Image Segmentation Using Compound Normal with Gamma Mixture Model

$$Normal(\mu,\sigma^2)\underset{\sigma^{-2}}{\Lambda}Gamma(c\chi_v^2)$$

Viziananda Row Sanapala[1], Sreenivasa Rao Kraleti[2], and Srinivasa Rao Peri[3]

[1]Associate Professor, Dept of Computer Science & Systems Engg, Andhra University, Visakhapatnam, India

[2]Professor, Dept of Statistics, Andhra University, Visakhapatnam, India

[3]Professor, Dept of Computer Science & Systems Engg, Andhra University, Visakhapatnam, India

Abstract

In this paper, we present the study of image segmentation using compound normal with gamma mixture(CNGM) distribution, derivation of the estimates of model parameters, construction of Expectation-Maximization(EM) algorithm and its performance analysis in comparison to the current model(statistical) and structural(organization and relationship among pixels) based image analysis approaches. In this model, the pixel intensities in each image region are assumed to follow Gaussian distribution in which the rate parameter(σ^{-2}) is random and follows gamma distribution. Hence it has the advantage of dealing with a mixture of mesokurtic and leptokurtic distributions that may be inherent in some images. Thus, it generally fares better than a finite normal mixture(NM) based model since the latter assumes that all components follow normal distribution and normal distribution is mesokurtic in general. Also it includes image segmentation based on Gaussian mixture model as a particular case.

Keywords: $Normal(\mu,\sigma^2)\underset{\sigma^{-2}}{\Lambda}Gamma(c\chi_v^2)$, *EM algorithm, statistical image analysis, image segmentation, pixel classification*

1. Introduction

Image analysis falls under the broad category of pattern analysis of which image segmentation is an important step to subdivide an image into its constituent parts. Pixel intensity value discontinuity and similarity in the local neighborhood are the basis for edge detection and separating parts of the whole image respectively[1],[2]. Region segmentation is considered more useful than edge detection since regions contain more information than edges and thus has been still actively pursued by research community in general[3].

However, image segmentation is an ill posed problem because it may be subjective or objective within the perspective of the end application. For example, homogeneous regions can be easily separated and thus segmentation here is more objective and is easily accomplished as is done in medical imaging. But the more general problem of image segmentation involving natural images is more subjective since for natural images region homogeneity is not well defined due to natural and environmental reasons[3],[4].

Ground truth segmentation conditions for natural images can not be precisely defined and these may vary from one perceptual view to the other. If we consider humans to be the qualitative judges for segmentation, we may consider the segmentation data produced by them as a benchmark against which any proposed segmentation method is compared and its segmentation quality assessed[4],[6],[9].

As stated in [3], image analysis techniques can be classified into two major groups: 1)statistical, which uses probability distribution functions of pixels and regions to characterize the image, and 2) structural, which analyzes the image in terms of organization and relationship of pixels and regions by the specified relations.

The current literature on statistical image segmentation techniques mostly assumes the data describing the image as a mixture of components each of which following normal distribution i.e., $N(\mu,\sigma^2)$ with some weight. And the whole image is thought of as following the weighted distribution where weighted distribution implies weighted average of the constituent components[3],[6],[7],[8]. However, in many natural images, the pixel intensity distributions may not be mesokurtic as we have noticed in several image data distributions collected from BSD image dataset[5]. This may be due to random nature of scale parameter involved in the normal mixture model. If we assume that the scale parameter of the normal model in each image region is random and follows a gamma distribution, the pixel intensities can be well characterized by compounding normal distribution with gamma distribution.

In our work, we propose the usefulness of compounding normal distribution with gamma distribution i.e.,

$Normal(\mu,\sigma^2) \underset{\sigma^{-2}}{\Lambda} Gamma(c\chi_v^2)$which keeps intact the characteristics of normal distribution and at the same time accommodates deviations from normal distribution controlled by scale(c) and shape(v) parameters in addition to location (μ). This is normally referred to as being equivalent to Pearson's type VII distributions. Thus, the proposed model effectively deals with mesokurtic and leptokurtic deviations as shown in Fig. 1 of normal distribution[10]. These factors weigh the proposed model's usefulness for segmentation of natural images more than the prevalent normal distribution. Accordingly we rephrase the previous paragraph to suggest this change. Thus, we assume the image data as a mixture of components each of which following compound normal with gamma distribution i.e., $Normal(\mu,\sigma^2) \underset{\sigma^{-2}}{\Lambda} Gamma(c\chi_v^2)$with some weight. And the whole image is thought of as following the weighted distribution where weighted distribution implies weighted average of the constituent components. We bring to the notice of the reader that in [11], the authors addressed the estimation of generalized mixtures in Pearson's system of distributions in general.

This paper is organized as follows. Section 2 covers compound normal with gamma mixture(CNGM) distribution, introduces estimation of model parameters via Expectation Maximization, and explains the derivation of estimates of the model parameters. In section 3, we treat image segmentation using EM framework, construct EM algorithm[12] for computing updated parameter estimates and discuss the experimental setup and give the results. In section 4, we evaluate the segmentation performance of the model vis-à-vis Gaussian mixture(NM) model and K-means clustering using different performance metrics[4],[14]. In section 5, we place our concluding remarks and suggest for further study.

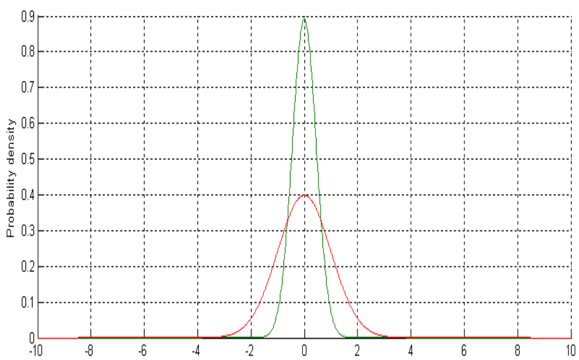


Fig. 1

2. The Model

2.1 Compound normal with gamma mixture(CNGM) distribution

As given in[10], compound normal with gamma distribution or $Normal(\mu,\sigma^2) \underset{\sigma^{-2}}{\Lambda} Gamma(c\chi_v^2)$ is formed by ascribing a distribution to σ^2 i.e., variance by considering it as a random variable and fitting a new distribution. The corresponding distribution is defined to have a density function as

$$f(x) = (2c)^{\frac{-v}{2}} \left[\Gamma\left(\frac{v}{2}\right)\right]^{-1} \int_0^{\infty} \left[\sqrt{2\pi}\sigma\right]^{-1} (\sigma^{-2})^{\left(\frac{v}{2}-1\right)} . \, exp[-(2c\sigma^2)^{-1}(2\sigma^2)^{-1}(x-\mu)^2] d\sigma^{-2}$$

$$= \frac{1}{c^{1/2}B(1/2, v/2)} \left[1 + \frac{(x-\mu)^2}{c}\right]^{-(v+1)/2} \quad (1)$$

We use the above distribution model as the basis for our work and now define a mixture model based on this.

The probability density function of the mixture model[12] is

$$p(x|\theta) = \sum_{l=1}^{M} \alpha_l \, p_l(x|\theta_l) \quad (2)$$

where the parameters are $\Theta = (\alpha_1, \dots, \alpha_M, \theta_1, \dots, \theta_M)$ such that
$\alpha_l = P(\theta = \theta_l)$
with $0 < \alpha_l < 1$ and

$$\sum_{l=1}^{M} \alpha_l = 1 \quad (3)$$

and each p_l is probability density function parameterized by θ_l where $\theta_l = (\mu_l, c_l, v_l)$. In other words, we assume we have M component densities mixed together with M mixing coefficients or weights α_l .

The probability density function p_l is defined according to (1) as

$$p_l(x|\theta_l) = \frac{1}{c_l^{1/2} B(1/2, v_l/2)} \left[1 + \frac{(x-\mu_l)^2}{c_l}\right]^{-(v_l+1)/2} \quad (4)$$

where x refers to each observation(individual pixel intensity here), μ_l, c_l, v_l are location, scale, and shape parameters of lth component of the mixture respectively and $B(1/2, v_l/2)$ is the beta function.

The above distribution has the following characteristics
$Mean = \mu$ (5)
Second central moment about mean, μ_2 i.e., variance is
$$\mu_2 = Variance = \frac{c}{(v-2)} ,$$
$v \geq 3 \, for \; finite \; variance$ (6)
Third central moment about mean, μ_3 is 0 since all odd moments about mean are zero.

Fourth central moment about mean, μ_4 is

$$\mu_4 = \frac{3c^2}{(v-4)(v-2)} \tag{7}$$

Pearson's β coefficients:

Coefficient of skewness

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0 \tag{8}$$

Coefficient of kurtosis

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3(v-2)}{(v-4)} \tag{9}$$

$v > 4$ for finite kurtosis.

For different values of v, this distribution has different shapes of frequency curves. As v increases β_2 tends to 3; thus it includes mesokurtic distribution.

2.2 Estimation of Model Parameters via Expectation Maximization

Before we attempt estimation of model parameters via Expectation Maximization for our model, we shall briefly discuss Maximum-likelihood and Expectation Maximization. The interested readers are encouraged to refer to [12] and also the references therein for a rigorous treatment of EM algorithm.

The likelihood of the ML estimate of the model parameters is given by

$$\mathcal{L}(\Theta|X) = \prod_{i=1}^{N} p(x_i|\Theta) \tag{10}$$

Often we maximize log-likelihood of the model parameters, $log(\mathcal{L}(\Theta|X))$ because it is analytically easier.
If we assume that $Z = (X, Y)$ is *complete data* where X is known and Y is unknown, a joint density function may be defined for z as

$$p(z|\Theta) = p(x, y|\Theta) = p(y|x, \Theta)p(x|\Theta) \tag{11}$$

which further leads to define the complete-data likelihood function,

$$\mathcal{L}(\Theta|Z) = \mathcal{L}(\Theta|X, Y) = p(X, Y|\Theta) \tag{12}$$

Expectation step:

The EM algorithm first finds the expected value of the complete-data log-likelihood $\log p(X, Y|\Theta)$ with respect to the unknown data Y given the observed data X and the current parameter estimates.
That is we define

$$Q\left(\Theta, \Theta^{(i-1)}\right) = E\left[\left(\log p(X, Y|\Theta)\right|X, \Theta^{(i-1)})\right]$$
$$= \int_{y \in \Upsilon} \log\ p(X, y|\Theta) f\left(y|X, \Theta^{(i-1)}\right) dy \tag{13}$$

where $\Theta^{(i-1)}$are the current parameters estimates that we use to evaluate the expectation and Θ are the new parameters that we optimize to increase Q. Note that $f\left(y|X, \Theta^{(i-1)}\right)$is the marginal distribution of the unobserved data and is dependent on both the observed data X and on the current parameters, and Υ is the space of values y can take on.

Maximization step:

Here we maximize the expectation we computed in the previous step. That is we find:

$$\Theta^i = \underset{\Theta}{\text{argmax}}\, Q\left(\Theta, \Theta^{(i-1)}\right) \tag{14}$$

The above two steps are repeated as necessary. Each iteration is guaranteed to increase the log-likelihood and the algorithm is guaranteed to converge to a local maximum of the likelihood function.

2.3 Derivation of estimates of model parameters:

Equation (12) takes the form

$$Q(\Theta, \Theta^g) = \sum_{y \in \Upsilon} \log\left(\mathcal{L}(\Theta|X, y)\right) p(y|X, \Theta^g)$$

solution of which yields

$$Q(\Theta, \Theta^g) = \sum_{l=1}^{M} \sum_{i=1}^{N} \log\left(\alpha_l p_l(x_i|\theta_l)\right) p(l|x_i, \Theta^g)$$

$$= \sum_{l=1}^{M} \sum_{i=1}^{N} \log(\alpha_l)\, p(l|x_i, \Theta^g) + \sum_{l=1}^{M} \sum_{i=1}^{N} \log\left(p_l(x_i|\theta_l)\right) p(l|x_i, \Theta^g) \tag{15}$$

where α_l is the prior probability of lth component, $p_l(x_i|\theta_l)$ is the conditional probability of x_i belonging to l and is defined for our model as in (4) , and $p(l|x_i, \Theta^g)$ is the posterior probability of component l given x_i and current estimates of parameters Θ^g and is defined as

$$p(l|x_i, \Theta^g) = \frac{\alpha_l p_l(x_i|\theta_l)}{\sum_{l=1}^{M} \alpha_l p_l(x_i|\theta_l)} \tag{16}$$

To maximize (15), we maximize the term containing α_l and the term containing θ_l independently since they are not related.

Analytical expressions for α_l and θ_l :

To find the expression for α_l, we introduce the Lagrange multiplier λ(optimization constrained by α_l) with the constraint that $\sum_l \alpha_l = 1$, and solve the following equation

$$\frac{\partial}{\partial \alpha_l}\left[\sum_{l=1}^{M} \sum_{i=1}^{N} \log(\alpha_l)\, p(l|x_i, \Theta^g) + \lambda\left(\sum_{l} \alpha_l - 1\right)\right] = 0 \tag{17}$$

Solution of which yields

$$\alpha_l = \frac{1}{N} \sum_{i=1}^{N} p(l|x_i, \Theta^g) \tag{18}$$

To find the expression for $\mu_l, c_l,$ *and* v_l we maximize the term containing $\theta_l(\mu_l, c_l, v_l)$ in (15) and solve by taking partial derivatives with respect to $\mu_l, c_l,$ *and* v_l and equate them to zero. That is we have to solve the following three equations

$$\frac{\partial}{\partial \mu_l}\left[\sum_{l=1}^{M}\sum_{i=1}^{N}\log\left(p_l(x_i|\theta_l)\right)p(l|x_i,\Theta^g)\right] = 0 \quad (19)$$

$$\frac{\partial}{\partial c_l}\left[\sum_{l=1}^{M}\sum_{i=1}^{N}\log\left(p_l(x_i|\theta_l)\right)p(l|x_i,\Theta^g)\right] = 0 \quad (20)$$

$$\frac{\partial}{\partial v_l}\left[\sum_{l=1}^{M}\sum_{i=1}^{N}\log\left(p_l(x_i|\theta_l)\right)p(l|x_i,\Theta^g)\right] = 0 \quad (21)$$

The solution for equations (19), (20), and (21) is derived in **Appendix** and is given as

$$\mu_l = \frac{\sum_{i=1}^{N} x_i p(l|x_i,\Theta^g)}{N\alpha_l} \quad (22)$$

$$c_l = \frac{(v_l+1)}{N\alpha_l}\sum_{i=1}^{N}(x_i-\mu_l)^2 p(l|x_i,\Theta^g) \quad (23)$$

$$v_l = \frac{N\alpha_l}{\sum_{i=1}^{N}\log\left[1+\frac{(x_i-\mu_l)^2}{c_l}\right]p(l|x_i,\Theta^g)} - 1 \quad (24)$$

3. Experimentation

3.1 Image segmentation

In this section, we describe how image segmentation is performed using EM algorithm for the mixture model defined by compound normal with gamma distribution i.e., $Normal(\mu,\sigma^2)\underset{\sigma^{-2}}{\Lambda}Gamma(c\chi_v^2)$. The basic steps here are

Step1: Decide M, the number of segments based on the number of components of the mixture i.e., fix $\Theta = (\alpha_1, \ldots, \alpha_M, \theta_1, \ldots, \theta_M)$.

Here we decide M, the number of segments, by looking at the number of peaks in the histogram of the image data. The decision on M is highly subjective for natural images. Though our approach for deciding M is purely heuristic in nature and involves some degree of randomness, we have relied on this approach since our main goal is to study the efficacy of the proposed model in comparison to the established methods under similar experimental settings. However, it is possible to obtain the number of image regions by optimizing model criterion like minimum description length(MDL)[3] which will be considered elsewhere.

Step2: Initialize Θ.

Here we initialize Θ using K-means clustering where K is set to M as decided in step 1. Since K-means gives mean and variance for all K, we use this variance , σ^2 , and some initial value for v_l to compute initial value for c_l by using (6).

Step3: Invoke EM algorithm.

EM Algorithm:

E-step: Compute the expectation as

$$p^{(q)}(l|x_i,\Theta^g) = \frac{\alpha_l^{(q)} p_l\left(x_i|\theta_l^{(q)}\right)}{\sum_{l=1}^{M}\alpha_l^{(q)} p_l\left(x_i|\theta_l^{(q)}\right)} \quad (q = 0,1,2,\ldots)$$

M-step: Compute the updated parameter estimates

$$\alpha_l^{(q+1)} = \frac{1}{N}\sum_{i=1}^{N} p^{(q)}(l|x_i,\Theta^g)$$

$$\mu_l^{(q+1)} = \frac{\sum_{i=1}^{N} x_i p^{(q)}(l|x_i,\Theta^g)}{N\alpha_l^{(q+1)}}$$

$$v_l^{(q+1)} = \frac{N\alpha_l^{(q+1)}}{\sum_{i=1}^{N}\log\left[1+\frac{\left(x_i-\mu_l^{(q+1)}\right)^2}{c_l^{(q)}}\right]p^{(q)}(l|x_i,\Theta^g)} - 1$$

$$c_l^{(q+1)} = \frac{\left(v_l^{(q+1)}+1\right)}{N\alpha_l^{(q+1)}}\sum_{i=1}^{N}\left(x_i-\mu_l^{(q+1)}\right)^2 p^q(l|x_i,\Theta^g)$$

$(q = 0,1,2,\ldots)$

The stopping criterion is

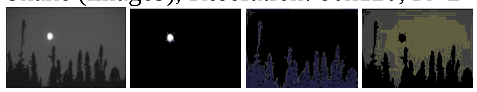
$$\left|\log\mathcal{L}^{(q+1)} - \log\mathcal{L}^{(q)}\right| < \epsilon$$

where $\mathcal{L}$, given by (10), is the likelihood of the parameter estimates, ϵ is error tolerance.

3.2 Implementation

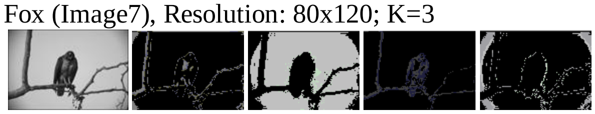
We have implemented the EM algorithm[12], [13] for our model in MATLAB and tested its efficacy for image segmentation using a collection of randomly picked images of 80x120 and 120x80 resolution from Berkeley Segmentation Image Dataset[5]. We have obtained fruitful segmentation results as shown in Fig. 2 and Table 1. In our experiment, we have used initial value for v_l as 100. Fig. 2 shows the original and its constituent segments for the images under study as obtained by our model based approach. Table 1 lists the corresponding EM estimates of $\alpha_l, \theta_l, \log\mathcal{L}$, and probability density plots for our model under varying conditions.

Snake (Image5), Resolution: 80x120; K=2

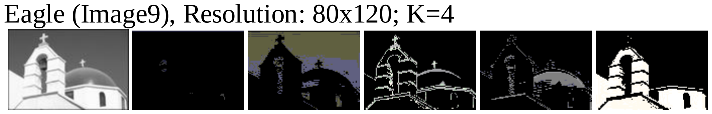


Sunset (Image2), Resolution: 80x120; K=3

Fox (Image7), Resolution: 80x120; K=3

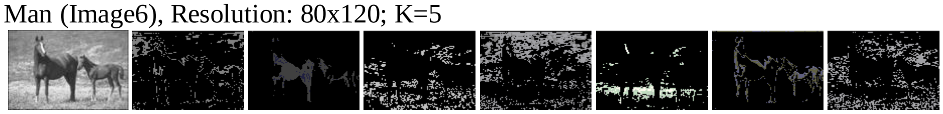


Eagle (Image9), Resolution: 80x120; K=4

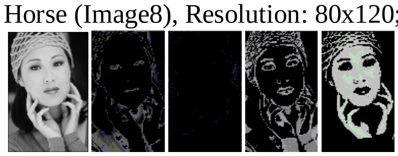


Church (Image1), Resolution: 80x120; K=5

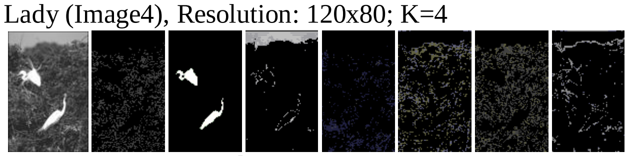
Man (Image6), Resolution: 80x120; K=5



Horse (Image8), Resolution: 80x120; K=7



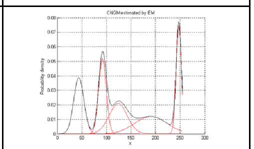
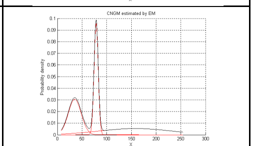
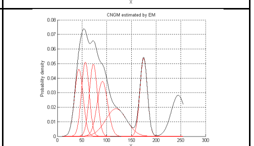
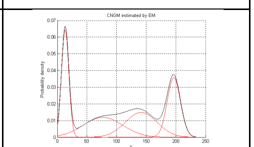
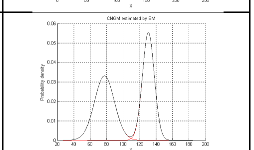
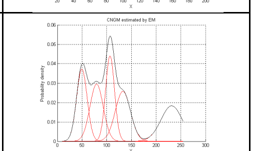
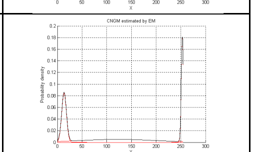
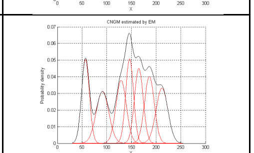
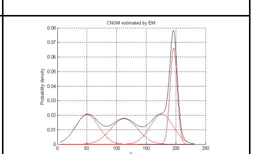
Lady (Image4), Resolution: 120x80; K=4



Crane (Image3), Resolution: 120x80; K=7

Fig. 2

Table 1: EM Estimates of CNGM

Image#	Image	α_l	Θ_l			Log L	Probability Density Plot
			μ_l	c_l	v_l		
1	Church	0.0295	43.5217	12225.3623	116.4324	-46062.0042 (K=5)	
		0.4354	91.2925	6728.5108	113.7220		
		0.1269	191.0480	120591.9695	110.2610		
		0.1232	123.7056	40050.4402	112.7671		
		0.2850	247.3454	3337.0373	115.9901		
2	Sunset	0.0072	158.3989	580883.8869	106.2480	-37155.9503 (K=3)	
		0.3231	35.1944	18903.8470	111.2617		
		0.6697	78.6328	2050.9579	118.1042		
3	Crane	0.2048	56.6862	7283.4899	119.4648	-46956.8852 (K=7)	
		0.0352	244.8573	25907.0627	130.2524		
		0.0953	174.8296	7281.1083	132.4431		
		0.1390	43.6292	9333.7853	125.0380		
		0.1750	91.3805	13429.9703	120.4027		
		0.2567	72.9626	7739.0791	119.6287		
		0.0940	119.0308	52827.3095	120.1271		
4	Lady	0.2230	77.0914	129189.4002	116.3851	-50285.3551 (K=4)	
		0.2072	13.2193	4654.2967	121.8693		
		0.2498	141.7001	83307.6606	117.1899		
		0.3200	196.6597	14763.2779	118.7747		
5	Snake	0.0611	77.3038	14822.8433	102.3654	-35050.7860 (K=2)	
		0.9389	130.6413	5321.0720	103.1120		
6	Man	0.1315	79.5459	19409.0749	106.4885	-46671.8175 (K=5)	
		0.5040	106.7774	8763.3540	107.0155		
		0.0808	231.5389	50469.6160	106.7767		
		0.0950	132.7610	26062.3880	109.0858		
		0.1887	49.7040	12322.8847	107.9518		
7	Fox	0.3179	253.2695	606.8151	123.5622	-37555.2564 (K=3)	
		0.5726	13.6695	2698.5352	120.9752		
		0.1095	122.5523	681368.7795	107.1093		
8	Horse	0.1428	130.8440	12285.8134	111.6455	-48789.2925 (K=7)	
		0.0969	56.9269	6979.7405	112.4993		
		0.1185	188.0797	10154.2553	111.2398		
		0.2483	147.7071	6948.1514	110.8624		
		0.0904	212.5953	15868.9668	114.9683		
		0.1233	91.8103	19171.3258	110.8071		
		0.1798	166.9797	8381.3243	110.8406		
9	Eagle	0.1352	112.9531	64739.0988	126.1231	-44258.9502 (K=4)	
		0.5649	196.2245	4861.7583	134.3927		
		0.1208	50.4070	46687.0443	122.4253		
		0.1791	176.1212	48108.4469	129.7354		

4. Performance Analysis of the model

4.1 Segmentation performance of the model

In this section, we compare the segmentation performance of the model studied against Gaussian mixture model and K-means clustering using three performance metrics, viz., Global Consistency Error(GCE), Probabilistic Rand Index(PRI), and Variation of Information(VoI) which are widely used in the current literature[4],[14]. All the three metrics compare two segmentation techniques by finding classification errors in the proposed method with respect to pixel labeling in the bench mark classifier.

In particular, the GCE is a measure of overlap between the segments produced by two segmentation methods. It is the average of minimum of the sums of the local refinement error(LRE) in both directions of the two segmentation methods for all pixels, where LRE is the degree of overlap between two sets; it is 0 if there is complete overlap in the given direction, and is 1 if there is no overlap at all. Thus GCE ranges between 0 and 1 and lower GCE values for test segmentation vis-à-vis the benchmark segmentation suggest their closeness with each other.

The PRI is defined based on the assumption that the test segmentation follows Bernoulli distribution given benchmark segmentation yielding values between 0 and 1, higher values being better.

The VoI is based on information theory and is a function of individual entropies of both the segmentations and mutual information between them yielding values between (0..VoI_max], lower values being better.

Mainly, we have considered the Gaussian mixture model and the agglomerative K-means clustering based segmentation methods to evaluate the performance of our model. The performance comparison is done on BSD image data set under varying conditions. Table 2 shows comparison of our model(CNGM) vis-à-vis Gaussian mixture(NM) and K-means clustering for different K values. Classification errors are almost zero in our model in comparison to the popular Gaussian mixture model. Even it is expected to perform better in case the pixel distribution in images is highly influenced by scale and shape parameters. Our model is shown to be more closer to K-means clustering than Gaussian mixture model. This is probably due to the reason that local distributions are more correctly modeled by our model than Gaussian mixture model. A scatter plot comparison of CNGM/NM/K-means for PRI, GCE, and VoI values on test images is shown in Fig. 3. These figures also testify the closeness of our model with Gaussian mixture model and relatively more closeness with K-means clustering than Gaussian mixture for the images we have considered for experimentation. Majority of the images, viz., Church, Sunset, Crane, Lady, Snake, and Fox have component distributions which are seemingly leptokurtic. For these images, we have observed that our model(CNGM) is more closer than Gaussian mixture(NM) model vis-à-vis K-means clustering. For the other images, viz., Man, Horse, and Eagle, NM is more closer to K-means. This may be due to the reason that CNGM is more correctly modeling leptokurtic distributions than NM.

4.2 Convergence Performance of the model

In our work, we have also addressed the convergence of EM algorithm of the model studied vis-à-vis Gaussian mixture model and found that the number of iterations taken by our model is almost always less than Gaussian mixture based one, and the Log-likelihood for our model is found to be always better. The segmentation time in seconds has been recorded for our model and Gaussian mixture model and found that our model is a competing one. These results are presented in Table 3 and Fig. 4.

Table 2: Segmentation Performance(CNGM/NM/K-means)

Image#	Image	CNGM/NM			CNGM/K-means			NM/K-means			K
		PRI	GCE	VoI	PRI	GCE	VoI	PRI	GCE	VoI	
1	Church	0.9545	0.1118	0.6019	0.8352	0.2156	1.2237	0.8210	0.2332	1.3803	5
2	Sunset	1	0	0	0.8941	0.0837	0.5760	0.8941	0.0837	0.5760	3
3	Crane	0.9908	0.0350	0.2147	0.9075	0.2167	1.0412	0.9092	0.2215	1.0633	7
4	Lady	0.9668	0.0603	0.3746	0.9169	0.1526	0.7778	0.8896	0.1948	0.9674	4
5	Snake	0.9988	0.0012	0.0125	0.9996	0.0004	0.0048	0.9983	0.0016	0.0160	2
6	Man	0.9026	0.1314	0.7178	0.9250	0.1406	0.7207	0.9232	0.1186	0.6891	5
7	Fox	0.9938	0.0088	0.0657	0.9350	0.0650	0.4812	0.9295	0.0676	0.5050	3
8	Horse	0.8685	0.3219	1.4234	0.9098	0.2198	1.1102	0.9241	0.2023	0.9998	7
9	Eagle	0.9622	0.0518	0.3236	0.9519	0.0734	0.4206	0.9869	0.0343	0.2144	4

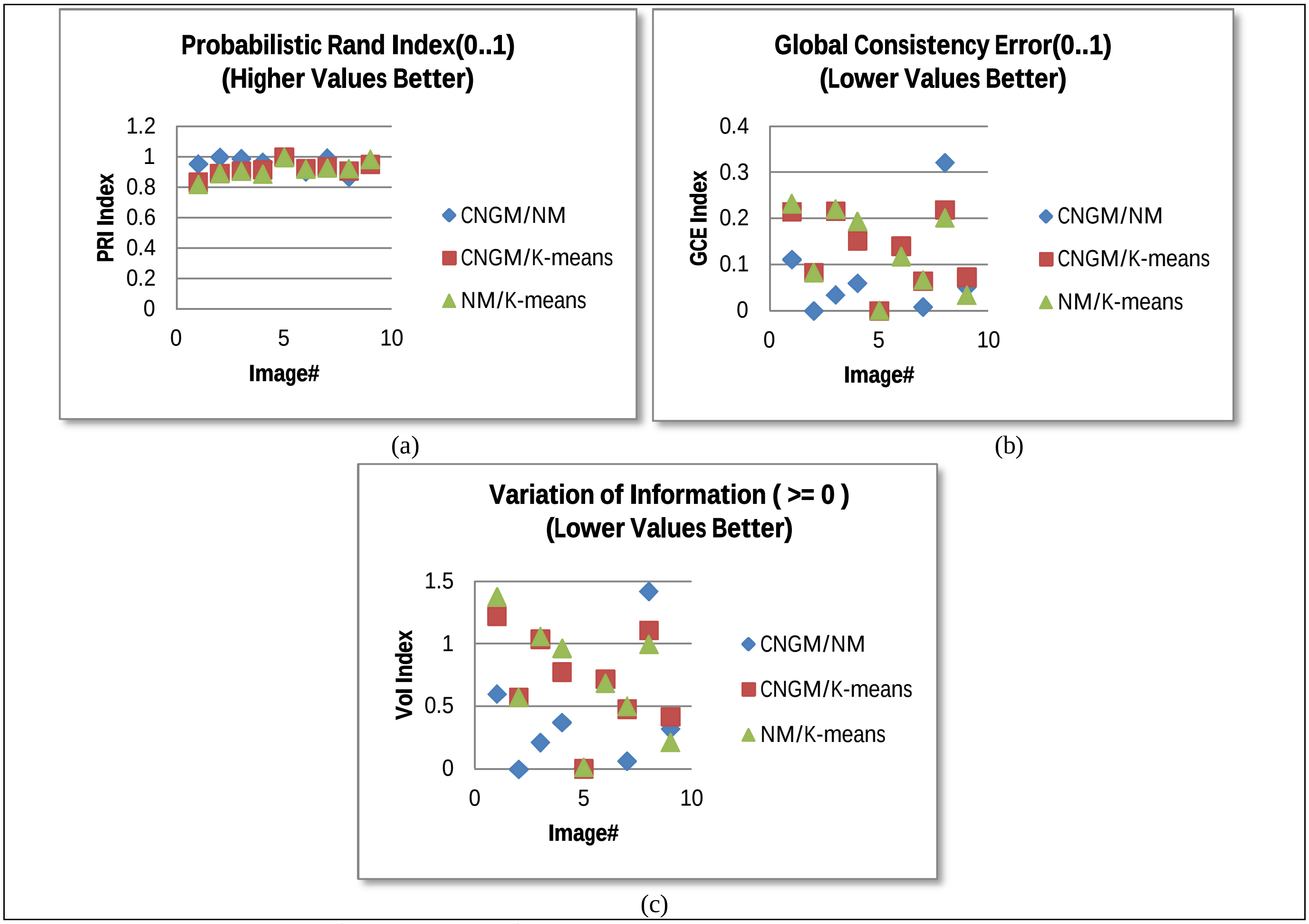


Fig. 3 Plots of Segmentation Performance: CNGM/NM/K-means; (a) PRI, (b) GCE, and (c) VoI

Table 3: Convergence Performance(CNGM/NM)

Image#	Image	No of Iterations		Log Likelihood		Segmentation Time (Seconds)		K
		CNGM	NM	CNGM	NM	CNGM	NM	
1	Church	9	16	-46062.0042	-1140379.3523	99.63	70	5
2	Sunset	8	10	-37155.9503	-162034.2116	52.03	28.16	3
3	Crane	14	23	-46956.8852	-307538.0757	210	131.7	7
4	Lady	13	27	-50285.3551	-658283.3970	113	95	4
5	Snake	1	4	-35050.7860	-53241.4023	5.9	6.95	2
6	Man	5	19	-46671.8175	-166267.4362	59.48	86.66	5
7	Fox	8	13	-37555.2564	-4016587.1177	53.4	34.55	3
8	Horse	7	49	-48789.2925	-81818.7168	109.48	298.94	7
9	Eagle	13	7	-44258.9502	-308081.3139	155.9	36.77	5

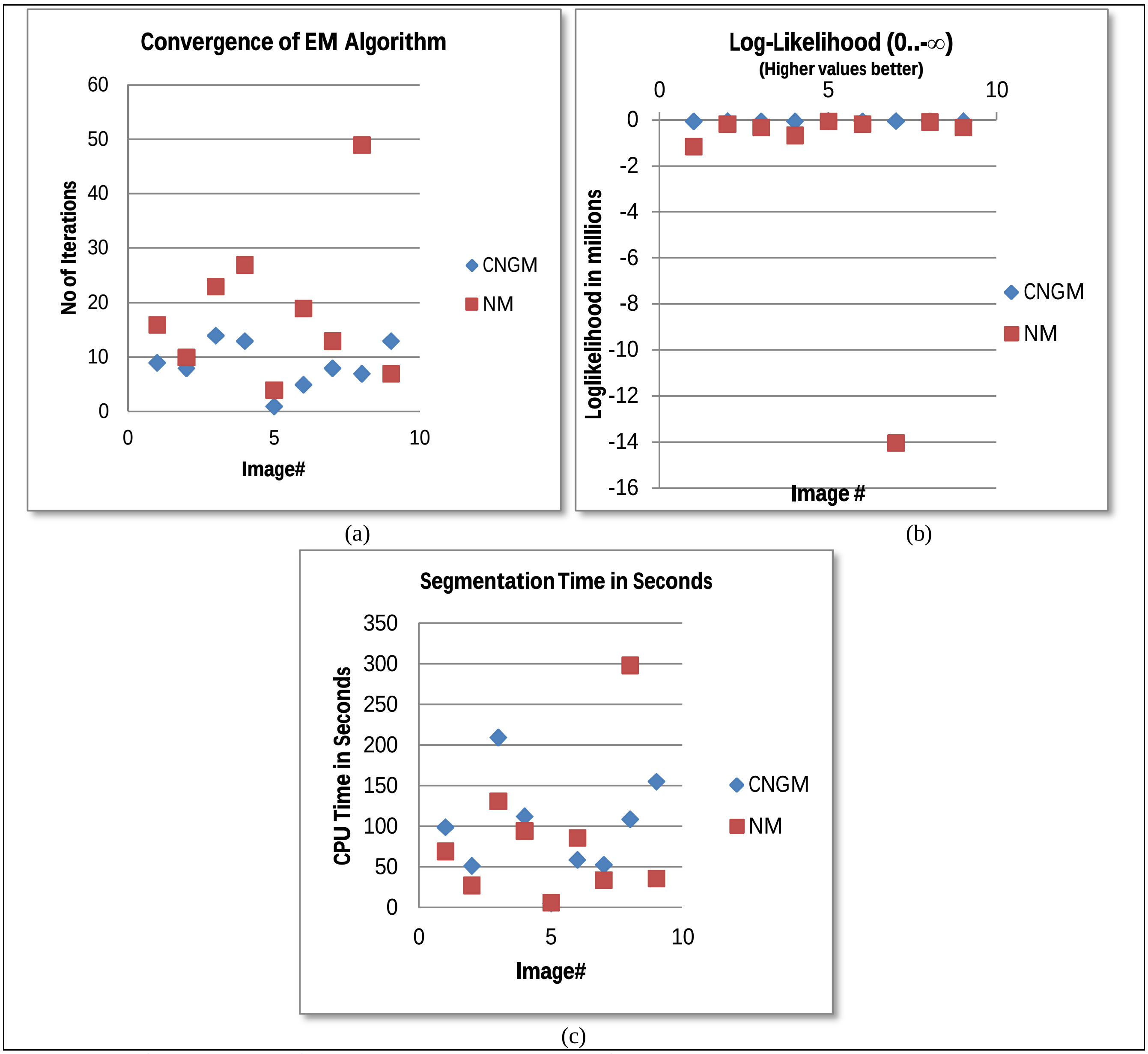


Fig. 4 Plots of Convergence Performance: CNGM/NM; (a) Number of Iterations,(b) Log-likelihood, and (c)Segmentation Time

5. Conclusions

In this work, we have studied the efficacy of the compound normal with gamma mixture distribution model for solving image segmentation problem. We have derived the updated equations for model parameters(closed form ones for $\mu_l, c_l, and\ v_l$; surprisingly for v_l which appears to be a possible novelty) and constructed EM algorithm for the proposed model. We experimented the model using MATLAB. We have rigorously tested the performance of the model vis-à-vis Gaussian mixture model and K-means clustering on several images collected from BSD image data set[5] under varying conditions. Classification errors are almost zero in our model in comparison to the popular Gaussian mixture model. Even it appears to perform better in case the pixel distribution in images is highly influenced by scale and shape parameters. Our model is shown to be more closer than Gaussian mixture model to K-means clustering. This is probably due to the reason that local distributions are more correctly modeled by our model than Gaussian mixture model. We have also studied convergence properties of EM algorithm based on our model with respect to number of iterations and log likelihood and found that it is relatively better than Gaussian mixture model. We have also given segmentation time that our algorithm takes in comparison to Gaussian mixture model based one and found that ours is a competing one. In our work, the decision on the number of segments, since it is more subjective for natural images, is based on identification of the number of peaks in the histogram plot of the image data. This approach is purely heuristic in nature and involves some degree of randomness. However, it is possible to obtain the number of image regions by optimizing model criterion like minimum description length(MDL)[3] which will be considered elsewhere.

Appendix

Aim: To derive solution for estimates of the model parameters, $\theta_l(\mu_l, c_l, v_l)$.

We know that

$$\alpha_l = \frac{1}{N}\sum_{i=1}^{N} p(l|x_i, \Theta^g) \quad (1)$$

Here, we present the derivation of model parameters' estimates, $\theta_l(\mu_l, c_l, v_l)$, by solving the following three equations.

$$\frac{\partial}{\partial \mu_l}\left[\sum_{l=1}^{M}\sum_{i=1}^{N} \log\big(p_l(x_i|\theta_l)\big)\, p(l|x_i, \Theta^g)\right] = 0 \quad (2)$$

$$\frac{\partial}{\partial c_l}\left[\sum_{l=1}^{M}\sum_{i=1}^{N} \log\big(p_l(x_i|\theta_l)\big)\, p(l|x_i, \Theta^g)\right] = 0 \quad (3)$$

$$\frac{\partial}{\partial v_l}\left[\sum_{l=1}^{M}\sum_{i=1}^{N} \log\big(p_l(x_i|\theta_l)\big)\, p(l|x_i, \Theta^g)\right] = 0 \quad (4)$$

where $p_l(x_i|\theta_l)$ is the probability density function governed by compound normal with Gamma distribution and is given by

$$p_l(x_i|\theta_l) = \frac{1}{c_l^{1/2} B(1/2\ , v_l/2)}\left[1 + \frac{(x_i-\mu_l)^2}{c_l}\right]^{-(v_l+1)/2}$$

After substitution of the RHS of the above equation in (2) and expanding *log* term, it yields

$$\frac{\partial}{\partial \mu_l}\left[\sum_{l=1}^{M}\sum_{i=1}^{N}\left[\log\frac{1}{c_l^{1/2}B(1/2\ , v_l/2)} - \left(\frac{v_l+1}{2}\right)\log\left[1 + \frac{(x_i-\mu_l)^2}{c_l}\right]\right] p(l|x_i, \Theta^g)\right] = 0$$

Taking linear term as approximation for the second '*log*' term expansion in the above equation, it becomes

$$\frac{\partial}{\partial \mu_l}\left[\sum_{l=1}^{M}\sum_{i=1}^{N}\left[\log\frac{1}{c_l^{1/2}B(1/2\ , v_l/2)} - \left(\frac{v_l+1}{2}\right)\left[\frac{(x_i-\mu_l)^2}{c_l}\right]\right] p(l|x_i, \Theta^g)\right] = 0 \quad (5)$$

since for any real no z that satisfies $0<z<2$, the following formula holds:

$$\ln(z) = (z-1) - \frac{(z-1)^2}{2} + \frac{(z-1)^3}{3} - \frac{(z-1)^4}{4} + ...$$

Taking linear term as approximation, $\ln z = (z-1)$.

Solution of (5) for component l after ignoring constant terms with respect to μ_l , it becomes

$$\sum_{i=1}^{N}\left[-\left(\frac{v_l+1}{2}\right)(-2)\left(\frac{x_i-\mu_l}{c_l}\right)\right] p(l|x_i, \Theta^g) = 0$$

The above equation leads to μ_l as

$$\mu_l = \frac{\sum_{i=1}^{N} x_i\, p(l|x_i, \Theta^g)}{N\alpha_l} \quad (6)$$

A similar treatment of (3) yields

$$\frac{\partial}{\partial c_l}\left[\sum_{l=1}^{M}\sum_{i=1}^{N}\left[\frac{1}{2}\log c_l + \log B(1/2\ , v_l/2) + \left(\frac{v_l+1}{2}\right)\left[\frac{(x_i-\mu_l)^2}{c_l}\right]\right] p(l|x_i, \Theta^g)\right] = 0$$

Solving the above equation for component l after ignoring constant terms with respect to c_l , it becomes

$$\sum_{i=1}^{N}\left[\frac{1}{c_l} - \left(\frac{(v_l+1)(x_i-\mu_l)^2}{c_l^2}\right)\right] p(l|x_i, \Theta^g) = 0$$

The above equation further leads to c_l defined as

$$c_l = \frac{(v_l+1)}{N\alpha_l}\sum_{i=1}^{N}(x_i-\mu_l)^2 p(l|x_i,\Theta^g) \quad (7)$$

A similar treatment of (4) yields

$$\frac{\partial}{\partial v_l}\left[\sum_{l=1}^{M}\sum_{i=1}^{N}\left[\frac{1}{2}\log c_l + \log B(1/2\,, v_l/2) + \left(\frac{v_l+1}{2}\right)\log\left[1 + \frac{(x_i-\mu_l)^2}{c_l}\right]\right] p(l|x_i,\Theta^g)\right] = 0$$

Solving the above equation for component l after ignoring constant terms with respect to v_l , it becomes

$$\sum_{i=1}^{N}\left[\frac{\partial \log B(1/2\,, v_l/2)}{\partial v_l} + \frac{\partial}{\partial v_l}\left(\frac{v_l}{2}\right)\log\left[1 + \frac{(x_i-\mu_l)^2}{c_l}\right]\right] p(l|x_i,\Theta^g) = 0 \quad (8)$$

The solution for the first term in the above equation is as follows:

$$\frac{\partial \log B(1/2\,, v_l/2)}{\partial v_l} = \frac{\frac{\partial}{\partial v_l}B(1/2\,, v_l/2)}{B(1/2\,, v_l/2)} \quad (9)$$

Since we know that

$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$, then

$$\frac{\partial}{\partial v_l}B(1/2\,, v_l/2) = \int_0^1 \frac{\partial}{\partial v_l}(1-x)^{-1/2}x^{(v_l-2)/2}dx$$

$$\Rightarrow \int_0^1 x^{(v_l-2)/2}\left(\frac{\log x}{2}\right)(1-x)^{-1/2}dx \quad (10)$$

$$\because \frac{\partial}{\partial v_l}x^{(v_l-2)/2} = x^{(v_l-2)/2}\left(\frac{\log x}{2}\right)$$

Taking linear approximation of *log* x as (x-1) in (10), it takes the form

$$-\int_0^1 x^{\frac{v_l}{2}-1}(1-x)^{\frac{3}{2}-1}dx \Rightarrow -B(v_l/2\,, 3/2)$$

Hence (9) takes the form

$\frac{-B(v_l/2\,,3/2)}{B(1/2\,,v_l/2)}$, which further reduces to $\frac{-1}{2(v_l+1)}$

Therefore, solution for (8) may be rewritten as

$$\sum_{i=1}^{N}\left[\frac{-1}{2(v_l+1)} + \left(\frac{1}{2}\right)\log\left[1 + \frac{(x_i-\mu_l)^2}{c_l}\right]\right] p(l|x_i,\Theta^g) = 0$$

Solving for v_l , the above equation yields

$$v_l = \frac{N\alpha_l}{\sum_{i=1}^{N}\log\left[1 + \frac{(x_i-\mu_l)^2}{c_l}\right]p(l|x_i,\Theta^g)} - 1 \quad (11)$$

References

[1] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification.* Second edition. Wiley 2007

[2] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Prentice Hall,2008

[3] T. Lei and J. K. Udupa. "Performance evaluation of finite normal mixture model-based image segmentaion techniques."*IEEE Trans on Image Processing*, vol 12, no 10, pp. 1153-1169, 2003.

[4] R. Unnikrishnan, C. Pantofaru, and M. Hebert. "Toward objective evaluation of image segmentation algorithms", *IEEE Trans on Pattern Anal. and Machine Intell.*, vol 29, no 6, pp. 929-944, 2007.

[5] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A Database of Human Segmented Natural Images and Its Application to Evaluating Segmentation Algorithms and Measuring EcologicalStatistics." Proc. Int'l Conf. Computer Vision, 2001.

[6] J. Zhang and J. M. Modestino, "A model fitting approach to cluster validity with application to stochastic model based image segmentation."*IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 10, pp.1009–1016, 1990.

[7] T. Lei and W. Sewchand, "Statistical approach to x-ray CT imaging and its applications in image analysis—Part 2: A new stochastic model based image segmentation technique for CT image." *IEEE Trans. Med.Imag.*, vol. 11, no. 1, pp. 62–69, 1992.

[8] Z. Liang and J. R. MacFall, "Parameter estimation and tissue segmentation of multispectral MR images." *IEEE Trans. Med. Imag.*, vol. 13, no.3, pp. 441–449, 1994.

[9] F. J. Estrada and A. D. Jepson, "Benchmarking Image Segmentation Algorithms", International Journal of Computer Vision (2009) 85: 167–181.

[10] Normal L. Johnson, Samuel Kotz and N. Balakrishnan. *Continuous Univariate Distributions,*vol-I , Second edition. John Wiley&Sons, 2007

[11] Y. Delignon, A. Marzouki, and W. Pieczynski, "Estimation of Generalized Mixtures and Its Application in Image Segmentation", *IEEE Trans on Image Processing*, vol. 6, no. 10, pp. 1364-1375, 1997.

[12] J. A. Bilmes, "A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models." International Computer Science Institute, Berkely CA, 94704

[13] MATLAB open source code for EM algorithm for Gaussian mixture model written by Patrick P C Tsui, PAMI research group, University of Waterloo,2006.

[14] MATLAB open source code for segmentation performance metrics written by John Wright, and Allen Y. Yang, University of California, Berkeley,2007.

Viziananda Row Sanapala received his B.Tech in Chemical Engg in 1983 and M Tech in Computer Science & Technology in 1987 both from Andhra University. He served as Lecturer in Computer Science in Acharya Nagarjuna University, India from 1988 to 1995 and presently is serving as Associate Professor in Computer Science & Systems Engg in Andhra University, India. His research interests include image processing, pattern recognition, artificial intelligence, data mining, distributed systems, and other areas. He is currently a fellow of institution of engineers(India).

Sreenivasa Rao Kraleti received his M Sc in Statistics in 1982 and Ph D in Statistics in 1990 both from Andhra University. He joined Andhra University in 1985 as Lecturer and is currently serving as Professor of Statistics in Andhra University, India. His research interests are data analysis, stochastic modeling, statistical inference, and other areas. He published several papers in journals and conferences at national and international level. He has guided a good number of doctoral level students and is presently guiding several others. He holds to his credit several administrative positions in his serving university. He is currently a fellow of Andhra Pradesh academy of sciences(India).

Srinivasa Rao Peri received his M Sc in Statistics in 1982 from Andhra University and Ph D in Computer Science in 1987 from IIT, Kharagpur, India. He joined Andhra University in 1987 as Lecturer and is presently serving as Professor in Computer Science & Systems Engg in Andhra University, India. His research interests include image processing, queueing theory, data mining, and other areas. He published several papers in journals and conferences at national and international level. He guided several doctoral level students and currently has a number of students pursuing their research under him. He holds to his credit several administrative positions in his serving university.