Image Segmentation Using Inverted Dirichlet Mixture Model and Spatial Information

Jai Puneet Singh *, Nizar Bouguila
CIISE, Concordia University, Montreal, Quebec, Canada
jaipuneet.singh@mail.concordia.ca,nizar.bouguila@concordia.ca

Abstract—**In this work, we present an unsupervised algorithm for Image segmentation using Inverted Dirichlet Mixture Model. The proposed approach uses image segmentation algorithm based on spatial information with the Inverted Dirichlet mixture model is presented. This method uses Markov Random Field to incorporate spatial information between neighboring pixels into a Inverted Dirichlet mixture model. The segmentation model is learned using Expectation Maximization (EM) algorithm based on Newton Raphson step. The obtained results using real image data set are more encouraging than those obtained using similar approaches.**

Keywords: Image Segmentation, Dirichlet, Mixture Model, Markov Random Field.

I. Introduction

The problem of image segmentation and grouping based on regions has remained as a great challenge in the field of computer vision. It is often the first step in variety of computer vision and image analysis tasks. There are various types of image we encounter in everyday life, for example, light intensity images, magnetic resonance image, etc [1]. There has been a large number of approaches proposed in previous years for image segmentation. The problem of image segmentation of noisy images or corrupt images is still an open challenge. Image segmentation is widely used for anomaly detection [2] [3] and medical image analysis [4]. Various statistical models have been proposed in the past.

In this paper, we develop an unsupervised approach for noisy image segmentation. Earlier similar unsupervised approaches have been used in research, for example, the author in [5] has used fuzzy c-means for medical image segmentation. In this paper, we had the focus on a particular statistical model which is finite mixture. The main problems faced were 1) To integrate Markov Random Field (MRF) with Inverted Dirichlet mixture model to achieve the segmentation of noisy images, 2) The estimation of parameters which is often difficult task in mixture models, and 3) Initialization of parameters where we have used moments method with k-means. Markov Random Field has been heavily used for modeling spatial information for medical image segmentation [6]. Markov Random Field is heavily used for semantic segmentation of images in supervised learning approaches which is defined as multi-label classification problem. An interesting approach to integrate spatial information has been proposed in [7] to integrate with Gaussian distribution. Other models has been also proposed such as Markov Random Field with Dirichlet [8] and generalized Dirichlet [9] distribution. As we know, Gaussian mixture is a popular model in the field of computer vision. The Gaussian mixture model is very restrictive model as we have seen from previous works [10] [11] where Dirichlet, and Inverted Dirichlet distributions have always shown better results.

To the best of our knowledge, the Inverted Dirichlet Distribution has never been considered with the Random Field for Image Segmentation. The main purpose of this paper is to show the practicality of a distribution, we had considered the Inverted Dirichlet Distribution for its flexibility, in contrast to the Gaussian distribution which permits multiple symmetric and asymmetric mode as shown in [12].

Hence, In this paper, we propose the integration of Markov Random Field in Inverted Dirichlet mixture model as this mixture model is flexible (i.e it is symmetric and asymmetric) [10] for data modeling. Experiments show that integrating spatial information into Inverted Dirichlet mixture model gives better results for noisy image segmentation.

The rest of the paper is organized as follows: In section II, the segmentation approach is explained in details giving the equations and describing the algorithms. In section III, experimental results with the proposed approach are shown. Finally, section IV gives the concluding remarks.

II. The Segmentation approach

In the following, we adopt the segmentation approach, based on Gaussian mixture models with Markov Random Field (MRF) proposed in [13] for the introduction of spatial information. This approach can be explained as follows. For each pixel $\vec{X}_i \in X$, there is a peer which has been arisen from the same cluster of $\vec{X}_i$. This spatial information can be used indirectly to estimate the number of clusters or regions in an image. Let $\mathcal{X} = \left\{\vec{X}_1, \vec{X}_2, ..., \vec{X}_N\right\}$ where each pixel is denoted by random vector $\vec{X}_i = (X_{i1}, X_{i2}, ..., X_{iD})$ and N is the number of pixels. Now, the random vector $\vec{X}_i$ follows Dirichlet mixture model and is considered to be independent from the label. The density function can be presented as:

$$p\left(\vec{X}_i|\theta\right) = \sum_{j=1}^{M} p_j p\left(\vec{X}_i|\vec{\alpha}_j\right) \tag{1}$$

where $\vec{\alpha}_j$ is the parameter vector of component j which can be represented as $\vec{\alpha}_j = (\alpha_{j1}, \alpha_{j2}, ..., \alpha_{jM})$. $\{p_j\}$ are the mixing proportions which should be positive and always sum to 1. $\theta = \{p_1, p_2, ..., p_M; \vec{\alpha}_1, \vec{\alpha}_2, ..., \vec{\alpha}_M\}$ is the complete set of parameters fully characterizing the mixture, $M \geq 1$ is the number of components.

$$p\left(\vec{X}_i|\vec{\alpha}_j\right) = \frac{\Gamma|\vec{\alpha_j}|}{\prod_{d=1}^{D+1}\Gamma\left(\alpha_{jd}\right)} \prod_{d=1}^{D} X_{id}^{\alpha_{jd}-1} \left(1+\sum_{d=1}^{D} X_{id}\right)^{-|\vec{\alpha}|} \tag{2}$$

where $X_{id} > 0$, $d = 1, 2, ..., D$, $X_{i1} + X_{i2}, ... + X_{iD} = 1$, $\vec{\alpha}_{jd} = (\alpha_{j1}, \alpha_{j2}, ..., \alpha_{jD})$, $|\vec{\alpha_j}| = \sum_{d=1}^{D+1} \alpha_{jd}$ and $\alpha_{jd} > 0$ represents parameter vector for j^{th} component. Let $\mathcal{X} = \left\{\vec{X}_1, \vec{X}_2, ..., \vec{X}_N\right\}$ be a data set of N D-dimensional positive vectors with a common, but unknown, probability density function $p(\vec{X}_i|\theta)$ as given in above equation. The vectors are modeled as statistically independent, the joint conditional density of data can be presented as:

$$f\left(\mathcal{X}|\theta\right) = \prod_{i=1}^{N} p\left(\vec{X}_i|\theta\right) = \prod_{i=1}^{N}\sum_{j=1}^{M} p_j p\left(\vec{X}_i|\vec{\alpha}_j\right) \tag{3}$$

As we have taken each pixel to be independent of pixel label, the spatial correlation between nearby pixels is not taken into an account. So, in this case, the image is relatively very sensitive to noise and illumination [14]. To overcome the problem of noise and illumination the MRF (Markov Random Field) distribution is used. It is used to create the spatial correlation between label values. The MRF distribution can be shown as:

$$f\left(\Pi\right) = Z^{-1} \exp\left\{-\frac{1}{T} U\left(\Pi\right)\right\} \tag{4}$$

In MRF, Z is considered as normalizing constant, T is a Temperature constant and $U\left(\Pi\right)$ is the smoothing prior where $\Pi = p_{ij}$. The Bayes rule for posterior probability can be represented as:

$$f\left(\theta|\mathcal{X}\right) \propto f\left(\mathcal{X}|\theta\right) f\left(\Pi\right) \tag{5}$$

The log likelihood is given as follow:

$$\begin{aligned} L\left(f\left(\theta|\mathcal{X}\right)\right) &= \log\left(f\left(\theta|\mathcal{X}\right)\right) \\ &= \sum_{i=1}^{N} \log\left\{\sum_{j=1}^{M} p_j p\left(\vec{X}_i|\vec{\alpha}_j\right)\right\} + \log f\left(\Pi\right) \\ &= \sum_{i=1}^{N} \log\left\{\sum_{j=1}^{M} p_j p\left(\vec{X}_i|\vec{\alpha}_j\right)\right\} - \log Z - \frac{1}{T} U\left(\Pi\right) \end{aligned}$$

(6)

There has been vast research which has already been conducted for determining smoothing prior of MRF distribution. The smoothing prior determined in most research is complex and requires lot of computation time when combined with mixture models. The example of such kind of smoothing prior is given by [15] which can be shown as follow:

$$U\left(\Pi\right) = \beta \sum_{i=1}^{N} \sum m \in \delta_i \left[1 + \left(\sum_{j=1}^{M} \left(p_{ij} - p_{mj}\right)^2\right)^{-1}\right]^{-1} \tag{7}$$

where β in this equation represents a constant value. In the above equation, Z and T are set to 1 ($Z = 1$ and $T = 1$). Due to the complexity of this equation, the M-step of EM algorithm cannot be applied directly to prior distribution p_{ij}. Various smoothing priors were proposed but the major drawback of all of them has been that they are not robust to noise. In order to overcome this difficulty prior distribution has been considered. A novel factor was proposed by [13] as follows:

$$G_{ij}^{t} = \exp\left[\frac{\beta}{2N_i} \sum \left(z_{mj}^{(t)} + p_{mj}^{(t)}\right)\right] \tag{8}$$

In this equation β is the temperature value and hence, changing the temperature value determines noise reduction of the image. It is used to determine neighborhood pixels around the pixel X_i. As proposed by authors [13] G_{ij} is only dependent on value of posteriors at previous step (t) and priors value.

The smoothing prior has been proposed in order to overcome the deficiencies which were seen previously in smoothing prior. Hence, smoothing prior is given by:

$$U\left(\Pi\right) = -\sum_{i=1}^{N}\sum_{j=1}^{M} G_{ij}^{(t)} \log p_{ij}^{(t+1)} \tag{9}$$

Maximizing equation (7) we get the expanded equation with the hidden variable z_{ij}

$$\begin{aligned} L\left(f\left(\theta|\mathcal{X}\right)\right) = \sum_{i=1}^{N}\sum_{j=1}^{M} z_{ij}^{t} \Big\{\log p_{ij}^{(t+1)} \\ + \log p\left(\vec{X}_i|\vec{\alpha}_j\right) - \log Z + \frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{M} G_{ij}^{t} \log p_{ij}^{(t+1)} \end{aligned} \tag{10}$$

The hidden variable can be expanded as

$$z_{ij}^{(t)} = \frac{p_{ij}^{(t)} p\left(\vec{X}_i|\vec{\alpha}_j\right)}{\sum_{k=1}^{K} p_{ik}^{(t)} p\left(\vec{X}_i|\vec{\alpha}_k^{t}\right)} \tag{11}$$

Hence, putting the value of $U\left(\Pi\right)$ in equation (9) and expanding the equation with Dirichlet distribution as well as setting normalizing constant Z and Temperature Value T to proportional over here (Z=1 and T=1) we get:

$$\begin{aligned} Q\left(f\left(\theta|\mathcal{X}\right)\right) = \sum_{i=1}^{N}\sum_{j=1}^{M} z_{ij}^{t} \Big\{\log p_{ij}^{(t+1)} + \log\Gamma|\vec{\alpha}| - \\ \sum_{d=1}^{D} \log\Gamma\left(\alpha_{jd}^{t+1}\right) + \sum_{d=1}^{D}\left(\alpha_{jd}^{t+1} - 1\right)\log X_{id} + \\ \left(-|\vec{\alpha}|\right)\log\left(1+\sum_{d=1}^{D} X_{id}\right) + \sum_{i=1}^{N}\sum_{j=1}^{M} G_{ij}^{(t)} \log p_{ij}^{(t+1)} \end{aligned} \tag{12}$$

Now, In M-Step of Expectation Maximization algorithm, $Q\left(\theta|\vec{X}\right)$ is maximized using Newton-Raphson approach as proposed in [16]. Hence, for the $\vec{\alpha}$ parameters we have:

$$\alpha_{jd}^{(t+1)} = \alpha_{jd}^{(t)} - H^{-1}\left(\alpha_{jd}^{(t)}\right) \times \left(\frac{\partial Q\left(\theta|\vec{X}\right)}{\partial \alpha_{jd}}\right)^{T} \tag{13}$$

In the above equation H is the Hessian matrix which requires the calculation of second and mixed derivatives as presented in [16]. To satisfy the condition of $\sum_{j=1}^{M} p_j = 1$, we use the Lagrangian multiplier Λ. Using the above methods of prior probability which gives:

$$p_{ij} = \frac{z_{ij}^{(t)} + G_{ij}^{(t)}}{\sum_{m=1}^{k}\left(z_{im}^{(t)} + G_{ik}^{(t)}\right)} \tag{14}$$

Now, we have done the integration of MRF into Dirichlet mixture model. Hence, we can see from the equation that MRF distribution is affecting the prior distribution which indirectly affects the estimation of the parameters. The algorithm is summarized in the following as:

Algorithm 1 EM Algorithm Inverted Dirichlet Mixture Model with MRF

1: Apply K-means [17] on image data points to obtain initial k clusters for segmentation.
2: The initial estimates for each mixture component j:
3: $\alpha_{jD+1} = \frac{E(X_d)^2+E(X_d)}{Var(X_d)} + 2$ and $\alpha_{jd} = E\left(X_d\right)\left(\alpha_{jD+1}-1\right)$,$d = 1, ..., D$.
4: Initialization of above equation using Method of Moments as proposed by the author in [12] to obtain the two parameters.
5: Use the image data points to update the mixture parameters.
6: E-Step: Compute the posterior probability $z_{ij}^{(t)}$
7: M-Step:
8: **repeat**:
9: Update priors p_j using equation 15 .
10: Update the parameters $\vec{\alpha}$ using Newton Raphson method [16].
11: **until** : $p_j \leq \epsilon$, discard j and go to E-Step.
12: if convergence test is passed then terminate, else go to E-Step.

III. Experimental results

The main goal of this section is to investigate the performance of proposed method of Inverted Dirichlet mixture model with Markov Random Field as compared with one developed by [8]. As in this paper, we have considered Inverted Dirichlet model. Evaluating segmentation results is an important problem and over here we are using NPR (Normalized Probabilistic Rand) [18] which can be given as follows:

$$\text{NPR Index} = \frac{\text{PR Index - Expected Index}}{\text{Maximum Index - Expected Index}} \tag{15}$$

The Expected value of PR Index can be given as follow:

$$E\left[PRI\left(S_{test}, \{S_k\}\right)\right] = \frac{1}{\binom{N}{2}} \sum_{\substack{i,j\\ i<j}} \left[p'_{ij}p_{ij}\left(1-p'_{ij}\right)\left(1-p_{ij}\right)\right] \tag{16}$$

This comparison model was proposed in [18] in order to provide comparison between image segmentation algorithms. In the experiment, the image is taken and converted to grayscale after that we have induced three types of noise in an image which are: Gaussian noise, Poisson Noise and Salt and Pepper. The Gaussian noise image denoising was earlier proposed by [19] who used soft threshold shrinkage method of sparse components. Adaptive median filter with specialized regularization method have been used to reduce salt and pepper impulsive noises [20]. We have observed that our method performs very well on both type of cases and image segmentation takes place without any difficulty and even the proposed method is better than median based filters.

In our experiment conducted, we have set the temperature value ($\beta = 10$). The other important factor in this equation is the determination of window size ($N_i = 25$). The data set used is Visual StoryTelling Dataset [21] which contains 13100 images, this dataset was made available to be used as sequential to language for visual storytelling. These are publicly available data sets and heavily used in the field of computer vision. It is difficult to calculate NPR index of every image as this process is really expensive so we have calculated NPR index of a limited number of images followed by different segmentation approaches. The experimental results show that there is a large difference between the two different mixture models used. Table 1 shows the NPR Index sample mean of images by different mixture models being performed on images. It can be seen that integrated model with Inverted Dirichlet performs way better than Modified Dirichlet mixture model. Fig 1 shows the image segmentation results obtained after applying different approaches on noisy images.

	DMM	IDMM	FRDMM	FRIDMM
NPR Index Sample Mean	0.4247	0.4390	0.6182	0.6793

TABLE I: NPR index sample for Dirichlet Mixture model (DMM), Inverted Dirichlet Mixture Model, Fast and Robust Dirichlet mixture model (FRDMM) and Fast and Robust Inverted Dirichlet mixture model (FRIDMM)

IV. Conclusion

In this paper, we performed image segmentation based on Inverted Dirichlet mixture model with MRF (Markov Random field) to integrate the spatial information. It gave us good results when compared with Dirichlet Mixture model. The selection of mixture model is motivated by its excellent results obtained when compared with other methodologies used in the past. The work can be extended for video segmentation being considered as another important application. The drawback of a mixture model is the initialization of parameters. If parameters are not initialized properly, it can give us bad results hence, proper initialization of parameters of Inverted Dirichlet mixture model helps us to improve further the results.

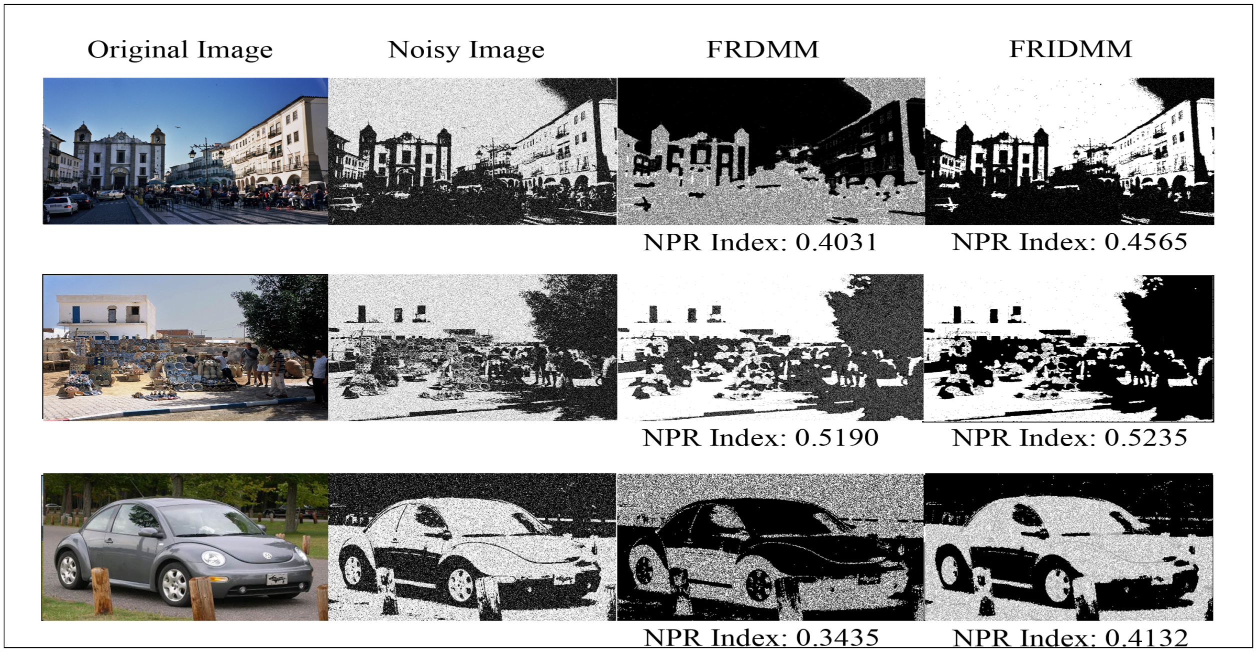

Fig. 1: Segmentation of images from Berkeley 500 database. Column 1 gives the original image, column 2: Noisy Image, Column 3: Segmentation with Markov Random Field with Dirichlet mixture model and Column 4: Proposed method with Inverted Dirichlet mixture model.

V. Acknowledgment

The authors would like to thank Mitacs Inc. for providing the research support and Concordia University for providing the facilities to conduct a research.

References

[1] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern recognition*, vol. 26, no. 9, pp. 1277–1294, 1993.

[2] M. Brand and V. Kettnaker, "Discovery and segmentation of activities in video," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 844–851, 2000.

[3] N. Bouguila and W. ElGuebaly, "A generative model for spatial color image databases categorization," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008, pp. 821–824.

[4] D. L. Pham, C. Xu, and J. L. Prince, "Current methods in medical image segmentation 1," *Annual review of biomedical engineering*, vol. 2, no. 1, pp. 315–337, 2000.

[5] K.-S. Chuang, H.-L. Tzeng, S. Chen, J. Wu, and T.-J. Chen, "Fuzzy c-means clustering with spatial information for image segmentation," *computerized medical imaging and graphics*, vol. 30, no. 1, pp. 9–15, 2006.

[6] H. Greenspan, A. Ruf, and J. Goldberger, "Constrained gaussian mixture model framework for automatic segmentation of mr brain images," *IEEE transactions on medical imaging*, vol. 25, no. 9, pp. 1233–1245, 2006.

[7] M.-S. Yang and H.-S. Tsai, "A gaussian kernel-based fuzzy c-means algorithm with a spatial bias correction," *Pattern recognition letters*, vol. 29, no. 12, pp. 1713–1725, 2008.

[8] J. P. Singh and N. Bouguila, "Spatially constrained non-gaussian mixture model for image segmentation," in *Electrical and Computer Engineering (CCECE), 2017 IEEE 30th Canadian Conference on*. IEEE, 2017, pp. 1–4.

[9] J. Singh and N. Bouguila, "Spatially constrained generalized dirichlet mixture model for image segmentation," in *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*. IEEE, Aug 2017, pp. 140–143.

[10] A. Sefidpour and N. Bouguila, "Spatial color image segmentation based on finite non-gaussian mixture models," *Expert Systems with Applications*, vol. 39, no. 10, pp. 8993–9001, 2012.

[11] N. Bouguila and D. Ziou, "Dirichlet-based probability model applied to human skin detection [image skin detection]," in *Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04). IEEE International Conference on*, vol. 5. IEEE, 2004, pp. V–521.

[12] T. Bdiri and N. Bouguila, "Positive vectors clustering using inverted dirichlet finite mixture models," *Expert Systems with Applications*, vol. 39, no. 2, pp. 1869–1882, 2012.

[13] T. M. Nguyen and Q. J. Wu, "Fast and robust spatially constrained gaussian mixture model for image segmentation," *IEEE transactions on circuits and systems for video technology*, vol. 23, no. 4, pp. 621–635, 2013.

[14] S. Sanjay-Gopal and T. J. Hebert, "Bayesian pixel classification using spatially variant finite mixtures and the generalized em algorithm," *IEEE Transactions on Image Processing*, vol. 7, no. 7, pp. 1014–1028, 1998.

[15] K. Blekas, A. Likas, N. P. Galatsanos, and I. E. Lagaris, "A spatially constrained mixture model for image segmentation," *IEEE Transactions on Neural Networks*, vol. 16, no. 2, pp. 494–498, 2005.

[16] N. Bouguila, D. Ziou, and J. Vaillancourt, "Novel mixtures based on the dirichlet distribution: application to data and image classification," in *International Workshop on Machine Learning and Data Mining in Pattern Recognition*. Springer, 2003, pp. 172–181.

[17] J. P. Singh and N. Bouguila, "Proportional data clustering using k-means algorithm: A comparison of different distances," in *Industrial Technology (ICIT), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1048–1052.

[18] R. Unnikrishnan, C. Pantofaru, and M. Hebert, "Toward objective evaluation of image segmentation algorithms," *IEEE transactions on pattern analysis and machine intelligence*, vol. 29, no. 6, pp. 929–944, 2007.

[19] A. Hyvärinen, P. O. Hoyer, and E. Oja, "Sparse code shrinkage: Denoising by nonlinear maximum likelihood estimation," *Advances in Neural Information Processing Systems*, pp. 473–479, 1999.

[20] R. H. Chan, C.-W. Ho, and M. Nikolova, "Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization," *IEEE Transactions on image processing*, vol. 14, no. 10, pp. 1479–1485, 2005.

[21] F. Ferraro, N. Mostafazadeh, I. Misra, A. Agrawal, J. Devlin, R. Girshick, X. He, P. Kohli, D. Batra, C. L. Zitnick *et al.*, "Visual storytelling," *arXiv preprint arXiv:1604.03968*, 2016.