A New Modified Gaussian Mixture Model for Color-Texture Segmentation

[1]M. Sujaritha and [2]S. Annadurai
[1]Department of Computer Science and Engineering,
J.J. College of Engineering and Technology, Trichy, India
[2]Directorate of Technical Education, Additional Director (Polytechnic), Chennai, India

Abstract: Problem statement: This study presents a new, simple and efficient modified Gaussian mixture model based clustering algorithm for color-texture segmentation. The proposed mixture model introduces a new component density function which incorporates spatial information and the weighting factor for neighborhood effect is fully adaptive to the image content. **Approach:** It enhances the smoothness towards piecewise-homogeneous segmentation and reduces the edge-blurring effect. An Expectation Maximization (EM) model fitting Maximum A Posteriori (MAP) algorithm segments the image by utilizing the pixel's color and texture features and the captured neighborhood relationships among them. **Results:** The algorithm simultaneously calculates the model parameters and segments the pixels iteratively in an interleaved manner. Finally, it converges to a solution where the model parameters and pixel labels are stabilized within a specified criterion. **Conclusion:**The experimental results with synthetic and natural images demonstrate that the proposed method is effective in improving the segmentation and it outperforms the Fusion of Clustering Results (FCR) algorithm, which is the recent popular segmentation approach.

Key words: Gaussian mixture model, Expectation Maximization (EM), Bayesian pixel classification, color texture segmentation, MAP estimation, EM algorithm, Maximum A Posteriori (MAP), Spatially Variant Finite Mixture Model (SVFMM), Markov Random Field (MRF), Probabilistic Rand Index (PRI), Boundary Displacement Errors (BDE), Variation of Information (VoI)

INTRODUCTION

Color image segmentation refers to partitioning an image into different regions that are homogeneous or "similar" in some image characteristics. Image segmentation has taken a central place in numerous applications, including, but not limited to, digital broadcasting, multimedia databases, color image and video transmission over the Internet, interactive TV, video-on-demand, computer-based training, distance education, video-conferencing, tele-medicine and, with the development of the hardware and communications infrastructure, to support visual applications. The field has become a prime area of research, not only in electrical engineering, but also in other academic disciplines, such as computer science, medical imaging, criminal justice, geography and remote sensing.

Clustering based image segmentation methods have been widely used on many applications (Safarinejadian *et al.* 2010; Roomi *et al.* 2010).

Unsupervised clustering method has high reproducibility because its results are mainly based on the information of image data itself and it requires little or no assumption of the model and the distribution of image data. However, pixel-based clustering algorithms cannot segment color texture images properly because they rely only on the intensity distribution of the pixels and disregard their geometric information. Due to the noise and intensity inhomogeneities introduced in imaging process, different color-textures at different locations may have similar intensity appearance, while the same color-texture at different locations may have a different intensity appearance. Therefore, the segmentation results would be totally wrong without the spatial information.

A popular method to incorporate the spatial context into clustering technique is modifying the objective function or adding the prior term in the Bayesian formulation. Bayesian method is particularly suitable for analyzing limited data as it allows for

Corresponding Author: M. Sujaritha, Department of Computer Science and Engineering,
J.J. College of Engineering and Technology, Trichy, India

updating of information by combining the current information with the prior belief. Bayesian classification system was used by several researchers for different applications. Some of the recent applications of Bayesian classifier are listed below. Moussaoui *et al*. (2006) implemented Bayesian training method to construct a series of hybrid Artificial Neural Network (ANN) structures to model hot rolling force prediction from real input/output data and empirical expressions. Mazouzi and Batouche (2007) used Bayesian estimation, based on some prior assumptions on the regions for Range image segmentation. Image priors were modeled by a new Markov Random Field (MRF) model. Saat *et al*. (2010) demonstrated the use of Bayesian methods to rank the severity of the disease, Apnea (daytime sleepiness) for 14 patients, based on the posterior mean of the rate of occurrence of Apnea.

In the problem of color texture segmentation recently, Islam *et al*. (2010) used multi dimensional spatially variant finite mixture model which is based on Markov Random Fields to model the color-texture images. The drawback of this method is that the parameter used to control the influence of the neighboring pixels is fixed and provided by the user. The disadvantage of this fixed spatial parameter is that it will blur the image feature like sharp edges while smoothing out noises.

This study presents a simple adaptive spatial Gaussian mixture model that uses a weighting function to weigh the influence of the central pixel and its neighbors. When the central pixel is in a homogeneous region, the component density function should be greatly dominated by its neighbors. The contrary holds for non-homogeneous regions. Also the weighting factor for neighborhood effect is adaptive to the image content to favor the solution of piecewise-homogeneous labeling.

MATERIALS AND METHODS

The block diagram of the proposed model based clustering algorithm is shown in Fig. 1. The proposed system consists of three consecutive steps. They are: (i) Generation of Feature set, (ii) Modeling the parameters using the proposed adaptive spatial finite mixture model and (iii) Bayesian pixel classification. These steps are described in the following subsections.

Generation of feature set: The color image is represented by a vector in a color space. The CIE spaces provide an approximately uniform chromaticity scale, which allows the use of Euclidean distance in

expressing the color difference of human perception and thus is especially efficient in the measurement of small color difference. The proposed approach uses CIE LUV color space; where L represents the luminance and U, V represent the chrominance information. The feature set is generated from the LUV color image. The dimension of the feature vector is six. The first two dimensions represent the chrominance information (U and V) and the next four dimensions represent the texture information. The texture features are generated by applying the wavelet decomposition on the luminance plane (L). The four texture features are namely, approximation information, horizontal details, vertical details and diagonal details.

Modeling the parameters using the proposed mixture model: The proposed mixture model with K>1 components in the feature space Z = {z_1, z_2,…,z_N}$\subset\Re^d$ for d $\geq$1 is defined as:

$$f(z_i \mid \theta) = \sum_{k=1}^{K} \alpha_k^i f(z_i \mid \theta_k) i = 1\ldots N \qquad (1)$$

Where:

$\alpha_1^i,\ldots,\alpha_K^i$	= The mixing proportions,
θ_k	= The set of parameters defining the kth component and
$\Theta = \{\theta_1,\ldots,\theta_K,\alpha_1^i,\ldots,\alpha_K^i\}$	= The complete set of parameters needed to specify the mixture

Being probabilities, the α_k^i must satisfy:

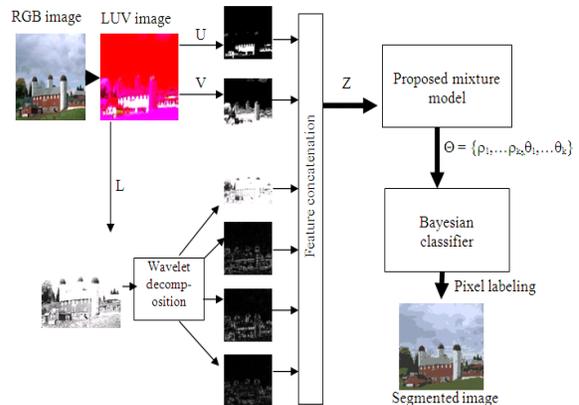$$\alpha_k^i > 0, k = 1,\ldots,K, \sum_{k=1}^{K} \alpha_k^i = 1 \qquad (2)$$

Fig. 1: Block diagram of the proposed system

In the existing Gaussian mixture model, each component density $f(z_i|\theta_k)$ is a normal probability distribution:

$$f(z_i \mid \theta_k) = f(z_i \mid \mu_k, \Sigma_k) =$$
$$\frac{\exp\{-[(z_i - \mu_k)^T \Sigma_k^{-1}(z_i - \mu_k)] / 2\}}{(2\pi)^{\frac{d}{2}} \det(\Sigma_k)^{\frac{1}{2}}} \quad (3)$$

where, $f(z_i|\mu_k, \Sigma_k)$ is a Gaussian distribution with mean μ_k and covariance Σ_k of kth component. The proposed method has incorporated the spatial relationship in calculating the density function so that the pixel z_i will be greatly influenced by its neighbors. The modified density function $f_s(z_i|\theta_k)$ is given by:

$$f_s(z_i \mid \theta_k) = \frac{\exp\{-[\dfrac{(1-\eta_i^k)\{(z_i - \mu_k)^T \Sigma_k^{-1}(z_i - \mu_k)\}}{2} + \dfrac{\eta_i^k}{8} \displaystyle\sum_{z_l \in V_{z_i}} \dfrac{\{(z_l - \mu_k)^T \Sigma_k^{-1}(z_l - \mu_k)\}}{2}]\}}{(2\pi)^{\frac{d}{2}} \det(\mathrm{diag}(\Sigma_k))^{\frac{1}{2}}} \quad (4)$$

Where:

η_i = The parameter that controls the neighbor's influence

V_{z_i} = The subset of neighborhood pixels of z_i in a 3×3 window

η_i is calculated using the following formula:

$$\eta_i^k = df_{std}^k(i) / z_{std}(i) \quad (5)$$

Where:

$$df_{std}^k(i) = (\frac{1}{9}[\sum_{z_l \in V_{z_i}} \{(df_{z_l}^k - \mu)^2\} + (df_{z_i}^k - \mu)^2])^{\frac{1}{2}} \quad (6)$$

where, μ is the mean value of df in the 3×3 window:

$$df_{z_i}^k = \frac{(z_i - \mu_k)^T \Sigma_k^{-1}(z_i - \mu_k)}{2} \quad (7)$$

In order to eliminate the unbalanced effect on the weighting functions between smooth and sharp edges, the df is divided by the standard deviation of all the pixels in the 3×3 window:

$$z_{std}(i) = \{\frac{1}{9}(\sum_{z_l \in V_{z_i}} (z_l - \hat{z})^2 + (z_i - \hat{z})^2)\}^{\frac{1}{2}} \quad (8)$$

where, $\hat{z}$ is the mean of the pixels in the 3×3 window. Note that η_i is positive and $\eta_i < 1$. The parameters of the proposed mixture model $\Theta = \{\theta_1, \ldots, \theta_K, \alpha_1^i, \ldots, \alpha_K^i\}$ are determined by Expectation-Maximization (EM) algorithm. The EM algorithm is general iterative technique for computing maximum a posteriori when the observed data can be regarded as incomplete. The EM algorithm consists of an E-step and an M-step.

E-step: Let $\Theta^{(t)}$ be the estimation of Θ obtained after the tth iteration of the algorithm. Then at the $(t+1)^{th}$ iteration, the E-step computes the expected a posteriori log density function $Q_{MAP}(\Theta, \Theta^{(t)})$:

$$Q_{MAP}(\Theta, \Theta^{(t)}) =$$
$$\sum_{i=1}^{N}\sum_{k=1}^{K} \{\log \alpha_k^i f_s(z_i \mid \theta_k)\} E(k \mid z_i, \Theta^{(t)}) \quad (9)$$

where, $E(k \mid p_i, \Theta^{(t)})$ is a posteriori probability and it is calculated by:

$$E\{k \mid z_i, \Theta^{(t)}\} = \frac{\alpha_k^{i(t)} f_s(z_i \mid \theta_k^{(t)})}{\sum_{l=1}^{K} \alpha_l^{i(t)} f_s(z_i \mid \theta_l^{(t)})} \quad (10)$$

M-Step: The M-step of the EM algorithm finds the $(t+1)^{th}$ estimation $\Theta^{(t+1)}$ of Θ by maximizing $Q_{MAP}(\Theta, \Theta^{(t)})$:

$$\alpha_k^{i(t+1)} = \frac{E\{k \mid z_i, \Theta^{(t)}\}}{\sum_{l=1}^{K} E\{l \mid z_i, \Theta^{(t)}\}} \quad (11)$$

$$\mu_k^{(t+1)} = \frac{\sum_{i=1}^{N} z_i E\{k \mid z_i, \Theta^{(t)}\}}{\sum_{i=1}^{N} E\{k \mid z_i, \Theta^{(t)}\}} \quad (12)$$

$$\Sigma_k^{(t+1)} = \frac{\sum_{i=1}^{N} E\{k \mid z_i, \Theta^{(t)}\}(z_i - \mu_k^{(t+1)})(z_i - \mu_k^{(t+1)})^T}{\sum_{i=1}^{N} E\{k \mid z_i, \Theta^{(t)}\}} \quad (13)$$

The E-step and M-step are repeated till the convergence of the parameters. In most of the cases, number of iterations taken by the EM algorithm for convergence is greater than 500. The parameters obtained after convergence are called optimal parameters. Bayesian classifier utilizes these optimal parameters for constructing the segmentation map. For every pixel it calculates posterior probabilities of classes. Then the label k (where k = 1…K) for which $E\{k \mid z_i, \hat{\Theta}\}$ (the posterior probability of class k) is maximum, is assigned to the pixel z_i.

RESULTS

The proposed model based classification system is tested on synthetic images formed using VisTex database (available online at http://vismod.media.mit.edu/vismod/imagery/VisionTexture/vistex.html) and natural images from Berkeley segmentation dataset. Berkeley dataset consists of 300 images of size 481×321. For each color image, a set of benchmark segmentation results, provided by human observers (between 4 and 7), is available and will be used to quantify the performance of the proposed segmentation algorithm. It is available online at http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/BSDS300/html.

In the proposed approach the perceptually uniform CIE-LUV color values are used as color features and a second level, wavelet pyramidal decomposition is performed to extract the texture features. Totally 6 [2 chromatic+4 (1 approx. & 3 detail)] features are extracted for every pixel and the proposed adaptive spatial finite mixture model is used for modeling the components and labeling the pixels.

Figure 2 shows the segmentation results of proposed classification system when it is applied on synthetic images. The number of pixels wrongly classified at the boundary is identified and the false percentage computed for both the images are 0.43 and 0.82 respectively.

Figure 3 compares the segmentation results of proposed system with that of recent clustering based segmentation algorithm FCR (Mignotte, 2008) (Fusion of Segmentation Results algorithm).
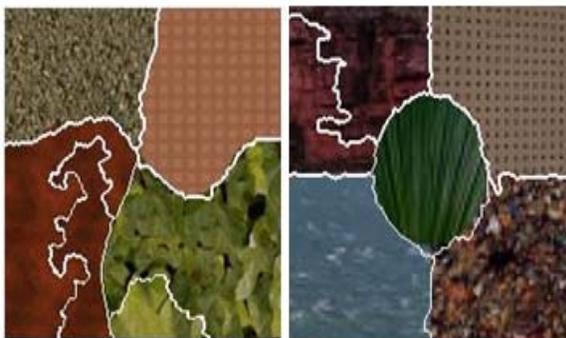

Fig. 2: Segmentation results of proposed algorithm when applied on synthetic images. The error percentage of the segmentation results are 0.43 and 0.82 respectively

DISCUSSION

The segmentation results in Fig. 3 indicate that the proposed system accurately determine the object borders between the regions that are characterized by similar color compositions. For example, in Fig. 3 (see the coconut tree image) the proposed system identifies the land, water and tree regions correctly and segments them properly. But the FCR algorithm could not identify the land under the cloud region and also the borders are not accurate. Similarly, in the bear image (Fig. 3) the proposed algorithm correctly classifies the illuminated bear region but FCR algorithm wrongly classifies the illuminated bear region as the background.

In order to quantitatively analyze the performance of the proposed algorithm, it is applied on all the 300 images in Berkeley segmentation dataset. Four performance measures namely, Probabilistic Rand Index (PRI), Global Consistency Measure (GCE), Variation of Information (VoI) and Boundary Displacement Errors (BDE) are used for comparison purpose as in (Ma *et al*., 2007; Mignotte, 2008). A good segmentation algorithm should give higher PRI and lower GCE, VoI and BDE measures.

Table 1 compares the average performance of the proposed system with recent popular segmentation algorithms (FCR (Mignotte, 2008), CTM (Ma *et al*., 2007)). The value of the parameters used in these algorithms is mentioned in the Table 1.
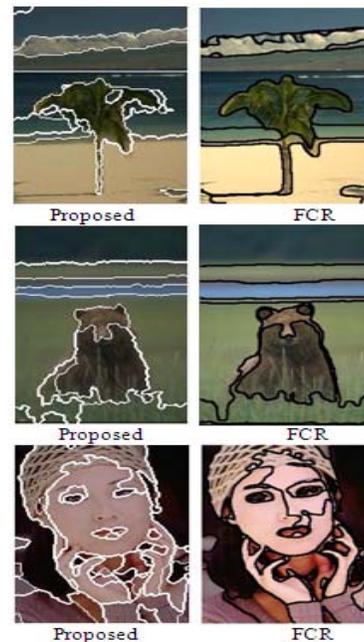

Fig. 3: Segmentation results of two algorithms (proposed, FCR) when applied on natural images from Berkeley Segmentation dataset.

Table 1: Average performance of the proposed algorithm and two recent state-of art algorithms

Method	PRI	GCE	VoI	BDE
Proposed Method	0.7917	0.2176	2.3096	7.9113
$FCR_{[K1=13|K2=6|\kappa=0.135]}$	0.7882	0.2115	2.3035	8.9951
$FCR_{[K1=13|K2=13|\kappa=0.145]}$	0.7849	0.1752	2.5494	8.7754
$CTM_{\eta=0.1}$	0.7561	0.1767	2.464	9.4211
$CTM_{\eta=0.2}$	0.7617	0.1877	2.0236	9.8962

CONCLUSION

A new, simple and adaptive spatial finite mixture model based classifier is introduced in this study. The adaptive weighting functions used with finite mixture model, aim to favor the solution of piecewise homogeneous labeling without affecting the edges. They are adaptive to the image content without any user interference. The efficiency of the proposed classifier is proved by comparing the proposed approach quantitatively and qualitatively with the recent state-of art algorithms.

REFERENCES

Islam, M., P. Vamplew and J. Yearwood, 2010. MRF model based unsupervised color textured image segmentation using multidimensional spatially variant finite mixture model. Technolog. Dev. Edu. Automation, 2010: 375-380. DOI: 10.1007/978-90-481-3656-8_68

Ma, Y., H. Derksen, W. Hong and J. Wright, 2007. Segmentation of multivariate mixed data via lossy data coding and compression. IEEE Trans. PAMI., 29: 1546-1562. DOI: 10.1109/TPAMI.2007.1085

Mazouzi, S. and M. Batouche, 2007. Range image segmentation by randomized region growing and Bayesian edge regularization. J. Comput. Sci., 3: 310-317. DOI: 10.3844/jcssp.2007.310.317

Mignotte, M., 2008. Segmentation by fusion of histogram-based *k*-means clusters in different color spaces. IEEE Trans. Image Process., 17: 780-787. DOI: 10.1109/TIP.2008.920761

Moussaoui, A., Y. Selaimia and H.A. Abbassi, 2006. Hybrid hot strip rolling force prediction using a Bayesian trained artificial neural network and analytical models. Am. J. Applied Sci., 3: 1885-1889. DOI: 10.3844/ajassp.2006.1885.1889

Roomi, S.M.M., R.J. Priya and H. Jayalakshmi, 2010. Hand gesture recognition for human-computer interaction. J. Comput. Sci., 6: 1002-1007. DOI: 10.3844/jcssp.2010.1002.1007

Saat, N.Z.M., K. Ibrahim and A.J. Jemain, 2010. Bayesian methods for ranking the severity of apnea among patients. Am. J. Applied Sci., 7: 167-170. DOI: 10.3844/ajassp.2010.167.170

Safarinejadian, B., M.B. Menhaj and M. Karrari, 2009. Distributed data clustering using expectation maximization algorithm. J. Applied Sci., 9: 854-864. DOI: 10.3923/jas.2009.854.864