A MULTIVARIATE GAUSSIAN MIXTURE MODEL OF LINEAR PREDICTION ERROR FOR COLOUR TEXTURE SEGMENTATION

Imtnan-Ul-Haque QAZI[1], *Fatima GHAZI*[2], *Olivier ALATA*[1], *J. C. BURIE*[3] *and C. F. MALOIGNE*[1]

[1]Laboratory XLIM/ Department SIC,
University of Poitiers, bat. SP2MI,
Av. Marie et Pierre Curie, 86960,
Chasseneuil-Futuroscope Cedex, France.
email: qazi@sic.sp2mi.univ-poitiers.fr
web: www.sic.sp2mi.univ-poitiers.fr

[2]LASTID Laboratory,
Ibn Tofail University,
BP133, 14000 Kénitra, Morocco.
web: www.univ-ibntofail.ac.ma

[3]Laboratory L3I,
University of La Rochelle,
avenue Michel Crepeau, 17042,
La Rochelle Cedex 1, France.
web: www.univ-lr.fr

ABSTRACT

This paper presents an algorithm for parametric supervised colour texture segmentation using a novel image observation model. The proposed segmentation algorithm consists of two phases: In the first phase, we estimate an initial class label field of the image based on a 2D multichannel complex linear prediction model. Information of both luminance and chrominance spatial variation feature cues are used to characterize colour textures. Complex multichannel version of 2D Quarter Plane Autoregressive model is used to model these spatial variations of colour texture images in CIE L*a*b* colour space. Overall colour distribution of the image is estimated from the multichannel prediction error sequence of this Autoregressive model. Another significant contribution of this paper is the modelling of this multichannel error sequence using Multivariate Gaussian Mixture Model instead of a single Gaussian probability. Gaussian parameters are calculated through Expectation Maximization on a training dataset. In second phase of the algorithm, initial class label field obtained through the first stage is spatially regularized by ICM algorithm to have the final segmented image. Visual and quantitative results for different number of components of Multivariate Gaussian Mixture Model are presented and discussed.

1. INTRODUCTION

The main goal of image segmentation is to partition the input image into different regions. Colour texture segmentation is useful in many real world applications. It gives the possibility to identify the regions of interest in scenes for subsequent image analysis. Various model based approaches for texture segmentation in grey level and colour images have been proposed in the recent years [7], [6], [9], [10]. There are many real world applications which demand for higher precision of segmentation along with stability under different constraints. Such applications normally use supervised segmentation methods.

An example of supervised colour texture segmentation based on the minimal cut/maximal flow algorithm in the representation graph of the image is presented in [8]. They incorporated the colour and texture feature cues into the whole graph used for segmentation. There exist a number of model based approaches for colour texture segmentation. A Markov Random Field *(MRF)* based image segmentation model has been proposed in [6]. The authors represented the different classes in the image by Multivariate Gaussian Mixture Model *(MGMM)*. They used Gabor filters as texture features whereas pixel values in CIE L*u*v* colour space are considered as colour feature cue. In [10], the authors used Gaussian Mixture Models (GMM) for autoregressive model features to classify colour textures. They have used several methods to represent the structure information of the colour images including wavelet coefficients, DCT coefficients and autoregressive model coefficients. Authors achieved colour texture classification by fusing this structure information with the pure colour information obtained through the mean and covariance information of the image. In both these approaches, the authors have not considered the fusion of chrominance structure information with luminance structure information as they did not take into account the chrominance structure information.

In this paper, we present a model based parametric algorithm for supervised colour texture segmentation through 2D multichannel complex random fields in the CIE L*a*b* colour space. We have not considered the other perceptual colour spaces because comparison of different colour spaces is beyond the scope of this paper due to the limited space. In the first phase of the algorithm, Two-Dimensional $(2D)$ multichannel complex version of Quarter Plane Autoregressive *(QP AR)* model is used to model the luminance and chrominance spatial variations of the multichannel complex image. Normal hypothesis for the multichannel prediction error sequence of this linear prediction model is that it can be modelled through a single Gaussian. This multichannel error sequence gives an approximation to the pure colour content of the image. Therefore, we modelled the same error sequence through Multivariate GMM *(MGMM)* to have a stable segmentation result. Parameters of MGMM are estimated through the Expectation Maximization (EM) algorithm. In the second phase of the algorithm, we use Potts model with an Iterative Conditional Mode *(ICM)* [3] to regularize the initial colour texture segmentation computed in the first phase.

Section 2 describes the colour space and image representation used. Image observation model, which is a 2D multichannel complex linear prediction model, is described in Section 3. MGMM based prediction error model and estimation of MGMM parameters through EM are explained in Section 4. Simulations and results are presented in Section 5 and finally Section 6 concludes the paper.

2. IMAGE REPRESENTATION

The RGB colour space is usually used for image processing and/or analyzing. However the representation of RGB components in a 3D polar coordinate system often reveals characteristics which are not visible in the rectangular representation. Our approach is fundamentally based on the modelling of spatial structure information of both luminance and chrominance channels of the image. As in RGB colour space it is very difficult to decorrelate and analyze the effects of luminance and chrominance channels separately, therefore we propose to use the polar CIE L*a*b* colour space. Other perceptual colour spaces like CIE L*u*v* or Improved Hue, Luminance and Saturation (*IHLS*) like in [4], may also be used.

2.1 CIE L*a*b* Colour Space

CIE L*a*b* is a uniform colour space based on human perceptual system defined by CIE in [1]. The transformation from the RGB colour space to the L*a*b* colour space is performed by the following equations:

$$
\begin{aligned}
L* &= 116 \times \left(\frac{Y}{Y_W}\right)^{\frac{1}{3}} - 16 \text{ for } \frac{Y}{Y_W} > 0.0008856 \\
L* &= 903.3 \times \frac{Y}{Y_W} \text{ for } \frac{Y}{Y_W} \leq 0.008856 \\
a* &= 500 \times \left(f\left(\frac{X}{X_W}\right) - f\left(\frac{Y}{Y_W}\right)\right) \\
b* &= 200 \times \left(f\left(\frac{Y}{Y_W}\right) - f\left(\frac{Z}{Z_W}\right)\right) \\
&\text{with } f(\mu) = \mu^{\left(\frac{1}{3}\right)} \text{ for } \mu > 0.008856 \\
f(\mu) &= 7.787 \times \mu + \frac{16}{116} \text{ for } \mu \leq 0.008856
\end{aligned}
\tag{1}
$$

where X_W, Y_W and Z_W are the CIE XYZ tristimulus values of the reference white point [1]. As we did not have the knowledge of the reference white points of the test images used in simulation section (cf. Section 5), we took the reference white point obtained when $R = 1$, $G = 1$ and $B = 1$.

2.2 Two Channel Complex Image

We use the colour information obtained through RGB to L*a*b* transformation to build a two channel image that contains pure luminance values in one channel and complex chrominance values in the other channel. We define this chrominance value as a complex function depending upon two chrominance variables i.e. a* and b* in case of L*a*b*. This exponential being independent of the luminance values shall give us the pure information about the colour variations in the spatial domain. We define the combined chrominance for CIE L*a*b* as:

$$C = a^* + j \times b^* \tag{2}$$

where a^* and b^* are two chroma variables obtained from RGB to L*a*b* transformation. We obtain a complex representation of chrominance content of the image whose spectrum is interesting to analyze from a colorimetric point of view. Now the image to be analyzed consists of two 2D channels in which first channel contains the luminance information and second is complex valued channel containing combined chrominance information and is expressed as:

$$x_n = \begin{bmatrix} l_n^* \\ c_n \end{bmatrix} \tag{3}$$

where $n = (n_1, n_2) \in \Lambda \subset \mathbb{Z}^2$ in which Λ is the finite 2-D image lattice region of size $|\Lambda|$, $l_n^* \in \mathbb{R}$ and $c_n \in \mathbb{C}$.

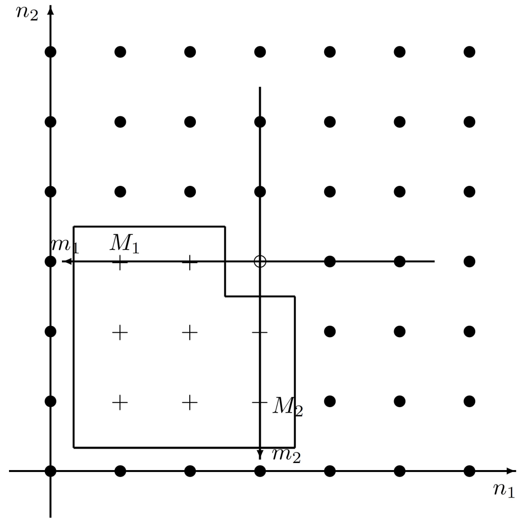


Figure 1: Two dimensional Quarter Plane Autoregressive model neighbour support region of order (M_1, M_2), where $M_1 = 2$ and $M_2 = 2$ for first quarter plane.

3. IMAGE OBSERVATION MODEL

In this section we present the 2D multichannel complex linear prediction model which is used to model the observed colour image.

3.1 Linear Prediction Model

A 2D multichannel random centered process represented by a vector sequence $X = \{X_n\}_{n\in\mathbb{Z}^2}$ with dimension P representing the number of channels following a linear prediction model can be defined through the prediction sequence:

$$\hat{X}_n = -\sum_{m\in D} A_m X_{n-m} \tag{4}$$

as

$$X_n = \hat{X}_n + E_n. \tag{5}$$

where $m = (m_1, m_2) \in D \subset \mathbb{Z}^{*2}$ is a point inside neighbour support region defined by D. $A_m, m \in D$, are $P \times P$ coefficient matrices and $E = \{E_n\}_{n\in\mathbb{Z}^2}$ is the prediction error sequence which is supposed to be a multichannel stationary process. We used 2D multichannel complex version of QP AR model as the image observation model. Details of QP AR are given in following subsection, while the details of other linear prediction models can be found in [11].

3.2 2D QP AR Model

2D multichannel complex QP AR model is defined by Equations 4 and 5 for which the causal QP_1 neighbourhood support region is defined as follows:

$$
\begin{aligned}
D_{M_1,M_2}^{QP_1} = \{&(m_1, m_2) / 0 \leq m_1 \leq M_1, \\
&0 \leq m_2 \leq M_2, (m_1, m_2) \neq (0,0)\}
\end{aligned}
\tag{6}
$$

where $(M_1, M_2) \in \mathbb{N}^2$ is the model order. In case of 2D multichannel complex QP AR model, A_m, $m \in D$, are $P \times P$ complex coefficient matrices. These 2D complex QP_1 AR parameters are estimated by a matrix solution of a system of normal equations. An example of such a neighbourhood support region for 2D QP_1 AR model is shown in Figure 1.

A common hypothesis about the multichannel stationary error sequence $E = \{E_n\}_{n\in\mathbb{Z}^2}$ in Equation 5 follows a Gaussian distribution [2]. Thus it can be characterized by a single Gaussian probability distribution with a zero mean value and having a $P \times P$ covariance matrix denoted by Σ_e.

4. PREDICTION ERROR MODEL

In this paper we assume that the multichannel linear prediction error sequence E can be modelled through a mixture of Gaussians. Details of MGMM and its parameter estimation using EM algorithm is explained in the next subsections.

4.1 Multivariate Gaussian Mixture Model

Gaussian mixture model for the multichannel error sequence E is defined as:

$$p(e_n|\theta) = \sum_{k=1}^{K} \alpha_k p(e_n|\theta_k) \quad (7)$$

where $\alpha_1, \ldots, \alpha_K$ are the prior probabilities of each Gaussian component of the mixture, and $K > 1$ is the number of components of MGMM. Each θ_k is the set of model parameters defining the kth component of the mixture model. The prior probability values must satisfy following conditions:

$$\alpha_k > 0, k = 1, \ldots, K \quad (8)$$

and

$$\sum_{k=1}^{K} \alpha_k = 1 \quad (9)$$

For the 2D complex error sequence E, MGMM can be conceived by considering the real and imaginary parts of the complex error sequence as two variates of the mixture model. For MGMM, each component density $p(e_n|\theta_k)$ is a normal probability distribution with $P = 3$ in our case:

$$p(e_n|\theta_k) = \frac{(2\pi)^{-p/2}}{\sqrt{det(\Sigma_{k,e})}} \exp\left[-\frac{(e_n-\mu_{k,e})^T (\Sigma_{k,e})^{-1} (e_n-\mu_{k,e})}{2}\right] \quad (10)$$

where $\mu_{k,e}$ is the mean and $\Sigma_{k,e}$ is the covariance matrix of kth component of the mixture. Thus the complete set of MGMM parameters is $\Theta = \{\theta_1, \ldots, \theta_K, \alpha_1, \ldots, \alpha_K\}$ where $\theta_k = \{\mu_{k,e}, \Sigma_{k,e}\}$, $k = 1, \ldots, K$.

4.2 Expectation Maximization

The most widely used approach for the estimation of the MGMM parameter set Θ, from a given dataset is to use Maximum Likelihood Estimation *(MLE)*:

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmax}}\, p(e|\theta) \quad (11)$$

where $f(\theta) = p(e|\theta)$ is the likelihood function. The EM algorithm is a general iterative technique for computing MLE when observed data can be considered as incomplete. The

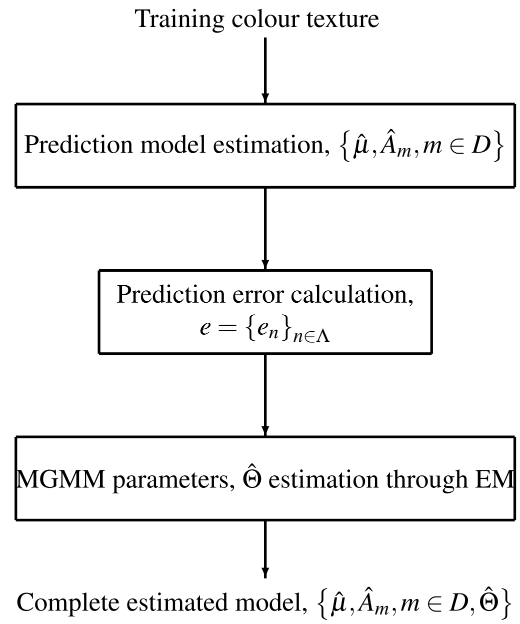


Figure 2: Block Diagram of the model parameter estimation for each texture class in the colour image.

algorithm consists of two steps: An E-step and an M-step. The EM algorithm produces a sequence of estimates $\Theta^{(t)}$, $t = 0, 1, 2, \ldots$ by repeating these two steps. If $\Theta^{(t)}$ denotes the estimated MGMM parameter set at iteration t, then at iteration $(t+1)$ the E-step computes the expected complete data log-likelihood function:

$$Q\left(\Theta, \Theta^{(t)}\right) = \sum_{n\in\Lambda} \sum_{k=1}^{K} \{\log \alpha_k p(e_n|\theta_k)\} P\left(k|e_n; \Theta^{(t)}\right) \quad (12)$$

where $P\left(k|e_n; \Theta^{(t)}\right)$ is the *a posteriori* probability and is computed as:

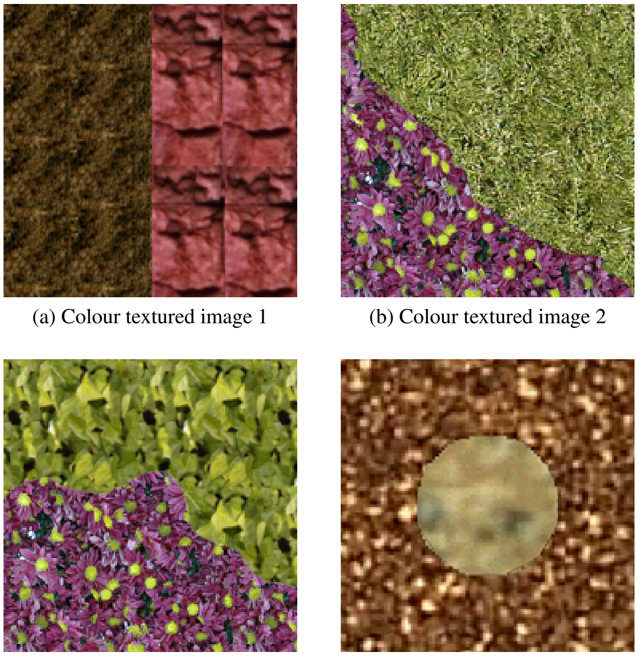
$$P\left(k|e_n; \Theta^{(t)}\right) = \frac{\alpha_k{}^{(t)} p\left(e_n|\theta_k{}^{(t)}\right)}{\sum_{l=1}^{K} \alpha_l{}^{(t)} p\left(e_n|\theta_l{}^{(t)}\right)} \quad (13)$$

The M-step finds the estimate of Θ at iteration $(t+1)$, by maximizing the hidden variable $Q\left(\Theta, \Theta^{(t)}\right)$:

$$\alpha_k{}^{(t+1)} = \frac{1}{|\Lambda|} \sum_{n\in\Lambda} P\left(k|e_n; \Theta^{(t)}\right) \quad (14)$$

$$\mu_{k,e}{}^{(t+1)} = \frac{\sum_{n\in\Lambda} e_n P\left(k|e_n; \Theta^{(t)}\right)}{\sum_{n\in\Lambda} P\left(k|e_n; \Theta^{(t)}\right)} \quad (15)$$

$$\Sigma_{k,e}{}^{(t+1)} = \frac{\sum_{n\in\Lambda} P\left(k|e_n; \Theta^{(t)}\right) (e_n - \mu_{k,e}) (e_n - \mu_{k,e})^T}{\sum_{n\in\Lambda} P\left(k|e_n; \Theta^{(t)}\right)} \quad (16)$$



(a) Colour textured image 1 (b) Colour textured image 2

(c) Colour textured image 3 (d) Colour textured image 4

Figure 3: Test colour textured images.

EM algorithm is strongly dependent on initialization of parameter values. One way is to start with a number of random starts and then assigning the final value which gives the maximum-likelihood. This will increase the computation time, as evident. In our approach we have used well known K-means algorithm to compute the initial values of MGMM parameter set Θ. Figure 2 shows the block diagram of model parameter estimation proposed in our segmentation algorithm. Finally, a coarse class label set for colour texture segmentation by the first phase of our approach is assigned according to a global criterion which maximizes the sum of the probability of all the mixture components, and is written mathematically as:

$$\hat{l}_n = \underset{c=1,\dots,C}{\operatorname{argmax}} \left(p\left(e_n | \hat{\Theta}_c \right) \right) \tag{17}$$

where C is the total number of classes in the test colour image and parameter set used in Equation 17 are the ones we estimated at the end of EM algorithm. In the second phase we use Potts model for spatial regularization of this estimated class label field [2].

5. SIMULATIONS AND RESULTS

The ground truth data associated with complex natural images is difficult to estimate and its extraction is highly influenced by the subjectivity of the human operator. Thus, the evaluation of proposed texture segmentation algorithm was performed on natural as well as synthetic colour textures which possess unambiguous ground truth data. Test images were taken from the colour texture database used in [5]. The database was constructed using colour textures from Vistex and Photoshop databases.

In the first phase of the proposed supervised colour texture segmentation algorithm, a single sub image of size 32×32 was used as the training image for each class. Image observation model parameters, multichannel prediction error and MGMM parameter sets were computed for this sub image. Then these parameters are used to compute the initial class label field using Equation 17 for each of the four test textured colour images shown in Figure 3. This initial class label field is coarse in its spatial nature. Then, in the second phase of the algorithm, this coarse class label field is spatially regularized using Potts model. An iterative solution to the Potts model was computed through conventional ICM based on doubleton clique information.

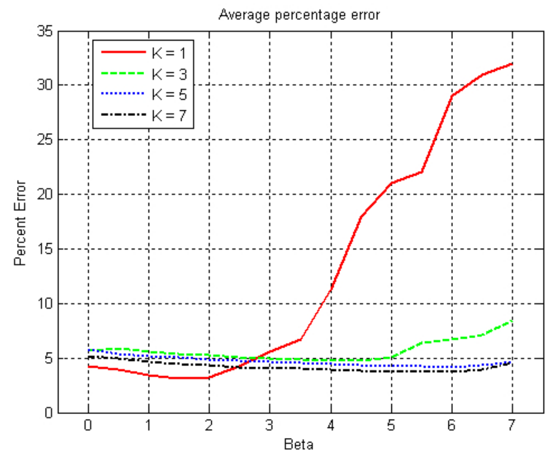


Figure 4: Average percentage error of four colour images for different values of K and β. Starting points of the curves at $\beta = 0$, show the average percentage pixel misclassification computed in first phase of the proposed algorithm.

We have evaluated the algorithm for all four textured colour images with different values of hyperparameter β of Potts model and for different number of components K in MGMM. In our experiments β was varied from 0 to 7 with a step of 0.5. Pixel misclassification percentage was computed for $K \in \{1,3,5,7\}$. When $K = 1$, linear prediction error sequence is modelled through a single Gaussian following the conventional model of the error sequence.

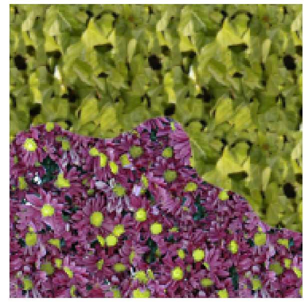
It is to note that the proposed method allows the description of a large family of distributions to model the pure colour content of the image. For this, we are modeling the multichannel prediction error of our combined luminance and chrominance structure model through MGMM to have the pure colour information of the image. The Gaussian distribution model (with a single component or a mixture model) for the multichannel error sequence is indirectly serving as the colour feature cue. One can observe that the percentage pixel misclassification obtained by the proposed method is minimum in the case of $K = 1$. This shows that the 2D multichannel complex version of QP AR linear prediction model, proposed in this paper can be used to efficiently model the textured colour images.

Figure 4 shows the graphical plot for average percentage pixel misclassification results for these four textures against different values of β and for different K. The values where $\beta = 0$, are the pixel misclassification percentage obtained without applying any spatial regularization technique (output of the first phase of our algorithm). The minimum average percentage pixel misclassification for $K = 1$ is 3.079% at

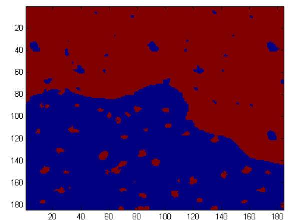
$\beta = 2$, for $K = 3$ is 4.754% at $\beta = 4.5$, for $K = 5$ is 4.18% at $\beta = 6$ and for $K = 7$ is 3.693% at $\beta = 5.5$.

Analyzing the curve shown in Figure 4, another important observation that could be made over the results is the stability of the results when $K > 3$ for $\beta \in [4,6]$. Contrarily, the results for $K = 1$ are unstable for $\beta > 3$. We can see that overall results obtained in the case when multichannel error sequence is modelled through a mixture of multivariate Gaussian distributions are relatively much stable than the case where it has been modelled through a single Gaussian distribution. Figure 5 shows the segmented images for test colour image 3. Higher values of K have better segmented results with high values of β.

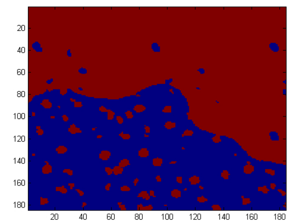
Even that we have several approaches for the estimation of optimal values of hyperparameters of spatial regularization techniques, adjustment or estimation of these parameters is an open and well known problem. Normally researchers adjust these values after extensive experimentation on their test databases or a value that gives the best results for a certain test dataset is chosen [2], [6]. The stable results obtained through our approach show considerable independence from any ***"to be adjusted or estimated"*** parameter (β in the case of ICM) of the spatial regularization techniques when $\beta \in [4,6]$.



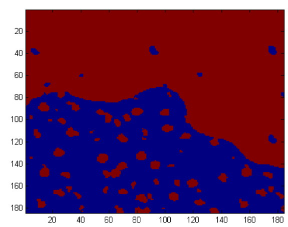
(a) Colour Image 3



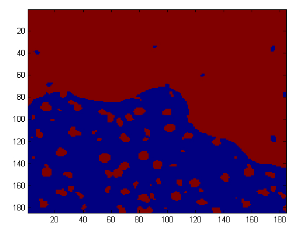
(b) $K = 1$, $\beta = 2.5$



(c) $K = 3$, $\beta = 7$



(d) $K = 5$, $\beta = 7$



(e) $K = 7$, $\beta = 7$

Figure 5: Segmented images for test colour image 3.

6. CONCLUSION

In this paper a multichannel complex image observation model for psychovisual colour spaces, is presented for supervised colour texture segmentation. The model shows success in modelling the complex colour textures with good percentage pixel misclassification results. Also, multivariate Gaussian mixture model for the multichannel complex linear prediction error is presented which is normally modelled by a single Gaussian distribution. The model has shown some stable results in terms of percentage pixel misclassification over changing values of adjustable regularization parameters. This percentage is computed for different number of components of mixture model and different values of spatial regularization hyperparameter β. The results show approximately same range of percentage pixel misclassification as compared to the single Gaussian model but a certain improvement in stability. Results obtained by this approach also show a certain degree of independence from the effect of manually set parameter for spatial regularization.

REFERENCES

[1] Colorimetry. Technical Report CIE 15.2, Commission Internationale de l'Eclairage, 1986.

[2] O. Alata and C. Ramananjarasoa. Unsupervised textured image segmentation using 2-d quarter plane autoregressive model with four prediction supports. *Pattern Recogn. Lett.*, 26(8):1069–1081, 2005.

[3] X. Guyon. *Random Fields on a Network - Modeling, Statistics and Application*. Probability and its Applications. Springer-Verlag, 1995.

[4] A. Hanbury and J. Serra. A 3d-polar coordinate colour representation well adapted to image analysis. In *Proceedings of the Scandinavian Conference on Image Analysis (SCIA)*, pages 804–811, 2003.

[5] D. Ilea and P. Whelan. Ctexan adaptive unsupervised segmentation algorithm based on color-texture coherence. *IEEE Transactions on Image Processing*, 17(10):1926–1939, 2008.

[6] Z. Kato and T.-C. Pong. A markov random field image segmentation model for color textured images. *Image Vision Comput.*, 24(10):1103–1114, 2006.

[7] B. Manjunath and R. Chellappa. Unsupervised texture segmentation using markov random field models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(5):478–482, 1991.

[8] B. Micusík and A. Hanbury. Supervised texture detection in images. In *CAIP*, pages 441–448, 2005.

[9] D. K. Panjwani and G. Healey. Markov random field models for unsupervised segmentation of textured color images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 17(10):939–954, 1995.

[10] H. Permuter, J. Francos, and I. Jermyn. A study of gaussian mixture models of color and texture features for image classification and segmentation. *Pattern Recogn.*, 39(4):695–706, 2006.

[11] I. Qazi, O. Alata, J. C. Burie, and C. Fernandez-Maloigne. Spatial structure characterization of textures in ihls colour space. In *To appear in Proceedings of ICASSP*, 2009.