Image Segmentation Method Based On Finite Doubly Truncated Bivariate Gaussian Mixture Model with Hierarchical Clustering

G.V.S. RAJKUMAR[1], K.SRINIVASA RAO[2] and P.SRINIVASA RAO[3]

[1] **Department of Information Technology, GITAM University, Visakhapatnam, Andhra Pradesh, INDIA**

[2] **Department of Statistics , Andhra University , Visakhapatnam, Andhra Pradesh, INDIA**

[3] **Department of Computer Science and Systems Engineering, Andhra University , Visakhapatnam, Andhra Pradesh, INDIA**

Abstract

Image segmentation is one of the most important area of image retrieval. In colour image segmentation the feature vector of each image region is 'n' dimension different from grey level image. In this paper a new image segmentation algorithm is developed and analyzed using the finite mixture of doubly truncated bivariate Gaussian distribution by integrating with the hierarchical clustering. The number of image regions in the whole image is determined using the hierarchical clustering algorithm. Assuming that a bivariate feature vector (consisting of Hue angle and Saturation) of each pixel in the image region follows a doubly truncated bivariate Gaussian distribution, the segmentation algorithm is developed. The model parameters are estimated using EM-Algorithm, the updated equations of EM-Algorithm for a finite mixture of doubly truncated Gaussian distribution are derived. A segmentation algorithm for colour images is proposed by using component maximum likelihood. The performance of the developed algorithm is evaluated by carrying out experimentation with six images taken form Berkeley image dataset and computing the image segmentation metrics like, Probabilistic Rand Index (PRI), Global Consistency Error (GCE), and Variation of Information (VOI). The experimentation results show that this algorithm outperforms the existing image segmentation algorithms.

Keywords: *Image Segmentation, Finite Doubly Truncated Bivariate Gaussian distribution, Hierarchical Clustering, Image Quality Metrics, EM algorithm.*

1. Introduction

Image segmentation is an important early vision task, where pixels with similar features are grouped into homogeneous regions (Zoltan Kato and Ting-Chuen Pong, 2006). Colour image segmentation has recently been an intensive topic in image processing and computer vision. It can be viewed as an generalization of grey level image segmentation (Cheng etal, 2001). The image segmentation methods can be classified into four categories namely cluster based segmentation (Roberts.S.J. (1997)), edge or contour detection based segmentation (Canny (1986)), Region or area extraction based segmentation (Adams and Bischof (1994)) and model based segmentation. Among these, model based image segmentation is more efficient compared to other methods (Lei T.and Udupa J. (2003)). In image segmentation, unsupervised clustering techniques have high reproducibility because its results 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 (Sujaritha M. and Annadurai S. (2010)).

Much work has been reported regarding image segmentation using Gaussian mixture model. In Gaussian mixture model, each image is divided into K regions and each region is characterized by a Gaussian distribution. For obtaining the number of regions, it is customary to consider the K-means algorithm (Prasad Reddy P.V.G.D. et.al (2007)). However, the K-means algorithm requires the initial number of clusters and it is semi-unsupervised. To overcome this problem one can consider hierarchical clustering. In hierarchical clustering, the clusters are formed by collecting similar clusters at different levels of detail into a single cluster using a multi branch tree structure. Such a system allows for selection of a set of clusters for further exploration. A number of Segmentation algorithms for building a hierarchical image representation have been proposed (Montanvert A. et al (1991), Nacken P.F.M. (1995) and Molina I. et al (2000)).

In colour images the pixel intensity distribution may not be able to characterize the image region as in the case of grey images. This problem can be resolved by considering the feature vector of the colour image. The colour features can be extracted from the images by utilizing CIE-L*U*V* colour space using the method given by (Sangwine and Horne (1998)). In this colour space the important features of each colour image region can be identified as Hue angle and Saturation. Even the value is also important it can be considered as a function of Hue and Saturation. This is also supported by experimental evidence. In addition to this the feature vector is having finite range and may not be symmetric and mesokurtic. To have an accurate modeling of this feature vector one has to consider the truncation of bivariate Gaussian distribution of each image region. This consideration characterizes the whole image as a finite doubly truncated bivariate Gaussian mixture model.

Hence, in this paper we develop an efficient image segmentation algorithm by assuming that the feature vector of the entire image follows a finite doubly truncated bivariate Gaussian mixture distribution and hierarchical clustering. The numbers of image regions are determined by using the hierarchical clustering. The model parameters are estimated with the updated equations of the EM algorithm. The EM algorithm (Expectation Maximization algorithm) has been extensively used to estimate the mixture parameters (McLanchlan G. and Krishnan T. (1997), Yiming Wu et al (2003), Lei T. et al (2003)).

In any image region the feature vector lie between two finite values and may be distributed as asymmetric. Neglecting the reality of finite range leads to serious falsification of model estimation. Let W = [X Y] is a bivariate random variable representing feature vector of an image region. The probability density function of the Doubly Truncated Bivariate Gaussian distribution is

$$g(x,y,\theta)=\frac{f(x,y)}{\int_{b_1}^{a_1}\int_{b_2}^{a_2} f(x,y)\,dxdy}\ ,\ b_1 < x < a_1\ ;\ b_2 < y < a_2 \tag{1}$$

where, (b_1, a_1) are the truncation points of the Hue angle and (b_2, a_2) are the truncation points of the Saturation, $f(x,y)$ is the probability density function of the bivariate Normal distribution is

$$f(x,y)=\frac{1}{2\pi\sqrt{1-\rho^2}\,\sigma_1\sigma_2}\exp\left\{\frac{-1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\right\}$$

$$-\infty<x<+\infty;\ -\infty<y<+\infty,\quad \sigma_1>0;\ \sigma_2>0;\ -1<\rho<1,\quad -\infty<\mu_1<+\infty;\ -\infty<\mu_2<+\infty \tag{2}$$

The value of $\left[1-\int_{b_1}^{a_1}\int_{b_2}^{a_2} f(x,y)\,dxdy\right]$ is significant based on the values of the parameters. This distribution includes the skewed, asymmetric and finite range bivariate distributions as particular cases for limiting and specific values of the parameters. This model also includes bivariate Gaussian distribution as a limiting case. The various shapes of the frequency curves of the doubly truncated bivariate Gaussian distribution are shown in Figure 1.

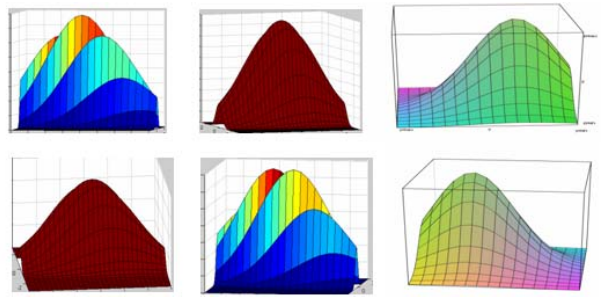

Figure 1: Shapes of doubly truncated bivariate Gaussian frequency surfaces

As a result of this finite range in the feature vector, it is needed to consider the feature vector of the entire image follows a finite doubly truncated bivariate Gaussian mixture distribution.

The performance of the developed segmentation algorithm is compared with finite Gaussian mixture model with K-Means through Image segmentation Measures like, Probabilistic Rand Index (PRI), the Variation of Information (VOI), and Global Consistency measure (GCE), and the Image Quality Measures like, Maximum distance, Image Fidelity, Mean Square Error, Signal to Noise Ratio, and Image Quality Index. Six images namely OSTRICH, POT, TOWER, BEARS, DEER and BIRD are used for experimentation.

2. Truncated Gaussian Mixture Model

The effect of truncation in bivariate Gaussian distribution has been discussed by several researchers (Norman L.Johnson, Samuel Kotz and Balakrishnan (1994)). Following the heuristic arguments given by Bengt Muthen (1990), the mean value of 'X'(hue) is obtained as

$$E(X) = \mu_1 + \sigma_1 A \tag{3}$$

where,

$$\begin{aligned}A=\Bigg[&-\phi\left(\frac{a_1-\mu_1}{\sigma_1}\right)\left[\Phi\left[\left(\left(\frac{a_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{a_1-\mu_1}{\sigma_1}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{a_1-\mu_1}{\sigma_1}\right)\right)c\right]\right]\\
&+\phi\left(\frac{b_1-\mu_1}{\sigma_1}\right)\left[\Phi\left[\left(\left(\frac{a_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{b_1-\mu_1}{\sigma_1}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{b_1-\mu_1}{\sigma_1}\right)\right)c\right]\right]\\
&-\rho\phi\left(\frac{a_2-\mu_2}{\sigma_2}\right)\left[\Phi\left[\left(\left(\frac{a_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{a_2-\mu_2}{\sigma_2}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{a_2-\mu_2}{\sigma_2}\right)\right)c\right]\right]\\
&+\rho\phi\left(\frac{b_2-\mu_2}{\sigma_2}\right)\left[\Phi\left[\left(\left(\frac{a_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{b_2-\mu_2}{\sigma_2}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{b_2-\mu_2}{\sigma_2}\right)\right)c\right]\right]\Bigg]\end{aligned}$$

and $c=\left(1-\rho^2\right)^{-1/2}$, ϕ, Φ are the ordinate and area of a standard Normal distribution.

Similarly the mean value of 'Y'(saturation) is

$$E(Y) = \mu_2 + \sigma_2 B \tag{4}$$

where,

$$
\begin{aligned}
B=\Bigg[& -\phi\left(\frac{a_2-\mu_2}{\sigma_2}\right)\left[\Phi\left[\left(\left(\frac{a_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{a_2-\mu_2}{\sigma_2}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{a_2-\mu_2}{\sigma_2}\right)\right)c\right]\right] \\
& +\phi\left(\frac{b_2-\mu_2}{\sigma_2}\right)\left[\Phi\left[\left(\left(\frac{a_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{b_2-\mu_2}{\sigma_2}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{b_2-\mu_2}{\sigma_2}\right)\right)c\right]\right] \\
& -\rho\phi\left(\frac{a_1-\mu_1}{\sigma_1}\right)\left[\Phi\left[\left(\left(\frac{a_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{a_1-\mu_1}{\sigma_1}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{a_1-\mu_1}{\sigma_1}\right)\right)c\right]\right] \\
& +\rho\phi\left(\frac{b_1-\mu_1}{\sigma_1}\right)\left[\Phi\left[\left(\left(\frac{a_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{b_1-\mu_1}{\sigma_1}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{b_1-\mu_1}{\sigma_1}\right)\right)c\right]\right]\Bigg]
\end{aligned}
$$

and $c=\left(1-\rho^2\right)^{-1/2}$

The Variance of X is
The Variance of X is

$$V(X) = \sigma_1^2 R - 2A\ \sigma_1\ A + A^2$$

$$= \sigma_1^2 R - A^2(2\sigma_1 - 1) \qquad (5)$$

where,

$$
\begin{aligned}
R=\Bigg[& \pi-\left(\frac{a_1-\mu_1}{\sigma_1}\right)\phi\left(\frac{a_1-\mu_1}{\sigma_1}\right)\left[\Phi\left[\left(\left(\frac{a_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{a_1-\mu_1}{\sigma_1}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{a_1-\mu_1}{\sigma_1}\right)\right)c\right]\right] \\
& +\left(\frac{b_1-\mu_1}{\sigma_1}\right)\phi\left(\frac{b_1-\mu_1}{\sigma_1}\right)\left[\Phi\left[\left(\left(\frac{a_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{b_1-\mu_1}{\sigma_1}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{b_1-\mu_1}{\sigma_1}\right)\right)c\right]\right] \\
& +\rho^2\left(\frac{a_1-\mu_1}{\sigma_1}\right)\phi\left(\frac{a_1-\mu_1}{\sigma_1}\right)\left[\Phi\left[\left(\left(\frac{a_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{a_1-\mu_1}{\sigma_1}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{a_1-\mu_1}{\sigma_1}\right)\right)c\right]\right] \\
& +c^{-1}\rho\phi\left(\frac{a_1-\mu_1}{\sigma_1}\right)\left[\phi\left[\left(\left(\frac{a_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{a_1-\mu_1}{\sigma_1}\right)\right)c\right]-\phi\left[\left(\left(\frac{b_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{a_1-\mu_1}{\sigma_1}\right)\right)c\right]\right] \\
& +\rho^2\left(\frac{b_1-\mu_1}{\sigma_1}\right)\phi\left(\frac{b_1-\mu_1}{\sigma_1}\right)\left[\Phi\left[\left(\left(\frac{a_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{b_1-\mu_1}{\sigma_1}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{b_1-\mu_1}{\sigma_1}\right)\right)c\right]\right] \\
& -c^{-1}\rho\phi\left(\frac{b_1-\mu_1}{\sigma_1}\right)\left[\phi\left[\left(\left(\frac{a_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{b_1-\mu_1}{\sigma_1}\right)\right)c\right]-\phi\left[\left(\left(\frac{b_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{b_1-\mu_1}{\sigma_1}\right)\right)c\right]\right]\Bigg]
\end{aligned}
$$

and $c=\left(1-\rho^2\right)^{-1/2}$ and A are given as in equation (3)

The Variance of Y is

$$V(Y) = \sigma_2^2 T - 2B\ \sigma_2\ B + B^2$$

$$= \sigma_2^2 T - B^2(2\sigma_2 - 1) \qquad (6)$$

where,

$$
\begin{aligned}
T=\Bigg[& \pi-\left(\frac{a_2-\mu_2}{\sigma_2}\right)\phi\left(\frac{a_2-\mu_2}{\sigma_2}\right)\left[\Phi\left[\left(\left(\frac{a_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{a_2-\mu_2}{\sigma_2}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{a_2-\mu_2}{\sigma_2}\right)\right)c\right]\right] \\
& +\left(\frac{b_2-\mu_2}{\sigma_2}\right)\phi\left(\frac{b_2-\mu_2}{\sigma_2}\right)\left[\Phi\left[\left(\left(\frac{a_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{b_2-\mu_2}{\sigma_2}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{b_2-\mu_2}{\sigma_2}\right)\right)c\right]\right] \\
& +\rho^2\left(\frac{a_2-\mu_2}{\sigma_2}\right)\phi\left(\frac{a_2-\mu_2}{\sigma_2}\right)\left[\Phi\left[\left(\left(\frac{a_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{a_2-\mu_2}{\sigma_2}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{a_2-\mu_2}{\sigma_2}\right)\right)c\right]\right] \\
& +c^{-1}\rho\phi\left(\frac{a_2-\mu_2}{\sigma_2}\right)\left[\phi\left[\left(\left(\frac{a_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{a_2-\mu_2}{\sigma_2}\right)\right)c\right]-\phi\left[\left(\left(\frac{b_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{a_2-\mu_2}{\sigma_2}\right)\right)c\right]\right] \\
& +\rho^2\left(\frac{b_2-\mu_2}{\sigma_2}\right)\phi\left(\frac{b_2-\mu_2}{\sigma_2}\right)\left[\Phi\left[\left(\left(\frac{a_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{b_2-\mu_2}{\sigma_2}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{b_2-\mu_2}{\sigma_2}\right)\right)c\right]\right] \\
& -c^{-1}\rho\phi\left(\frac{b_2-\mu_2}{\sigma_2}\right)\left[\phi\left[\left(\left(\frac{a_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{b_2-\mu_2}{\sigma_2}\right)\right)c\right]-\phi\left[\left(\left(\frac{b_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{b_2-\mu_2}{\sigma_2}\right)\right)c\right]\right]\Bigg]
\end{aligned}
$$

$c=\left(1-\rho^2\right)^{-1/2}$ and B are as given in equation (4)

The Covariance of (X, Y) is

$$COV(X, Y) = \sigma_1\ \sigma_2\ U - AB\,[\sigma_1 + \sigma_2 - 1] \qquad (7)$$

where,

$$
\begin{aligned}
U=\Bigg[& \rho\pi-\rho\left(\frac{a_1-\mu_1}{\sigma_1}\right)\phi\left(\frac{a_1-\mu_1}{\sigma_1}\right)\left[\Phi\left[\left(\left(\frac{a_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{a_1-\mu_1}{\sigma_1}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{a_1-\mu_1}{\sigma_1}\right)\right)c\right]\right] \\
& +c^{-1}\phi\left(\frac{a_1-\mu_1}{\sigma_1}\right)\left[\phi\left[\left(\left(\frac{a_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{a_1-\mu_1}{\sigma_1}\right)\right)c\right]-\phi\left[\left(\left(\frac{b_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{a_1-\mu_1}{\sigma_1}\right)\right)c\right]\right] \\
& +\rho\left(\frac{b_1-\mu_1}{\sigma_1}\right)\phi\left(\frac{b_1-\mu_1}{\sigma_1}\right)\left[\Phi\left[\left(\left(\frac{a_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{b_1-\mu_1}{\sigma_1}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{b_1-\mu_1}{\sigma_1}\right)\right)c\right]\right] \\
& -c^{-1}\phi\left(\frac{b_1-\mu_1}{\sigma_1}\right)\left[\phi\left[\left(\left(\frac{a_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{b_1-\mu_1}{\sigma_1}\right)\right)c\right]-\phi\left[\left(\left(\frac{b_2-\mu_2}{\sigma_2}\right)-\rho\left(\frac{b_1-\mu_1}{\sigma_1}\right)\right)c\right]\right] \\
& +\rho\left(\frac{a_2-\mu_2}{\sigma_2}\right)\phi\left(\frac{a_2-\mu_2}{\sigma_2}\right)\left[\Phi\left[\left(\left(\frac{a_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{a_2-\mu_2}{\sigma_2}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{a_2-\mu_2}{\sigma_2}\right)\right)c\right]\right] \\
& +\rho\left(\frac{b_2-\mu_2}{\sigma_2}\right)\phi\left(\frac{b_2-\mu_2}{\sigma_2}\right)\left[\Phi\left[\left(\left(\frac{a_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{b_2-\mu_2}{\sigma_2}\right)\right)c\right]-\Phi\left[\left(\left(\frac{b_1-\mu_1}{\sigma_1}\right)-\rho\left(\frac{b_2-\mu_2}{\sigma_2}\right)\right)c\right]\right]\Bigg]
\end{aligned}
$$

$c=\left(1-\rho^2\right)^{-1/2}$, A and B are as given in equations (3) and (4) respectively.

Since the entire image is a collection of regions, which are characterized by doubly truncated bivariate Normal distribution, it can be characterized through a *K*-Component finite doubly truncated bivariate Gaussian distribution and its probability density function is of the form

$$h(x,y)=\sum_{i=1}^{K}\alpha_i g_i(x_i,y_i;\theta) \qquad (8)$$

where, *K* is the number of regions, $0\le\alpha_i\le 1$; i = 1,2,…,*K* are weights such that $\sum_{i=1}^{K}\alpha_i=1$ and $\theta=\left\{\mu_{1k},\mu_{2k},\sigma_{1k}^2,\sigma_{2k}^2,\rho_k\right\}$ is the parametric set. $g_i(x,y/\theta_i)$ given in equation (1) represent the probability density function of the i[th] image region. α_i is the probability of occurrence of the i[th] component of the finite doubly truncated bivariate Gaussian mixture model (FDTBGMM) i.e., the probability that the feature belongs to the i[th] image region.

The mean vector representing the entire image is

$$E(W^T)=\begin{bmatrix}\sum_{i=1}^{K}\alpha_i E(X_i)\\ \sum_{i=1}^{K}\alpha_i E(Y_i)\end{bmatrix} \qquad (9)$$

where, $E_i(X)$ and $E_i(Y)$ are given in equations (3) and (4) for the i[th] image region.

3. Estimation of the Model Parameters by EM Algorithm

To obtain the estimation of model the model parameters, we utilized the EM-Algorithm by maximizing the expected likelihood function for carrying out the EM-Algorithm. The likelihood function of bivariate observations $(x_1,y_1),(x_2,y_2),(x_3,y_3),\ldots,(x_N,y_N)$ drawn from an image is

$$L(\theta) = \prod_{s=1}^{N} h(x_s, y_s; \theta)$$

$$= \prod_{s=1}^{N} \left(\sum_{i=1}^{K} \alpha_i g_i(x_s, y_s; \theta)\right) \qquad (10)$$

$$= \prod_{s=1}^{N}\left(\sum_{i=1}^{K}\alpha_i \frac{\exp\left\{\frac{-1}{2(1-\rho_i^2)}\left[\left(\frac{x_s-\mu_{1i}}{\sigma_{1i}}\right)^2 - 2\rho_i\left(\frac{x_s-\mu_{1i}}{\sigma_{1i}}\right)\left(\frac{y_s-\mu_{2i}}{\sigma_{2i}}\right)+\left(\frac{y_s-\mu_{2i}}{\sigma_{2i}}\right)^2\right]\right\}}{2\pi\sqrt{1-\rho_i^2}\,\sigma_{1i}\sigma_{2i}\int_{b_1}^{a_1}\int_{b_2}^{a_2} f_i(x,y;\theta)\,dxdy}\right)$$

This implies $\log L(\theta) = \log \prod_{s=1}^{N}\left(\sum_{i=1}^{K}\alpha_i g_i(x_s, y_s; \theta)\right)$

$$(11)$$

The updated equations of EM algorithm for estimating the model parameters are

$$\alpha_k^{(l+1)} = \frac{1}{N}\sum_{s=1}^{N}\left[t_k(x_s, y_s; \theta^{(l)})\right]$$

$$= \frac{1}{N}\sum_{s=1}^{N}\left(\frac{\alpha_k^{(l)} g_k(x_s, y_s; \theta^{(l)})}{\sum_{i=1}^{K}\alpha_k^{(l)} g_i(x_s, y_s; \theta^{(l)})}\right) \qquad (12)$$

where, $g_k(x_s, y_s; \theta^{(l)})$ is as given in equation (1).

For updating μ_{1k} we have,

$$\mu_{1k}^{(l+1)}\sum_{s=1}^{N} t_k(x_s,y_s;\theta^{(l)}) - \sum_{s=1}^{N} t_k(x_s,y_s;\theta^{(l)})x_s + \sum_{s=1}^{N} t_k(x_s,y_s;\theta^{(l)})\sigma_{1k}^{(l)}\left[\frac{\rho_k^{(l)}(y_s-\mu_{2k}^{(l)})}{\sigma_{2k}^{(l)}} + \left[\mathrm{A}-\rho_k^{(l)}\mathrm{B}\right]\right]=0 \qquad (13)$$

For updating μ_{2k}, we have ,

$$\mu_{2k}^{(l+1)}\sum_{s=1}^{N} t_k(x_s,y_s;\theta^{(l)}) - \sum_{s=1}^{N} t_k(x_s,y_s;\theta^{(l)})y_s + \sum_{s=1}^{N} t_k(x_s,y_s;\theta^{(l)})\sigma_{2k}^{(l)}\left[\frac{\rho_k^{(l)}(x_s-\mu_{1k}^{(l)})}{\sigma_k^{(l)}} + \left[\mathrm{B}-\rho_k^{(l)}\mathrm{A}\right]\right]=0 \qquad (14)$$

where,

$$t_k(x_s, y_s; \theta^{(l)}) = \frac{\alpha_k^{(l)} g_k(x_s, y_s; \theta^{(l)})}{h(x_s, y_s; \theta^{(l)})} = \frac{\alpha_k^{(l)} g_k(x_s, y_s; \theta^{(l)})}{\sum_{i=1}^{K}\alpha_i^{(l)} g_i(x_s, y_s; \theta^{(l)})},$$

$$g_k(x_s, y_s; \theta^{(l)}) = \frac{\exp\left\{\frac{-1}{2(1-\rho_k^2)}\left[\left(\frac{x_s-\mu_{1k}}{\sigma_{1k}}\right)^2 - 2\rho_k\left(\frac{x_s-\mu_{1k}}{\sigma_{1k}}\right)\left(\frac{y_s-\mu_{2k}}{\sigma_{2k}}\right)+\left(\frac{y_s-\mu_{2k}}{\sigma_{2k}}\right)^2\right]\right\}}{2\pi\sigma_{1k}\sigma_{2k}\sqrt{1-\rho_k^2}\int_{b_1}^{a_1}\int_{b_2}^{a_2} f_k(x,y;\theta)\,dxdy},$$

A and B are as given in equations (3) and (4) respectively.

Solving the equations (13) and (14) we get the revised estimates of μ_{1k} and μ_{2k}.

For updating σ_{1k}^2, we have ,

$$\sum_{s=1}^{N} t_k(x_s,y_s;\theta^{(l)})\left[\left[\left(\frac{x_s-\mu_{1k}^{(l)}}{\sigma_{1k}^{(l+1)}}\right)^2 - \frac{\rho_k(x_s-\mu_{1k}^{(l)})(y_s-\mu_{2k}^{(l)})}{\sigma_{1k}^{(l+1)}\sigma_{2k}^{(l)}}\right] - [D] + \rho_k^{(l)}[E]\right] = 0 \qquad (15)$$

For updating σ_{2k}^2, we have ,

$$\sum_{s=1}^{N} t_k(x_s,y_s;\theta^{(l)})\left[\left[\left(\frac{y_s-\mu_{2k}^{(l)}}{\sigma_{2k}^{(l+1)}}\right)^2 - \frac{\rho_k(x_s-\mu_{1k}^{(l)})(y_s-\mu_{2k}^{(l)})}{\sigma_{1k}^{(l)}\sigma_{2k}^{(l+1)}}\right] - [G] + \rho_k^{(l)}[E]\right] = 0 \qquad (16)$$

For updating ρ_k, we have,

$$\sum_{s=1}^{N} t_i(x_s,y_s;\theta^{(l)})\left[-\frac{\rho_k}{(1-\rho_k^2)^2}\left[\left(\frac{x_s-\mu_{1k}}{\sigma_{1k}}\right)^2 + \left(\frac{y_s-\mu_{2k}}{\sigma_{2k}}\right)^2\right] - \frac{1+\rho_k^2}{(1-\rho_k^2)^2}\left[\left(\frac{x_s-\mu_{1k}}{\sigma_{1k}}\right)\left(\frac{y_s-\mu_{2k}}{\sigma_{2k}}\right)\right] + \frac{\rho_k}{(1-\rho_k^2)^2}[D+F] + \frac{1+\rho_k^2}{(1-\rho_k^2)^2}[E]\right] = 0 \qquad (17)$$

where,

$$\begin{aligned} D = {} & \pi\sigma_{1k}\sigma_{2k} + \sigma_{1k}\sigma_{2k}c^{-1}\rho_k\Big[\phi(\mathrm{A}_1)\big[\phi[((\mathrm{A}_1)\text{-}\rho_k(\mathrm{A}_1))c] - \phi[((\mathrm{B}_1)\text{-}\rho_k(\mathrm{A}_1))c]\big] \\ & - \phi(\mathrm{B}_1)\big[\phi[((\mathrm{A}_1)\text{-}\rho_k(\mathrm{B}_1))c] - \phi[((\mathrm{B}_1)\text{-}\rho_k(\mathrm{B}_1))c]\big]\Big] \\ & + \sigma_{2k}(\rho_k^2-1)(a_1\text{-}\mu_{1k})\phi(\mathrm{A}_1)\big[\Phi[((\mathrm{A}_1)\text{-}\rho_k(\mathrm{A}_1))c] - \Phi[((\mathrm{B}_1)\text{-}\rho_k(\mathrm{A}_1))c]\big] \\ & + \sigma_{2k}(\rho_k^2+1)(b_1\text{-}\mu_{1k})\phi(\mathrm{B}_1)\big[\Phi[((\mathrm{A}_1)\text{-}\rho_k(\mathrm{B}_1))c] - \Phi[((\mathrm{B}_1)\text{-}\rho_k(\mathrm{B}_1))c]\big] \end{aligned}$$

$$\begin{aligned} E = {} & \rho_k\pi\sigma_{1k}\sigma_{2k} + \sigma_{1k}\sigma_{2k}c^{-1}\Big[\phi(\mathrm{A}_1)\big[\phi[((\mathrm{A}_2)\text{-}\rho_k(\mathrm{A}_1))c] - \phi[((\mathrm{B}_2)\text{-}\rho_k(\mathrm{A}_1))c]\big] \\ & - \phi(\mathrm{B}_1)\big[\phi[((\mathrm{A}_2)\text{-}\rho_k(\mathrm{B}_1))c] - \phi[((\mathrm{B}_2)\text{-}\rho_k(\mathrm{B}_1))c]\big]\Big] \\ & + \rho_k\sigma_{1k}\Big[(a_2\text{-}\mu_{2k})\phi(\mathrm{A}_2)\big[\Phi[((\mathrm{A}_1)\text{-}\rho_k(\mathrm{A}_2))c] - \Phi[((\mathrm{B}_1)\text{-}\rho_k(\mathrm{A}_2))c]\big] \\ & + (b_2\text{-}\mu_{2k})\phi(\mathrm{B}_2)\big[\Phi[((\mathrm{A}_1)\text{-}\rho_k(\mathrm{B}_2))c] - \Phi[((\mathrm{B}_1)\text{-}\rho_k(\mathrm{B}_2))c]\big]\Big] \\ & + \rho_k\sigma_{2k}\Big[(b_1\text{-}\mu_{1k})\phi(\mathrm{B}_1)\big[\Phi[((\mathrm{A}_2)\text{-}\rho_k(\mathrm{B}_1))c] - \Phi[((\mathrm{B}_2)\text{-}\rho_k(\mathrm{B}_1))c]\big] \\ & + (a_1\text{-}\mu_{1k})\phi(\mathrm{A}_1)\big[\Phi[((\mathrm{A}_2)\text{-}\rho_k(\mathrm{A}_1))c] - \Phi[((\mathrm{B}_2)\text{-}\rho_k(\mathrm{A}_1))c]\big]\Big] \end{aligned}$$

$$\begin{aligned} G = {} & \pi\sigma_{1k}\sigma_{2k} + \sigma_{1k}\sigma_{2k}c^{-1}\rho_k\Big[\phi(\mathrm{A}_2)\big[\phi[((\mathrm{A}_2)\text{-}\rho_k(\mathrm{A}_2))c] - \phi[((\mathrm{B}_2)\text{-}\rho_k(\mathrm{A}_2))c]\big] \\ & - \phi(\mathrm{B}_2)\big[\phi[((\mathrm{A}_2)\text{-}\rho_k(\mathrm{B}_2))c] - \phi[((\mathrm{B}_2)\text{-}\rho_k(\mathrm{B}_2))c]\big]\Big] \\ & + \sigma_{1k}(\rho_k^2-1)(a_2\text{-}\mu_{2k})\phi(\mathrm{A}_2)\big[\Phi[((\mathrm{A}_2)\text{-}\rho_k(\mathrm{A}_2))c] - \Phi[((\mathrm{B}_2)\text{-}\rho_k(\mathrm{A}_2))c]\big] \\ & + \sigma_{1k}(\rho_k^2+1)(b_2\text{-}\mu_{2k})\phi(\mathrm{B}_2)\big[\Phi[((\mathrm{A}_2)\text{-}\rho_k(\mathrm{B}_2))c] - \Phi[((\mathrm{B}_2)\text{-}\rho_k(\mathrm{B}_2))c]\big] \end{aligned}$$

IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 4, No 2, July 2011
ISSN (Online): 1694-0814
www.IJCSI.org

and

$$
\begin{aligned}
F = \pi\sigma_{1k}\sigma_{2k} - \sigma_{1k}\Big[& (a_2-\mu_{2k})\phi(A_2)\big[\Phi\big[((A_2)-\rho_k(A_2))c\big]-\Phi\big[((B_2)-\rho_k(A_2))c\big]\big] \\
& -(b_2-\mu_{2k})\,\phi(B_2)\big[\Phi\big[((A_2)-\rho_k(B_2))c\big]-\Phi\big[((B_2)-\rho_k(B_2))c\big]\big]\Big] \\
+\sigma_{1k}\rho_k^{2}\Big[& (a_2-\mu_{2k})\phi(A_2)\big[\Phi\big[((A_2)-\rho_k(A_2))c\big]-\Phi\big[((B_2)-\rho_k(A_2))c\big]\big] \\
& +(b_2-\mu_{2k})\phi(B_2)\big[\Phi\big[((A_2)-\rho_k(B_2))c\big]-\Phi\big[((B_2)-\rho(B_2))c\big]\big]\Big] \\
+\sigma_{1k}\sigma_{2k}\rho_k c^{-1}\Big[& \phi(A_2)\big[\phi\big[((A_2)-\rho_k(A_2))c\big]-\phi\big[((B_2)-\rho_k(A_2))c\big]\big] \\
& -\phi(B_2)\big[\phi\big[((A_2)-\rho_k(B_2))c\big]-\phi\big[((B_2)-\rho_k(B_2))c\big]\big]\Big]
\end{aligned}
$$

where,

$$A_1 = \frac{a_1-\mu_{1k}}{\sigma_{1k}},\ A_2 = \frac{a_2-\mu_{2k}}{\sigma_{2k}},\quad B_1 = \frac{b_1-\mu_{1k}}{\sigma_{1k}},\ B_2 = \frac{b_2-\mu_{2k}}{\sigma_{2k}}$$

Solving the equations (15), (16) and (17) iteratively using we get the revised estimates of $\sigma_{1k}^2, \sigma_{2k}^2$ and ρ_k .

4. Initialization of Parameters by Hierarchical Clustering

To utilize the EM algorithm we have to initialize the parameters α_k , μ_{1k} , μ_{2k} , σ_{1k}^2 , σ_{2k}^2 , and ρ_k and , for k=1,2...,K. The Truncation Points (b_1, a_1) and (b_2, a_2) can be estimated with the values of the maximum and the minimum values of Hue and Saturation of the entire image respectively. The initial values of α_i can be taken as $\alpha_i = \frac{1}{K}$, where, K is the number of image regions obtained from the Hierarchical clustering algorithm. The steps involved in hierarchical clustering algorithm are as follows (S.C. Johnson (1967)).

Step 1: Start by assigning each item to a segment. Each of the N items, are associated with N segments, each containing just one item. Let the distances (similarities) between the segments be the same as the distances (similarities) between the items they contain.

Step 2: Find the closest (most similar) pair of segments and merge them into a single segment. The number of segments is now reduced by one. Compute distances (similarities) between the new segment and each of the old segments.

Step 3: Repeat the steps 2 and 3 until all items are segmented.

Step 3 can be done in different ways, namely i) Single-Linkage ii) Complete-Linkage and iii) Average-Linkage segmenting. We consider the Average -Linkage methodology.

In Average-Linkage segmenting (also called the unweighted pair-group method using arithmetic averages), is one of the most widely used hierarchical clustering algorithms. The average linkage algorithm (Richard A. Johnson and Dean W. Wichern (2009)) is obtained by defining the distance between two segments to be the average distance between a point in one segment and a point in the other segment. The algorithm is an agglomerative scheme that erases rows and columns in the proximity matrix as old segments are merged into new ones.

The proximity matrix is D = [d(i,j)]. The segments are assigned sequence numbers 0,1.., (n-1) and L(k) is the level of the K^{th} segment. A segment with sequence a number m is denoted (m) and the proximity between segments (r) and (s) is denoted d [(r),(s)].

The algorithm is composed of the following steps:

Begin with the disjoint segment having level L(0) = 0 and sequence number m = 0.

Find the average dissimilar pair of segments in the current segment, say pair (r), (s),
where the average of all pairs of segments in the current segment.

1. Increment the sequence number: m = m +1. Merge segments (r) and (s) into a single segment to form the next segmenting m. Set the level of this segmenting to L(m) = d[(r),(s)]

2. Update the proximity matrix, D, by deleting the rows and columns corresponding to segments (r) and (s) and adding a row and column corresponding to the newly formed segment. The proximity between the new segment, denoted (r,s) and old segment(K)is defined in this way.

$$d_{(r,s)K} = \frac{\sum_i \sum_j d(i,j)}{N_{(r,s)} N_K}$$

where, d(i,j) is the distance between object i in the cluster (r,s) and object j in the cluster K, and $N_{(r,s)}$ and N_K are the number of items in the clusters (r,s) and K respectively. The above procedure is repeated till the distance between two clusters is less than the specified threshold value.

After obtaining the final value for the number of regions K, we obtain the initial estimates of $\mu_{1k}, \mu_{2k}, \sigma_{1k}^2$, σ_{2k}^2 , and ρ_k for the k^{th} region using the segmented region values with the method of estimation given by Bengt Muthen (1990) for truncated bivariate Normal distribution with initial parameters. After getting these initial estimates for μ_{1k} , μ_{2k} , σ_{1k}^2 , σ_{2k}^2 , and ρ_k , we obtain the final refined estimates of the model parameters through EM algorithm given in section (3).

5. Segmentation Algorithm

After refining the parameters the prime step is image segmentation, by allocating the pixels to the segments. This operation is performed by Segmentation Algorithm. The image segmentation algorithm consists of four steps

IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 4, No 2, July 2011
ISSN (Online): 1694-0814
www.IJCSI.org

Step 1) Obtain the number of image regions using hierarchical clustering algorithm.
Step 2) obtain the initial estimates of the Model parameters using Hierarchical clustering and moment estimates for each image region as discussed in 4
Step 3) Using the initial estimates of the parameters the final estimates of the model parameters μ_{1k}, μ_{2k}, σ_{1k}^2, σ_{2k}^2, ρ_k and α_k for k= 1,2,…,K are obtained using the updated equations given in section 3.

The EM algorithm contributes to the segmentation algorithm by improving the parameters of the model.

Step 4) The image segmentation is carried out by assigning each pixel into a proper region(segment) according to the Maximum likelihood Estimate of the ith component (L_i).
i.e., (x_s, y_s) is assigned to the jth region for which L_j is maximum.

$$L_i = \max_{i \in k} \left\{ \frac{\exp\left\{ \frac{-1}{2(1-\rho_k^2)} \left[\left(\frac{x_s - \mu_{1k}}{\sigma_{1k}} \right)^2 - 2\rho_k \left(\frac{x_s - \mu_{1k}}{\sigma_{1k}} \right) \left(\frac{y_s - \mu_{2k}}{\sigma_{2k}} \right) + \left(\frac{y_s - \mu_{2k}}{\sigma_{2k}} \right)^2 \right] \right\}}{2\pi\sigma_{1k}\sigma_{2k}\sqrt{1-\rho_k^2} \int_{b_1}^{a_1} \int_{b_2}^{a_2} f_k(w,\theta)\, dw} \right\}$$

6. Experimental Results

To demonstrate the utility of the image segmentation algorithm developed in this chapter, an experiment is conducted with six images taken from Berkeley image data set (http://www.eecs.berkeley.edu /Research/Projects/CS/Vision/bsds/BSDS300/html) .The images namely, OSTRICH, POT, TOWER, BEARS, DEER and BIRD are considered for image segmentation. The feature vector consisting of hue and saturation values of the whole image is assumed that it follows a mixture of doubly truncated bivariate Gaussian distribution. That is the whole image is a collection of *K*-components and the feature vectors in each component follows a doubly truncated bivariate Gaussian distribution. The number of image regions of each image considered for experimentation is determined by hierarchical clustering algorithm.

The number of image regions for each image obtained through hierarchical clustering for the images under study are given in Table 1.

Table 1: Estimated value of K (Hierarchical Clustering Algorithm)

IMAGE	OSTRICH	POT	TOWER	BEARS	DEER	BIRD
Estimate of K	2	3	4	3	3	2

After assigning these initial values of K to each image data set, the K-Means algorithm is performed. The initial values of the model parameters $\mu_{1i}, \mu_{2i}, \sigma_{1i}^2, \sigma_{2i}^2, \rho_i$ and α_i for i=1,2,…,K for each image region of the images are computed by using the method given in section 3. Using these initial estimates, the refined estimates of the model parameters for each image region are obtained by using EM algorithm given in section 3. The computed values of the initial estimates and the final estimates of the model parameters K, $\mu_{1i}, \mu_{2i}, \sigma_{1i}^2, \sigma_{2i}^2, \rho_i$ and α_i for i=1,2…,K for each image are shown in tables -2.a, 2.b, 2.c, 2.d, 2.e and 2f.

Table-2a
Estimated Values of The Parameters For OSTRICH Image
Number of Image Regions (K=2)

Parameters	Estimation of Initial Parameters by Hierarchical clustering, Regions (i)		Estimation of Final Parameters by EM-Algorithm, Regions (i)	
	1	2	1	2
α_i	0.5	0.5	0.2429	0.7571
μ_{1i}	0.1782	0.1937	0.2070	0.2954
μ_{2i}	0.3218	0.7571	0.1281	0.1088
σ_{1i}^2	0.0017	0.0004	0.4769	0.4120
σ_{2i}^2	0.0110	0.0213	0.5630	0.0317
ρ_i	-0.4663	0.7077	0.0820	0.0310

Minimum Hue Value (b_1) = 0.0662 Maximum Hue Value (a_1) = 0.3611
Minimum Saturation Value (b_2) =0.0930 Maximum Saturation Value (a_2) = 1.0000

Table-2b
Estimated Values of The Parameters For POT Image
Number of Image Regions (K =3)

Parameters	Estimation of Initial Parameters by Hierarchical clustering, Regions (i)			Estimation of Final Parameters by EM-Algorithm, Regions (i)		
	1	2	3	1	2	3
α_i	1/3	1/3	1/3	0.8900	0.0200	0.0900
μ_{1i}	0.5450	0.1434	0.4028	0.3887	0.0691	0.1533
μ_{2i}	0.1987	0.1262	0.0574	0.2670	0.2911	0.3004
σ_{1i}^2	0.0008	0.0018	0.0066	0.0567	0.0456	0.2617
σ_{2i}^2	0.0020	0.0033	0.0006	0.1326	0.5541	0.1341
ρ_i	0.4066	-0.7347	0.2008	0.1694	-0.4310	-0.2805

Minimum Hue Value (b_1) = 0 Maximum Hue Value (a_1) = 0.6154
Minimum Saturation Value (b_2) = 0 Maximum Saturation Value (a_2) = 0.3456

Table 2c
Estimated Values of The Parameters For TOWER Image
Number of Image Regions (K =4)

Parameters	Estimation of Initial Parameters by Hierarchical clustering, Regions (i)				Estimation of Final Parameters by EM-Algorithm, Regions (i)			
	1	2	3	4	1	2	3	4
α_i	1/4	1/4	1/4	1/4	0.1605	0.0601	0.0577	0.7218
μ_{1i}	0.1596	0.1496	0.5741	0.5671	0.3473	0.3912	0.6910	0.8760
μ_{2i}	0.8680	0.3166	0.7724	0.2813	0.1517	0.7157	0.9689	0.5634
σ_{1i}^2	0.0033	0.0021	0.0006	0.0077	0.0436	0.1214	0.0312	0.2539
σ_{2i}^2	0.0130	0.0380	0.0054	0.0161	0.1064	0.1497	0.1683	0.1328
ρ_i	0.2443	0.0878	0.0695	0.0208	0.6539	0.4380	-0.9496	0.1391

Minimum Hue Value (b_1)= 0 Maximum Hue Value (a_1)= 0.9815
Minimum Saturation Value (b_2)= 0 Maximum Saturation Value (a_2)= 1.0000

Table-2d
Estimated Values of The Parameters For BEARS Image
Number of Image Regions (K =3)

Parameters	Estimation of Initial Parameters by Hierarchical clustering, Regions (i)			Estimation of Final Parameters by EM-Algorithm, Regions (i)		
	1	2	3	1	2	3
α_i	1/3	1/3	1/3	0.0963	0.2699	0.6338
μ_{1i}	0.8570	0.2265	0.3414	0.1146	0.7238	0.3333
μ_{2i}	0.1267	0.8706	0.3431	0.3937	0.9347	0.8023
σ_{1i}^2	0.0075	0.0286	0.0239	0.0947	0.1072	2.3094
σ_{2i}^2	0.0041	0.0092	0.0238	0.1207	0.1117	0.1922
ρ_i	0.5473	0.0417	0.2715	0.2741	-0.2107	-0.3248

Minimum Hue Value (b_1) = 0 Maximum Hue Value (a_1) = 0.9815
Minimum Saturation Value (b_2) = 0 Maximum Saturation Value (a_2) = 1.0000

Table-2e						
Estimated Values of The Parameters For DEER Image Number of Image Regions (K =4)						
Parameters	Estimation of Initial Parameters by Hierarchical clustering			Estimation of Final Parameters by EM-Algorithm		
	Regions(i)			Regions(i)		
	1	2	3	1	2	3
α_i	1/3	1/3	1/3	0.0035	0.9536	0.0429
μ_{1i}	0.1159	0.2657	0.1275	0.1715	0.5025	0.3551
μ_{2i}	0.4536	0.1872	0.7304	0.3694	0.2064	0.8299
σ_{1i}^2	0.0001	0.0025	0.0004	0.0330	0.1524	0.0041
σ_{2i}^2	0.0033	0.0042	0.0095	1.1451	0.6753	2.9450
ρ_i	-0.1138	-0.4516	-0.0293	0.4580	0.0681	-0.6859
Minimum Hue Value (b_1) = 0 Minimum Saturation Value (b_2) = 0				Maximum Hue Value (a_1) = 0.5152 Maximum Saturation Value (a_2) = 1.0000		

Table-2f				
Estimated Values of The Parameters For BIRD Image Number of Image Regions (K =2)				
Parameters	Estimation of Initial Parameters by Hierarchical clustering		Estimation of Final Parameters by EM-Algorithm	
	Regions(i)		Regions(i)	
	1	2	1	2
α_i	1/2	1/2	0.0571	0.9429
μ_{1i}	0.1351	0.4639	0.5676	0.7221
μ_{2i}	0.9920	0.1767	0.0096	0.7937
σ_{1i}^2	0.0088	0.0460	0.1499	0.1018
σ_{2i}^2	0.0005	0.0250	0.0168	0.6469
ρ_i	0.1100	-0.6184	-0.0949	-0.0561
Minimum Hue Value (b_1) = 0 Minimum Saturation Value (b_2) = 0			Maximum Hue Value (a_1) = 0.9583 Maximum Saturation Value (a_2) = 1.0000	

Substituting the final estimates of the model parameters, the probability density function of the feature vector of each image are estimated. Using the estimated probability density functions and the image segmentation algorithm given in section 5, the image segmentation is done for each of the six images under consideration. The original and segmented images are shown in Figure 2.

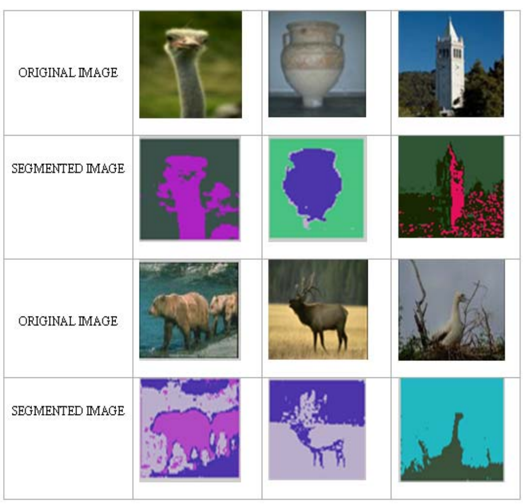

Figure 2: Original and Segmented Images

7. Performance Evaluation

The developed algorithm has been tested on a variety of colour images using Berkeley image data set. The performance of the developed algorithm is compared with the segmentation method based on FGMM with K-means given by Prasad Reddy P.V.G.D. et.al (2007). The comparison is based on three performance measures namely, Probabilistic Rand Index (PRI) given by R.Unnikrishnan and et.al (2007), the Variation of Information (VOI) given by Meila M. (2005), and Global Consistency measure (GCE) given by Martin D. and et al (2001). The objective of the segmentation methods are based on regional similarity measures in relations to their local neighborhood.

The performance of developed algorithm using finite doubly truncated bivariate Gaussian mixture model with hierarchical clustering(FDTBGMM-H) is studied by computing the segmentation performance measures namely, PRI, GCE and VOI for the six images under study. The computed values of the performance measures for the developed algorithm and the earlier existing Finite Gaussian Mixture Model (FGMM) with K-Means algorithm are presented in Table 3 for a comparative study.

Table 3: Segmentation performance measures

IMAGES	METHOD	PERFORMANCE MEASURES		
		PRI	GCE	VOI
OSTRICH	GMM	0.9234	0.4317	2.2761
	FDTBGMM-H	0.9812	0.3531	0.8829
POT	GMM	0.9456	0.5281	2.5973
	FDTBGMM-H	0.9810	0.3855	1.6067
TOWER	GMM	0.9615	0.4469	3.7121
	FDTBGMM-H	0.9829	0.3614	1.6189
BEARS	GMM	0.9121	0.4418	3.2693
	FDTBGMM-H	0.9835	0.4315	2.5895
DEER	GMM	0.9774	0.4829	2.2863
	FDTBGMM-H	0.9861	0.3318	0.4119
BIRD	GMM	0.9673	0.4671	2.7197
	FDTBGMM-H	0.9705	0.3572	1.0251

From the above table 3, It is observed that the developed algorithm is having high PRI and low GCE and VOI compared to finte Gausiisn mixture model (GMM). Therefore the segmentation method proposed outperforms the existing algorithms.

The developed image segmentation method can also used to reconstruct the image. The Performance Evaluation of the retrieved image is done by Subjective Image Quality testing or by Objective Image Quality testing. The Objective Image Quality testing methods are often used since the numerical results of an objective measure are readily computed and allow a consistence comparison of different algorithms. There are several Image Quality measures available for Performance Evaluation of the Image Segmentation method. An extensive survey of Quality Measures is given by Eskicioglu A.M. and Fisher P.S. (1995). For the Performance Evaluation of the developed Segmentation algorithm, we consider the Image Quality Measures namely (a) Maximum Distance, (b) Image Fidelity, (c) Mean Square Error, (d) Signal to Noise Ratio and (e) Image Quality Index. Using the estimated probability density functions of the images under consideration, the retrieved images are obtained and shown in Figure 3.

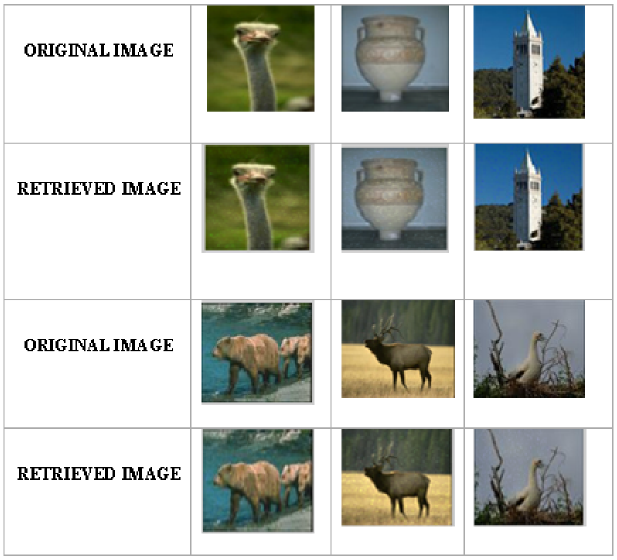

Figure 3: Original and Retrieved Images

There are several Image Quality measures available for Performance Evaluation of the Image Segmentation method. An extensive survey of Quality Metrics is given by Eskicioglu A.M. and Fisher P.S. (1995). Using the formulas of Image Quality Metrics namely, Maximum Distance, Image Fidelity, Mean Square Error, Signal to Noise Ratio and Image Quality index are computed for all the Six images with respect to the developed method and earlier methods and presented in Table- 4.

Table 4: Comparative study of Image Quality Metrics

IMAGE	Quality Metrics	GMM	FDTBGMM with-H	Standard Limits	Optimal Criteria
OSTRICH	Maximum Distance	0.5013	0.4209	-1 to +1	Close to 1
	Image Fidelity	0.7910	0.9139	0 to +1	Close to 1
	Mean Square Error	0.0932	0.0320	0 to ∞	Close to 0
	Signal to Noise Ratio	13.3781	15.2211	0 to ∞	As big as possible
	Image Quality Index	0.8102	0.8928	-1 to +1	Close to 1
POT	Maximum Distance	0.3290	0.3850	-1 to +1	Close to 1
	Image Fidelity	0.6729	0.6786	0 to +1	Close to 1
	Mean Square Error	0.0738	0.0459	0 to ∞	Close to 0
	Signal to Noise Ratio	11.7401	13.0036	0 to ∞	As big as possible
	Image Quality Index	0.6075	0.6179	-1 to +1	Close to 1
TOWER	Maximum Distance	0.8481	0.9832	-1 to +1	Close to 1
	Image Fidelity	0.5217	0.8173	0 to +1	Close to 1
	Mean Square Error	0.2101	0.0580	0 to ∞	Close to 0
	Signal to Noise Ratio	8.8724	11.5633	0 to ∞	As big as possible
	Image Quality Index	0.6271	0.8303	-1 to +1	Close to 1
BEARS	Maximum Distance	0.5387	0.6961	-1 to +1	Close to 1
	Image Fidelity	0.4277	0.6269	0 to +1	Close to 1
	Mean Square Error	0.0872	0.0430	0 to ∞	Close to 0
	Signal to Noise Ratio	12.9217	10.5372	0 to ∞	As big as possible
	Image Quality Index	0.5951	0.6161	-1 to +1	Close to 1
DEER	Maximum Distance	0.6217	0.6135	-1 to +1	Close to 1
	Image Fidelity	0.3982	0.8724	0 to +1	Close to 1
	Mean Square Error	0.0828	0.0154	0 to ∞	Close to 0
	Signal to Noise Ratio	10.0629	15.3428	0 to ∞	As big as possible
	Image Quality Index	0.3763	0.8366	-1 to +1	Close to 1
BIRD	Maximum Distance	0.8429	0.9500	-1 to +1	Close to 1
	Image Fidelity	0.1920	0.7668	0 to +1	Close to 1
	Mean Square Error	0.2013	0.0289	0 to ∞	Close to 0
	Signal to Noise Ratio	8.9231	11.8708	0 to ∞	As big as possible
	Image Quality Index	0.3481	0.7985	-1 to +1	Close to 1

From the Table 4, it is observed that all the image quality metrics for the six images are meeting the standard criteria. This implies that using the proposed algorithm the images are retrieved accurately. A comparative study of proposed algorithm with that of algorithm based on Finite Gaussian Mixture Model (FGMM) and Finite doubly truncated bivariate Gaussian mixture model (FDTBGMM) with K-means reveals that the Mean Square Error of the proposed model is less than that of the FGMM and FDTBGMM. Based on all other quality metrics also it is observed that the performance of the proposed model in retrieving the images is better than the finite Gaussian mixture model with K-means .

8. Conclusion

In this paper a new image segmentation algorithm is developed and analyzed based on Finite doubly truncated bivariate Gaussian mixture distribution. Here two important characteristics of the colour image namely, Hue and Saturation are considered as feature vector. Using EM algorithm the parameters are estimated, the hierarchical algorithm is used to obtain the initial estimates. The segmentation algorithm is developed with component maximum likelihood. The experimentation with Berkeley colour images reveals that this algorithm outperforms the existing algorithms in both image segmentation and image retrievals. The image quality metrics also supported the utility of the proposed algorithm. It is possible to develop image segmentation algorithm with finite mixture of doubly truncated multivariate Gaussian distribution with more image features which require further investigations.

9. References

[1] Adolfo Martine et al, (2006) "Unsupervised Image Segmentation Using Hierarchical clustering Selection Process", IEEE Transactions on Image Processing, Vol.3.

[2] Adams.R. and Bischof L. (1994)" Seeded region growing," IEEE Trans. On PAMI, Vol.8, no.6, pp.679-698.

[3] Bengt Muthen (1990) "Moments of the censored and truncated bivariate normal distribution", British Journal of Mathematical and Statistical psychology, No.43, pp.131-143.

[4] Canny.J.F. (1986), " A computational approach to edge detection," IEEE Trans.on PAMI, vol.8,no.6,pp.679-698.

[5] Cheng H. , X.Jiang, Y.Sun, and J.Wang (2001) " Color image segmentation: advances and prospects", Pattern Recognition , Vol.34, pp2259-2281.

[6] Eskicioglu M.A. and Fisher P.S (1995) "Image Quality Measures and their Performance", IEEE Transactions On Communications, Vol.43, No.12.

[7] Johnson S.C. (1967), "A Tutorial on Clustering Algorithms",http://home.dei.polimi.it/matteucc /Clustering /tutorial_html/hierarchical.html.

[8] Lei T. and Udupa J. (2003) "Performance evaluation of finite normal mixture model-based image segmentation techniques," *IEEE Transactions on Image Processing*, vol. 12, no.10, pp. 1153–1169.

[9] Martin D. , Fowlkes C., D. Tal, and J.Malik (2001) " A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring

ecological statistics," in proc. 8th Int. Conf. Computer vision, vol.2, pp.416-423.

[10] Molina I. et al, (2000.) " Hierarchical Image Segmentation Based on Nearest Neighbor Region Chains", IEEE Electronic Letters, Vol 36, issue 12.

[11] Montanvert A. et al (1991.), " Hierarchical image analysis using irregular tessellations", Transactions on Pattern Analysis and Machine Learning, Vol 13, No.4.

[12] Mclanchlan G. And Krishnan T(1997)., " The EM Algorithm and Extensions", John Wiley and Sons, New York -1997.

[13] Meila M. (2005) " Comparing Clustering – An axiomatic view," in proc.22nd Int. Conf. Machine Learning, pp. 577-584.

[14] Nacken P.F.M. (1995), "Image segmentation by connectivity preserving relinking in hierarchical graph structures", Pattern Recognition, 907-920.

[15] Norman L.Johnson, Kortz and Balakrishnan(1994), "Continuous Univariate Distributions"Volume-I, John Wiley and Sons Publications,Newyork.

[16] Pal S.K. and Pal N.R. (1993), " A Review On Image Segmentation Techniques", Pattern Recognition, Vol.26, N0.9, pp 1277-1294.

[17] Prasad Reddy P.V.G.D, Srinivas Rao.K and Srinivas.Y(2007), " Unsupervised image segmentation Method based On Finite Generalized Gaussian Distribution with EM and k-menas algorithm", International Journal of Computer Science and Network Security, VOL.7. No.4, pp.317-32.

[18]Richard A. Johnson and Dean W. Wichern (2009), "Applied multivariate Statistical Analysis- 5th edition", Eastern economy edition, PHI.

[19] Roberts.S.J.(1997), " Parametric and non-parametric unsupervised cluster analysis," Pattern Recognition, Vol.30, no.2, pp.261-272.

[20] Srinivas. Y and Srinivas. K (2007), "Unsupervised image segmentation using finite doubly truncated gaussian mixture model and Hierarchial clustering ",Journal of Current Science,Vol.93,No.4, pp.507-514.

[21] Sangwine S.J. and R.E.N.Horne (1998), "The Colour Image Processing Hand Book," Chapmann and Hall, UK.

[22] Sujaritha M. and Annadurai S. (2010), Color image segmentation using Adaptive Spatial Gaussian Mixture Model", International journal of signal processing 6:1, pp. 28-32.
[23] "The Berkeley segmentation dataset" http://www.eecs.berkeley.edu/ Research/ Projects/CS/ vision/ bsds/BSDS300/html/dataset/images.html.

[24] Unnikrishnan R., C,Pantofaru, and M.Hernbert,(2007), " Toward objective evaluation of image segmentation algorithms," IEEE Trans.Pattern Anal.Mach.Intell. Vol.29, no.6, pp.929-944.

[25] Wei Yu. , Jason Fritts et al, (2002) "An Hierarchical Image Segmentation Algorithm', IEEE International Conference on Multimedia and Expo, pp.105.

[26] Yiming WU. et al(2003), "Unsupervised Color Image Segmentation based on Gaussian Mixture Models" Proceedings of 2003 Joint Conference at the 4th International Conference on Information, Communication and Signal Processing, vol1, pp. 541-544.

[27] Zoltan Kato, Ting-Chuen Pong (2006), "A Markov random field image segmentation model for color textured images", Image and Computing Vision, 24(10), pp 1103-1114.

Author's profile:

Mr. GVS RAJKUMAR is presently working as Associate professor in the Department of Information Technology, GITAM University, Visakhapatnam. He presented several research papers in national and International conferences and seminars. He published a good number of papers in national and International journals. He guided several students for getting their M.Tech degrees in Information Technology. His current research interests are Image Processing, Internet Technologies and Network security.

Dr. K.Srinivasa Rao is presently working as Professor and Head, Department of Statistics, Andhra University, Visakhapatnam. He is elected chief editor of Journal of ISPS and elected Vice-President of Operation Research of India. He guided 28 students for Ph.D in Statistics, Mathematics, Computer Science, Electronics and Communications Engineering, Industrial Engineering and Operations Research. He published 90 research papers in national and International journals with high reputation. His research interests are Image Processing, Communication Systems, Data Mining and stochastic models.

Dr. Peri. Srinivasa Rao is presently working as Professor in the Department of Computer Science and Systems Engineering, Andhra University, Visakhapatnam. He got his Ph.D degree from Indian Institute of Technology, Kharagpur in Computer Science in 1987. He published several research papers and delivered invited lectures at various conferences, seminars and workshops. He guided a number of students for their Ph.D and M.Tech degrees in Computer Science and Engineering and Information Technology. His current research interests are Image Processing, Communication networks, Data Mining and Computer Morphology.