BOUNDED ASYMMETRIC MIXTURE MODEL FOR MEDICAL IMAGE SEGMENTATION

Thanh Minh Nguyen[1], *Q. M. Jonathan Wu*[1], *Senior Member, IEEE, Dibyendu Mukherjee*[1], *and Hui Zhang*[1,2]

[1]Department of Electrical and Computer Engineering, University of Windsor
[2]School of Computer & Software, Nanjing University of Information Science & Technology
{nguyen1j, jwu, mukherjd, nrzh}@uwindsor.ca

ABSTRACT

Segmentation of medical image based on the modeling and estimation of the tissue intensity probability density functions via Gaussian mixture model (GMM) has recently received great attention. However, Gaussian distribution is unbounded and symmetrical around its mean. This study presents a new bounded asymmetric mixture model for analyzing both univariate and multivariate data. The advantage of the proposed model is that it has the flexibility to fit different shapes of observed data such as non-Gaussian, non-symmetric, and bounded support data. Another advantage is that each component of the proposed model has the ability to model the observed data with different bounded support regions, which is suitable for application on image segmentation. Our method is intuitively appealing, simple, and easy to implement. We also propose a new method to estimate the model parameters in order to minimize the higher bound on the data negative log-likelihood function. Numerical experiments are presented where the proposed model is tested in various images from simulated to real $3D$ medical ones.

Index Terms— Medical image segmentation, non-Gaussian, non-symmetric, bounded support regions.

1. INTRODUCTION

Automatic segmentation plays an important role in the field of medical imaging. During the last decades, there has been a growing research interest in Bayesian technique based on the modeling of the probability density function of the data via finite mixture model. Among the algorithms based on Bayesian technique, Gaussian mixture model (GMM) [1, 2] is a well-known method used in most applications. Many researchers have used it to study a number of key problems in the area of image segmentation [3, 4].

In order to improve the robustness of the algorithm, the Student's-t mixture model (SMM) has been proposed in [5–7]. Compared to the GMM, each component of SMM has one more parameter called the degrees of freedom (v). When v tends to infinity, the Student's-t distribution approaches the Gaussian distribution. Another way to model data with different shapes is to use the generalized Gaussian mixture models (GGMM) [8, 9]. Compared to GMM, each component of GGMM has one more parameter (λ). This parameter controls the tails of the distribution. However, the major disadvantage of GGMM is that, this model assumes that the dimensions of the observed data are independent. Therefore, it is not suitable for analyzing correlated data.

In many practical problems, we observe that the intensity distribution of each tissue type in an image need not be symmetric. In contrast, the Gaussian distribution, Student's-t distribution, and generalized Gaussian distribution are symmetrical around their mean, and these curves have a single peak. To overcome these problems, a model has been proposed in [10, 11]. However, these methods are only applied for univariate data. In [12], an asymmetrical distribution was proposed to solve the problem of non-Gaussian multivariate data. One drawback of all the above-mentioned mixture models is that their distributions are unbounded with support range $(-\infty, +\infty)$. We observe in many real applications, the observed data are always in the bounded support regions.

Motivated by the aforementioned observations, we introduce in this paper a new bounded asymmetric mixture model for analyzing both univariate and multivariate data. Our approach differs from those discussed above by the following statements. Firstly, a new bounded asymmetric distribution, which has the flexibility to fit different shapes of observed data such as non-Gaussian, non-symmetric, and bounded support data, is proposed in this paper. The advantage of this distribution is that it is simple, and intuitively appealing. Secondly, each component of our model has the ability to model the observed data with different bounded support regions. Finally, to estimate the parameters of the proposed model, we propose a new method in order to minimize the higher bound on the data negative log-likelihood function. We demonstrate through extensive simulations that the proposed model is superior to other methods based on the modeling of the probability density function of the data via finite mixture model.

The remainder of this paper is organized as follows: section 2 describes the proposed method in detail; section

This research has been supported in part by the Canada Research Chair Program, AUTO21 NCE, and the NSERC Discovery grant.

This work was supported in part by the National Natural Science Foundation of China under Grant 61105007.

3 presents the parameter estimation; section 4 sets out the experimental results; and section 5 presents our conclusions.

2. PROPOSED METHOD

Let us consider the problem of estimating the posterior probability of x_i belonging to label Ω_j. The finite mixture [1, 2] assumes the density function at a pixel x_i is given by:

$$f(\mathrm{x}_i|\Theta) = \sum_{j=1}^{K} \pi_j p(\mathrm{x}_i|\Omega_j) \tag{1}$$

where π_j is the prior probability that pixel x_i is in label Ω_j. Each distribution $p(\mathrm{x}_i|\Omega_j)$ is called a component of the mixture. Note that, $p(\mathrm{x}_i|\Omega_j)$ can be any kind of distribution. As shown in (1), the main goal of statistical modeling is to establish a model that can best describe the statistical properties of the underlying source. The existing mixture models [1, 2, 5, 7–9] have relied on $p(\mathrm{x}_i|\Omega_j)$ for modeling the underlying distributions. However, the distributions of all above-mentioned mixture models are unbounded with support range $(-\infty, +\infty)$. In many real applications, the observed data are always in the bounded support regions. For example, in the area of signal processing, the power spectrum is semi-bounded. In the area of image computer vision, the pixels are usually not symmetric, and in the limited range. Also, in the area of MRI segmentation, each tissue label is in different bounded support regions. Motivated by the aforementioned observations, we propose a bounded asymmetric mixture model (BAMM) with bounded support region, non-Gaussian, non-symmetric distribution. The advantage of the proposed model is that it is simple, intuitively appealing, and has the ability to analyze both univariate and multivariate data.

First, for each label Ω_j, we define ∂_{Ω_j} to be the bounded support region in $\mathbb{R}^D$, and the indicator function as:

$$\mathrm{H}(\mathrm{x}_i|\Omega_j) = \begin{cases} 1 & \text{IF } \mathrm{x}_i \in \partial_{\Omega_j} \\ 0 & \text{Otherwise} \end{cases} \tag{2}$$

We define a multivariate Gaussian distribution $\Phi(\mathrm{x}_i|\mu_{jk}, \Sigma_{jk})$ as follow:

$$\Phi(\mathrm{x}_i|\mu_{jk}, \Sigma_{jk}) = \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma_{jk}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathrm{x}_i - \mu_{jk})^T \Sigma_{jk}^{-1} (\mathrm{x}_i - \mu_{jk})\right\} \tag{3}$$

In (3), i=(1,2,...,N), j=(1,2,...,K), and k=(1,2,...,K_j). The D-dimensional vector μ_{jk} is the mean. The DxD matrix Σ_{jk} is the covariance, and $|\Sigma_{jk}|$ denotes the determinant of Σ_{jk}. With the indicator function $\mathrm{H}(\mathrm{x}_i|\Omega_j)$ in (2) and the distribution $\Phi(\mathrm{x}_i|\mu_{jk}, \Sigma_{jk})$ in (3), we define a bounded multivariate Gaussian distribution $\Psi(\mathrm{x}_i|\mu_{jk}, \Sigma_{jk})$:

$$\Psi(\mathrm{x}_i|\mu_{jk}, \Sigma_{jk}) = \frac{\Phi(\mathrm{x}_i|\mu_{jk}, \Sigma_{jk})\mathrm{H}(\mathrm{x}_i|\Omega_j)}{\int_{\partial_{\Omega_j}} \Phi(\mathrm{x}|\mu_{jk}, \Sigma_{jk})d\mathrm{x}} \tag{4}$$

In (4), $\int_{\partial_{\Omega_j}} \Phi(\mathrm{x}|\mu_{jk}, \Sigma_{jk})d\mathrm{x}$ is the normalization constant, and it identified as the share of $\Phi(\mathrm{x}_i|\mu_{jk}, \Sigma_{jk})$ that belongs to the support region ∂_{Ω_j}. The idea to define the distribution $\Psi(\mathrm{x}_i|\mu_{jk}, \Sigma_{jk})$ in (5) is based on a fact that the observed data of an image is digitalized and have bounded support. We want to assign $\Psi(\mathrm{x}_i|\mu_{jk}, \Sigma_{jk})$ equal to $\Phi(\mathrm{x}_i|\mu_{jk}, \Sigma_{jk})$ in the support region ∂_{Ω_j}, and zero outside.

Next, in order to fit different shapes of observed data such as non-Gaussian, non-symmetric, and bounded support data, we define a new D-dimensional bounded asymmetric distribution $p(\mathrm{x}_i|\Omega_j)$ in this paper. Each component density in our model is modeled with multiple D-dimensional bounded asymmetric distribution $\Psi(\mathrm{x}_i|\mu_{jk}, \Sigma_{jk})$. The proposed distribution $p(\mathrm{x}_i|\Omega_j)$ is defined as:

$$p(\mathrm{x}_i|\Omega_j) = \sum_{k=1}^{K_j} \eta_{jk} \Psi(\mathrm{x}_i|\mu_{jk}, \Sigma_{jk}) \tag{5}$$

where, K_j is the number of the bounded multivariate Gaussian distribution that is used to model the label Ω_j. And η_{jk} is called the weighting factor that satisfies the following constraints:

$$\eta_{jk} \geq 0 \text{ and } \sum_{k=1}^{K_j} \eta_{jk} = 1 \tag{6}$$

The idea to define the distribution in (7) is based on a fact that non-Gaussian, non-symmetric, and bounded support data can be approximated by multiple bounded multivariate Gaussian distributions. It is worth mentioning that the proposed distribution in (7) will always satisfy the conditions of the probability density [1]:

$$p(\mathrm{x}_i|\Omega_j) \geq 0 \text{ and } \int_{-\infty}^{\infty} \sum_{k=1}^{K_j} \eta_{jk} \Psi(\mathrm{x}|\mu_{jk}, \Sigma_{jk})d\mathrm{x} = 1 \tag{7}$$

Given the distribution $p(\mathrm{x}_i|\Omega_j)$ in (5), the log-likelihood function is written in the form.

$$L(\Theta) = \sum_{i=1}^{N} \log \sum_{j=1}^{K} \pi_j \sum_{k=1}^{K_j} \eta_{jk} \frac{\Phi(\mathrm{x}_i|\mu_{jk}, \Sigma_{jk})\mathrm{H}(\mathrm{x}_i|\Omega_j)}{\int_{\partial_{\Omega_j}} \Phi(\mathrm{x}|\mu_{jk}, \Sigma_{jk})d\mathrm{x}} \tag{8}$$

From (8), we can see that each component of the proposed model has the ability to model the observed data with different bounded support regions ∂_{Ω_j}. We can define any shape of ∂_{Ω_j} based on the prior knowledge about the observed data.

3. PARAMETER LEARNING

Thus far, the discussion has focused on $p(\mathrm{x}_i|\Omega_j)$ in Eq.(7) for modeling the underlying distributions. In order to determine the label Ω_j to which the pixel x_i should be assigned, we need to adjust the parameters $\Theta = \{\pi_j, \eta_{jk}, \mu_{jk}, \Sigma_{jk}\}$ in order to maximize the likelihood function in (8). Since the logarithm

is a monotonically increasing function, it is more convenient to consider the negative logarithm of the likelihood function [6, 14], as an error function

$$J(\Theta) = -L(\Theta) = -\sum_{i=1}^{N} \log \sum_{j=1}^{K} \pi_j p(\mathrm{x}_i|\Omega_j) \tag{9}$$

Therefore, maximizing the likelihood $L(\Theta)$ in (8) is then equivalent to minimizing $J(\Theta)$ in (9). In order to minimize the error function $J(\Theta)$, we define two variables $z_{ij}^{(t)}$ and $y_{ijk}^{(t)}$. The variables $z_{ij}^{(t)}$ is defined as:

$$z_{ij}^{(t)} = \frac{\pi_j \sum_{k}^{K_j} \eta_{jk} \Psi(\mathrm{x}_i|\mu_{jk}, \Sigma_{jk})}{\sum_{m=1}^{K} \pi_m \sum_{k}^{K_m} \eta_{mk} \Psi(\mathrm{x}_i|\mu_{mk}, \Sigma_{mk})} \tag{10}$$

and the value of $y_{ijk}^{(t)}$ is defined as:

$$y_{ijk}^{(t)} = \frac{\eta_{jk} \Psi(\mathrm{x}_i|\mu_{jk}, \Sigma_{jk})}{\sum_{m}^{K_j} \eta_{jm} \Psi(\mathrm{x}_i|\mu_{jm}, \Sigma_{jm})} \tag{11}$$

Because the values $z_{ij}^{(t)}$ in Eq.(10), and $y_{ijk}^{(t)}$ in Eq.(11) always satisfy the conditions $\sum_{j=1}^{K} z_{ij}^{(t)} = 1$ and $\sum_{k=1}^{K_j} y_{ijk}^{(t)} = 1$, we can apply Jensen's inequality [14] in the form $\log(\sum_{j=1}^{K} z_{ij}^{(t)} s) \geq \sum_{j=1}^{K} z_{ij}^{(t)} \log(s)$ and $\log(\sum_{k=1}^{K_j} y_{ijk}^{(t)} s) \geq \sum_{k=1}^{K_j} y_{ijk}^{(t)} \log(s)$ to the error function in (9) to give

$$\begin{aligned} J(\Theta) \leq -\sum_{i=1}^{N}\sum_{j=1}^{K} z_{ij}^{(t)} \{\log \pi_j + \sum_{k=1}^{K_j} y_{ijk}^{(t)} \{ \\ \log \eta_{jk} + \log \Psi(\mathrm{x}_i|\mu_{jk}, \Sigma_{jk})\}\} \end{aligned} \tag{12}$$

Then, minimizing the negative log-likelihood function in Eq.(12), is equivalent to minimizing the error function $E(\Theta)$:

$$\begin{aligned} E(\Theta) = \\ -\sum_{i=1}^{N}\sum_{j=1}^{K} z_{ij}^{(t)} \{\log \pi_j + \sum_{k=1}^{K_j} y_{ijk}^{(t)} \{\log \eta_{jk} + \log \mathrm{H}(\mathrm{x}_i|\Omega_j) \\ + \log \Phi(\mathrm{x}_i|\mu_{jk}, \Sigma_{jk}) - \log \int_{\partial\Omega_j} \Phi(\mathrm{x}|\mu_{jk}, \Sigma_{jk}) d\mathrm{x}\}\} \end{aligned} \tag{13}$$

To minimize this function, we consider the derivation of the error function $E(\Theta)$ with the means μ_{jk}, and Σ_{jk}^{-1} at the (t+1) iteration step. After some manipulation [13], we have the estimates of μ_{jk} and Σ_{jk} at the (t+1) step

$$\mu_{jk}^{(t+1)} = \frac{\sum_{i=1}^{N} z_{ij}^{(t)} y_{ijk}^{(t)} \mathrm{x}_i}{\sum_{i=1}^{N} z_{ij}^{(t)} y_{ijk}^{(t)}} - \frac{1}{P_{jk} M} \sum_{m=1}^{M} (\mu_{jk}^{(t)} - \mathrm{s}_{mjk}) \mathrm{H}(\mathrm{s}_{mjk}|\Omega_j) \tag{14}$$

$$\begin{aligned} \Sigma_{jk}^{(t+1)} = \frac{\sum_{i=1}^{N} z_{ij}^{(t)} y_{ijk}^{(t)} (\mathrm{x}_i - \mu_{jk})(\mathrm{x}_i - \mu_{jk})^T}{\sum_{i=1}^{N} z_{ij}^{(t)} y_{ijk}^{(t)}} \\ - \frac{\sum_{m}^{M} ((\mathrm{s}_{mjk} - \mu_{jk}^{(t)})(\mathrm{s}_{mjk} - \mu_{jk}^{(t)})^T - \Sigma_{jk}^{(t)}) \mathrm{H}(\mathrm{s}_{mjk}|\Omega_j)}{P_{jk} M} \end{aligned} \tag{15}$$

where, $\mathrm{s}_{mjk} \sim \Phi(\mathrm{x}|\mu_{jk}^{(t)}, \Sigma_{jk}^{(t)})$ denotes the random vector that is drawn from the probability distribution $\Phi(\mathrm{x}|\mu_{jk}^{(t)}, \Sigma_{jk}^{(t)})$. $P_{jk} \approx \frac{1}{M} \sum_{m=1}^{M} \mathrm{H}(\mathrm{s}_{mjk}|\Omega_j)$. And M is the number of random vectors s_{mjk}. We use $M = 10^6$ for the experiments in this work. The next step is to update the estimate of the prior probability π_j and the weighting factor η_{jk}. The constraint $\sum_{j=1}^{K} \pi_j = 1$ and $\sum_{k=1}^{K_j} \eta_{jk} = 1$ enables

$$\pi_j = \frac{1}{N} \sum_{i=1}^{N} z_{ij}^{(t)} \text{ and } \eta_{jk} = \frac{\sum_{i=1}^{N} z_{ij}^{(t)} y_{ijk}^{(t)}}{\sum_{i=1}^{N} z_{ij}^{(t)} \sum_{m=1}^{K_j} y_{ijm}^{(t)}} \tag{16}$$

So far, the discussion has focused on estimating $\Theta = \{\pi_j, \eta_{jk}, \mu_{jk}, \Sigma_{jk}\}$ of the model. In the next section, we will demonstrate the robustness and accuracy of the proposed model, as compared with other approaches.

4. EXPERIMENTS

In this section, the performance of BAMM is compared to the GMM [1, 2], SMM [5, 7], GGMM [8, 9], EMS [10], and SPM [11].

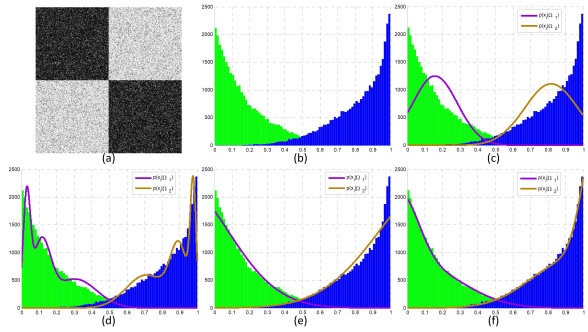

Fig. 1. The first experiment, (a): Noisy image, (b): Ground truth distributions, (c): The estimated distributions of BAMM without bounded support regions $\partial\Omega_j \in (-\infty, +\infty)$ and K_j=1, (d): The estimated distributions of BAMM without bounded support regions $\partial\Omega_j \in (-\infty, +\infty)$ and K_j=3, (e): The estimated distributions of BAMM with $\partial\Omega_j \in (0, 1)$ and K_j=1, (f): The estimated distributions of BAMM with $\partial\Omega_j \in (0, 1)$ and K_j=3.

In the first experiment, in order to explain in the details

why the performance of the proposed distribution $p(\mathrm{x}_i|\Omega_j)$ with bounded support regions is better than the regular distribution, we show a noisy image in Fig. 1(a). The ground truth distributions is shown in Fig. 1(b). In Fig. 1(c), we show the performance of with $K_j = 1$, and without support range $[\partial_{\Omega_j} \in (-\infty, +\infty)]$. Note that, in this case, BAMM will become the GMM model. As shown in Fig. 1(c), the distribution without bounded support range is very poor in this situation. In Fig. 1(d), we use the same conditions as Fig. 1(c). The only difference between them is that we increase the number of the multivariate Gaussian distribution to model each label. In Fig. 1(d), we set $K_j = 3$. As you can see, compared with Fig. 1(c), the estimated distributions of BAMM in Fig. 1(d) is improved very much. However, a closer look at Fig. 1(d) (in the range $0 < \mathrm{x}_i < 0.1$ or $0.9 < \mathrm{x}_i < 1$) clearly shows that it cannot estimate the ground truth distribution. In Fig. 1(e) the result of the proposed model by using the bounded support regions $[\partial_{\Omega_j} \in (0, 1)]$ with $K_j = 1$, is shown. A visual inspection of the result indicates that Fig. 1(e) yields a better result compared to the two previous results in Fig. 1(c) and (d). However, some small portions of pixels have been estimated. As shown in Fig. 1(e), the BAMM by using the bounded support regions $[\partial_{\Omega_j} \in (0, 1)]$ with $K_j = 3$, can better estimate the observed data in comparison to the three previous results.

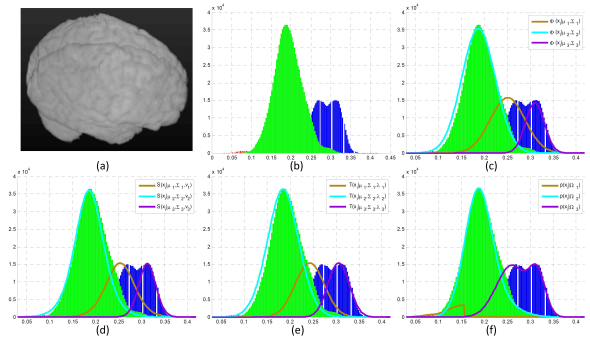

Fig. 2. The second experiment (IBSR18, $3D$), (a): original image, (b): Histogram of the image, (c): GMM (MCR=31.01%), (d): SMM (MCR=31.52%), (e): GGMM (MCR=29.81%), (f): BAMM (MCR=13.36%).

In second experiment, another dataset (IBSR18), as shown in Fig. 2(a), is used. In this experiment, we choose the bounded support regions for GM and WM as [GM: $\partial_{\Omega_2} \in (a, b)$, WM: $\partial_{\Omega_3} \in (a, b)$]. And the bounded support region for CSF is chosen as CSF: $\partial_{\Omega_1} \in (a, c)$. Where c denote the pixel value at position of 10% on the histogram image. The chosen bounded support region for CSF is based on a fact that tissue label CSF is a very small amount compared to GM and WM. As shown in Fig. 2(b), the observed data are always in the bounded support regions. The intensity distribution of each label type of the data does not exhibit exactly a Gaussian shape and are not symmetric. The GMM, SMM, and GGMM methods are very poor in this situation with MCR=31.01%, MCR=31.52% and MCR=29.81%, respectively. As evident from the results, on average, the proposed method outperforms other methods.

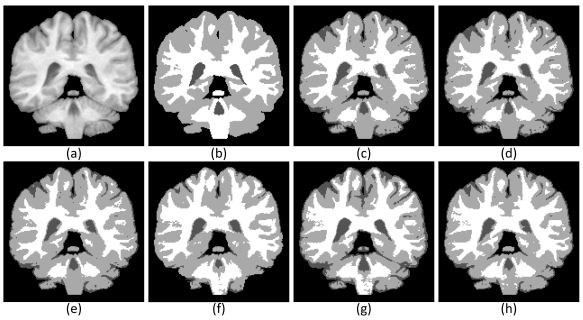

Fig. 3. The third experiment (IBSR02, index=45), (a): Original image, (b) Ground truth image, (c): GMM (MCR=24.42%), (d): SMM (MCR=22.44%), (e): GGMM (MCR=23.49%), (f): EMS (MCR=19.36%), (g): SPM (MCR=22.48%), (h): BAMM (MCR=17.46%).

In the next experiment, real dataset (IBSR02, 256x256x 128) is used. Fig. 3(a) shows one example slice (index=45) of this $3D$ image. The objective is to segment the $3D$ image into three labels: CSF, GM and WM. The image shown in Fig. 3(b) is the ground truth of the original image. Fig. 3(c) to Fig. 3(h) show the results obtained by implementing GMM, SMM, GGMM, EPS, SPM, and BAMM. A low contrast between the GM and the CSF increases the complexity of the image. The segmentation accuracy of GMM in Fig. 3(c) is quite poor. SMM, and GGMM in Fig. 3(d)-(e) slightly improve the segmentation result. EMS (default settings: Bias-order=4, MRF=yes, $3D$), and SPM (default settings:SPM8) methods in Fig. 3(f)-(g) yield a better result with a lower MCR compared to the two previous methods. We rank in the top among these methods with the lowest MCR.

5. RELATION TO PRIOR WORK AND CONCLUSIONS

We introduce a new bounded asymmetric distribution, which has the flexibility to fit different shapes of observed data such as non-Gaussian, non-symmetric, and bounded support data. Our model can be used for analyzing both univariate and multivariate data. Each component of the proposed model has the ability to model the observed data with different bounded support regions. The proposed distribution has the flexibility to fit different shapes of observed data. We also present a new way to estimate the model parameters in order to minimize the higher bound on the data negative log-likelihood function. The proposed method is compared to other finite mixture models for medical image segmentation, and demonstrates a competitive performance in terms of quality.

6. REFERENCES

[1] Bishop C. M., "*Pattern Recognition and Machine Learning*", Springer, 2006.

[2] McLachlan G., and Peel D., "*Finite Mixture Models*", New York, Wiley, 2000

[3] Thanh M. N., and Wu Q. M. J, "Gaussian Mixture Model Based Spatial Neighborhood Relationships for Pixel Labeling Problem," *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 42, no. 1, pp. 193–202, 2012.

[4] Carson C., Belongie S., Greenspan H., and Malik J., "Blobworld: Image segmentation using expectation-maximization and its application to image querying," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1026–1038, 2002.

[5] Peel D., and McLachlan G., "Robust Mixture Modeling Using the t Distribution," *Statistics and Computing*, vol. 10, no. 4, pp. 339–348, 2000.

[6] Thanh M. N., and Wu Q. M. J., "Robust Student's-t Mixture Model with Spatial Constraints and Its Application in Medical Image Segmentation," *IEEE Transactions on Medical Imaging*, vol. 31, no. 1, pp. 103–116, 2012.

[7] Shoham S., "Robust clustering by deterministic agglomeration EM of mixtures of multivariate t-distributions," *Pattern Recognition*, vol. 35, no. 55, pp. 1127–1142, 2000.

[8] Allili M. S., Ziou D., Bouguila N., and Boutemedjet S., "Image and Video Segmentation by Combining Unsupervised Generalized Gaussian Mixture Modeling and Feature Selection," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 10, pp. 1373–1377, 2010.

[9] Tarek E., and Nizar B., "Bayesian learning of finite generalized Gaussian mixture models on images," *Signal Processing*, vol. 91, no. 4, pp. 801–820, 2011.

[10] Van Leemput K. V., Maes F., Vandermeulen D., and Suetens P., "Automated model-based tissue classification of MR images of the brain," *IEEE Transactions on Medical Imaging*, vol. 18, no. 10, pp. 897–908, 1999.

[11] Ashburner J., and Friston K. J., "Unified segmentation," *NeuroImage*, vol. 26, no. 3, pp. 839–851, 2005.

[12] Thanh M. N., and Wu Q. M. J, "A Nonsymmetric Mixture Model for Unsupervised Image Segmentation," *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. PP, no. 99, pp. 1–15, 2012.

[13] Hedelin P., and Skoglund J., "Vector Quantization Based on Gaussian Mixture Models," *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 4, pp. 385–401, 2000.

[14] Thanh M. N., Wu Q. M. J., and Ahuja S., "An Extension of the Standard Mixture Model for Image Segmentation," *IEEE Transactions on Neural Networks*, vol. 21, no. 8, pp. 1326–1338, 2010.