Image Segmentation by Using Modified Spatially Constrained Gaussian Mixture Model

SivaNagiReddy Kalli, Dr. Bhanu Murthy Bhaskara

Abstract—For the probabilistic data analysis generally mixture models, graphical models, Markov random fields and hidden Markov models are used. In the above models for density estimation via Gaussian mixture modeling has been successfully applied to image/video segmentation, speech processing and other fields relevant to clustering analysis and Probability density function (PDF) modeling. Finite Gaussian mixture model is usually used in practice and the selection of number of mixture components is a significant problem in its application. In this paper a modified Gaussian mixture model (MGMM) presented which is based on Markov random field (MRF) and a dependable spatial constraint with EM (Expectation -Maximization) was proposed in order to enhance the segmentation performance. In this method making the M step of the expectation maximization (EM) algorithm cannot be directly applied to the prior distribution for maximization of the log likelihood with respect to the corresponding parameters, the proposed method will results fast and accuracy.

Index Terms— MGMM, Probability Density Function, Clustering, Log-likely hood, MRF Distribution, Similarity.

1. INTRODUCTION

For the video segmentation, first the video converted in to frames or images and it is the partition of image pixels into non-overlap clusters containing each pixel with similar attributes is the called segmentation [1]. Image segmentation is an initial step to understanding an image and it is a one of important technology for image processing for many applications such as identifying objects in a scene for object-based measurements like size and shape recognizing objects in a moving scene. There are so many methods have been presented to deal with image segmentation. Although these methods may seem quite different, most of them belong to two categories, one based on clustering pixels and the other edge detection. In both categories the model based (clustering) segmentation algorithms will be more efficient compared to non-parametric methods. The application of clustering methods to image segmentation has the particular characteristic that spatial information should be taken into account. That is, apart from the intensity values, the pixel location must also be used to determine the cluster to which each pixel is assigned. In recent years, many clustering algorithms such as hierarchical clustering, partition based clustering; mixture densities-based clustering and search techniques-based clustering, have been applied to image segmentation problem and achieved the satisfying results [2-5].

- *SivaNagiReddy Kalli is currently pursuing Ph.D degree program in electrics and communication engineering JNTUH, India, E-mail: sivanagireddykalli@gmail.com*
- *Bhanu Murthy Bhaskara, is currently Professor and Vice Dean, faculty of Integrative Science and Technology, Quest International University, Perak, Malaysia. E-mail: bhau.b@gmail.com*

In this paper, we focused on those clustering algorithms based on Gaussian Mixture Model (GMM) [6]. Which is a popular clustering method because of its simple mathematical form and the closed form expressions for its parameters, but it produces an unacceptable segmentation on noise-corrupted image due to no consideration of spatial information. To overcome this drawback, modified algorithm has been proposed by embedding Markov Random Field (MRF) spatial dependence into GMM. The Expectation-Maximization framework constitutes an efficient method for GMM training based on likelihood maximization Following this formulation, a likelihood term which is based exclusively on the data captures the pixel intensity information, while a prior biasing term that uses a Markov Random Field (MRF) captures the spatial location information. In abovementioned models based MRF; 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 rugged with respect to noise, more accurate and effective as compared to other GMM based methods.

The rest of this paper is organized as follows: In section 2 we describe The GMM Formulation based on MRF. In section 3 we present our modified method. In section 4 provide experimental results and finally in section 5 explain conclusion.

2. THE GMM FORMULATION BASED ON MRF

The value of a pixel in an image (i.e. the intensity or the color) can be taking as a random variable. Since every

random variable has a probability distribution then pixel values also have probability distribution. The Gaussian mixture distribution is a good probability distribution for pixel values of an image. Let x_i, $i = (1, 2. . . N)$, denote an observation at the ith pixel of an image with dimension D. The neighborhood of the ith pixel is presented by ∂i. Labels are denoted by $(1, 2. . . K)$. in order to partition an image consisting of N pixels into K labels, GMM [10] assumes that each observation xi is considered independent of the label j. The probability density function f $(x_i | \Pi, \Theta)$ at an observation x_i is given by

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

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 xi is in label j, which satisfies the constraints

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

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

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

Where $j = \{\mu j, \Sigma j\}$, $j = (1, 2. . . K)$. The D-dimensional vector μ_j is the mean, the $D \times D$ matrix j is the covariance, and $| \Sigma j|$ denotes the determinant of Σj. Note that the observation xi in (1) is modeled as statistically independent, the joint conditional density of the data set $X = (x1, x2, \dots, x_N)$ can be modeled as

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

From the observation xi 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 [10]. To overcome this problem, MRF distribution [9] is applied to incorporate the spatial correlation among label values

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

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

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

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

$$L(\Pi,\Theta|X) = log(p(\Pi,\Theta|X))$$
$$= \sum_{i=1}^{N} log\left\{\sum_{j=1}^{k}\pi_{ij}\Phi(x_i|\Theta_j)\right\} + logp(\Pi)$$
$$= \sum_{i=1}^{N} log\{\sum_{j=1}^{k}\pi_{ij}\Phi(x_i|\Theta_j)\} - logZ - \frac{1}{T}U(\Pi) \qquad (7)$$

Depending on the type of energy U (Π) selected in (7), from the Bayesian kinds of model [9], [14], the function $U(\Pi)$ is chosen to incorporate the spatial correlation as

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

Where α_{ij} and β_{ijm} form the parameter set. In this model, in order to maximize the log-likelihood function, we need to optimize many parameters and disadvantage is that its segmentation is not sufficiently robust to noise. Other mixture models based on MRF have been successfully applied to image segmentation [10], [11] and different ways are adopted to select the energy U (Π). From [10], Z and T in (7) are set to one, and U (Π) is given by

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

While in an another MRF model-based method [30], U (Π) is given by

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

Where β is constant value in (9) and (10).

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

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

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). β_{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 *and*, 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 πij. Note that the prior distribution πij 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 πij.

3. PROPOSED METHODOLOGY

Various mixture models differ based on the way they derive the strength of the smoothing prior U (Π). In [11], given in (8), the smoothing prior U (Π) has a simple form, thus, it is easy to optimize the parameter set {Π, Θ} to maximize the log-likelihood function. However, one of its main drawbacks is that the segmentation result is not robust to noise. Models in [10], [11], 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 C_{ij} defined as

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

Where z_{mj} is the posterior probability and β is the temperature value that controls the smoothing prior. In this paper, it has been set to 12 (β = 12). ∂i is the neighborhood of the ith pixel, including itself. A square window of size 5 × 5 is used in this paper. Ni is the number of neighboring pixels around the pixel xi in this window (Ni = 25). By taking a closer look at (12), it can be visualized that the factor C_{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 π_{ij} in an image with the average value of their neighbors, including themselves. Note that the factor C_{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 C_{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 (Π) is given by

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

The intuition of (13) that the derivative of the smoothing prior $U(\Pi)$ with respect to prior distribution π_{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 (Π) in (5) is given by

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

Given the MRF distribution p(Π) , the log-likelihood function in (7) is written in the form as

$$L(\Pi,\Theta|X) = \sum_{i=1}^{N} log\left\{\sum_{j=1}^{K}\pi_{ij}^{(t+1)}\Phi\left(x_i|\Theta_j^{(t+1)}\right)\right\} - logZ + 1Ti=1Nj=1KGij(t)log\pi ij(t+1) \qquad (15)$$

Application of the complete data condition in [10], maximizing the log-likelihood function $L(\Pi,\Theta\ |X)$ in (15) will lead to an increase in the value of the objective function

$$J(\Pi,\theta\ |X)\ = J(\Pi,\Theta|X) = \sum_{i=1}^{N}\sum_{j=1}^{K} z_{ij}^{(t)}\left\{log\,\pi_{ij}^{(t+1)} + \right.$$
$$\left. log\,\Phi xi/\Theta jt+1 - logZ+1Ti=1Nj=1KGij(t)log\pi ij(t+1) \qquad (16)\right.$$

The conditional expectation values zij of the hidden variables can be computed as follows:

$$z_{ij}^{(t)} = \frac{\pi_{ij}^{(t)}\Phi\left(x_i|\Theta_j^{(t)}\right)}{\sum_{j=1}^{K}\pi_{ij}^{(t)}\Phi\left(x_i|\Theta_j^{(t)}\right)} \qquad (17)$$

The next objective is to optimize the parameter set {Π, Θ} in order to maximize the objective function $J(\Pi,\Theta\ |X)$ in (16). Similar to the MRF-based methods [10], [11], Z and T in (16) are set proportional to one. From (16), the new objective function is given by

$$J(\Pi,\Theta|X) = \sum_{i=1}^{N}\sum_{j=1}^{K} z_{ij}^{(t)}\left\{log\,\pi_{ij}^{(t+1)} + log\,\Phi\left(x_i|\Theta_j^{(t+1)}\right)\right\} +$$
$$i=1Nj=1KGij(t)log\pi ij(t+1) \qquad (18)$$

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

$$J(\Pi,\Theta|X) = \sum_{i=1}^{N}\sum_{j=1}^{K} z_{ij}^{(t)}\left\{log\,\pi_{ij}^{(t+1)} - \frac{D}{2}log(2\pi) - \right.$$
$$12log\Sigma jt+1 \qquad +$$
$$i=1Nj=1Kzij(t)-12xi-\mu jt+1T\Sigma i-1(t+1)xi-\mu jt+1+i=1Nj=1$$
$$KGijtlog\pi ijt+1 \qquad (19)$$

To maximize this function, the EM algorithm [10], [11], [12], [13] is applied. Let us now consider the derivation of the function J (Π, Θ |X) with the means μj at the (t+1) iteration step. We have

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

The solution of $\partial J/\partial\mu j = 0$ yields the minimize of μj at the (t+1) step

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

Thus, setting the derivative of the function in (18) with respect to Σ_j^{-1} at the (t+1) iteration step, we have

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

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

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 η_i for each data point

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

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

$$\frac{z_{ij}^{(t)}}{\pi_{ji}^{(t+1)}} + \frac{C_{ij}^{(t)}}{\pi_{ji}^{(t+1)}} - \eta_i = 0 \tag{25}$$

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

$$\eta_i = 1 + \sum_{j=1}^{K} C_{ij}^{(t)} \tag{26}$$

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

$$\pi_{ji}^{(t+1)} = \frac{z_{ij}^{(t)} + C_{ij}^{(t)}}{\sum_{k=1}^{K} \left(z_{ij}^{(t)} + C_{ij}^{(t)} \right)} \tag{27}$$

So far, the discussion has focused on estimating {Π, Θ} of the model in order to assign a label Σ j to the pixel xi. The various steps of the proposed mixture model incorporating spatial information based on MRF can be summarized as follows.

Step 1: *Initialize the parameters* {Π, Θ}: the means μj, *covariance* values Σj, and prior distributions π_{ij}.
Step 2: E *step*. a) Evaluate the values z_{ij} in (17) using the current parameter values. b) Update the factor C_{ij} by using (12).
Step 3: *M step*: Re-estimate the parameters {Π, Θ}. a) Update the means μj by using (21). b) Update covariance values Σ j by using (23). c) Update prior distributions π_{ij} by using (27).

Step 4: Evaluate the log-likelihood in (14) 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 xi is assigned to the label with the largest posterior probability z_{ij}

$X_i \in \Omega$ j: IF $z_{ij} \geq z_{ik}$; j, k = (1, 2... K).
(28)

4. EXPERIMENTAL RESULTS

For the performance analysis fist we have to find out specificity analysis with respect to the ground truth image depend on True-positive (tp) pixels, True-negative pixels (tn), False-positive pixels (fp), and False-negative pixels (fn). True-positive pixels (tp) are the correctly detected pixels by the algorithm of the moving object. By using sensitivity values we can find out the following parameters.

The relevant pixels of the detected object can be found out by using recall formula it is given below:

$$Recall = \frac{tp}{tp + tn} \tag{29}$$

Irrelevant pixels can be determined by using precision, the formula for precision is given below:

$$Precision = \frac{tp}{tp + fp} \tag{30}$$

$$Similarity = \frac{tp}{tp + fp + fn} \tag{31}$$

$$False\ Measure\ (FM): \frac{2\ Recall * Precision}{Recall + Precision} \tag{32}$$

Percentage of correct classification (*PCC*):

$$PCC = \frac{tp + tn}{tp + fp + fn + tn} \tag{33}$$

True positive rate and True negative Rate:

$$Precision = \frac{tp}{tp + fn} \tag{34}$$

$$Recall = \frac{fp}{fp + tn} \tag{35}$$

Table 1 shows the performance of algorithm

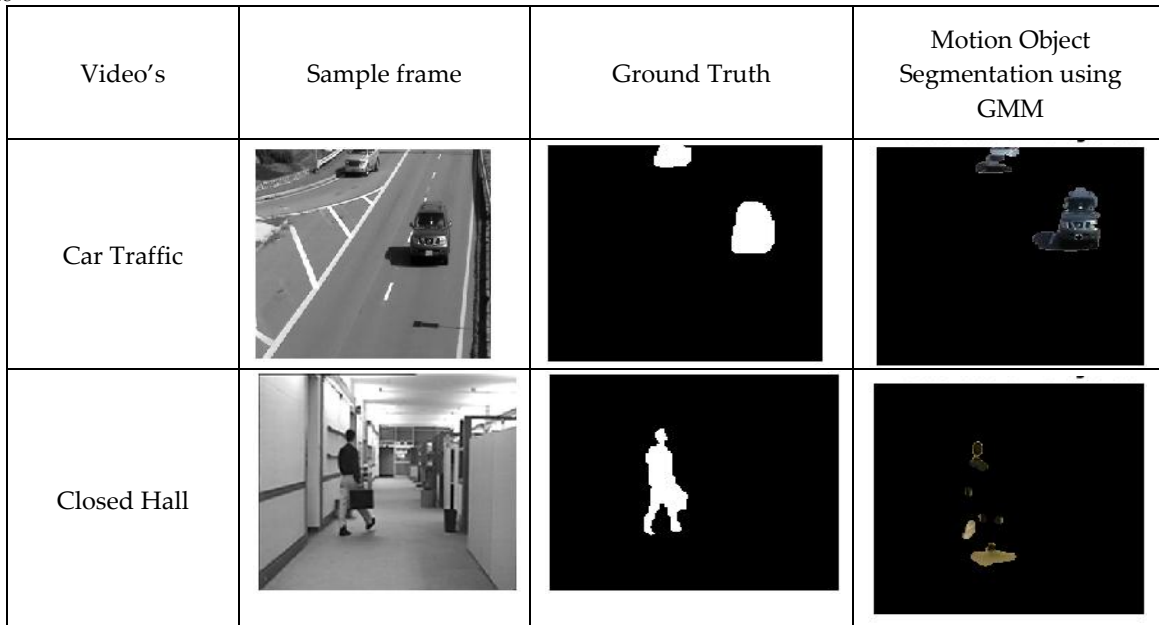| Video's	Sample frame	Ground Truth	Motion Object Segmentation using GMM
Car Traffic			
Closed Hall			

Fig.1. Motion mask generated by the proposed method

Sequences	Evaluation	MGMM
Traffic	Recall	0.9039
	Precision	0.5316
	Similarity	0.502
	FM	0.668
	PCC	0.9773
	TRP	0.9039
	FRP	0.8892
Hall	Recall	0.5384
	Precision	0.6356
	Similarity	0.4155
	FM	0.5744
	PCC	0.9752
	TRP	0.5384
	FRP	0.4046

TABLE 1

PERFORMANCE ANALYSIS WITH DEFENDING VIDEOS

5. CONCLUSION

In this paper we presented a new mixture model for image segmentation that incorporated the spatial relationships based on MRF. The proposed method directly applied the EM algorithm to optimize the parameters by making it simple, fast, and easy to implement and it is applied real world videos, thereby producing excellent performance in noisy conditions.

REFERENCES

[1] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", Addison-Wesley Longman, Boston, MA, USA, 2002.

[2] Songcan Chen and Daohong Li, "Image Binarization Focusing on Objects", Neural Computing,vol. 69, no. 16-18, pp. 2411-2415, 2006.

[3] Hassana Grema Kaganami, Zou Beiji, M Sami Soliman, "Optimal Color Image Enhancement Using Wavelet and K-means Clustering", JDCTA, vol. 5, no. 1, pp. 112- 122, 2011.

[4] Ahmed M.Nazif, Martin D. Levine, "Low Level Image Segmentation: An Expert System", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 6, no. 5, pp. 555-577, 2009.

[5] Jae Hyup Kim, Kyung Hyun Jang, Young Shik Moon, "Image Segmentation Method Using Hybrid Region Flow", JCIT, vol. 3, no. 2, pp. 39-44, 2008.

[6] Geoffrey McLachlan, David Peel, "Finite Mixture Models", Wiley, New York, USA, 2000.

[7] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

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

[9] S. Z. Li, *Markov Random Field Modeling in Image Analysis*. New York: Springer-Verlag, 2009.

[10] 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.

[11] K. Blekas, A. Likas, N. P. Galatsanos, and I. E. Lagaris, "A spatially constrained mixture model for image segmentation,"

IEEE Trans. Neural Networks., vol. 16, no. 2, pp. 494–498, Mar. 2005.

[12] G. J. McLachlan and T. Krishnan, "*The EM Algorithm and Extensions*" (Wiley Series in Probability and Statistics). New York, Wiley, 1997.

[13] 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.

[14] J. Besag, "spatial interaction and the statistical analysis of lattice systems," *j. Royal statist. Soc. Ser. B*, vol. 36, pp. 192–236, jan.1974.

IJSER