

Grayscale image segmentation by spatially variant mixture model with student's t-distribution

Taisong Xiong · Zhang Yi · Lei Zhang

Published online: 10 January 2013
© Springer Science+Business Media New York 2013

Abstract A spatially variant finite mixture model with Student's t-distribution component function is proposed for grayscale image segmentation. This model employs a new weight function which contains the information along the different spatial directions indicating the relationship of the pixels in the neighborhood. The label probability proportions are explicitly represented as probability vectors in the model. Gradient descend method is used to update the unknown parameters. The proposed model contains fewer parameters and it is easy to be implemented compare with the Markov random field (MRF) models. Comprehensive experiments on synthetic and natural images are carried out to demonstrate that the proposed model outperforms some other related ones.

Keywords Spatially variant finite mixture model · Student's t-distribution · Image segmentation · Gradient descent

T. Xiong
School of Computer Science and Engineering,
University of Electronic Science and Technology of China,
Chengdu, 610054, People's Republic of China
e-mail: xiongtaisong@gmail.com

Z. Yi (✉) · L. Zhang
Machine Intelligence Laboratory, College of Computer Science,
Sichuan University, Chengdu, 610065, People's Republic of China
e-mail: zhangyi@scu.edu.cn

L. Zhang
e-mail: leizhang@scu.edu.cn

1 Introduction

Image segmentation is one of the most important steps for image analysis and image understanding. Its key goal is to divide an image into segments according to the features of the image, such as the intensity and texture, where each region has a reasonably homogeneous visual appearance or corresponds to objects or parts of objects [8].

In the field of image segmentation, the clustering method has become one of the hot topics over the last few years. The finite mixture model (FMM) [2, 16, 25], a well known clustering model, provides a flexible method to model many random phenomena. The FMM can be used for image segmentation because its k components can be used to denote k classes. In the FMM, the mixture density is used to realize the data points. The probabilistic clustering is derived according to the fitted posterior probabilities of component membership. However, it is difficult to compute the parameters of the FMM by maximum likelihood (ML) estimation. The FMM supposes the attribute of these pixels is generated by a specific finite mixture distribution. The component function of the FMM can be various statistical distributions. When the component function of the FMM is Gaussian distribution, it is called Standard Gaussian mixture model (GMM). One drawback of the FMM in image segmentation is that it supposes the pixels are independent and does not take into account the spatial relationship among the pixels.

In order to overcome the aforementioned drawback in the FMM, a spatially variant finite mixture model (SVFMM) was proposed by Sanjay et al. [20] for image segmentation where the spatial information was considered in SVFMM. The spatial information of the model was captured by MRF [9, 13]. Then the Expectation-Maximization (EM) [7] algorithm was used through ML to compute the parameters of the model. Subsequently, some models based on MRF have been proposed and applied to medical imaging segmentation [27–29]. In order to estimate the parameters of these models, EM algorithm, Markov Chain Monte Carlo (MCMC) and Variational Bayes (VB) [2, 12] approximation inference algorithm are utilized in these models, respectively.

However, the label probability proportions (the probabilities of each pixel belonging to some classes) in the SVFMM can not be obtained in closed form in the M-step of the EM algorithm. The label probability proportions are formulated as a constrained optimization problem. The results of the label probability proportions require a projection onto the unit simplex (non-negative and summing up to one components). A gradient projection algorithm is used in [20]. Convex quadratic programming is proposed for the SVFMM in [1] and it is shown to produce both better segmentations and values of the criterion function than [20]. One of the shortcomings of SVFMM in [20] and [1] is that the MRF prior lacks flexibility and it can not learn from the data. Nikou et al. [17] proposes a Gauss–Markov random field prior probability which is imposed on the label probability proportions. The parameters of the model can be estimated from the data adaptively. Concurrently, it utilizes the information in different spatial directions of the neighborhood. However, in the M-step, the label probability proportions still don't satisfy the constraint of probability vectors.

To preserve the region boundaries in image segmentation, two synthetical models combining the class prior information with line processes [9] are proposed in [21].

In the two models, MRF smoothness priors are imposed on the label probability proportions of a spatially varying Gaussian mixture model. In the first model, the local differences between the label probability proportions are supposed to follow Gaussian distribution, at the same time the Bernoulli distribution is imposed on the binary line processes (BLP). It is a discrete line process model and the VB algorithm is used to estimate the unknown parameters of the BLP model. In the second model, the local differences between the label probability proportions are supposed to follow the continuous univariate Student's *t*-distribution and the continuous line process (CLP) random variables follow a Gamma distribution. EM algorithm is used to estimate the parameters of the CLP model. But the results of the label probability proportions of the two models can not satisfy the constraint that the label probability proportions must be probability vectors. A new projection method based on a quadratic approximation is proposed to solve this problem. In order to avoid the reparatory computations in the M step of the EM, two models based on the GMM are proposed in [24] and [23], respectively. The label probability proportions in the two models are explicitly modelled as probability vectors. They both use the gradient method to estimate the models' parameters based on minimizing the higher bound on the data negative log-likelihood. A new approach which incorporates the spatial relationships between the neighboring pixels is proposed to model the label probability proportions in [24]. We regard the model proposed in [24] as an extension of the GMM (EGMM). Considering the information that the neighboring pixels in an image are similar in some sense, the pixel value is represented by the average of its neighborhood pixels in the weight function of the label probability proportions of the model in [23]. The model in [23] is called spatial neighborhood GMM (SNGMM). A novel model for image segmentation based on the maximum likelihood estimator is proposed in [14] which is different from the pattern of SVFMM in utilizing the spatial information. The model in [14] assumes that all the pixels values are independent and identically distributed samples of generative models. The neighborhood of a pixel is regarded as its multiple observations. The likelihood of a pixel is assumed to be a weight sum of the neighborhood's likelihood where its coefficient depends on their spatial distances to the pixel. The model in [14] based on the maximum of spatial likelihoods (MSL) is well applied to image segmentation.

The component functions of the aforementioned SVFMMs are Gaussian distribution. However, the tails of the Gaussian distribution are too short to satisfy the needs for many applications. The Student's *t*-distribution has heavier tails compared with the Gaussian distribution. It is a robust alternative to the Gaussian distribution [18, 22]. The Student's *t*-distribution has been successfully used in many applications, such as signal analysis [5], sequential data modelling setting [6] and speaker identification [4].

In this paper, combining the advantages of the SVFMM with robustness of the Student's *t*-distribution, a spatially varying Student's *t*-distribution mixture model is proposed. In the proposed model, a weight function is defined to denote the spatial relationship between the pixels, which contains the directional information of a pixel's neighborhood. The label probability proportions of the pixels are explicitly modelled as probability vectors, they automatically satisfy the conditions of the probability vectors. Therefore the reparatory computations are not needed during the process of inference. Based on the maximum of the log-likelihood function, the

gradient method is used to estimate the parameters of the proposed model. Numerical experiments have been conducted on synthetic and natural grayscale images to compare our proposed algorithm with other image segmentation algorithms. The segmentation accuracy is quantified in terms of the misclassification ratio (MCR) [29] and the probabilistic rand (PR) index [26]. The experimental results demonstrate the effectiveness and robustness of the proposed model compared with the other models.

The remainder of this paper is organized as follows. A brief review of the SVFMM and the Student’s t-distribution is given in Section 2. In Section 3, the proposed model is described in detail. Section 4 describes a number of experiments conducted to evaluate the efficiency of the proposed model. Finally, conclusions are given in Section 5.

2 Preliminary

In this section, the SVFMM is introduced firstly, then the Student’s t-distribution is reviewed briefly. In this paper x_n denotes the pixel in the n th position of an image.

2.1 Spatially variant finite mixture model

The SVFMM [1, 17, 20] gives a modification of the FMM. It uses the parameter of label probability proportion π_{nk} to denote the probability that the n th pixel belongs to the k th class. We assume that there are K classes in the model. The n th pixel’s label probability vector is denoted by π^n . The parameter vector is $\pi^n = (\pi_{n1}, \pi_{n2}, \dots, \pi_{nK})^T$, where T indicates the transpose of the vector. The variable π_{nk} satisfies the constraints

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

Let $\Pi = \{(\pi^1)^T, (\pi^2)^T, \dots, (\pi^N)^T\}$ be the set of probability vectors, and $\Theta = \{\theta_1, \theta_2, \dots, \theta_K\}$ be the set of component parameters. The density function of the n th pixel is given by

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

where $p(x_n|\theta_k)$ is a Gaussian function with parameters $\theta_k = \{\mu_k, \sigma_k^2\}$, μ_k and σ_k^2 are the mean and the variance, respectively. A data set X comprises observations $\{x_n\}$, where $n = 1, \dots, N$. The observations are supposed to be statistically independent. The X ’s probability density function (pdf) [20] is given by

$$f(X|\Pi, \Theta) = \prod_{n=1}^N f(x_n|\Pi, \Theta). \tag{1}$$

A prior density based on the Gibbs distribution for the parameter set Π is introduced in [1],

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

where Z is a normalizing constant, β is regarded as regularization parameter and the function $V_{N_n}(\Pi)$ denotes the clique potential function within the neighborhood N_n . The drawback of this prior probability is that the parameter β can not be estimated from the data in an easy way.

In order to overcome the aforementioned drawback, a Gauss–Markov random field prior probability for parameter Π is proposed in [17]

$$p(\Pi) \propto \prod_{k=1}^K \beta_k^{-N} \exp \left[-\frac{1}{2} \frac{\sum_{n=1}^N \sum_{m \in N_n} (\pi_{nk} - \pi_{km})^2}{\beta_k^2} \right],$$

where the parameter β of the function is used to capture the spatial smoothness of cluster and it can be easily estimated from the data.

When the prior density is given, a posteriori pdf is given by

$$p(\Pi|X; \Theta) \propto \prod_{n=1}^N p(\Pi) f(x_n|\Pi, \Theta).$$

Then the log-density function can be obtained as follows [1]:

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

In order to estimate the parameters $\Omega = \{\mu_k, \sigma_k^2, \pi_{nk}\}$ in (2), the EM algorithm is in general used to maximize the log of the likelihood function of the given data set. However, the result of the π_{nk} obtained in the M-step of the EM algorithm is usually not satisfied with the constraints: $0 \leq \pi_{nk} \leq 1$ and $\sum_{k=1}^K \pi_{nk} = 1$. Therefore, the computation of the algorithm becomes very complex to maximize the log of the likelihood function in (2). A reparatory computation is introduced in the M-step to enforce the label probability proportion π_{nk} to satisfy these constraints. For example, gradient projection algorithm is introduced in [20] and quadratic programming is

used in [1]. After the optimization of the parameters, the label of every pixel can be obtained according to the posterior probability $p(\theta_k|x_n)$ in (3)

$$p(\theta_k|x_n) = \frac{\pi_{nk}p(x_n|\theta_k)}{\sum_{j=1}^K \pi_{nj}p(x_n|\theta_j)}. \quad (3)$$

2.2 The student's t-distribution

A random variable x follows a univariate Student's t-distribution $x \sim St(x | \mu, \lambda, \nu)$, the univariate Student's t-distribution takes the following form [2]

$$St(x|\mu, \lambda, \nu) = \frac{\Gamma(\nu/2 + 1/2)}{\Gamma(\nu/2)} \left(\frac{\lambda}{\pi \nu} \right)^{1/2} \left[1 + \frac{\lambda(x - \mu)^2}{\nu} \right]^{-\nu/2 - 1/2}, \quad (4)$$

where μ and λ are the mean and the precision of the t-distribution, respectively. The value of λ is equal to the inverse of the variance. $\Gamma(\cdot)$ is the gamma function. The parameter ν is called the t-distribution's degrees of freedom. When $\nu = 1$ the t-distribution becomes the Cauchy distribution. The parameter ν is used to tune the tail of t-distribution. if $\nu > 1$, μ is the mean of x , As ν tends to infinity, the t-distribution tends to a Gaussian distribution with mean μ and precision λ [2]. Moreover, the Student's t-distribution can be regarded as an infinite mixture of scaled Gaussian distribution, which can be rewritten in the form [2]

$$St(x|\mu, \lambda, \nu) = \int_0^{\infty} \mathcal{G}(x|\mu, (u\lambda)^{-1}) Gam(u|\nu/2, \nu/2) du,$$

where $\mathcal{G}(x|\mu, (u\lambda)^{-1})$ denotes the Gaussian distribution, and $Gam(u|\nu/2, \nu/2)$ is the gamma distribution. Therefore, the t-distribution provides a heavy-tailed alternative to the Gaussian distribution. The t-distribution has been successfully used as a robust alternative to the Gaussian distribution in [18, 22].

3 The proposed model

In this section, we propose a spatially varying finite mixture model whose component function is the Student's t-distribution. A second-order neighborhood system (8 neighbors for each interior pixel) is adopted in the proposed model. There are four different adjacency types in the second-order neighborhood system. They are horizontal, vertical and two diagonal directions, as shown in Fig. 1. Let the parameter D stand for the number of adjacency types. Thus, the value of the parameter D is equal to 4.

Fig. 1 Second-order neighborhood system. The numbers indicate the adjacency types

3	2	4
1	X	1
4	2	3

Firstly, we define a function that represents the weight of the n th pixel belonging to the k th class

$$\xi_k(x_n) = \sum_{d=1}^D \sum_{x_i \in N_n} J_{nd}(x_n, x_i) \exp\left(-\frac{(x_i - c_{kd})^2}{2b_{kd}^2}\right) \tag{5}$$

where N_n represents the neighborhood of the n th pixel. $J_{nd}(x_n, x_i)$ is an identity function, its definition is given by

$$J_{nd}(x_n, x_i) = \begin{cases} 1, & \text{if } x_i \in N_n \text{ and } x_i \text{ is } x_n\text{'s } d\text{th adjacency type} \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

Then the n th pixel’s label probability proportion π_{nk} is defined as follows:

$$\pi_{nk} = \frac{\xi_k(x_n)}{\sum_{j=1}^K \xi_j(x_n)} = \frac{\sum_{d=1}^D \sum_{x_i \in N_n} J_{nd}(x_n, x_i) \exp\left(-\frac{(x_i - c_{kd})^2}{2b_{kd}^2}\right)}{\sum_{j=1}^K \sum_{d=1}^D \sum_{x_i \in N_n} J_{nd}(x_n, x_i) \exp\left(-\frac{(x_i - c_{jd})^2}{2b_{jd}^2}\right)} \tag{7}$$

Obviously, the label probability proportion π_{nk} in (7) is nonnegative and subject to $\sum_{k=1}^K \pi_{nk} = 1$. The density function of the n th pixel is given by

$$f(x_n) = \sum_{k=1}^K \pi_{nk} St(x_n|\theta_k), \tag{8}$$

where the pdf $St(x_n|\theta_k)$ is the Student’s t-distribution defined in (4), whose parameters are $\theta_k = \{\mu_k, \lambda_k\}$.

The log-likelihood function is given by

$$L(\Theta) = \sum_{n=1}^N \log f(x_n) = \sum_{n=1}^N \log \left(\sum_{k=1}^K \pi_{nk} St(x_n|\theta_k) \right). \tag{9}$$

There are four unknown parameters in (9). They are μ_j , λ_j , c_{jd} and b_{jd}^2 . The derivatives of the function L with respect to μ_j , λ_j , c_{jd} , b_{jd}^2 , respectively, can be calculated by

$$\frac{\partial L}{\partial \mu_j} = \sum_{n=1}^N \frac{\pi_{nj}}{\sum_{k=1}^K \pi_{nk} St(x_n|\theta_j)} \cdot \frac{\lambda_j(\nu + 1)(x_n - \mu_j)}{\nu + \lambda_j(x_n - \mu_j)^2} \cdot St(x_n|\theta_j). \tag{10}$$

$$\frac{\partial L}{\partial \lambda_j} = \sum_{n=1}^N \frac{\pi_{nj}}{\sum_{k=1}^K \pi_{nk} St(x_n|\theta_j)} \cdot \frac{\nu - \nu \times \lambda_j(x_n - \mu_j)^2}{2\lambda_j[\nu + \lambda_j(x_n - \mu_j)^2]} \cdot St(x_n|\theta_j). \tag{11}$$

$$\begin{aligned} \frac{\partial L}{\partial c_{jd}} &= \sum_{n=1}^N \frac{1}{\sum_{k=1}^K \pi_{nj} St(x_n|\theta_j)} \cdot \frac{1}{\left[\sum_{p=1}^K \xi_p(x_n) \right]^2} \\ &\cdot \left\{ \sum_{p=1, p \neq j}^K \xi_p(x_n) \cdot St(x_n|\theta_j) - \sum_{p=1, p \neq j}^K \xi_p(x_n) \cdot St(x_n|\theta_p) \right\} \\ &\cdot \left[\sum_{x_i \in N_n} J_{id}(x_n, x_i) \cdot \exp\left(-\frac{(x_i - c_{jd})^2}{2b_{jd}^2}\right) \frac{x_i - c_{jd}}{b_{jd}^2} \right]. \end{aligned} \tag{12}$$

$$\begin{aligned} \frac{\partial L}{\partial b_{jd}^2} &= \sum_{n=1}^N \frac{1}{\sum_{k=1}^K \pi_{nj} St(x_n|\theta_j)} \cdot \frac{1}{\left[\sum_{p=1}^K \xi_p(x_n) \right]^2} \\ &\cdot \left\{ \sum_{p=1, p \neq j}^K \xi_p(x_n) \cdot St(x_n|\theta_j) - \sum_{p=1, p \neq j}^K \xi_p(x_n) \cdot St(x_n|\theta_p) \right\} \\ &\cdot \left[\sum_{x_i \in N_n} J_{nd}(x_n, x_i) \exp\left(-\frac{(x_i - c_{jd})^2}{2b_{jd}^2}\right) \left(\frac{(x_i - c_{jd})^2}{2b_{jd}^4}\right) \right]. \end{aligned} \tag{13}$$

After the optimization of the parameters, the posterior probability is derived as follows:

$$St(\theta_k|x_n) = \frac{\pi_{nk} St(x_n|\theta_k)}{\sum_{j=1}^K \pi_{nj} St(x_n|\theta_j)}. \tag{14}$$

Then according to a maximum a posteriori, a k th class label is assigned to the n th pixel by solution of

$$\arg \max_k \{St(\theta_k|x_n)\}. \tag{15}$$

We refer to the proposed model as the directional spatially varying Student's t-distribution mixture model (DSVStMM). The gradient descend method is used to estimate the unknown parameters of the proposed model. Finally, in order to mitigate the effects of the noise, we utilize a simple competitive selection method [24] to optimize the result. The operation of the method is defined as follows: all the neighbors of one pixel belong to the same class, then set the pixel to the neighbors' class. The method can slightly mitigate the effects of the noise.

The proposed model DSVStMM is summarized as follows.

Algorithm 1 DSVStMM

Initialize:

Use K -means to obtain the mean u_j and the variance σ_j^2 . Then set $\mu_j = u_j$, $\lambda_j = 1/\sigma_j^2$, $c_{jd} = u_j$ and $b_{jd}^2 = \sigma_j^2$.

Step 1:

Compute the Student's t-distribution $St(x_i|\theta_j)$ in (4).

Calculate the weight function $\xi_j(x_i)$ in (5).

Update the label probability proportion π_{ij} using (7).

Compute the posterior probability $St(\theta_j|x_i)$ using (14).

Step 2:

Update the parameters $\Phi = (\mu_j, \lambda_j, c_{jd}, b_{jd}^2)^T$ using the gradient descend method

$$\Phi^{(t+1)} = \Phi^{(t)} - \eta \nabla L(\Phi^{(t)}) \quad (16)$$

where η is the learning rate. We set $\eta = 10^{-5}$. $\nabla L(\Phi)$ is the derivative of the function L with respect to Φ , where $\nabla L(\Phi) = [\partial L/\partial \mu_j, \partial L/\partial \lambda_j, \partial L/\partial c_{jd}, \partial L/\partial b_{jd}^2]^T$.

Step 3:

If the value of (9) changes significantly, set $\Phi^{(t)} = \Phi^{(t+1)}$, and return to step 1.

Step 4:

Evaluate the posterior probability $St(\theta_j|x_i)$ given in (15) to get the class labels of the pixels. Use the competitive selection method to alleviate the effects of noise.

4 Numerical experiments

We conduct some comprehensive experiments to evaluate the performance of our proposed model, compared with some state-of-the-art image segmentation models. These models are the standard GMM [2] which assumes that the relationships between the pixels are independent; the MRF-based methods which are similar to the proposed model, such as SVFMM [1], DCASV [17], CLP, BLP [21], EGMM [24] and SNGMM [23]; MSL [14] which bases on the maximum likelihood estimator and computes the likelihood of the given pixel's neighborhood; the K -means which is used to initialize GMM, DCASV, EGMM, SNGMM and the proposed model. These models are divided into two types, one type considers that the pixels are independent of each other, the other type considers the spatial relationships of the pixels which are based on MRF and are similar to the proposed model. All of the models are implemented in MATLAB. The source code of SVFMM can be downloaded from <http://www.cs.uoi.gr/kblekas/sw/MApsegmentation.html>. The source code of

standard GMM, DCASV, CLP and BLP is available at <http://www.cs.uoi.gr/cnikou/>. We implement the algorithms for K -means, MSL, EGMM, SNGMM and the proposed model (DSVStMM) in MATLAB. For the DSVStMM model, we have selected $\nu = 0.001$. According to (4), the value of λ must be nonnegative, so we use the absolute value of λ in (16). The loop of DSVStMM is terminated when the percentage changed of (9) between two consecutive iterations is less than 10^{-5} , i.e., $|\Delta L/L| < 10^{-5}$. These models are tested on both synthetic and natural grayscale images.

To quantify segmentation results, the misclassification ratio (MCR) [29] is adopted to evaluate the segmentation accuracy of synthetic images, the probabilistic rand (PR) index [26] is used to evaluate the image segmentation performance of the natural images. The MCR is defined as follows:

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

The value of MCR is bounded by the interval $[0, 1]$, where lower values indicate better segmentation results.

The PR index measures consistency between two segmentation labels via an overlapping fraction, its definition is given by

$$\text{PR}(S_{\text{test}}, \{S_k\}) = \frac{2}{N(N-1)} \sum_{i < j} [c_{ij} p_{ij} + (1 - c_{ij})(1 - p_{ij})].$$

where $c_{ij} = 1$ when the pixels i and j belong to the same cluster in the test image S_{test} , otherwise $c_{ij} = 0$. The value of the p_{ij} can be obtained by using sample mean estimator defined as follows:

$$p_{ij} = \frac{1}{K} \sum_{k=1}^K \Pi(l_i^{S^k} = l_j^{S^k}).$$

where K indicates the number of the ground-truth sets, $\Pi(l_i^{S^k} = l_j^{S^k})$ is an identity function, whose value is equal to 1 when the pixels i and j are in the same class in the ground-truth set k , otherwise is equal to 0. The value of PR is bounded by the interval $[0, 1]$, where the larger PR values indicate better segmentation results.

4.1 Piecewise synthetic images

To demonstrate the robustness of our model against noise, we first test the proposed model on a synthetic image (128×128 image resolution) similar to the one used in [29]. The three classes synthetic image shown in Fig. 2a is sampled from the MRF model using a Gibbs sampler. The intensities for the image with three classes are 55, 115 and 225, respectively. Figure 2b shows the original image corrupted by the mixed noise. Gaussian noise ($\mu = 0$, $\sigma^2 = 0.005$) is added first, then the salt & pepper noise

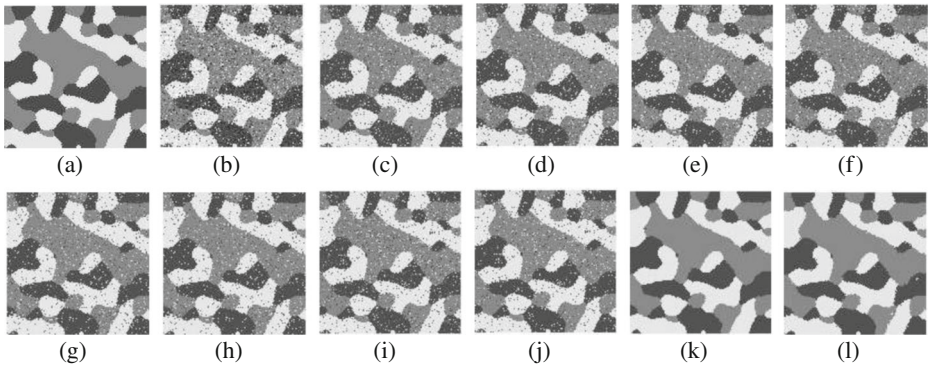


Fig. 2 First experiment (128×128 image resolution). **a** The original image. **b** Noise image with Gaussian noise (0 mean, 0.005 variance) and salt & pepper ($sp = 0.08$). **c** K -means (MCR = 8.422 %). **d** GMM (8.136 %). **e** SVFMM (MCR = 6.769 %). **f** DCASV (MCR = 7.251 %). **g** CLP (MCR = 6.323 %). **h** BLP (MCR = 6.158 %). **i** MSL (MCR = 8.441 %). **j** SNGMM (MCR = 6.586 %). **k** EGMM (MCR = 1.318 %). **l** DSVStMM (MCR = 0.842 %)

($sp = 0.08$) is added to corrupt the synthetic image. The segmentation results are shown in Fig. 2c–l, respectively. As can be seen, the SVFMM can get better results than the standard GMM. However it requires $K \cdot (N + 2)$ parameters. That is 49158 in the first example [24]. The proposed model, DSVStMM only needs $2 \times K \cdot (D + 1)$ parameters. In our implementation of the proposed model, $D = 4$, if $K = 3$, then the number of parameters in our model is only 30. The number of parameters is much fewer than the SVFMM. As shown in Fig. 2k, the EGMM reduces the effect of noise significantly and can get a better image segmentation result. However, DSVStMM can reduce the effect of noise and obtain a better result than the EGMM. Its MCR is lower 5.77 % than SNGMM’s and 0.476 % than EGMM’s. We can see that the noise is reduced clearly and the edge information of the image is kept very well in Fig. 11. In order to further test the image segmentation effectiveness of the proposed model. The experimental results obtained with various levels of mixed noise are given in Table 1. In order to reduce the randomness of the added noise and the sensitivity of the initialization, the tests are executed 10 times, then the average of the results is

Table 1 The comparison of the MCR for the first experiment

Methods	Gaussian noise (0 mean,var 0.005) + salt & pepper (sp)					Mean
	sp = 0.10	sp = 0.12	sp = 0.14	sp = 0.16	sp = 0.18	
K -means	10.225 %	11.561 %	13.021 %	14.943 %	16.180 %	13.186 %
GMM	11.562 %	12.557 %	13.412 %	14.775 %	15.631 %	13.587 %
SVFMM	8.240 %	9.424 %	10.663 %	12.225 %	13.226 %	10.756 %
DCASV	8.909 %	10.106 %	11.320 %	12.958 %	13.999 %	11.458 %
CLP	7.915 %	9.174 %	10.465 %	12.133 %	13.181 %	10.574 %
BLP	7.767 %	9.044 %	10.322 %	12.001 %	13.063 %	10.439 %
MSL	9.733 %	11.157 %	12.482 %	13.964 %	15.267 %	12.521 %
SNGMM	7.877 %	9.570 %	11.214 %	12.888 %	14.678 %	11.245 %
EGMM	2.048 %	2.865 %	4.024 %	6.055 %	7.324 %	4.463 %
DSVStMM	1.021 %	1.436 %	1.870 %	2.673 %	3.489 %	2.098 %

reported in Table 1. The results in Table 1 indicate that our proposed model has a lower MCR compared with the other models on each noise level. Its average MCR is lower 2.365 % than the EGMM's. It shows that the proposed model is more robust against noise than EGMM. The experiments prove that the proposed model gets better segmentation results than any other model, the student's t-distribution is more robust against noise because it owns heavy-tailed than Gaussian distribution.

In the second experiment, we use a synthetic four classes ($K = 4$) image (128×128 image resolution) as shown in Fig. 3a. The intensities of the image are 0, 85, 170 and 255, respectively. The corrupted image is shown in Fig. 3b. Mixed noise is added to corrupt the image which is Gaussian noise ($\mu = 0$, $\sigma^2 = 0.003$) and salt & pepper ($sp = 0.15$). The segmentation results obtained by K -means, GMM, SVFMM, DCASV, CLP, BLP, MSL, SNGMM, EGMM and DSVStMM are shown in Fig. 3c–l, respectively. Amongst these models, the EGMM and DSVStMM get the lowest MCR, but the DSVStMM does better than the EGMM. It can be seen from the Fig. 3l that the edges and contours of the image are kept very clearer than other segmentation results. In the segmentation result of EGMM shown in Fig. 4k, the edge information is lost. The MCR of the proposed model is lower 1.190 % than EGMM. In Table 2, the results obtained (executing the test 10 times and the average of the results) for these models are presented with various mixed noise. Our proposed model gets the lowest MCR in every noise level. The MCR of the proposed model is lower 2 % or so on each noise level than EGMM. From these synthetic image tests, we can inference that the proposed model is more robust than other models against noise.

Filtering can effectively reduce the noise. In order to test the impact of filtering before applying segmentation, a synthetic image is first corrupted by salt & pepper noise and Gaussian noise, then median filtering is used to reduce the noise, at last, all models are used to segment the image obtained by median filtering. Salt & pepper noise represents itself as randomly occurring white and black pixels. Median filtering is very effective to reduce the salt & pepper noise. Figure 4a shows the

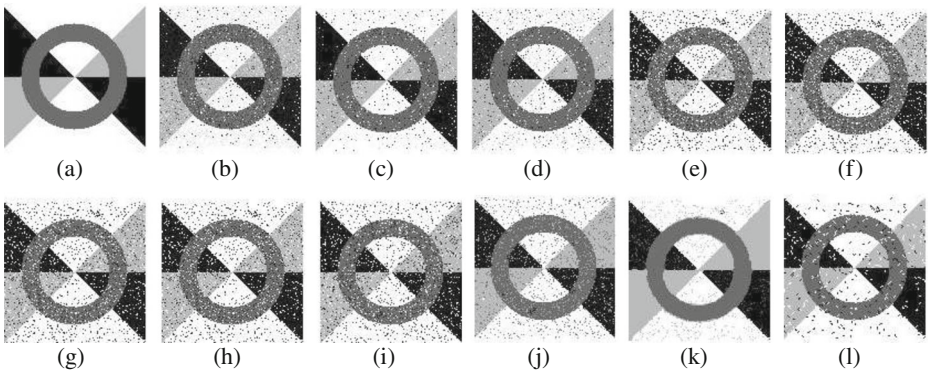


Fig. 3 Second experiment (128×128 image resolution). **a** The original image. **b** Noise image with Gaussian noise (0 mean, 0.003 variance) and salt & pepper ($sp = 0.15$). **c** K -means (MCR = 11.029 %). **d** GMM (11.664 %). **e** SVFMM (MCR = 11.420 %). **f** DCASV (MCR = 11.401 %). **g** CLP (MCR = 11.078 %). **h** BLP (MCR = 11.151 %). **i** MSL (MCR = 11.169 %). **j** SNGMM (MCR = 8.490 %). **k** EGMM (MCR = 6.561 %). **l** DSVStMM (MCR = 5.371 %)

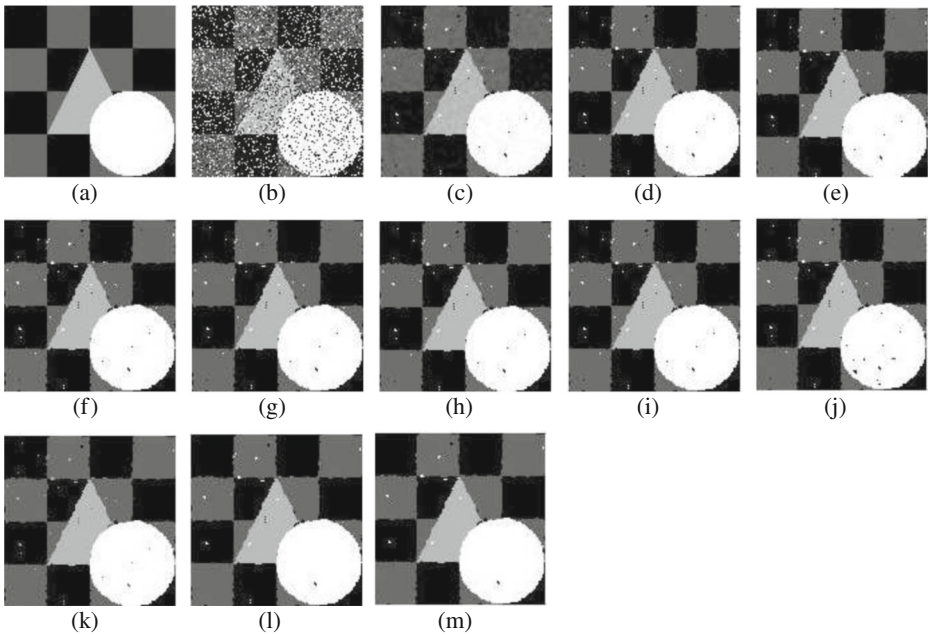


Fig. 4 Third experiment (128×128 image resolution). **a** The original image. **b** Noise image with Gaussian noise (0 mean, 0.001 variance) and salt & pepper ($sp = 0.25$). **c** image obtained by median filtering. **d** *K*-means (MCR = 1.959 %). **e** GMM (1.959 %). **f** SVFMM (MCR = 1.959 %). **g** DCASV (MCR = 1.959 %). **h** CLP (MCR = 1.959 %). **i** BLP (MCR = 1.990 %). **j** MSL (MCR = 2.057). **k** SNGMM (MCR = 1.984). **l** EGMM (MCR = 1.825 %). **m** DSVStMM (MCR = 1.752 %)

synthetic image (128×128 image resolution) which consists of four classes ($K = 4$) with luminance values $[0, 85, 170, 255]$, used in the third experiment. The image shown in Fig. 4b is obtained by corrupting the original image with the salt & pepper noise ($sp = 0.25$) and Gaussian noise (0 mean, 0.001 variance). The image shown in Fig. 4c is obtained by median filtering acting on the image shown in Fig. 4b. It can be seen that the median filtering reduces the noise dramatically. All models

Table 2 The comparison of the MCR for the second experiment

Methods	Gaussian noise (0 mean, var) + salt & pepper (sp)					Mean
	var = 0.004 sp = 0.16	var = 0.005 sp = 0.17	var = 0.006 sp = 0.18	var = 0.007 sp = 0.19	var = 0.008 sp = 0.20	
<i>K</i> -means	11.934 %	13.406 %	14.735 %	16.290 %	17.902 %	14.853 %
GMM	26.593 %	35.907 %	36.705 %	37.596 %	38.299 %	35.020 %
SVFMM	12.750 %	14.303 %	15.702 %	17.544 %	18.816 %	15.823 %
DCASV	13.214 %	17.322 %	21.798 %	26.757 %	29.903 %	21.799 %
CLP	11.931 %	13.265 %	14.359 %	15.842 %	17.430 %	14.565 %
BLP	11.882 %	12.955 %	13.821 %	14.899 %	16.174 %	13.946 %
MSL	12.422 %	13.769 %	15.147 %	16.821 %	18.416 %	15.315 %
SNGMM	9.550 %	10.747 %	11.679 %	13.117 %	14.394 %	11.897 %
EGMM	7.459 %	8.969 %	9.540 %	10.529 %	10.892 %	9.478 %
DSMM	5.862 %	6.678 %	7.287 %	8.219 %	8.914 %	7.392 %

Table 3 The comparison of the MCR for the third experiment

Methods	Gaussian noise (0 mean, var 0.001) + salt & pepper (sp)					Mean
	sp = 0.30	sp = 0.35	sp = 0.40	sp = 0.45	sp = 0.50	
<i>K</i> -means	2.545 %	3.827 %	5.676 %	7.672 %	9.735 %	5.891 %
GMM	2.545 %	3.827 %	5.676 %	7.672 %	9.735 %	5.891 %
SVFMM	2.545 %	3.827 %	5.676 %	7.672 %	9.735 %	5.891 %
DCASV	2.545 %	3.827 %	5.676 %	7.672 %	9.735 %	5.891 %
CLP	2.545 %	3.827 %	5.676 %	7.672 %	9.735 %	5.891 %
BLP	2.557 %	3.845 %	5.676 %	7.660 %	9.729 %	5.893 %
MSL	2.747 %	4.010 %	5.963 %	7.715 %	9.918 %	6.071 %
SNGMM	2.429 %	3.589 %	5.310 %	7.428 %	9.283 %	5.608 %
EGMM	2.368 %	3.510 %	5.408 %	7.379 %	9.723 %	5.678 %
DSVStMM	2.252 %	3.388 %	5.078 %	7.086 %	8.911 %	5.343 %

are then used to segment the image obtained by median filtering. In Fig. 4d–m, we present the segmentation results of *K*-means, GMM, SVFMM, DCASV, CLP, BLP, MSL, SNGMM, EGMM and DSVStMM, respectively. It can be seen that the filtering affects the segmentation results obviously. The segmentation results show that most models get the same MCR. The MCR obtained by the EGMM model is lower than other models except the DSVStMM model. DSVStMM model get lower MCR than any other model. The image in Fig. 4a is corrupted by Gaussian noise (0 mean, 0.001 variance) and various levels of the salt & pepper noise, then median filtering is used to reduce the noise. The segmentation results obtained by all methods after application of the median filtering are given in Table 3. It can be seen that most models get the same MCR, while the DSVStMM gets lower MCR than any other model. It shows that the DSVStMM can effectively reduce the effects of the noise than any other model.

4.2 Natural images

Segmenting natural images is very difficult and challenging. It is hard to provide a good model to segment objects such as humans, animals, trees, buildings, etc. In this series of experiments, we test the proposed model on the large standard databases Amsterdam Library of Object Images (ALOI) [10] visually and the Berkeley image segmentation database [15] visually and quantitatively. In these natural images segmentation experiments, a set of natural grayscale images with and without artificial noise are used to evaluate the performance of the proposed model (DSVStMM), compared with GMM, SVFMM, DCASV, CLP, BLP, MSL, EGMM and SNGMM.

The ALOI database consists of 1000 categories, each of which contains dozens of pictures of a specific object. The ALOI database can be used to object recognition, color constancy, invariant feature extraction and image retrieval, it also offer visual evaluation of image segmentation [11]. The images shown in Fig. 5 are taken from ALOI database. In Fig. 5, the first to the third column of the first row are the original images, the fourth to the seventh column of the first row are the images corrupted by mixture noise. These images are first corrupted by Gaussian noise (0 mean, 0.005 variance) and then added salt & pepper noise ($sp = 0.10$). The typical segmentation results are shown from the second row to the tenth row in Fig. 5 for the different models. For the segmentation results of images which are not corrupted by noise,

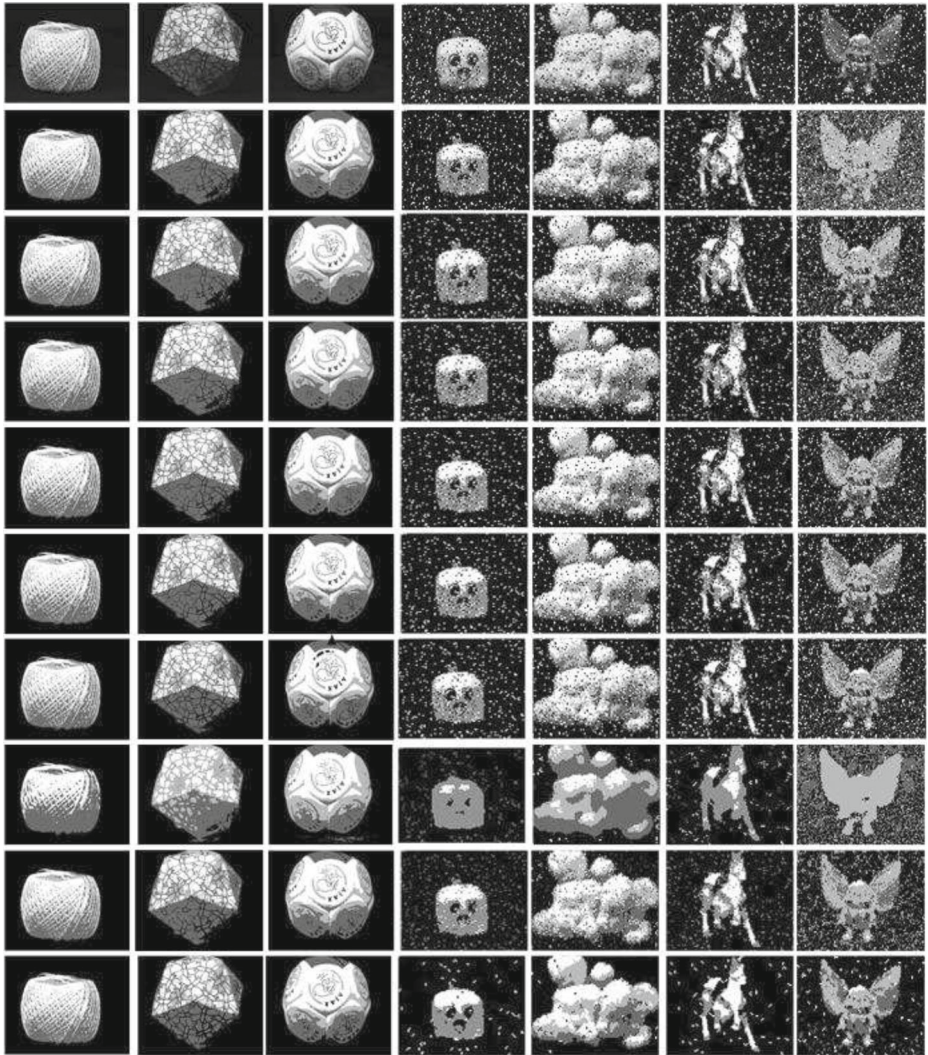


Fig. 5 Segmentation examples of different methods based on the ALOI database. The *first row*, from the *first column* to the *third column*: the original images, from the *fourth column* to the *seventh column*, the original images corrupted by Gaussian noise ($\mu = 0$, $\sigma^2 = 0.005$) and salt & pepper ($sp = 0.10$). From the *second row* to the *tenth row*, each row shows: GMM, SVMM, DCASV, CLP, BLP, MSL, EGMM, SNGMM, DSVStMM

we can see that DSVStMM yields better segmentation results than other models. The results obtained by DSVStMM own more homogeneous segmentation regions and sharper segment boundaries than other models. For the segmentation results of the corrupting images, the effect of reducing the noise by GMM model is not clear. The contours are obviously blurring of the last segmentation result of GMM. The EGMM and DSVStMM reduce the noise clearly, but EGMM is sensitive to the texture appearance. For the segmentation results of noise image, the EGMM reduce

the noise effectively, but the edge and contour information of its segmentation results is blurred. However, DSVStMM yields better segmentation results than any other model. The DSVStMM not only degrades side effect of the noise for noise images but also its segmentation results keep better edge and contour information than any other model. The segmentation results obtained by DSVStMM provide the detailed information of the contours and edges. Since the component function of DSVStMM is Student's t-distribution which has heavier tails than Gaussian distribution, the results obtained by DSVStMM show that DSVStMM can effectively process complex texture images and is more robust against noise than other models.

The Berkeley image segmentation database is considered to be used for image segmentation which consists of 300 natural color and grayscale images. Each image of the image database has 481×321 pixels. Several manual segmentation ground-truth images are provided for each image. Some natural grayscale images taken from the Berkeley image segmentation database with and without artificial noise are used to evaluate the segmentation performance visually and quantitatively. For the natural image segmentation, the proposed model is compared with the GMM, SVFMM DCASV, CLP, BLP, MSL, EGMM and SNGMM. The visual segmentation results are shown in Fig. 6. Three representative example images without noise are shown in the first column to the third column of the first row in Fig. 6; another three images are corrupted by mixture noise are shown in the fourth column to the sixth column of the first row in Fig. 6. The last three images are first corrupted by Gaussian noise (0 mean, 0.003 variance) and then added salt & pepper noise ($sp = 0.05$). The typical segmentation results are shown from the second row to the tenth row of Fig. 6 for the different models. According to the segmentation results of these images without noise, the differences of the segmentation results of different models are not very clear. The EGMM produces very smooth segmentations, while certain texture regions may be merged into one piece. The DSVStMM favors a more contiguous segmentation for the texture region, and the information of the edges is preserved very well compared with the other models. From the segmentation results of the images with noise, we can see that the EGMM and DSVStMM both reduce the noise significantly. About the fourth image, there is a plane in the sky. The segmentation results of the GMM and DCASV show that the noise is reduced poorly and the details of the plane are lost. The EGMM model can reduce noise dramatically, but the contours and edges of the segmentation results are not kept clearly. For the EGMM, much region of the sky is wrongly segmented into the plane. The segmentation result obtained by DSVStMM shows that the noise is reduced dramatically and the details of the plane are preserved very well. For the segmentation results of the fifth image, the DSVStMM not only reduces the effect of noise significantly, but also preserves the details of the balcony better than the other models. However, the information of the window and balcony is lost in the segmentation result of the EGMM for the fifth image. For the last noise image, the EGMM and DSVStMM reduce the effect of noise obviously compared with the other models, however, the tree can not be segmented correctly by the EGMM model. The tree, background and the mountain can be distinguished obviously for the segmentation results of the DSVStMM.

Next, we evaluate the performance of the DSVStMM from the objective view, as well as those of the GMM, SVFMM, DCASV, CLP, BLP, MSL, EGMM and SNGMM. In this experiment, 50 grayscale images with and without artificial noise are



(a)

Fig. 6 Segmentation examples of different methods based on the Berkeley segmentation grayscale database. The *first row*, from *first column* to the *third column*: the original images, from the *fourth column* to the *sixth column*, the original images corrupted by Gaussian noise ($\mu = 0, \sigma^2 = 0.003$) and salt & pepper ($sp = 0.05$). From the *second row* to the *tenth row*, each row shows: GMM, SVM, DCASV, CLP, BLP, MSL, EGMM, SNGMM, DSVStMM

chosen from the Berkeley image segmentation database. The images are corrupted by mixture noise (mn) which is Gaussian noise (0 mean, 0.005 variance) and salt & pepper noise ($sp = 0.10$). The PR value is used to quantify segmentation results. The PR values of image segmentation results are given in Table 4. It can be seen from the Table 4 that the differences of the segmentation results of the different models are relatively little for the images without noise. But under the noise conditions, the

Table 4 Comparison of image segmentation results based on berkeley grayscale images: pr index

Image	N	K	GMM	SVM	DCASV	CLP	BLP	MSL	EGMM	SNGMM	DVStMM
55067	-	4	0.8838	0.8932	0.8855	0.8896	0.8940	0.8787	0.8933	0.8963	0.8988
148089	-	5	0.7083	0.7065	0.7101	0.7098	0.7098	0.7083	0.7062	0.7118	0.7111
291000	-	6	0.6760	0.6765	0.6776	0.6789	0.6797	0.6761	0.6616	0.6819	0.6811
169012	-	7	0.7843	0.7814	0.7858	0.7859	0.7853	0.7854	0.7739	0.7874	0.7887
159091	-	3	0.6291	0.6338	0.6326	0.6344	0.6349	0.6304	0.6281	0.6351	0.6356
157055	-	5	0.7387	0.7409	0.7404	0.7397	0.7395	0.7383	0.7386	0.7399	0.7433
323016	-	3	0.7257	0.7345	0.7281	0.7313	0.7319	0.7317	0.7364	0.7354	0.7366
66053	-	6	0.6808	0.6840	0.6825	0.6830	0.6835	0.6819	0.6841	0.6858	0.6843
86016	-	7	0.6962	0.7020	0.6967	0.6977	0.6975	0.6978	0.6415	0.7011	0.7018
229036	-	5	0.6838	0.6822	0.6865	0.6877	0.6890	0.6817	0.6753	0.6886	0.6928
123074	-	6	0.6890	0.6912	0.6905	0.6895	0.6906	0.6907	0.6912	0.6911	0.6920
147091	mn	3	0.6761	0.6774	0.6786	0.6825	0.6841	0.6715	0.7108	0.6816	0.7194
219090	mn	4	0.6902	0.7124	0.6984	0.7118	0.7124	0.7062	0.6980	0.6969	0.7300
41033	mn	3	0.6825	0.6854	0.6844	0.6858	0.6862	0.6809	0.6742	0.6854	0.7106
249061	mn	5	0.7630	0.7468	0.7650	0.7679	0.7800	0.7218	0.7611	0.7516	0.8030
189080	mn	4	0.7570	0.7646	0.7634	0.7657	0.7665	0.7554	0.7771	0.7711	0.7945
126007	mn	3	0.5917	0.6275	0.6093	0.6285	0.6308	0.6090	0.6123	0.6170	0.6463
38092	mn	3	0.7004	0.7131	0.7119	0.7156	0.7172	0.6985	0.7507	0.7207	0.7516
113016	mn	3	0.6023	0.6012	0.6074	0.6077	0.6092	0.5948	0.6438	0.6128	0.6520
117054	mn	3	0.5972	0.5973	0.6031	0.6039	0.6054	0.5902	0.6144	0.6110	0.6359
2092	mn	3	0.6545	0.6849	0.6835	0.7102	0.7138	0.6589	0.7116	0.6948	0.7536
3096	mn	2	0.8211	0.8704	0.8426	0.8747	0.8758	0.8290	0.8194	0.8668	0.9444
15088	mn	2	0.8078	0.8118	0.8099	0.8145	0.8153	0.7926	0.8971	0.8161	0.8909
24063	mn	3	0.7502	0.7760	0.7640	0.7763	0.7774	0.7559	0.8080	0.7678	0.8237
28075	mn	3	0.5508	0.5532	0.5517	0.5593	0.5609	0.5425	0.6103	0.5626	0.6009

46076	mm	4	0.7326	0.7543	0.7420	0.7462	0.7456	0.7442	0.7020	0.7526	0.7702
55075	mm	4	0.7310	0.7354	0.7336	0.7312	0.7317	0.7203	0.7471	0.7238	0.7582
61060	mm	4	0.5996	0.6175	0.6081	0.6254	0.6288	0.5984	0.4996	0.6181	0.6653
66075	mm	2	0.6222	0.6393	0.6212	0.6352	0.6338	0.6322	0.6605	0.6374	0.6681
67079	mm	2	0.6363	0.6545	0.6460	0.6579	0.6599	0.6334	0.7176	0.6581	0.7025
105019	mm	3	0.7473	0.7447	0.7552	0.7652	0.7672	0.7383	0.8817	0.8248	0.8920
138078	mm	5	0.7177	0.7527	0.7506	0.7604	0.7635	0.7409	0.6071	0.7599	0.7901
155060	mm	3	0.7080	0.7163	0.7126	0.7188	0.7202	0.7107	0.6910	0.7251	0.7619
167062	mm	2	0.8614	0.8619	0.8611	0.8613	0.8615	0.8628	0.8887	0.8776	0.9200
176039	mm	3	0.7374	0.7619	0.7521	0.7648	0.7675	0.7316	0.8486	0.7595	0.8201
196073	mm	3	0.8346	0.8335	0.8346	0.7640	0.8346	0.8303	0.9157	0.8388	0.8877
198054	mm	5	0.6908	0.7070	0.7013	0.7098	0.7126	0.6951	0.6828	0.7018	0.7366
216041	mm	4	0.6986	0.7019	0.7106	0.7212	0.7279	0.6840	0.7798	0.7177	0.7738
225017	mm	4	0.6497	0.6830	0.6659	0.6752	0.6769	0.6584	0.5702	0.6658	0.6934
253036	mm	2	0.7931	0.7888	0.7966	0.8136	0.8175	0.7828	0.9090	0.7936	0.8871
268002	mm	2	0.6656	0.6774	0.6721	0.6817	0.6831	0.6858	0.7453	0.6835	0.7542
271008	mm	3	0.6449	0.6370	0.6459	0.6474	0.6491	0.6326	0.6714	0.6396	0.6726
271031	mm	3	0.7612	0.7959	0.7692	0.7957	0.7977	0.7199	0.8997	0.7930	0.8834
310007	mm	3	0.6249	0.6728	0.6478	0.6779	0.6792	0.6395	0.4860	0.6864	0.7321
135037	mm	4	0.6362	0.6356	0.6394	0.6445	0.6440	0.6270	0.6422	0.6488	0.6725
106024	mm	3	0.6299	0.6394	0.6356	0.6420	0.6432	0.6323	0.6633	0.6401	0.6795
144067	mm	4	0.7143	0.7187	0.7161	0.7227	0.7238	0.7161	0.5914	0.7163	0.7415
181091	mm	4	0.7242	0.7273	0.7271	0.7291	0.7295	0.7256	0.6951	0.7399	0.7515
286092	mm	5	0.7700	0.7821	0.7755	0.7849	0.7864	0.7749	0.7773	0.7707	0.8035
22093	mm	3	0.7060	0.7167	0.7125	0.7184	0.7192	0.7114	0.6773	0.6952	0.7423
mean	–	–	0.7036	0.7141	0.7104	0.7165	0.7195	0.7027	0.7174	0.7176	0.7517

differences of the DSVStMM between any other model are obviously larger. At the same time the average PR of the DSVStMM is also larger than any other model. It can be seen from Table 4, the MRF-based models obtain better segmentation results than non MRF-based model under the noise conditions. The average PR of the proposed model is larger 0.04 or so than the other models. The experimental results indicate that the DSVStMM is robust against noise and produces better segmentation results. There may be two reasons for the DSVStMM producing better segmentation results. One is that the component function of DSVStMM is Student's *t*-distribution which owns heavier tails than Gaussian distribution and can meet the needs of image segmentation, the other is that the models of the DSVStMM's label probability proportions effectively represent the neighborhood relationships of the pixels.

The execution time of the program is affected by many factors, such as the dimensions of the image and the number of segments. The dimension of the image is the most important factor. The tests given in Table 4 are performed on 481×321 images with a PC with a 2.20 GHz dual core processor and 2 GB RAM. The DSVStMM model converges after iterating averagely 84.0 times and the execution time is 4.7 s per iteration on the images in Table 4.

5 Conclusions

In this paper, we introduce an image segmentation algorithm which takes into account the spatial relationships among pixels. At the same time the directional information of the neighborhood is incorporated in the model. The number of parameters needed in the proposed model is fewer than that in other MRF-based models, so it is easier to be implemented. Furthermore, the pixels' label probability proportions are explicitly modelled as probability vectors, thus a reparatory project step is avoided. The gradient descend method is used to estimate the parameters of the proposed model instead of the EM algorithm which is widely applied in most models. The experiments on both the synthetic and natural images show that the proposed model obtains better results than the other models. To extend the applications of the proposed model, we plan to apply the proposed model to medical imaging and bioinformatics in the future.

However, there are several improved aspects in our model. One limitation of our model is that it can be only used for grayscale images, not for color images. Another limitation is the choice of the parameter K . The number of model components K is currently set by user based on prior knowledge. This is an open question in mixture models. How to extend our model to apply in color images and adaptively choose K from the image data are our future works. In the context of image segmentation, we only use a 1-dimensional (grayscale color) feature to describe the grayscale image data. To improve the quality of image segmentation, more sophisticate features for natural images such as the Blobworld features [3] and the MRF texture features [19] will be considered to integrate into our model in the future researches.

Acknowledgements The authors would like to thank the anonymous reviewers for their valuable comments and suggestions, which greatly helped to improve both the technical content and the presentation quality of the paper. This work was supported by National Basic Research Program of China (973 Program) under Grant 2011CB302201, and National Nature Science Foundation of China under grant No. 60931160441 and No. 61003042.

References

1. Blekas K, Likas A, Galatsanos N, Lagaris I (2005) A spatially constrained mixture model for image segmentation. *IEEE Trans Neural Netw* 16:494–498. doi:[10.1109/TNN.2004.841773](https://doi.org/10.1109/TNN.2004.841773)
2. Bishop C (2006) *Pattern recognition and machine learning*. Springer, Berlin Heidelberg New York
3. Carson C, Belongie S, Greenspan H, Malik J (2002) Blobword:image segmentation using expectation-maximization and its application to image querying. *IEEE Trans Pattern Anal Mach Intell* 24:1026–1038. doi:[10.1109/TPAMI.2002.1023800](https://doi.org/10.1109/TPAMI.2002.1023800)
4. Chatzis SP, Kosmopoulos DI (2011) A variational Bayesian methodology for hidden Markov models utilizing student's-t mixtures. *Pattern Recogn* 44:295–306. doi:[10.1016/j.patcog.2010.09.001](https://doi.org/10.1016/j.patcog.2010.09.001)
5. Chatzis SP, Kosmopoulos DI, Varvarigou TA (2008) Signal modeling and classification using a robust latent space model based on t distributions. *IEEE Trans Signal Process* 56:949–963. doi:[10.1109/TSP.2007.907912](https://doi.org/10.1109/TSP.2007.907912)
6. Chatzis SP, Kosmopoulos DI, Varvarigou TA (2009) Robust sequential data modeling using an outlier to tolerant hidden Markov model. *IEEE Trans Pattern Anal Mach Intell* 31:1657–1669
7. Dempster P, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via EM algorithm. *J R Stat Soc* 39:1–38
8. Forsyth DA, Ponce J (2003) *Computer vision: a modern approach*. Prentice Hall, Englewood Cliffs
9. Geman S, Geman D (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans Pattern Anal Mach Intell* 6:721–741
10. Geusebroek JM, Burghouts GJ, Smeulders AW (2005) The Amsterdam library of object images. *Int J Comput Vis* 61:103–112. doi:[10.1023/B:VISI.0000042993.50813.60](https://doi.org/10.1023/B:VISI.0000042993.50813.60)
11. Ismail B, Amar M (2008) A region merging prior for variational level set image segmentation. *IEEE Trans Image Process* 17:2301–2311. doi:[10.1109/TIP.2008.2006425](https://doi.org/10.1109/TIP.2008.2006425)
12. Jordan MT, Ghabramani Z, Jaakkola TS, Saul LK (1999) An introduction to variational methods for graphical models. *Mach Learn* 37:183–233. doi:[10.1023/A:1007665907178](https://doi.org/10.1023/A:1007665907178)
13. Li SZ (2009) *Markov random field modeling in image analysis*. Springer, Berlin Heidelberg New York
14. Mariano R, Oscar D, Washington M, Alonso R (2012) Spatial sampling for image segmentation. *Comput J* 55:313–324. doi:[10.1093/comjnl/bxr032](https://doi.org/10.1093/comjnl/bxr032)
15. Martin D, Fowlkes C, Tal D, Malik J (2001) A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: *Proc IEEE conf computer vision (ICCV2001)*, pp 416–423
16. McLachlan G, Peel D (2000) *Finite mixture models*. Wiley, New York
17. Nikou C, Galatsanos N, Likas A (2007) A class-adaptive spatially variant mixture model for image segmentation. *IEEE Trans Image Process* 16:1121–1130. doi:[10.1109/TIP.2007.891771](https://doi.org/10.1109/TIP.2007.891771)
18. Peel D, McLachlan GJ (2000) Robust mixture modelling using the t distribution. *Stat Comput* 10:335–344. doi:[10.1023/A:1008981510081](https://doi.org/10.1023/A:1008981510081)
19. Ren X, Malik J (2003) Learning a classification model for segmentation. In: *Proc IEEE conf computer vision (ICCV2003)*, pp 10–17
20. Sanjay GS, Hebert TJ (1998) Bayesian pixel classification using spatially variant finite mixtures and the generalized EM algorithm. *IEEE Trans Image Process* 7:1014–1028. doi:[10.1109/83.701161](https://doi.org/10.1109/83.701161)
21. Sfikas G, Nikou C, Galatsanos N, Heinrich C (2010) Spatially varying mixtures incorporating line processes for image segmentation. *J Math Imaging Vis* 36:91–110. doi:[10.1007/s10851-009-0174-x](https://doi.org/10.1007/s10851-009-0174-x)
22. Svensen M, Bishop C (2005) Robust Bayesian mixture modelling. *Neurocomputing* 64:235–252. doi:[10.1016/j.neucom.2004.11.018](https://doi.org/10.1016/j.neucom.2004.11.018)
23. Thanh MN, Wu QM, (2012) Gaussian-mixture-model-based spatial neighborhood relationships for pixel labeling problem. *IEEE Trans Syst Man Cybern, Part B, Cybern* 42:193–202. doi:[10.1109/TSMCB.2011.2161284](https://doi.org/10.1109/TSMCB.2011.2161284)
24. Thanh MN, Wu QM, Ahuja S (2010) An extension of the standard mixture model for image segmentation. *IEEE Trans Neural Netw* 21:1326–1338. doi:[10.1109/TNN.2010.2054109](https://doi.org/10.1109/TNN.2010.2054109)
25. Titterton DM, Smith AFM, Makov UE (1985) *Statistical analysis of finite mixture distributions*. Wiley, New York
26. Unnikrishnan R, Pantofaru C, Hebert M (2007) Toward objective evaluation of image segmentation algorithms. *IEEE Trans Pattern Anal Mach Intell* 29:929–944. doi:[10.1109/TPAMI.2007.1046](https://doi.org/10.1109/TPAMI.2007.1046)

27. Woolrich M, Behrens T (2006) Variational Bayes inference of spatial mixture models for segmentation. *IEEE Trans Med Imag* 25:1380–1391. doi:[10.1109/TMI.2006.880682](https://doi.org/10.1109/TMI.2006.880682)
28. Woolrich M, Behrens T, Bechmann C, Smith S (2005) Mixture models with adaptive spatial regularisation for segmentation with an application to FMRI data. *IEEE Trans Med Imag* 24:1–11. doi:[10.1109/TMI.2004.836545](https://doi.org/10.1109/TMI.2004.836545)
29. Zhang Y, Brady M, Smith S (2001) Segmentation of brain MR images through a hidden Markov random field model and the expectation maximization algorithm. *IEEE Trans Med Imag* 20:45–57. doi:[10.1109/42.906424](https://doi.org/10.1109/42.906424)



Taisong Xiong was born in Sichuan, China. He received the B.S. degree and the M.S. degrees in computer science and engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2000 and 2006, respectively. He is currently working toward the Ph.D. degree in the School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China. His current research interests include computer vision and machine learning.



Zhang Yi (SM'09) received the Ph.D. degree in mathematics from the Institute of Mathematics, Chinese Academy of Science, Beijing, China, in 1994. He is currently a Professor at Sichuan University, Chengdu, China. He is the Dean of the College of Computer Science at Sichuan University. He is the founder of the Machine Intelligence Laboratory, College of Computer Science, Sichuan University. He has published more than 100 refereed articles. He is the co-author of three books: *Convergence Analysis of Recurrent Neural Networks* (Norwell, MA: Kluwer, 2004); *Neural Networks: Computational Models and Applications* (Heidelberg, Germany: Springer, 2007); *Subspace Learning of Neural Networks* (CRC Press, Taylor & Francis Group, 2010). He edited six volumes of books: *Advances in Neural Networks* (Springer, 2005, 2006). His current research interests include machine intelligence, neural networks, data mining, machine learning, and neocortex computing.



Lei Zhang (M'09) received the B.S. and Masters degrees in mathematics and the Ph.D. degree in computer science from the University of Electronic Science and Technology of China, Chengdu, China, in 2002, 2005, and 2008, respectively. She was a Post-Doctoral Research Fellow in the Department of Computer Science and Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong, from 2008 to 2009. Currently, she is a Professor at Sichuan University, Chengdu, China. Her current research interests include theory and applications of Machine Intelligence.