International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:4, No:1, 2010

Color Image Segmentation using Adaptive Spatial Gaussian Mixture Model

M.Sujaritha and S. Annadurai

Abstract—An adaptive spatial Gaussian mixture model is proposed for clustering based color image segmentation. A new clustering objective function which incorporates the spatial information is introduced in the Bayesian framework. The weighting parameter for controlling the importance of spatial information is made adaptive to the image content to augment the smoothness towards piecewise-homogeneous region and diminish the edge-blurring effect and hence the name adaptive spatial finite mixture model. The proposed approach is compared with the spatially variant finite mixture model for pixel labeling. The experimental results with synthetic and Berkeley dataset demonstrate that the proposed method is effective in improving the segmentation and it can be employed in different practical image content understanding applications.

Keywords—Adaptive; Spatial; Mixture model; Segmentation ; Color.

I. Introduction

SEGMENTATION methods are based on some pixel or region similarity measure in relation to their local neighborhood. A variety of different methods have been proposed for image segmentation such as boundary-based segmentation, region-based segmentation and pixel labeling. Boundary–based methods search for the most dissimilar pixels which represent discontinuities in the image, while region based methods on the contrary search for the most similar areas. Pixel labeling (Clustering) algorithm is composed of the hidden label process and the observable noisy image process. The goal is to find an optimal labeling which maximizes the posterior probability that is the maximum a posteriori (MAP) estimate. Recently, Expectation-Maximization algorithm has attracted considerable interest to compute the maximum likelihood estimates, when the observations are unlabeled.

Unsupervised clustering techniques [9] have high reproducibility because its result are mainly based on the intensity information of image data itself. They do not require training data, but they do require an initial segmentation and they rely only on the intensity distribution of the pixels and disregard their geometric information. For example, relaxation labeling [5], [8] methods make little or no use of the observed image, except perhaps to initialize the label configuration for the iterative algorithm.

Several model based clustering approaches have been proposed by the researchers for the unsupervised segmentation problem. Model based approaches uses Gibbs random fields, Gauss-Markov random fields, Gaussian autoregressive random fields, and univariate and multivariate Gaussian densities to model the image. And the posterior functions developed by these models are optimized using some optimization techniques such as simulated annealing [8], Iterated conditional modes(ICM) [3], gradient projection [16] etc.

The application of Gaussian mixture models to labeling is based on the assumption that the intensity (grey-scale or color) value of each pixel in the observed image is a sample from a Gaussian mixture distribution. . Gaussian Mixture models constitute a well-known probabilistic neural network model. The Expectation –Maximization framework constitutes an efficient method for GMM training based on likelihood maximization. Expectation-Maximization algorithms are used for parameter estimation and maximization which consists of two steps: the estimation of a tentative labeling of the data followed by updating the parameter values based on the tentatively labeled data. Upon estimating the parameters of this distribution a suitable labeling rule is applied to assign labels to the pixels of an image.

In work [20] T.Yamazaki considered the color image as a mixture of multi-variant densities and the parameters are estimated using EM algorithm. The segmentation is completed by clustering each pixel into a component according to the Maximum likelihood (ML) estimation. He proved that the method is stable and useful for color image segmentation. The drawback of that method is the number of mixture components is assumed known as prior, so it cannot be considered as totally unsupervised segmentation.

The major drawback of pixel labeling is, due to the noise and intensity in-homogeneities introduced in imaging process, dissimilar regions at different locations may have same intensity appearance, while the similar regions at different locations may have different intensity appearance. Hence, the segmentation results would be totally wrong without the spatial information.

The problem of traditional finite mixture models is lack of spatial correlation into the labeling process. One can reintroduce spatial correlation into the labeling process using the Gaussian mixture model by imposing dependence structures in the form of Markov chains and two-dimensional (2-D) MRF models on the complete data of the EM algorithm. But again this is a non-optimal approach, and the resulting algorithms have been found to be considerably complicated and computationally intractable.

Several researchers have suggested modifications to the Finite mixture model to address the problem related to the lack of spatial information. Sanjay-Gopal and Herbert [16] proposed a Bayesian framework using spatially variant mixture

M.Sujaritha is with J.J College of Engineering and Technology, Trichy, India, Asia. email: sujaritha2009@gmail.com
S. Annadurai is the Additional Director (Polytechnic), Directorate of Technical Education, Chennai, India, email: anna_prof@yahoo.co.in

models (SVFMM) by introducing a process based on Markov Random Fields (MRF) which is able to capture the spatial relationship among neighboring pixels. MRF introduces a prior distribution that takes into account the neighborhood dependency or relationship among the neighboring pixels. His a posteriori density function consists of likelihood term and biasing term. The likelihood term which is exclusively based on the intensity distribution of the data, captures the pixel intensity information, while a priori biasing term that uses a Markov Random Field (MRF) captures the spatial location information. He incorporates the local correlations between the neighboring parameters, through the application of suitable prior density function that models the local correlation in the parameter estimation process. Also he implemented the M-step of the EM algorithm using gradient projection method (GEM-MAP) to overcome the difficulty which arises due to the introduction of the prior term in the E-step.

Blekas et al., [4] modified the M-step of SVFMM–GP and named as spatially –constrained Mixture model. Zoltan Kato and T.C.Pong [12] incorporates the spatial information by adding the prior term in the posterior energy function like Sanjay et al. But instead of gradient decent technique , simulated annealing with Gibbs sampler and Iterated conditional modes were used to find the global minimum.

The three important issues in the above MRF based mixture models are the initialization of mixture weights and component parameters, the number of pixel classes and the specification of parameter (β) encouraging neighbors to have similar characteristics.

Since EM algorithms have convergence to local maxima, initialization is very important. Yiming Wu et al., [19] circumvented the initialization problem of EM algorithm by using K-means algorithm to initialize the Gaussian mixture parameters. But the draw backs of k-means algorithm are it is sensitive to initialization and convergence problem.

Islam et al., [10] used SNOB and cluster Ensembles for finding the number of components in a given image. Zoltan et al. [11] used Reversible Jump Markov Chain Monte Carlo algorithm for the model dimension switching. Fauzi et al., [7] used mean shift algorithm to identify the number of clusters. Yiming Wu et al., [19] introduced MML (Minimum Message Length) criterion into the EM algorithm to automatically determine the number of mixture components.

The proposed method tries to overcome the first issue by initializing the label parameters using K-means algorithm and density parameters using the histograms of the feature vector. The second issue is solved by applying MML model selection criterion [2] to estimate the number of components, due to its efficiency and simplicity. To beat the third issue the weighting factor for neighborhood effect is made adaptive to the image content.

This paper elaborates on a pixel labeling (clustering) technique based on a new adaptive spatial Gaussian Mixture Models and EM algorithm, which introduces the spatial information into the clustering process. The spatial factors are adaptive to the image content to favor the solution of piecewise homogeneous labeling. The idea has come up from [18] by Wang et al., who have extended deterministic annealing clustering technique to Adaptive Spatial Deterministic Annealing (ASDA) technique.

This paper is organized as follows: Section 2 describes the proposed Adaptive Spatial Gaussian Mixture Model (AS-GMM) and Section 3 presents the results and Section 4 concludes the paper.

II. The Proposed method

The proposed method modifies the Gaussian mixture model and EM algorithm to incorporate the spatial information into the model. The following subsection states the essence of GMM and EM algorithm.

A. Gaussian Mixture Models and EM algorithm

Color image is represented by a vector in a color space. Several color spaces have been proposed for different contexts of image processing. RGB, HSI, YIQ and CIE spaces etc. [6], [13], [15], [17] . Although lots of discussions have been made so far, the selection of the best color space is still a difficult problem for the color image segmentation. The RGB space is quite commonly used because of its simplicity in implementation. A better color space than the RGB space in representing the colors of human perception is the HSI space, in which the color information is represented by hue and saturation values while brightness is represented by an intensity value [13]. The YIQ is obtained by a linear transformation on the RGB space, where the Y component is a measurement of the luminance and is argued to be a likely candidate for edge detection while the I and Q components jointly describe the hue and the saturation of the image. 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 [17].

Gaussian Mixture Model:

Despite the difference among those color spaces, the input color image is expressed in general by an expression $z = \{z_1, z_2, \ldots, z_N\} \subset \Re^d$, where N is the number of input data patterns, d is the dimension of input color space. To define a Gaussian mixture model with $K > 1$ components in $\Re^d$ for $d \geq 1$, let z_i denote the observation at the ith pixel of an image. The density function $p(z|\theta)$ at any observation z_i is given by:

$$p(z_i \mid \theta) = \sum_{k=1}^{K} \rho_k f(z_i \mid \theta_k) \tag{1}$$

where $\rho_1, \ldots, \rho_K$ are the mixing weights, $f(.)$ is a Gaussian distribution with the parameters $\theta_k = (\mu_k, \Sigma_k)$, each θ_k defining the kth component and $\Theta = \{\rho_1, \ldots, \rho_K, \theta_1, \ldots, \theta_K\}$ is the complete set of parameters needed to specify the mixture. Let the mixture weights satisfy the following conditions

$$\rho_k \geq 0, \quad k = 1, \ldots, K, \quad \sum_{k=1}^{K} \rho_k = 1 \tag{2}$$

For the Gaussian mixtures, each component density $f(z_i \mid \theta_k)$ is a Gaussian probability distribution with μ_k and covariance Σ_k:

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

EM Algorithm:
Expectation-Maximization algorithm is used to determine the parameters Θ of a Gaussian mixture model from a given image. The EM algorithm is general iterative technique for computing maximum-likelihood when the observed data can be regarded as incomplete. To define this complete data let m_i, $i = 1, 2, \ldots, N$ denote $K \times 1$ random indicator vectors, each of which takes a value from the set of vectors Λ defined by

$$\Lambda = \{\nu^j; \quad \nu^j_{l=j} = 1, \quad \nu^j_{l\neq j} = 0, \quad 1 \leq l, j \leq K\} \tag{4}$$

Also m^i_j is a discrete random variable with probability function defined by $\Pr ob(m^i = \nu^j) = \rho^i_j, \forall i, j$. The complete data $y^T = (z^T, m^{1^T}, \ldots, m^{N^T})$ with superscript T denoting vector transpose. Here, z and $\{m^i\}$ are defined to be statistically independent So that

$$p(y \mid \Theta) = \prod_{i=1}^{N}\prod_{k=1}^{K}[\rho^i_k f(z_i \mid \theta_k)]^{m^i_k} \tag{5}$$

The usual EM algorithm consists of an E-step and an M-step. Let $\Theta^{(t)}$ denote the estimation of Θ obtained after the tth iteration of the algorithm. For the case of Gaussian component densities $f(.)$, the E-step computes the following expected log-likelihood function at the $(t+1)$th iteration.

$$Q(\Theta, \Theta^{(t)}) = \sum_{i=1}^{N}\sum_{k=1}^{K}\{\ln \rho_k f(z_i \mid \theta_k)\}E\{m^i_k \mid z_i, \Theta^{(t)}\} \tag{6}$$

where $E\{m^i_k \mid z_i, \Theta^{(t)}\}$ is a posterior probability and is computed as

$$E\{m^i_k \mid z_i, \Theta^{(t)}\} = \frac{\rho^{(t)}_k f(z_i \mid \theta^{(t)}_k)}{\sum_{l=1}^{K}\rho^{(t)}_l f(z_i \mid \theta^{(t)}_l)} \tag{7}$$

The M-step finds the $(t+1)$th estimation $\Theta^{(t+1)}$ of Θ by maximizing $Q(\Theta, \Theta^{(t)})$.

$$\rho^{i(t+1)}_k = \frac{E\{m^i_k \mid z_i, \Theta^{(t)}\}}{\sum_{k=1}^{K}E\{m^i_k \mid z_i, \Theta^{(t)}\}} \tag{8}$$

$$\mu^{(t+1)}_k = \frac{\sum_{i=1}^{N} z_i E\{m^i_k \mid z_i, \Theta^{(t)}\}}{\sum_{i=1}^{N}E\{m^i_k \mid z_i, \Theta^{(t)}\}} \tag{9}$$

$$\Sigma^{2(t+1)}_k = \frac{\sum_{i=1}^{N}E\{m^i_k \mid z_i, \Theta^{(t)}\}(z_i - \mu^{(t+1)}_k)(z_i - \mu^{(t+1)}_k)^T}{\sum_{i=1}^{N}E\{m^i_k \mid z_i, \Theta^{(t)}\}} \tag{10}$$

EM algorithm is highly dependent on initialization. Here, K-means algorithm is used for initializing ρ and histograms are used for initializing μ and Σ. Initially the histograms of K levels over the range of feature values are computed and then the mean and covariance at each level are considered as the initial values for μ_k and Σ_k.

B. Adaptive Spatial Gaussian Mixture Model

In the finite mixture model, the labeling of a pixelz_i, with the kth component only depends on the density function $p(z_i \mid \theta)$. If a noisy image in which intensity values of some data points have been changed is considered the finite mixture model does has a solution to overcome this problem. The solution is to employ the neighborhood relationships among the neighboring pixels.

We have incorporated the spatial relationship in calculating the density function in (11), so that the pixel z_i will be greatly influenced by its neighbors.

$$\begin{aligned} f_s(z_i \mid \theta_k) = & \frac{1}{(2\pi)^{\frac{d}{2}}\det(\Sigma_k)^{\frac{1}{2}}} \\ & \times \Big[\exp\{-[\frac{\eta^k_i(z_i-\mu_k)^T(z_i-\mu_k)}{2\Sigma_k} \\ & + \frac{\eta^k_i}{8}\sum_{Z_l \in V_{z_i}}\frac{(z_l-\mu_k)^T(z_l-\mu_k)}{2\Sigma_k}]\}\Big] \end{aligned} \tag{11}$$

where η_i is the parameter that controls the neighbor's influence and V_{z_i} is the subset of neighborhood pixels of z_i in a 3x3 window. η_i is calculated using the following formula:

$$\eta^k_i = df^k_{std}(i)/z_{std}(i) \tag{12}$$

where

$$df^k_{std}(i) = (\frac{1}{9}[\sum_{z_l \in V_{z_i}}\{(df^k_{z_l} - \mu)^2\} + (df^k_{z_i} - \mu)^2])^{\frac{1}{2}} \tag{13}$$

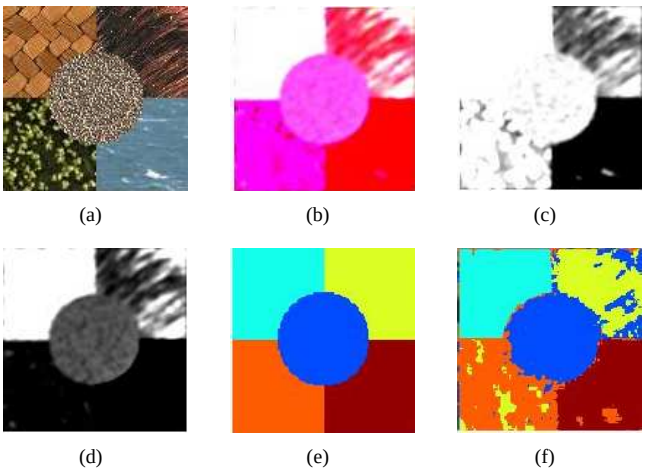

Fig. 1. Segmentation of synthetic image–1 using proposed method. (a) Original image (b) CIE-Luv color image (c)Chromatic v-feature (d) Chromatic u-feature (e) Ground truth image (f) segmented image.

μ is the mean value of df in the 3×3 window

$$df_{z_i}^k = \frac{(z_i - \mu_k)^T (z_i - \mu_k)}{2\Sigma_k} \tag{14}$$

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

$$z_{std}(i) = \{(\frac{1}{9} \sum_{z_l \in V_{z_i}} (z_l - \hat{z})^2 + (z_i - \hat{z})^2)\}^{\frac{1}{2}} \tag{15}$$

$\hat{z}$ is the mean of the pixels in the 3x 3 window. Note that η_i is positive and $\eta_i < 1$.
Now the posterior probability (7) is modified as

$$E_s\{m_k^i \mid z_i, \Theta^{(t)}\} = \frac{\rho_k^{(t)} f_s(z_i \mid \theta_k^{(t)})}{\sum_{l=1}^{K} \rho_l^{(t)} f_s(z_i \mid \theta_l^{(t)})} \tag{16}$$

and the M-step is modified as

$$\rho_k^{i(t+1)} = \frac{E_s\{m_k^i \mid z_i, \Theta^{(t)}\}}{\sum_{k=1}^{K} E_s\{m_k^i \mid z_i, \Theta^{(t)}\}} \tag{17}$$

$$\mu_k^{(t+1)} = \frac{\sum_{i=1}^{N} z_i E_s\{m_k^i \mid z_i, \Theta^{(t)}\}}{\sum_{i=1}^{N} E_s\{m_k^i \mid z_i, \Theta^{(t)}\}} \tag{18}$$

$$\Sigma_k^{2(t+1)} = \frac{\sum_{i=1}^{N} E_s\{m_k^i | z_i, \Theta^{(t)}\}(z_i - \mu_k^{(t+1)})(z_i - \mu_k^{(t+1)})^T}{\sum_{i=1}^{N} E_s\{m_k^i \mid z_i, \Theta^{(t)}\}} \tag{19}$$

III. Experimental Results and Discussions

The proposed algorithm has been tested on a variety of color images including synthetic images (Figures 1 and 2), Berkeley dataset images [1] (Figure 3). MIT's Vistex database has been used to compose the synthetic color images. It has been proved [11] that the performance of segmentation is improved when the color and texture features are combined. Therefore, the perceptually uniform CIE-LUV color values are used as color features and a 3 level, wavelet pyramidal decomposition is performed to extract the texture features. Totally 12 (10 (approx & detail) +2 chromatic) features are extracted for every pixel and the proposed ASGMM method is used for modeling the components and labeling the pixels.

Figures 1 and 2 shows the segmentation of synthetic texture images which consists of bark, fabric, metal, water, misc and Leaves. The number of classes in both the images is five. The figures show the original image, its CIE-Luv color image, chromatic features, ground truth image and the segmented image using the proposed ASGMM method.

Initially, Gaussian smoothing is performed on those color texture images. The parameters of Gaussian smoothing are 1.5 and 5. The classification error rate of the synthetic color texture images shown in Figures 1 and 2 are 0.1123 and 0.1411 respectively.

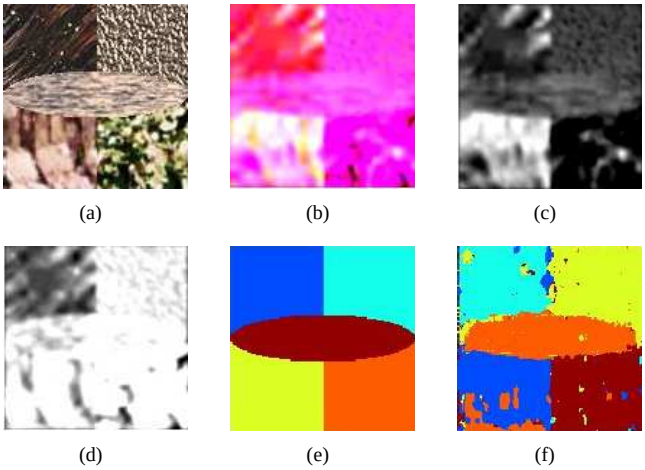

Fig. 2. Segmentation of synthetic image–2 using proposed method. (a) Original image (b) CIE-Luv color image (c)Chromatic u-feature (d) Chromatic v-feature (e) Ground truth image (f) segmented image.

TABLE I
Performance measures for the Clustering results obtained by SVFMM and the proposed method on the two synthetic images presented in Figures 1 and 2.

Images	Segmentation Techniques	Performance Measures			
		PRI	VOI	GCE	BDE
Synthetic image (Figure1)	SVFMM	0.8132	2.2967	0.1536	8.9969
	Proposed ASGMM	**0.8877**	**1.9621**	**0.1423**	**6.3612**
Synthetic image (Figure2)	SVFMM	0.8010	2.6964	**0.1897**	9.3143
	Proposed ASGMM	**0.8589**	**1.9808**	0.1903	**7.1635**

The proposed method is compared with (SVFMM) Spatially Variant Finite Mixture model [16] with $\beta = 10$ and number of iterations 45. The comparison is based on four performance measures namely, Probabilistic Rand Index (PRI), the Variation of Information (VOI), Global Consistency measure (GCE) and the boundary displacement error,(BDE) following Max Mignotte [14]. PRI calculates the consistency between the computed segmentation and the ground truth. The VOI measures the amount of randomness in one segmentation which cannot be explained by other. The GCE measures the extent to which one segmentation map can be viewed as a refinement of another segmentation. BDE measures the average displacement error of one boundary pixel and the closest boundary pixels in the other segmentation. Table I shows the performance measures for the obtained segmentation results by SVFMM and the proposed method on the two synthetic images presented in Figures 1 and 2. Higher is better for PRI and lower is better for VOI, GCE and BDE.

Figure 3 shows the comparison result for the same number of iterations. In Figure 3, first column shows the images from Berkeley dataset, second column shows the segmentation result of SVFMM and the third column shows the segmentation result of the proposed ASGMM. The superior performance of our ASGMM algorithm is noticeable. The adaptive weighting functions are able to put more spatial constraint on the homogeneous regions as compared to those

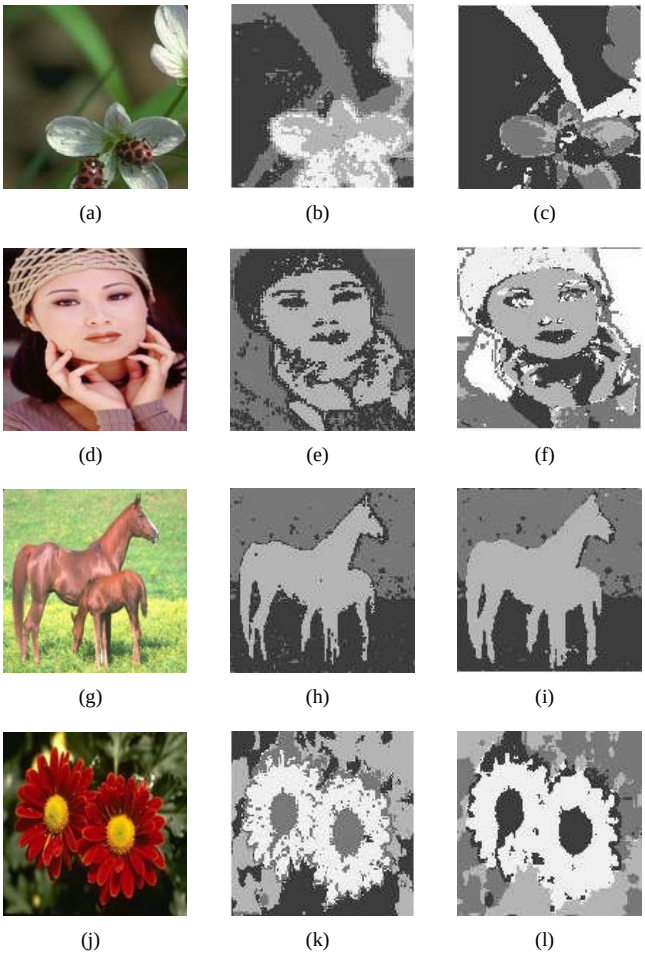

Fig. 3. Comparison of Segmentation Results.
First column (a, d, g, j)—original images.
Second column (b, e, h, k)—SVFMM method,
Third column (c, f, i, l)—Proposed method.

with fixed neighborhood effect. Furthermore, the adaptive weighting functions are obtained automatically from the image content, no user interference is required. This mechanism is very useful for developing fully unsupervised system.

IV. Conclusion

The problem of color image segmentation using color and texture is addressed. CIE lab color spaces and wavelet transformations are applied to extract color and texture features respectively. A novel Adaptive Spatial Gaussian mixture model which incorporates the spatial information into a mixture model, with a weighting factor which is adaptive to image information is proposed. The number of components is automatically identified using MML algorithm. Very good segmentation results are obtained for both synthetic and Berkeley dataset using the proposed method.

References

[1] "The Berkely segmentation dataset," http//www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/BSDS300/html.
[2] "MML clustering of multi-state, poisson von mises circular and gaussian distributions," *Statistics and Computing*, vol. 10, pp. 73–83, 2000.
[3] J. Besag, "On the statistical analysis of dirty images," *Journal of Royal Statistics Society*, vol. 48, pp. 259–302, 1986.
[4] K. Blekas, N. Likas A. Galatsanos, and I. Lagaris, "A spatially-constrained mixture model for image segmentation," *IEEE Trans. On Neural Networks*, vol. 162, pp. 494–498, 2002.
[5] P. Chou and C. Brown, "The theory and practice of Bayesian image labeling," *International Journal of Computer Vision*, vol. 4, pp. 185–210, 1990.
[6] D. Comaniciu and P. Meer, "Robust analysis of feature spaces color image segmentation," in *IEEE Conference on Computer Vision and Pattern Recognition*, 1997, pp. 750–755.
[7] M. F. A. Fauzi and P. H. Lewis, "Automatic texture segmentation for content-based image retrieval application," *Patt. Anal. Applic.*, vol. 9, pp. 307–323, 2006.
[8] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 6, pp. 721–741, 1984.
[9] M. Haindl and S. Mikes, *Unsupervised Texture Segmentation, Pattern Recognition Techniques, Technology and Applications*. I-Tech, Vienna, Austria, Nov. 2008, ISBN 978-953-7619-24-4.
[10] M. Islam, P. Vamplew, and J. Yearwood, "Unsupervised color textured image segmentation using cluster ensembles and MRF model advances in computer and information sciences and engineering," in *Proceedings of the 2007 International Conference on Systems, Computing Sciences and Software Engineering SCSS*, vol. 1, 2007, pp. 323–328.
[11] Z. Kato and T.-C. Pong, "A Markov random field image segmentation model using combined color texture features," in *CAIP 2001, LNCS 2124*, 2001, pp. 547–554.
[12] ——, "A Markov random field image segmentation model for color textured images," *Image and Vision Computing*, vol. 24, pp. 1103–1114, 2006.
[13] W. Kim and R. Park, "Color image palette construction based on the HSI color system for minimizing the reconstruction error," *IEEE International Conference on Image Processing*, vol. 3, pp. 1041–1044, 1996.
[14] M. Mignotte, "Segmentation by fusion of histogram-based k-means clusters in different color spaces," *IEEE Transactions on image Processing*, vol. 17, pp. 780–787, 2008.
[15] M. Pietikainen, S. Nieminen, E. Marszalec, and T. Ojala, "Accurate color discrimination with classification based on feature distributions," in *International Conference on Pattern Recognition*, vol. 3, 1996, pp. 833–838.
[16] S. Sanjay Gopal and T. Herbert, "Bayesian pixel classification using spatially variant finite mixtures and generalized algorithm," *IEEE Trans. On Image Processing*, vol. 7, no. 7, pp. 207–216, July 1998.
[17] C. Tseng, D and C. C.H., "Color, segmentation using perceptual attributes," in *Proc. 11th IEEE International Conference on Pattern Recognition*, vol. 3, 1992, pp. 228–231.
[18] Z. M. Wang, Q. Song, and Y. C. Soh, "MRI brain image segmentation by adaptive spatial deterministic annealing clustering," in *3rd IEEE International Symposium on Biomedical Imaging*, no. 6–9, 2006, pp. 299–302.
[19] Y. Wu and X. Y. K. L. Chan, "Unsupervised color image segmentation based on gaussian mixture model," in *Proc. of the 2003 Joint Conf. of the 4th Intl. Conf. on Information, Communications and Signal Processing*, vol. 1, 2003, pp. 541–544.
[20] T. Yamazaki, "Introduction of EM algorithm into color image segmentation," in *Proceedings of ICIRS'98*, 1998, pp. 368–371.