Image Segmentation Based on a Finite Generalized New Symmetric Mixture Model with K – Means

[1]M.Seshashayee [2]K. Srinivas Rao [3]Ch. Satyanarayana [4]P. Srinivasa Rao

[1]Department of CS, GITAM University, Visakhapatnam-5300045, India

[2]Department of Statistics, Andhra University, VIsakhapatnam-530003, India

[3]Department of CSE, JNTUK, Kakinada, A.P., India

[4]Department of CS&SE, College of Engineering, Andhra Univerity, Visakhapatnam-530003, India

Abstract

In this paper a novel image segmentation and retrieval method based on finite new symmetric mixture model with K-means clustering is developed. Here it is considered that pixel intensities in each image region follow a new symmetric distribution. The new symmetric distribution includes platy-kurtic and meso-kurtic distributions. This also includes Gaussian mixture model as a particular case. The number of components (image regions) is obtained through K-means algorithm. The model parameters are estimated by deriving the updated equations of the EM algorithm. The segmentation of the image is done by maximizing the component likelihood. The performance of the proposed algorithm is studied by computing the segmentation performance metrics like, PRI, VOI, and GCE. The ability of this method for image retrieval is demonstrated by computing the image quality metrics for five images namely HORSE, MAN, BIRD, BOAT and TOWER. The experimental results show that this method outperforms the existing model based image segmentation methods.

Keywords: Image segmentation, EM algorithm, New Symmetric Distribution. Image Quality Metrics

1. Introduction

Segmentation is the main consideration for image analysis and image retrieval. With segmentation it is possible to identify the regions of interest and objects which are highly useful. Image segmentation is defined as the process of dividing the image into different image regions such that each region is homogeneous. Image segmentation can be classified into two categories namely, parametric and non-parametric image segmentation. A more comprehensive discussion on image segmentation is given by (S.K.Pal and N.R.Pal (1993), Jahne (1995), and Cheng et al (2001)). There does not exist a single algorithm that works for all applications.

Model based image segmentation is more efficient compared to the non-parametric methods of segmentation. Recently, much emphasis is given for image analysis through Finite Gaussian Mixture Model (Yamazaki et al. (1998), T.Lie et al.(1993), N.Nasios et al.(2006), Z.H.Zhang et al.(2003)). In Finite Gaussian Mixture Model each image region is characterized by a Gaussian distribution and the entire image is considered to be a mixture of these Gaussian components. Here it is assumed that the whole image is characterized by Gaussian mixture model in which the pixel intensities of each image region follow a Gaussian distribution. For gray level images the pixel intensity is the most suitable feature for segmenting the image (S.K.Pal and N.R.Pal, (1993)).

However, in finite Gaussian mixture model the pixel intensities of the image region are considered to be meso-kurtic and symmetric. But in some images the pixel intensities of the image region may not be distributed as meso – kurtic even though they are symmetric. To have a more close approximation to the pixel intensities of each image region it is needed to consider that the pixel intensities of each region follow a more general symmetric distribution. Srinivasa Rao, et al., (1997) have introduced a new symmetric distribution which is capable of portraying several platy – kurtic distributions. It also includes Gaussian as a particular case for a specific value of the index parameter. Hence, in this chapter an image segmentation algorithm is developed and analyzed with the assumption that the whole image is characterized by a finite mixture of new symmetric distribution in which the pixel intensities of each image region follows a new symmetric distribution.

In mixture models one of the important factors is the number of components K (regions). Usually the number of components are assumed to be known as apriori . This will generally effect the segmentation results. If this number deviates from true value of K then the misclassification of pixels in the image is very high. To have a more accurate analysis of the number of regions in the whole image, the K value is identified through the K – Means algorithm (Rose H.Turi, (2001)) along with the histogram of the pixel intensities.

Using the Expectation Maximization (EM) algorithm the model parameters are estimated. The segmentation algorithm is developed through maximizing the component likelihood. The performance of the segmentation algorithm is evaluated by obtaining performance measures like PRI, GCE and VOI by applying them on five images HORSE, MAN, BIRD, BOAT and TOWER. The performance of this algorithm is compared with the image segmentation algorithm based on Finite Gaussian Mixture Model with K-Means. The efficiency of it in image retrievals is also studied through obtaining the image quality metrics like, average difference, maximum distance, image fidelity, mean square error, signal to noise ratio and image quality index and comparing it with earlier algorithms.

2. Finite Mixture Of New Symmetric Distribution

In low level image analysis the entire image is considered as a union of several image regions. In each image region the image data is quantized by pixel intensities. The pixel intensity z = f(x , y) for a given point (pixel) (x , y) is a random variable, because of the fact that the brightness measured at a point in the image is influenced by various random factors like vision, lighting, moisture, environmental conditions etc,. To model the pixel intensities the image region it is assumed that the pixel intensities of the region follows a new symmetric distribution given by Srinivasa Rao et al., (1997).

The probability density function of the pixel intensity is

$$f(Z,\mu,\sigma^2,r)=\frac{\left(2r+\left(\frac{z-\mu}{\sigma}\right)^2\right)^r e^{\frac{-1}{2}\left(\frac{z-\mu}{\sigma}\right)^2}}{\sigma(2r)^r(2\Pi)^{\frac{1}{2}}+\sum_{j=1}^{r}\binom{r}{j}(2r)^{r-j}2^{j+\frac{1}{2}}\Gamma(j+\frac{1}{2})\sigma},$$

$$-\infty<Z<\infty,-\infty<\mu<\infty,\sigma>0 \quad (1)$$

For different values of the parameters the various shapes of probability curves associated with new symmetric distribution are shown in Figure 1.

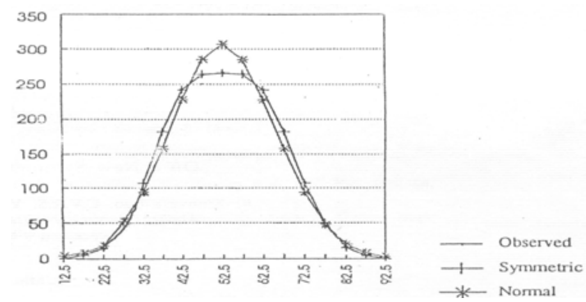


Figure1. Frequency curves of new symmetric distribution

Each value of the shape parameter 's' (= 0,1,2,3,…,) gives a bell shaped distribution. For r = 0 the equation reduces to a normal probability density function with parameter μ and σ .

Its central moments are

$$\mu_{2n}=\left[\frac{\Gamma(n+\frac{1}{2})+\sum_{j=1}^{r}\binom{r}{j}r^{-j}\Gamma(n+j+\frac{1}{2})}{(\pi)^{\frac{1}{2}}+\sum_{j=1}^{r}\binom{r}{j}r^{-j}\Gamma(j+\frac{1}{2})}\right]2^n\sigma^{2n} \quad (2)$$

And $\mu_{2n+1}=0$

The kurtosis of the distribution is

$$\beta_2=\frac{\left[\frac{3}{4}\pi^{\frac{1}{2}}+\sum_{j=1}^{r}\binom{r}{j}r^{-j}(j+\frac{1}{2})(j+\frac{3}{2})\Gamma(j+\frac{1}{2})\right]\left[\pi^{\frac{1}{2}}+\sum_{j=1}^{r}\binom{r}{j}r^{-j}\Gamma(j+\frac{1}{2})\right]}{\left[\frac{\pi^{\frac{1}{2}}}{2}+\sum_{j=1}^{r}\binom{r}{j}r^{-j}\Gamma(j+\frac{1}{2})\right]^2} \quad (3)$$

The entire image is a collection of regions which are characterized by new symmetric distribution. Here, it is assumed that the pixel intensities of the whole image follows a K – component mixture of new symmetric distribution and its probability density function is of the form

$$p(z)=\sum_{i=1}^{K}\alpha_i f_i(z/\mu_i,\sigma_i^2,r_i) \qquad (4)$$

where, K is number of regions , $0\le\alpha_i\le 1$ are weights such that $\sum\alpha_i=1$ and $f_i(z,\mu,\sigma^2,r)$ is as given in equation (1). α_i is the weight associated with i[th] region in the whole image.

In general the pixel intensities in the image regions are statistically correlated and these correlations can be reduced by spatial sampling (Lie. T and Sewehand. W (1992)) or spatial averaging (Kelly P.A. et al.,(1998)) . After reduction of correlation the pixels are considered to be uncorrelated and independent. The mean pixel intensity of the whole image is $E(Z)=\sum_{i=1}^{K}\alpha_i\mu_i$.

3. Estimation of the Model Parameter by EM Algorithm

In this section we derive the updated equations of the model parameters using Expectation Maximization (EM) algorithm. The likelihood function of the observations z_1, z_2, z_3,…,z_N drawn from an image is

$$L(\theta)=\sum_{s=1}^{N}\left(\sum_{i=1}^{K}\alpha_i f_i(z_s,\theta_i)\right),$$

$$L(\theta)=\sum_{s=1}^{N}\left[\sum_{i=1}^{K}\frac{\alpha_i\left(2r_i+\left(\frac{z_s-\mu_i}{\sigma_i}\right)^2\right)^{r_i}e^{\frac{-1}{2}\left(\frac{z_s-\mu_i}{\sigma_i}\right)^2}}{\sigma_i(2r_i)^{r_i}(2\Pi)^{\frac{1}{2}}+\sum_{j=1}^{r_i}\binom{r_i}{j}(2r_i)^{r_i-j}2^{j+\frac{1}{2}}\Gamma(j+\frac{1}{2})\sigma_i}\right] \qquad (5)$$

where, $\theta=(\mu_i,\sigma_i^2,r_i,\alpha_i;i=1,2,...,K)$ is the set of parameters.

The expectation of the log likelihood function of the sample is

$$Q\left(\theta;\theta^{(l)}\right)=E_{\theta^{(l)}}\left[\log L(\theta)/\overline{z}\right]$$

This implies

$$Q\left(\theta;\theta^{(l)}\right)=\sum_{i=1}^{K}\sum_{s=1}^{N}\left(t_i(z_s,\theta^{(l)})\left(\log f_i(z_s,\theta)+\log\alpha_i\right)\right) \qquad (6)$$

The updated equation of α_i at $(l+1)^{\text{th}}$ iteration is

$$\alpha_i^{(l+1)}=\frac{1}{N}\sum_{s=1}^{N}t_i(z_s,\theta^{(l)})$$

$$=\frac{1}{N}\sum_{s=1}^{N}\left[\frac{\alpha_i^{(l)}f_i(z_s,\theta^{(l)})}{\sum_{i=1}^{K}\alpha_i^{(l)}f_i(z_s,\theta^{(l)})}\right] \qquad (7)$$

The updated equation of μ_i at $(l+1)^{\text{th}}$ iteration is

$$\mu_i^{(l+1)}=\frac{\sum_{s=1}^{N}z_s t_i(z_s,\theta^{(l)})-\sum_{s=1}^{N}t_i(z_s,\theta^{(l)})\left(\frac{2r_i\sigma_i^{2(l)}\left(z_s-\mu_i^{l}\right)}{\left(2r_i\sigma_i^{2(l)}+\left(z_s-\mu_i^{l}\right)^2\right)}\right)}{\sum_{s=1}^{N}t_i(z_s,\theta^{(l)})} \qquad (8)$$

where,

$$t_i(z_s,\theta^{(l)})=\frac{\alpha_i^{(l+1)}f_i(z_s,\mu_i^{(l)},\left(\sigma_i^2\right)^{l},r_i)}{\sum_{i=1}^{K}\alpha_i^{(l+1)}f_i(z_s,\mu_i^{(l)},\left(\sigma_i^2\right)^{l},r_i)}$$

The updated equation of σ_i^2 at $(l+1)^{\text{th}}$ iteration is

$$\sigma_i^{2(l+1)}=\frac{2\sum_{s=1}^{N}(z_s-\mu_i^{(l+1)})^2\left(\frac{1}{2}-\frac{r_i\left(\sigma_i^2\right)^{(l)}}{\left(2r_i\sigma_i^{2(l)}+\left(z_s-\mu_i^{(l+1)}\right)^2\right)^2}\right)\left(t_i(z_s,\theta^{(l)})\right)}{\sum_{s=1}^{N}t_i(z_s,\theta^{(l)})} \qquad (9)$$

where, $t_i(z_s,\theta^{(l)})=\dfrac{\alpha_i^{(l+1)}f_i(z_s,\mu_i^{(l+1)},\left(\sigma_i^2\right)^{(l)},r_i)}{\sum_{i=1}^{K}\alpha_i^{(l+1)}f_i(z_s,\mu_i^{(l+1)},\left(\sigma_i^2\right)^{(l)},r_i)}$

4. Initialization of the Parameters K – Means

The efficiency of the EM algorithm in estimating the parameters is heavily dependent on the number of regions in the image. The number of mixture components initially taken for K – Means algorithm is by plotting the histogram of the pixel intensities of the whole image. The number of peaks in the histogram can be taken as the initial value of the number of regions K.

The mixing parameters α_i and the model parameters μ_i, σ_i^2, r_i are usually considered as known apriori. A commonly used method in initializing parameters is by drawing a random sample from the entire image (Mclanchan G and Peel D (2000)). This method performs well if the sample size is large and its computational time is heavily increased. When the sample size is small, some small regions may not be sampled. To overcome this problem we use a K – Means algorithm to divide the whole image into various homogeneous regions.

After determining the final values of K (number of regions) , we obtain the initial estimates of μ_i, σ_i^2, r_i and α_i for the i^{th} region using the segmented region pixel intensities with the method given by Srinivasa Rao etal.,(1997) for new symmetric distribution .The initial estimate α_i is taken as $\alpha_i = \frac{1}{K}$, where, i=1,2,...,K. The shape parameter r_i can be estimated through sample kurtosis by using the following equation,

$$\left[\frac{3\sqrt{\pi}}{4}+\sum_{j=1}^{r_i}\binom{r_i}{j}r_i^{-j}\left(j+\frac{1}{2}\right)\left(j+\frac{3}{2}\right)\Gamma\left(j+\frac{1}{2}\right)\right]\frac{(\pi)^{1/2}+\sum_{j=1}^{r_i}\binom{r_i}{j}r_i^{-j}\Gamma\left(j+\frac{1}{2}\right)}{\left[\frac{\sqrt{\pi}}{2}+\sum_{j=1}^{r_i}\binom{r_i}{j}_i^{-j}\left(j+\frac{1}{2}\right)\Gamma\left(j+\frac{1}{2}\right)\right]^2}$$

$$=\frac{\left[\frac{1}{n}\sum_{i=1}^{n}(z_i-\bar{z})^4\right]}{\left[\frac{1}{n}\sum_{i=1}^{n}(z_i-\bar{z})^2\right]}$$

By Rockies theorem there is one and only one real root to this equation. We can take the nearest integer to this real root as an estimate to the shape parameter r_i . Knowing the shape parameter r_i we can obtain the estimates of the parameter μ_i and σ_i^2 by method of moments as

$\mu_i = \bar{z}$ and

$$\sigma_i^2 = \frac{\text{n}}{\text{(n-1)}}\left[1+\frac{\sum_{j=1}^{r_i}\binom{r_i}{j}r_i^{-j}\,\Gamma\left(j+\frac{1}{2}\right)}{2\sum_{j=1}^{r_i}\binom{r_i}{j}_i^{-j}r_i^{-j}\left(j+\frac{1}{2}\right)\Gamma\left(j+\frac{1}{2}\right)}\right]S^2$$

where, S^2 is the sample variance.

5. Segmentation Algorithm

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

Step 1) Plot the histogram of the whole image.

Step 2) Obtain the initial estimates of the model parameters using K-Means algorithm and moment estimators as discussed in section 3.

Step 3) Obtain the refined estimates of the model parameters by using the EM algorithm with the updated equations given by (7), (8) and (9) respectively.

Step 4) Assign each pixel into the corresponding j^{th} region (segment) according to Maximum likelihood of the j^{th} component (L_j).

That is

$$L_j = \max_{j\in k}\left\{\frac{\left(2r+\left(\frac{z_s-\mu_j}{\sigma_j}\right)^2\right)^r e^{\frac{-1}{2}\left(\frac{z_s-\mu_j}{\sigma_j}\right)^2}}{\sigma_j(2r)^r(2\Pi)^{\frac{1}{2}}+\sum_{j=1}^{r}\binom{r}{j}(2r)^{r-j}2^{j+\frac{1}{2}}\Gamma(j+\frac{1}{2})\sigma_j}\right\},$$

$$-\infty < z_s < \infty,\ -\infty < \mu_j < \infty, \sigma_j > 0$$

6. Experimental Results

To demonstrate the utility of the image segmentation algorithm developed in this chapter, an experiment is conducted with five images taken from Berkeley images dataset (http://www.eecs.berkeley.edu/Research/Projects/CS/Vision/bsds/BSDS300/html). The images HORSE, MAN , BOAT and TOWER are considered for image segmentation. The pixel intensities of the whole image are taken as feature. The pixel intensities of the image are assumed to follow a mixture of new symmetric distribution. That is, the image contains K regions and pixel intensities in each image region follows a new symmetric distribution with different parameters. The number of segments in each of the five images considered for experimentation is determined by the histogram of pixel intensities. The histograms of the pixel intensities of the five images are shown in figure 2.

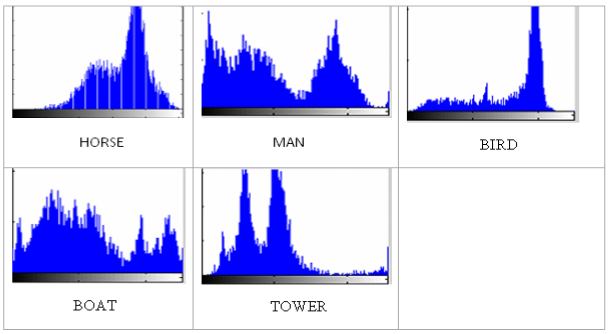


Figure 2: HISTOGRAMS OF THE IMAGES

The initial estimates of the number of the regions K in each image are obtained and given in Table1.

Table 1. INITIAL ESTIMATES OF K

IMAGE	HORSE	MAN	BIRD	BOAT	TOWER
Estimate of K	2	4	3	4	3

From Table 1, we observe that the image HORSE has two segments, images TOWER and BIRD have three segments each and images MAN and BOAT have four segments each. The initial values of the model parameters μ_i , σ_i^2 , r_i and α_i for i = 1,2,…,K for each image region are computed by the method given in section 3.

Using these initial estimates and the updated equations of the EM Algorithm given in Section 3 the final estimates of the model parameters for each image are obtained and presented in tables 2.a, 2.b, 2.c, 2.d and 2.e for different images.

Table-2.a

ESTIMATED VALUES OF THE PARAMETERS FOR HORSE IMAGE

		Estimation of Initial Parameters		Estimation of Final Parameters by EM Algorithm	
		Number of Image Regions (K =2)		Number of Image Regions (K =2)	
Regions(i)		1	2	1	2
Weights	α_i	0.5	0.5	0.31361	0.68639
Means	μ_i	121.47	187.91	145.01	170.99
Variances	σ_i^2	609.82	426.21	1943.1	1034.3
Estimated r values	r	2	1	2	1

Table-2.b

ESTIMATED VALUES OF THE PARAMETERS FOR MAN IMAGE

		Estimation of Initial Parameters				Estimation of Final Parameters by EM Algorithm			
		Number of Image Regions (K =4)				Number of Image Regions (K =4)			
Regions(i)		1	2	3	4	1	2	3	4
Weights	α_i	1/4	1/4	1/4	1/4	0.18275	0.5234	0.23545	0.081981
Means	μ_i	36.503	75.342	126.54	203.31	64.46	31.213	183.49	113.5
Variances	σ_i^2	298.16	161.79	393.55	897.3	1549.3	835.35	529.85	3104
Estimated r values	r_i	4	4	1	2	4	4	1	2

Table-2.d

ESTIMATED VALUES OF THE PARAMETERS FOR BOAT IMAGE

		Estimation of Initial Parameters				**Estimation of Final Parameters by EM Algorithm**			
		Number of Image Regions (K =4)				Number of Image Regions (K =4)			
Regions(i)		1	2	3	4	1	2	3	4
Weights	α_i	1/4	1/4	1/4	1/4	0.2540	0.2670	0.23038	0.2485
Means	μ_i	34.98	81.14	131.13	216.5	37.644	80.467	130.62	213.83
Variances	σ_i^2	342.0	249.4	372.13	618.0	632.41	565.83	684.41	775.17
Estimated r values	r_i	5	2	2	3	5	2	2	3

Table-2.c

ESTIMATED VALUES OF THE PARAMETERS FOR BIRD IMAGE

		Estimation of Initial Parameters			Estimation of Final Parameters by EM Algorithm		
		Number of Image Regions (K =3)			Number of Image Regions (K =3)		
Regions (i)		1	2	3	1	2	3
Weights	α_i	1/3	1/3	1/3	0.12722	0.21391	0.65886
Means	μ_i	53.491	124.05	193.19	65.184	127.31	192.79
Variances	σ_i^2	503.23	43.05	148.42	1556.5	2447.1	86.651
Estimated r values	r_i	3	3	1	3	3	1

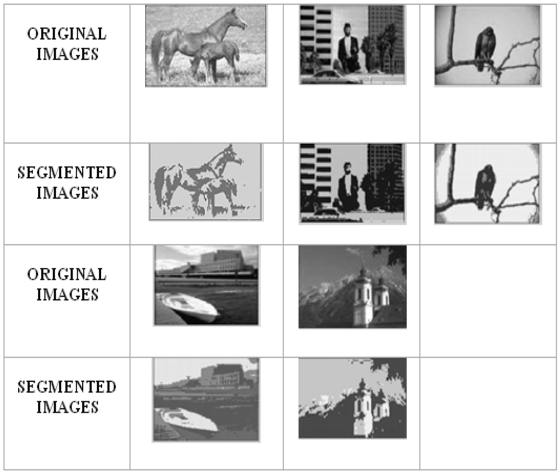
Table-2.e

ESTIMATED VALUES OF THE PARAMETERS FOR TOWER IMAGE

		Estimation of Initial Parameters			Estimation of Final Parameters by EM Algorithm		
		Number of Image Regions (K =3)			Number of Image Regions (K =3)		
Regions (i)		1	2	3	1	2	3
Weights	α_i	1/3	1/3	1/3	0.71336	0.1941	0.092544
Means	μ_i	64.363	107.54	185.83	71.14	110.03	145.58
Variances	σ_i^2	358.2	393.11	1253.6	602.94	3140	2975.9
Estimated r values	r_i	3	2	3	3	2	3

Substituting the final estimates of the model parameters, the probability density function of pixel intensities of each image is 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 five images under consideration. The original and segmented images are shown in figure 3

Figure 3 Original and segmented Images



7. Performance Evaluation

After conducting the experiment with the image segmentation algorithm developed in this chapter, its performance is studied. The performance evaluation of the segmentation technique is carried by obtaining the four performance measures namely,(i) Probabilistic Rand Index (PRI), (ii) Variation Of Information (VOI) and (iii) Global Consistence Error (GCE). The Rand index given by Unnikrishnan et al (2005) counts the fraction of pairs of pixels whose labeling are consistent between the computed segmentation and the ground truth. This quantitative measure is easily extended to the Probabilistic Rand index (PRI) given by Unnikrishnan and et al (2007). The variation of information (VOI) metric given by Meila (2005) is based on relationship between a point and its cluster. It uses mutual information metric and entropy to approximate the distance between two clustering across the lattice of possible clustering. It measures the amount of information that is lost or gained in changing from one clustering to another. The Global Consistency Error (GCE) given by D.Martin and et al (2001) measures the extent to which one segmentation map can be viewed as a refinement of segmentation. For a perfect match, every region in one of the segmentations must be identical to, or a refinement (i.e., a subset) of, a region in the other segmentation.

The performance of developed algorithm using finite new symmetric distribution mixture model (FNSDMM) is studied by computing the segmentation performance measures namely, PRI, GCE and VOI for the five 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 PERFORMACE MEASURES

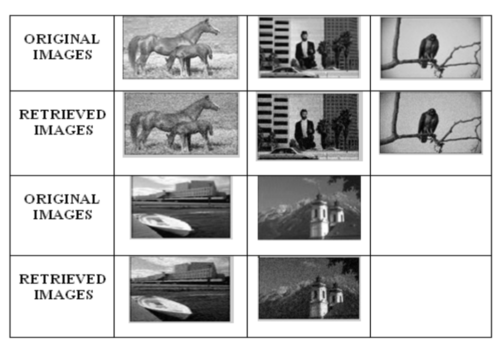
IMAGES	METHOD	PERFORMACE MEASURES		
		PRI	GCE	VOI
HORSE	FGMM	0.9403	0.8827	9.4277
	FNSDMM	0.9321	0.8759	9.3763
MAN	FGMM	0.9779	0.9111	9.1134
	FNSDMM	0.9727	0.9069	9.3946
BIRD	FGMM	0.8734	0.7273	7.5725
	FNSDMM	0.8358	0.6809	7.4998
BOAT	FGMM	0.9765	0.8826	8.8952
	FNSDMM	0.9713	0.8831	9.1621
TOWER	FGMM	0.9010	0.7308	8.0586
	FNSDMM	0.9020	0.7421	8.2153

From table 3 it is observed that the PRI values of the proposed algorithm for the five images considered for experimentation are less than that of the values from the segmentation algorithm based on Finite Gaussian Mixture Model with K-means. Similarly GCE and VOI values of the proposed algorithm are less than that of Finite Gaussian Mixture Model. This reveals that the proposed algorithm outperforms the existing algorithm based on the Finite Gaussian Mixture Model. When the kurtosis parameter of each component of the model is zero, the model reduces to Finite Gaussian Mixture Model and even in this case the algorithm performs well.

After developing the image segmentation method it is needed to verify the utility of segmentation in the model building of the image for image retrieval. The performance evaluation of the retrieved image can be 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 allow a consistent 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 a) Average Difference, b) Maximum Distance, c) Image Fidelity, d) Mean Square Error, e) Signal to Noise Ratio and f) Image Quality Index.

Using the estimated probability density functions of the images under consideration the retrieved images are obtained and are shown in figure 4.

Figure 4: The Original and Retrieved Images



The image quality measures are computed for the five retrieved images HORSE, MAN, BIRD, BOAT and TOWER using the proposed model and FGMM with K-means and their values are given in the table 4.

Table 4: Comparative Study of Image Quality Metrics

IMAGE	Quality Metrics	FGMM	FNSDMM with K-Means	Standard Limits
HORSE	Average Difference	0.5011	0.4089	Close to 1
	Maximum Distance	1.0000	1.0000	Close to 1
	Image Fidelity	1.0000	1.0000	Close to 1
	Mean Square Error	0.5011	0.4090	Close to 0
	Signal to Noise Ratio	5.6542	6.0957	As big as possible
	Image Quality Index	1.0000	1.0000	Close to 1
MAN	Average Difference	0.4858	0.4907	Close to 1
	Maximum Distance	1.0000	1.0000	Close to 1
	Image Fidelity	1.0000	1.0000	Close to 1
	Mean Square Error	0.4946	0.4995	Close to 0
	Signal to Noise Ratio	5.6828	5.6615	As big as possible
	Image Quality Index	1.0000	1.0000	Close to 1
BIRD	Average Difference	0.4939	0.5050	Close to 1
	Maximum Distance	1.0000	1.0000	Close to 1
	Image Fidelity	1.0000	1.0000	Close to 1
	Mean Square Error	0.5050	0.4939	Close to 0
	Signal to Noise Ratio	5.6861	5.6376	As big as possible
	Image Quality Index	1	1.0000	Close to 1
BOAT	Average Difference	0.5039	0.5043	Close to 1
	Maximum Distance	1.0000	1.0000	Close to 1
	Image Fidelity	1.0000	1.0000	Close to 1
	Mean Square Error	0.5070	0.5064	Close to 0
	Signal to Noise Ratio	5.6318	5.6291	As big as possible
	Image Quality Index	1	1.0000	Close to 1
TOWER	Average Difference	0.4936	0.5074	Close to 1
	Maximum Distance	1.0000	1.0000	Close to 1
	Image Fidelity	0.9999	0.9999	Close to 1
	Mean Square Error	0.5076	0.4936	Close to 0
	Signal to Noise Ratio	5.6870	5.6264	As big as possible
	Image Quality Index	1.0000	1.0000	Close to 1

From the Table 4, it is observed that all the image quality measures for the five images are meeting the standard criteria. This implies that using the proposed algorithm the images are retrieved accurately. A comparative of study of proposed algorithm with that of algorithm based on Finite Gaussian Mixture Model reveals that the MSE of the proposed model is less than that of the Finite Gaussian Mixture Model. 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.

8. Conclusions

In this paper we propose an unsupervised image segmentation algorithm based on finite new symmetric mixture model with K-means clustering. The finite mixture of new symmetric distribution includes Finite Gaussian Mixture Model as a particular case when the kurtosis parameter equals to zero. This includes several platy-kurtic mixture distributions as particular cases. As a result of this generic nature this algorithm can handle a wide variety of images. An EM algorithm is developed and used for estimating the model parameters. In our experimentation with five images taken from Berkeley image data set, it is observed that the developed algorithm performs better with respect to the image segmentation metrics and the image quality metrics. The hybridization of model based approach with K-means has improved the accuracy of retrieval. This algorithm can be utilized for image analysis and retrieval of grey and colour images more accurately.

References:

[1] Cheng et al (2001) "Color Image Segmentation: Advances and Prospects" Pattern Recognition, Vo1.34, pp 2259-2281.

[2] D. Martin, C. Fowlkes,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 proceedings of 8th International Conference on Computer vision, vol.2, pp.416-423.

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

[4] Jahne (1995), " A Practical Hand Book on Image segmentation for Scientific Applications, CRC Press.

[5] Jeff A.Bilmes (1997), " A Gentle Tutorial of the EM Algorithm and its application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models", Technical Report, University of Berkeley, ICSI-TR-97-021.

[6] Lie. T and Sewehand.W(1992), " Statistical approach to X-ray CT imaging and its applications in image analysis", IEEE Trans. Med. Imag. Vol.11, No.1, Pg 53-61.

[7] Kelly .P.A et al(1988), " Adaptive segmentation of speckled images using a hierarchical random field model". IEEE Transactions Acoust. Speech. Signal Processing, Vol.36, No.10, Pg.1628-1641.

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

[9] Mclanchlan G. and Peel.D(2000), " The EM Algorithm For Parameter Estimations", John Wiley and Sons, New York -2000.

[10] M.Meila, (2005) " Comparing Clustering – An axiomatic view," in proceedings of the 22nd International Conference on Machine Learning, pp. 577-584.

[11] Nasios, N., A.G.Bors,(2006) " Variational learning for Gaussian Mixtures", IEEE Transactions on Systems, Man, and Cybernatics - Part B : Cybernatics, Vol36(4), pp849-862.

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

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

[14] RoseH.Turi.(2001), " Cluster Based Image Segmentation", phd Thesis, Monash University, Australia.

[15] Srinivas Rao. K,C.V.S.R.Vijay Kumar, J.Lakshmi Narayana, (1997) " On a New Symmetrical Distribution ", Journal of Indian Society for Agricultural Statistics, Vol.50(1), pp 95-102.

[16] Srinivas. Y and Srinivas Rao. 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.

[17] "The Berkeley segmentation dataset" http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/BSDS300/html/dataset/images.html.

[18] T. Yamazaki (1998), " Introduction of EM algorithm into color image segmentation," in Proceedings of ICIRS'98, pp. 368–371.

[19] T.Lie et al.(1993), " Performance evaluation of Finite normal mixture model based image segmentation, IEEE Transactions on Image processing, Vol .12(10), pp 1153-1169.

[20] Z.H.Zhang et al (2003)., " EM Algorithms for Gaussian Mixtures with Split-and- merge Operation", Pattern Recognition, Vol. 36(9),pp1973-1983