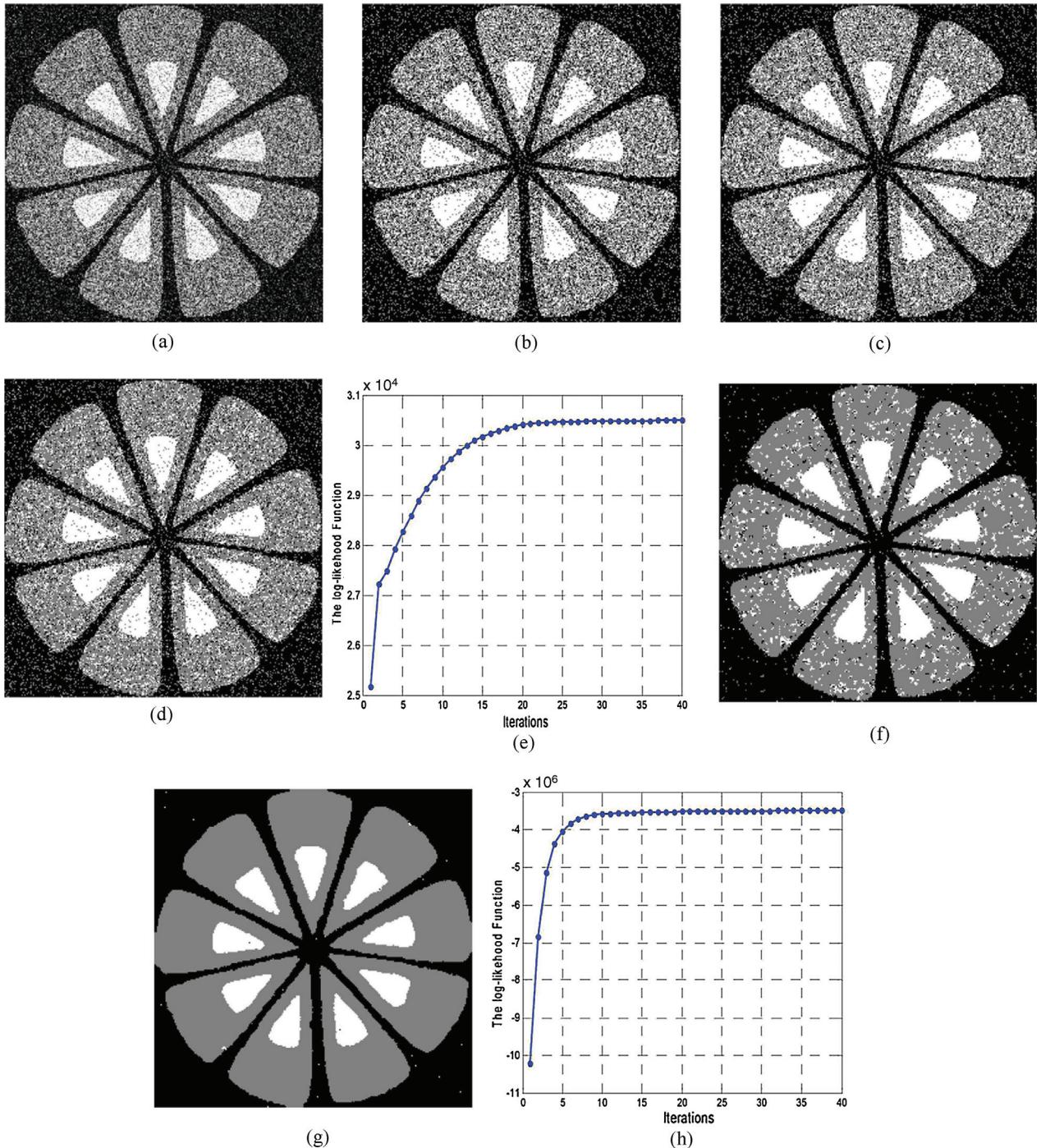


Fig. 8. Segmentation of 256 × 256 synthetic image. (a) Noisy image. (b) Initialization. (c) SVFMM at the second iteration (MCR = 27.80%). (d) SVFMM at the 15th iteration (MCR = 22.15%). (e) Log-likelihood function of SVFMM (time = 465.51 s). (f) Our method at the second iteration (MCR = 10.31%). (g) Our method at the 15th iteration (MCR = 1.25%). (h) Log-likelihood function of our method (time = 10.31 s).

conditions, the segmentation accuracy of the standard GMM method is quite poor. SIMF in Fig. 1(h) and MEANF in Fig. 1(i) demonstrate better classification and segment the image well. However, the proposed method in Fig. 1(j) segments the image better with the lowest MCR. Moreover, due to the inherent simplicity of the proposed algorithm, it has low computational cost. In this experiment, for ICM, SIMF, and MEANF methods, we used the software implementation (for Linux environment, C++ Language) publicly available at <http://mistis.inrialpes.fr/software/SEMMS.html>. Standard

GMM, SVFMM, CA-SVFMM, and the proposed method were implemented using MATLAB in the Windows environment. All experiments were performed on a PC (Pentium 4, running at 3 GHz with 1 GB of RAM) until convergence. Table I lists the computation time and the MCR for each of the aforementioned methods. As shown in Table I, standard GMM takes the least amount of time for segmentation, while the slowest one is CA-SVFMM. Although the proposed method comes second (convergence after 4.9 s and 50 iterations, as shown in Fig. 2) in terms of

the speed, it has the lowest MCR and demonstrates a higher degree of robustness with respect to noise.

In Fig. 3, we show the segmentation results of a synthetic image corrupted by Gaussian noise. The original image [38] with luminance values [0, 0.5, 1] is shown in Fig. 3(a). The image in Fig. 3(b) is obtained by adding Gaussian noise (0 mean, 0.05 variance) to the original image. Fig. 3(d)–(i) shows the segmentation results for standard GMM, SVFMM, CA-SVFMM, ICM, SIMF, MEANF, and the proposed method, respectively. Among these methods, the proposed method demonstrates a higher degree of robustness to a specified noise level. All algorithms were initialized using the  $K$ -means algorithm, as shown in Fig. 3(c). Results in Table II confirm that the proposed method is quite fast with the highest accuracy.

To test the details of the results obtained by different methods on varying noise level, a synthetic image with luminance values [0, 1/3, 2/3, 1] from Fig. 4(a) is used. This original image is corrupted with varying levels of noise. The goal is to segment the corrupted image into four labels. The results are presented in Table III. ICM method works well when an image is corrupted by low level of noise; however, the segmentation results is poor when the noise level is increased. The effect of noise on the performance of SIMF and MEANF methods is much less when we compared to ICM. However, comparing to the proposed method, both SIMF and MEANF lose most of the sharpness and details in the segmented image, as shown in Fig. 4. We also compared the performance of the proposed algorithm with others when the image is corrupted with varying levels of mixed noise. First, the original image is corrupted with salt & pepper noise, and later Gaussian noise is added. As shown in Table III, the proposed method demonstrates a higher degree of robustness with respect to the given noise level.

A image with five labels ( $K = 5$ ) with luminance values [0, 1/4, 2/4, 3/4, 1], as shown in Fig. 5(a), was used in the fourth experiment. The image shown in Fig. 5(b) is obtained by corrupting the original image with Gaussian noise (0 mean, 0.01 variance). The results for the GMM, auto-logistic [11], [35], SVFMM, and the proposed method are shown in Fig. 5(c)–(f), respectively. In this paper, parameter  $\alpha_{ij}$  and  $\beta_{ijm}$  in auto-logistic algorithms are assigned a value of 1 ( $\alpha_{ij} = 1$ ,  $\beta_{ijm} = 1$ ). As compared with SVFMM and the proposed method, the segmentation result obtained by employing the auto-logistic, as shown in Fig. 5(d), demonstrates low degree of robustness with respect to the given level of noise (MCR = 17.89%). As can be seen, the proposed method has lowest MCR (MCR = 0.73%) compared with the other methods.

Fig. 6(a) shows the original image used in the next experiment. This original image is corrupted with Gaussian noise (0 mean, 0.02 variance). The objective is to segment the noisy image in Fig. 6(b) into four labels. All methods are initialized with the same initial condition. Fig. 6(c) and (d) shows the segmentation results of SIMF method with 8-neighbor and 24-neighbor, respectively. In this example, SIMF method with 8-neighbor takes 81.3 s to segment this image with a value of MCR = 2.53%. Compared to SIMF method with 8-neighbor, SIMF method with 24-neighbor

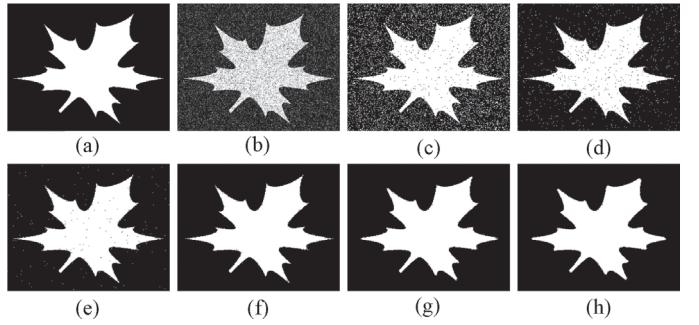


Fig. 9. Proposed method with different temperature value  $\beta$ . (a) Original image. (b) Noisy image. (c)  $\beta = 2$  (MCR = 10.15%). (d)  $\beta = 6$  (MCR = 1.53%). (e)  $\beta = 10$  (MCR = 0.38%). (f)  $\beta = 12$  (MCR = 0.17%). (g)  $\beta = 14$  (MCR = 0.30%). (h)  $\beta = 18$  (MCR = 0.35%).

obtained a slightly better segmentation result (MCR = 2.19%). However, it has high computational cost (109.6 s). As shown in Fig. 6(e), the proposed method takes only 0.81 s and performs very well in segmenting this image with the lowest MCR.

In the next experiment, we will look inside the factor  $G_{ij}^{(t)}$  to see how it works. One image with two labels as shown in Fig. 7(a) is used in this experiment. The original image is corrupted with Gaussian noise (0 mean, 0.07 variance). The objective is to segment the noisy image in Fig. 7(b) into two labels. Fig. 7(c) shows the segmentation result of the proposed method. As can be seen, the segmentation accuracy of our method is very high. In Fig. 7(d) and (e), we show the posterior probability  $z_{ij}^{(t)}$  and the prior distribution  $\pi_{ij}^{(t)}$  at the initial condition (the first iteration iteration,  $t = 1$ ), respectively. As seen from the red box in Fig. 7(d) and (e), the central pixel is corrupted by noise, while its neighboring pixels are homogenous (not corrupted by noise). The normalized result obtained by employing the factor  $G_{ij}^{(t)}$  in (12) is shown in Fig. 7(f). As shown, the factor  $G_{ij}^{(t)}$  keeps the similar value to the central pixel and ignores the influence of the noise.

In order provide more details about the comparison in terms of implementation, in the experiment of Fig. 8, we show the log-likelihood function for segmentation by employing SVFMM and the proposed method, respectively. In this experiment, the objective is to segment the noisy image in Fig. 8(a) into three labels. All methods are initialized with the same initial condition as shown in Fig. 8(b). Fig. 8(c) and (d), and (f) and (g) shows the evolution of the segmentation as the log-likelihood function iterates toward the final segmentation result. A closer look at the log-likelihood functions in Fig. 8(e) and (h) shows that our method converted at the 15th iteration. Compared to the SVFMM method, the log-likelihood function is converted at the 25th iteration that is lower than that of the proposed method. Next, in Table IV, we will look inside each iteration of SVFMM and our method to see how it works. As shown in this table, at each iteration, the computation time in the E-step of the EM algorithm of two methods are almost the same. However, the M-step of SVFMM cannot be applied directly to the prior distribution. Therefore, it takes a long time to maximize the log-likelihood with respect to the parameters. It is worth mentioning that these two methods are performed in the MATLAB environment under the same conditions without any particular code optimization.

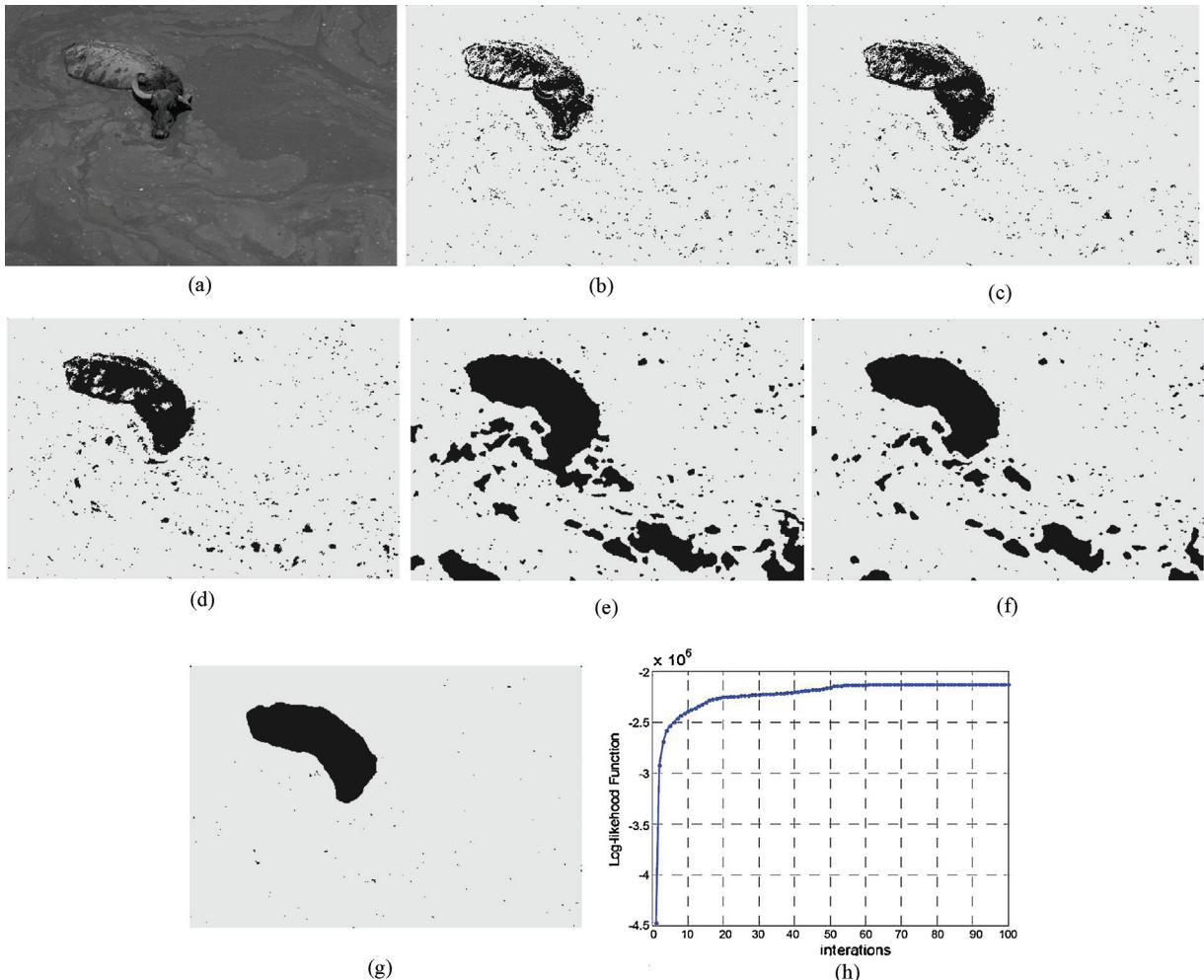


Fig. 10. Grayscale natural image segmentation (80 099). (a) Original image. (b) Standard GMM ( $VI = 0.059$ ,  $GCE = 0.642$ ,  $NPR = 0.834$ ). (c) SVFMM ( $VI = 0.058$ ,  $GCE = 0.630$ ,  $NPR = 0.841$ ). (d) CA-SVFMM ( $VI = 0.057$ ,  $GCE = 0.624$ ,  $NPR = 0.842$ ). (e) SIMF ( $VI = 0.083$ ,  $GCE = 0.783$ ,  $NPR = 0.791$ ). (f) MEANF ( $VI = 0.052$ ,  $GCE = 0.589$ ,  $NPR = 0.846$ ). (g) Proposed method ( $VI = 0.019$ ,  $GCE = 0.413$ ,  $NPR = 0.881$ ). (h) Maximization progress of the log-likelihood of the proposed method.

TABLE IV  
COMPUTATIONAL COST (IN SECONDS) COMPARISON AT EACH  
ITERATION FOR THE EXPERIMENT IN FIG. 8

Methods	At the Second Iteration		At the 15th Iteration		Total (40 Iterations)	
	E-Step	M-Step	E-Step	M-Step	E-Step	M-Step
	0.261	8.621	0.264	8.635	9.928	343.05
Our method	0.220	0.032	0.218	0.031	8.902	1.229

In the last part of this section, in order to give some implementation details about the temperature parameter  $\beta$ , an image with two labels with luminance values  $[0, 1]$ , as shown in Fig. 9(a), is used. Fig. 9(b) represents the noisy image that is obtained by corrupting the original image with the Gaussian noise (0 mean, 0.1 variance). In Fig. 9(c)–(h), we show the obtained results by the proposed methods when temperature parameter  $\beta$  is varied with 2, 6, 10, 12, 14, and 18. By taking a closer look at Fig. 9, it can be visualized that the temperature parameter  $\beta$  in the proposed method is chosen small enough to prevent the image from losing much

of its sharpness and details. In other words, the value of  $\beta$  has to be chosen large enough to tolerate the noise.

### B. Segmentation of Grayscale Natural Images

A real-world grayscale image, as shown in Fig. 10(a), from the Berkeley's image segmentation dataset [41] was used to compare the proposed algorithm with other algorithms. The objective is to segment the image into two labels: "buffalo" and "water." The main difficulty in this experiment is that the boundaries between the two labels are difficult to detect. Fig. 10(c)–(g) shows the results obtained by implementing SVFMM, CA-SVFMM, SIMF, MEANF, and the proposed method. The initialization for all these algorithms was carried out by using the standard GMM algorithm, as shown in Fig. 10(b). As can be seen, the segmentation accuracy of SVFMM, CA-SVFMM, SIMF, and MEANF methods, along the object boundaries is quite poor. In this experiment, only the proposed was able to successfully segment the image into "buffalo" and "water" regions. The maximization progress of the log-likelihood of the proposed method is shown in Fig. 10(h).

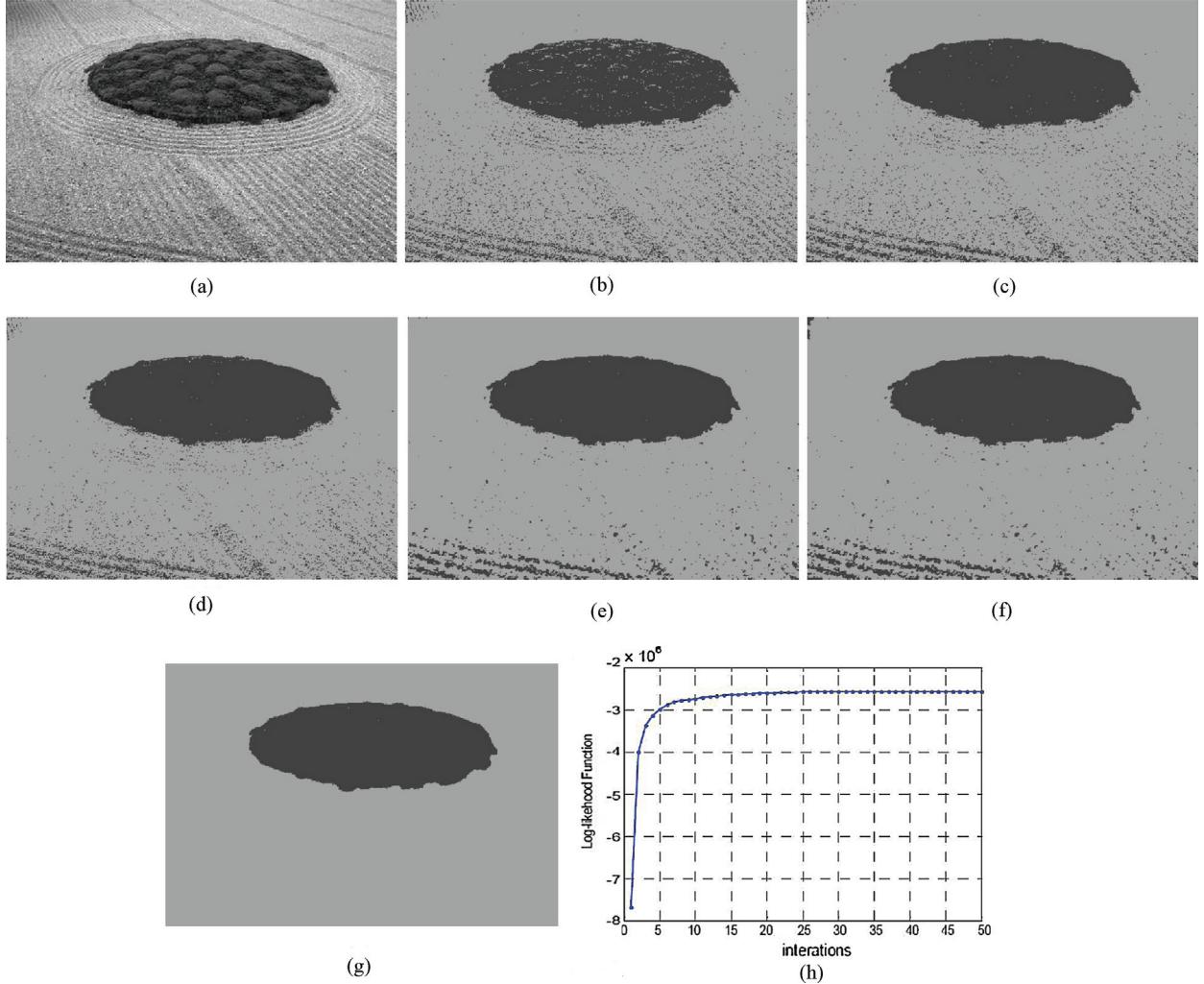


Fig. 11. Grayscale natural image segmentation (86016). (a) Original image. (b) Standard GMM (VI = 0.128, GCE = 2.717, NPR = 0.507). (c) SVFMM (VI = 0.097, GCE = 2.578, NPR = 0.515). (d) CA-SVFMM (VI = 0.087, GCE = 2.542, NPR = 0.515). (e) SIMF (VI = 0.050, GCE = 2.403, NPR = 0.507). (f) MEANF (VI = 0.066, GCE = 2.455, NPR = 0.514). (g) Proposed method (VI = 0.011, GCE = 2.222, NPR = 0.501). (h) Maximization progress of the log-likelihood of the proposed method.

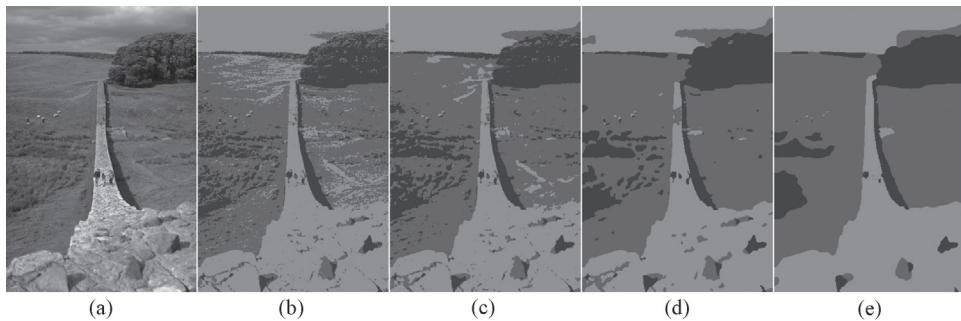


Fig. 12. Grayscale natural image segmentation (374067). (a) Original image. (b) Standard GMM (VI = 0.328, GCE = 2.731, NPR = 0.676). (c) SVFMM (VI = 0.287, GCE = 2.482, NPR = 0.706). (d) SIMF (VI = 0.247, GCE = 2.314, NPR = 0.713). (e) Proposed method (VI = 0.203, GCE = 2.169, NPR = 0.721).

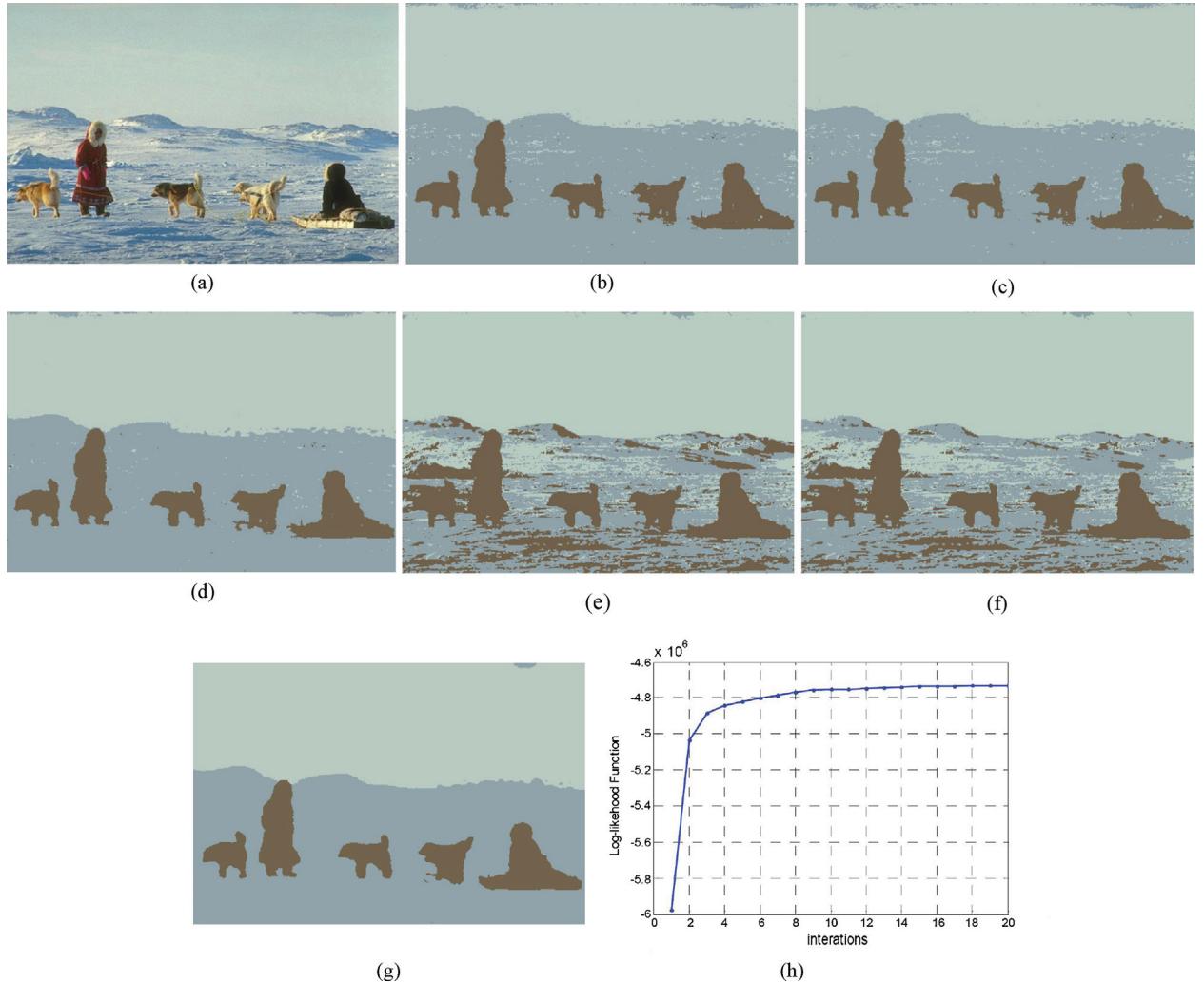


Fig. 13. Color image segmentation (310 007). (a) Original image. (b) Standard GMM ( $VI = 0.057$ ,  $GCE = 0.600$ ,  $NPR = 0.949$ ). (c) SVFMM ( $VI = 0.051$ ,  $GCE = 0.546$ ,  $NPR = 0.956$ ). (d) CA-SVFMM ( $VI = 0.495$ ,  $GCE = 0.538$ ,  $NPR = 0.959$ ). (e) SIMF ( $VI = 0.268$ ,  $GCE = 1.435$ ,  $NPR = 0.785$ ). (f) MEANF ( $VI = 0.254$ ,  $GCE = 0.138$ ,  $NPR = 0.789$ ). (g) Proposed method ( $VI = 0.034$ ,  $GCE = 0.443$ ,  $NPR = 0.970$ ). (h) Maximization progress of the log-likelihood of the proposed method.

Another real-world grayscale image is used to test the efficiency and effectiveness of our proposed algorithm. We observe that the spot (noise) in the bottom left-hand side of this image is high. The objective is to segment the original image, as shown in Fig. 11(a), into two labels. As seen from the results in Fig. 11, standard GMM, SVFMM, CA–SVFMM, SIMF, and MEANF could not successfully segment this image, and there is a high degree of misclassification. Compared with these methods, the proposed method, as shown in Fig. 11(g), successfully segments the image.

A grayscale real-world image, shown in Fig. 12(a) [41], was segmented into three labels: “road,” “grass,” and “tree.” A low segmentation accuracy is obtained with standard GMM, as shown in Fig. 12(b). This result is used as the initialization step for all the remaining methods. SVFMM and SIMF methods in Fig. 12(c) and (d) yields good segmentation results. However, the spots obtained in the segmentation remain quite high. Compared with these methods, we find that the proposed method, shown in Fig. 12(f) successfully segments all objects, and the effect of noise on the final segmented image is quite low.

### C. Segmentation of Colored Images

In Fig. 13, we show the segmentation results of a real-world color image from the Berkeley’s image segmentation dataset [31]. The original image, as shown in Fig. 13(a), was segmented into three labels: “snow,” “sky,” and “others.” The initialization for all algorithms was carried out by using the standard GMM as shown in Fig. 13(b). Extraction accuracies of the SVFMM, CA–SVFMM, SIMF, and MEANF methods are shown in Fig. 13(c)–(f). As shown in Fig. 13, the segmentation accuracy for SIMF and MEANF methods is quite poor. CA–SVFMM can produce a better segmentation, however, the edge between the “snow” and the “sky” is lost. A closer inspection of the “sky” area indicates that a small portion of pixels have been misclassified. The proposed method in Fig. 13(g) can be better classified with more detail along the sharp edge between “snow” and the “sky,” as compared with the CA–SVFMM method. Fig. 13(h) shows the maximization progress of the log-likelihood of the proposed method in this experiment.

A set of real-world color images [41] are used to evaluate the performance of the proposed method against SVFMM,

TABLE V  
COMPARISON OF IMAGE SEGMENTATION RESULTS ON BERKELEY'S COLOR IMAGE SEGMENTATION DATASET: NPR INDEX

Image	Labels	SVFMM	CA-SVFMM	ICM	SIMF	MEANF	Proposed Method
86 000	4	0.811	0.814	0.794	0.773	0.785	0.825
101 085	2	0.598	0.600	0.572	0.566	0.569	0.603
100 080	5	0.802	0.801	0.778	0.768	0.767	0.811
105 053	2	0.778	0.782	0.756	0.650	0.689	0.817
241 004	5	0.839	0.841	0.838	0.841	0.842	0.841
175 043	2	0.800	0.801	0.798	0.713	0.722	0.802
374 067	4	0.815	0.814	0.801	0.781	0.783	0.818
24 063	3	0.840	0.839	0.838	0.839	0.839	0.841
106 025	4	0.832	0.833	0.824	0.796	0.803	0.844
147 091	3	0.813	0.815	0.770	0.773	0.775	0.824
277 095	3	0.811	0.812	0.773	0.637	0.690	0.814
113 009	3	0.671	0.670	0.645	0.614	0.622	0.676
296 007	3	0.839	0.840	0.841	0.832	0.841	0.841
189 080	4	0.861	0.862	0.814	0.805	0.807	0.867
38 092	4	0.871	0.873	0.820	0.820	0.817	0.879
41 004	3	0.908	0.909	0.891	0.848	0.872	0.917
Mean	-	0.806	0.807	0.785	0.754	0.764	0.814

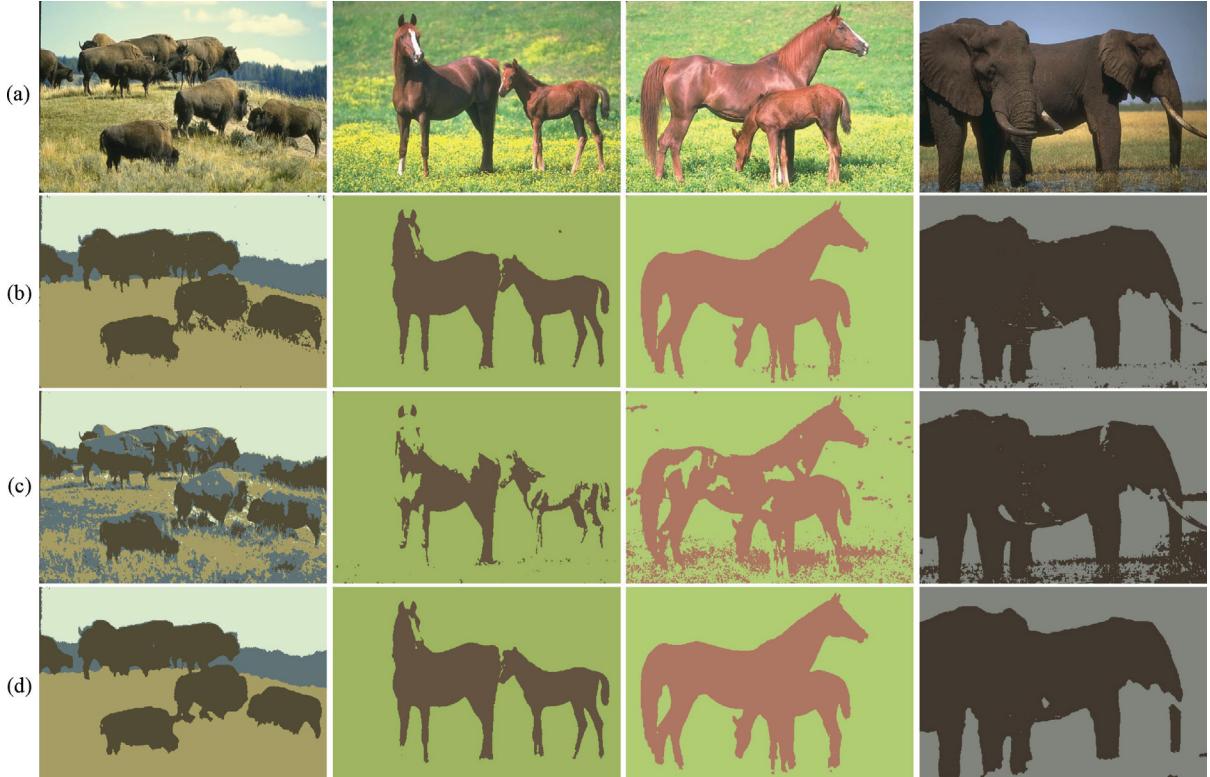


Fig. 14. Color image segmentation(38 092 113 016, 113 044, 296 059). (a) Original image. (b) SVFMM. (c) MEANF. (d) Proposed method.

CA-SVFMM, ICM, SIMF, and MEANF methods. Table V contains the cumulative results obtained for all methods, for the given set of real-world images. As evident from the results, on average, the proposed method outperforms other methods with a higher NPR.

Fig. 14 shows some of the other real-world images used for segmentation by employing SVFMM, MEANF, and the proposed method, respectively. The first row shows the original images, followed by the corresponding segmented images in the second, third, and the last row. Fig. 14 clearly indicates that our proposed method achieves a better segmentation accuracy.

## V. CONCLUSION

We presented a new mixture model for image segmentation that incorporated the spatial relationships based on MRF. Compared with other MRF-based mixture models, our proposed method directly applied the EM algorithm to optimize the parameters, making it simple, fast, and easy to implement. The proposed method was tested with many synthetic and real-world grayscale and colored images, thereby demonstrating excellent performance in noisy conditions, compared to other mixture model-based approaches.

One limitation of this work was the temperature value  $\beta$ . This parameter was chosen small enough to prevent the image

from losing much of its sharpness and details. In other words, the value of  $\beta$  was chosen large enough to tolerate the noise. In our paper, it was manually set to 12. One possible extension of this work is to use a different value of  $\beta$  throughout the image and automatically optimize this parameter. Another limitation was that the proposed method, which was based on the EM algorithm, performed only local optimization. Thus, it depended on the initial starting point. Bad initialization can lead to bad results. One possible solution to this problem is to apply the global optimization in order to estimate the model parameters. This is an open question and remains the subject of our research.

#### ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers and the Associate Editor for their insightful comments that significantly improved the quality of this paper.

#### REFERENCES

- [1] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.
- [2] P. F. Felzenswalb and D. Huttenlocher, "Efficient graph-based image segmentation," *Int. J. Comput. Vision*, vol. 59, no. 2, pp. 167–181, Sep. 2004.
- [3] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603–619, May 2002.
- [4] S. Philippe and G. Luis, "Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval," *IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 561–576, Apr. 2000.
- [5] L. Huihai, C. W. John, and G. Mohammed, "Binary partition tree for semantic object extraction and image segmentation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 3, pp. 378–383, Mar. 2007.
- [6] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
- [7] W. Cai, S. Chen, and D. Zhang, "Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation," *Patt. Recognit.*, vol. 40, no. 3, pp. 825–838, Mar. 2007.
- [8] G. McLachlan and D. Peel, *Finite Mixture Models*. New York: Wiley, 2000.
- [9] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Oxford Univ. Press, 1995.
- [10] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer, 2006.
- [11] S. Z. Li, *Markov Random Field Modeling in Image Analysis*. New York: Springer-Verlag, 2009.
- [12] M. N. Thanh and Q. M. J. Wu, "Maximum likelihood neural network based on the correlation among neighboring pixels for noisy image segmentation," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2008, pp. 3020–3023.
- [13] 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, Aug. 2010.
- [14] P. Haim, F. Joseph, and J. Ian, "A study of Gaussian mixture models of color and texture features for image classification and segmentation," *Patt. Recognit.*, vol. 39, no. 4, pp. 695–706, Apr. 2006.
- [15] C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blobworld: Image segmentation using expectation-maximization and its application to image querying," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 24, no. 8, pp. 1026–1038, Aug. 2002.
- [16] K. S. Jae, G. J. Ho, L. Gen, and K. Jaihie, "Mixture of Gaussians-based background subtraction for Bayer-pattern image sequences," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 3, pp. 365–370, Mar. 2011.
- [17] M. S. Allili, D. Ziou, N. Bouguila, and S. Boutemedjet, "Image and video segmentation by combining unsupervised generalized Gaussian mixture modeling and feature selection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 10, pp. 1373–1377, Oct. 2010.
- [18] D. M. Titterington, A. F. M. Smith, and U. E. Makov, *Statistical Analysis of Finite Mixture Distributions*. Hoboken, NJ: Wiley, 1985.
- [19] A. K. Jain, R. P. W. Duin, and J. C. Mao, "Statistical pattern recognition: A review," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 22, no. 1, pp. 4–37, Jan. 2000.
- [20] G. K. Wu and T. R. Reed, "Image sequence processing using spatiotemporal segmentation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 5, pp. 798–807, Aug. 1999.
- [21] S. W. Chee, "A block-based MAP segmentation for image compressions," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 5, pp. 592–601, Sep. 1998.
- [22] H. S. Choi, D. R. Haynor, and Y. Kim, "Partial volume tissue classification of multichannel magnetic resonance images: A mixed model," *IEEE Trans. Med. Imag.*, vol. 10, no. 3, pp. 395–407, Sep. 1991.
- [23] F. Forbes and N. Peyrard, "Hidden Markov random field model selection criteria based on mean field-like approximations," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1089–1101, Sep. 2003.
- [24] G. Celeux, F. Forbes, and N. Peyrard, "EM procedures using mean field-like approximations for Markov model-based image segmentation," *Patt. Recognit.*, vol. 36, no. 1, pp. 131–144, Jan. 2003.
- [25] J. M. Laferte, P. Perez, and F. Heitz, "Discrete Markov image modeling and inference on the quadtree," *IEEE Trans. Image Process.*, vol. 9, no. 3, pp. 390–404, Mar. 2000.
- [26] M. Robinson, M. Azimi-Sadjadi, and J. Salazar, "A temporally adaptive classifier for multispectral imagery," *IEEE Trans. Neural Netw.*, vol. 15, no. 1, pp. 159–165, Jan. 2004.
- [27] R. Fjortoft, Y. Delignon, W. Pieczynski, M. Sigelle, and F. Tupin, "Unsupervised classification of radar images using hidden Markov chains and hidden Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 3, pp. 675–686, Mar. 2003.
- [28] J. L. Marroquin, E. A. Santana, and S. Botello, "Hidden Markov measure field models for image segmentation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 25, no. 11, pp. 1380–1387, Nov. 2003.
- [29] G. S. Sanjay and 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, Jul. 1998.
- [30] 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, Mar. 2005.
- [31] C. Nikou, N. Galatsanos, and A. Likas, "A class-adaptive spatially variant mixture model for image segmentation," *IEEE Trans. Image Process.*, vol. 16, no. 4, pp. 1121–1130, Apr. 2007.
- [32] A. Diplaros, N. Vlassis, and T. Gevers, "A spatially constrained generative model and an EM algorithm for image segmentation," *IEEE Trans. Neural Netw.*, vol. 18, no. 3, pp. 798–808, May 2007.
- [33] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions* (Wiley Series in Probability and Statistics). New York, Wiley, 1997.
- [34] M. A. T. Figueiredo and A. K. Jain, "Unsupervised learning of finite mixture models," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 24, no. 3, pp. 381–396, Mar. 2002.
- [35] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. Royal Statist. Soc. Ser. B*, vol. 36, pp. 192–236, Jan. 1974.
- [36] P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via EM algorithm," *J. Royal Statist. Soc.*, vol. 39, no. 1, pp. 1–38, 1977.
- [37] J. Besag, "On the statistical analysis of dirty pictures," *J. Royal Statist. Soc. Ser. B*, vol. 48, no. 3, pp. 259–302, 1986.
- [38] 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. Imag.*, vol. 20, no. 1, pp. 45–57, Jan. 2001.
- [39] Y. Fisher, *Fractal Image Compression: Theory and Application to Digital Images*. New York: Springer-Verlag, 1995.
- [40] M. Meila, "Comparing clusterings: An axiomatic view," in *Proc. Int. Conf. Mach. Learn.*, 2005, pp. 577–584.
- [41] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. Int. Conf. Comput. Vis.*, vol. 2, 2001, pp. 416–423.

- [42] R. Unnikrishnan, C. Pantofaru, and M. Hebert, "A measure for objective evaluation of image segmentation algorithms," in *Proc. IEEE Int. Conf. Comput. Vis. Patt. Recog.*, vol. 3. Jun. 2005, pp. 34–41.



**Thanh Minh Nguyen** received the B.E. (Hons.) degree in electrical and electronic engineering from the Ho Chi Minh City University of Technology, Ho Chi Minh City, Vietnam, in 2004, the M.S. degree from the Department of Electrical Engineering, Dayeh University, Changhua, Taiwan, in 2006, and the Ph.D. degree from the Department of Electrical Engineering, University of Windsor, Windsor, ON, Canada, in 2011.

He is currently a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Windsor. His current research interests include computer vision and pattern recognition.



**Q. M. Jonathan Wu** (M'92–SM'09) received the Ph.D. degree in electrical engineering from the University of Wales, Swansea, U.K., in 1990.

He was with the National Research Council of Canada for 10 years, where he later became a Senior Research Officer and a Group Leader. He is currently a Professor with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON, Canada. He has published more than 250 peer-reviewed papers in computer vision, image processing, intelligent systems, robotics, and integrated microsystems. His current research interests include 3-D computer vision, active video object tracking and extraction, interactive multimedia, sensor analysis and fusion, and visual sensor networks.

Dr. Wu is the Tier 1 Canada Research Chair of the Automotive Sensors and Information Systems. He is an Associate Editor of the *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS PART A* and the *International Journal of Robotics and Automation*. He was a Technical Program Committee Member and an International Advisory Committee Member of many prestigious conferences.