Journal of Signal and Image Processing
ISSN: 0976-8882 & E-ISSN: 0976-8890, Vol. 2, Issue 1, 2011, pp-19-25
Available online at http://www.bioinfo.in/contents.php?id=48

SEGMENTATION OF BRAIN MRI IMAGES USING SKEW GAUSSIAN DISTRIBUTION WITH K-MEANS AND EM ALGORITHMS

NAGESH VADAPARTHI[1], SRINIVAS YARRAMALLE[2], SURESH VARMA P.[3], MURTHY P.S.R.[4]
[1]Department of IT, MVGR College of Engineering, Vizianagaram, India, itsnageshv@gmail.com
[2] Department of IT, GITAM University, Visakhapatnam, India, ysrinivasit@rediffmail.com
[3] Department of Computer Science, Aadikavi Nannaya University, Rajahmundry, India, vermaps@yahoo.com
[4] Department of Mathematics, GITAM University, Visakhapatnam, India, drmurtypsr@yahoo.com
*Corresponding author. E-mail: itsnageshv@gmail.com

Received: Received: July 31, 2011; Accepted: August 13, 2011

Abstract- In this paper, an efficient approach for medical image segmentation based on Skew Gaussian distribution using EM algorithm is proposed. It is necessary to classify the brain voxels into one of the 3 main tissues mainly Gray matter (GM), White matter (WM) and Cerebro Spinal fluid (CSF) in any brain MRI image. Quantization of Gray & White matter is a topic of concern in neuro-degenerative disorders. Viz., Alzheimer disease and Parkinson's diseases. Hence, it is necessary to identify the tissue more efficiently. In this approach we used Skew Gaussian distribution to classify the tissue voxels and the updated parameters are obtained using EM algorithm. The outputs generated are evaluated using the medical image quality metrics. Experimentation is carried out on two different T_1 weighted brain images.
Keywords- Segmentation, Skew Gaussian distribution, Classification, medical image quality metrics, Finite Gaussian Mixture Model

Introduction

The field of medical imaging improved significantly with recent advancements in technology. The wide spread availability of suitable detectors have helped for the rapid development of new technologies for monitoring, diagnosing and as well as treatment of the patients. Many models were utilized to identify the diseases, but MRI brain segmentation has gained popularity over the other models because of the non-ionizing radiation that is being used. Many researchers have developed models for the medical image segmentation, in particular to the brain segmentation [1-5]. Among these models, brain segmentation based on Gaussian Mixture models has gained significant importance [2, 3], this is due to the fact of the basic assumption that the pixel intensities inside the image regions of a medical image follow a bell shaped distribution and hence to identify the patterns of the pixels inside these image regions, a bell shaped distribution i.e., Gaussian distribution is utilized [6]. These approximations of using a Gaussian Mixture models for the brain images is a crude approximation, since, in reality the pattern of the pixels inside the image regions may be Mesokurtic or Leptokurtic i.e., the shape of the pixels may be either symmetric or asymmetric. Hence, assuming that the pixel intensities to follow a bell shaped distribution is a crude approximation [7]. Therefore, in order to segment the medical images more approximately, it is needed to have a better distribution that contains Gaussian distribution as a particular case. In this paper, Skew Gaussian mixture model is utilized for segmenting the medical images. The advantage of this model is that, it contains Gaussian distribution as a particular case. The performance evaluation of the developed method is analyzed using quality metrics like Jaccard coefficient (JC) and Volume Similarity (VS). A comparative study with respect to GMM is also presented. The experimentation is carried out using both T_1 and T_2 weighted brain medical images.
The rest of the paper is organized as follows: Section-2 describes the K-Means algorithm and the details of Skew Gaussian distribution is presented in section-3. Section-4 discusses about the initialization of parameters and updation of parameters is explained in section-5. In section-6, segmentation algorithm is proposed and the experimental results and performance evaluation is done in section-7 and Section-8. Finally, section-9 concludes the paper.

K – Means Algorithm

The main disadvantage of unsupervised learning algorithm is the conversion of heterogeneous to homogeneous data. Many segmentation algorithms have been developed and analyzed [1]. But, the main disadvantage of segmentation algorithm is that it differs from application to application and there exists no unique segmentation model which suits for all purposes [7]. In order to segment the unsupervised data, K-means algorithm is used. K-means algorithm is one of the simplest partition clustering methods. The main disadvantage of K-means algorithm is to identify the initial value of K. Hence, a histogram is utilized for the initialization of K. The K-means algorithm is given below.

Inputs:

P = { 1, 2, ... *k* } (Pixels to be clustered)

K (No of Clusters)

Outputs:

C = { 1, 2,........ *k* } (Cluster Centroids)

m: P -> {1, 2…K} (Cluster Membership)

Algorithm K-Means:

Set C to initial value (e.g. Random selection of P)

For each $p_i \in P$

$$m(p_i) = \underset{j \in \{1..n\}}{\operatorname{argmin}} \text{distance}(p_i, c_j)$$

End

While m has changed

For each $j \in \{1.....K\}$

Recompute c_i as the centroid of

$\{p \mid m(p) = i\}$

End

For each $p_i \in P$

$$m(p_i) = \underset{j \in \{1..n\}}{\operatorname{argmin}} \text{distance}(p_i, c_j)$$

End

End

End

Skew Gaussian distribution

The pixels intensities inside the medical images may not be symmetric or bell shaped due to several factors associated like part of the body, bone structure etc. In these cases, the pixels are distributed asymmetrically and follow a skew distribution. Hence, to categorize these sorts of medical images, Skew Gaussian distribution is well suited. Every image is a collection of several regions. To model the pixel intensities inside these image regions, we assume that the pixels in each region follow a Skew normal distribution, where the probability density function is given by

$$f(z) = 2.\phi(z).\Phi(\propto z); \qquad -\infty < z < \infty \tag{1}$$

$$where, \quad \Phi(\propto z) = \int_{-\infty}^{\propto z} \phi(t)\; dt \tag{2}$$

$$and, \qquad \phi(z) = \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} \tag{3}$$

Let, $y = \mu + \sigma z$

$$z = \frac{y-\mu}{\sigma} \tag{4}$$

Substituting equations (2), (3), and (4) in equation (1),

$$f(z) = \sqrt{\frac{2}{\pi}}.\quad e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}\left[\int_{-\infty}^{\alpha\left(\frac{y-\mu}{\sigma}\right)} \frac{e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}}{\sqrt{2\pi}}\,dt\right] \tag{5}$$

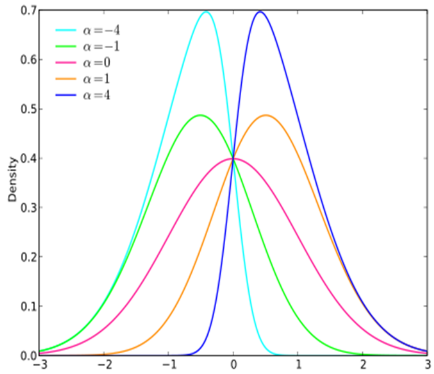


Fig. 1-Image Intensity Graph

Initialization of Parameters

In order to initialize the parameters, it is needed to obtain the initial values of the model distribution. The initial estimates of the Mixture model μ_i, σ_i, λ_i and α_i where i=1,2,.....,k are estimated using K-Means algorithm as proposed in section II. It is assumed that the pixel intensities of the entire image is segmented into a K component model π_i, i=1,2,..,k with the assumption that $\pi_i = 1/k$ where k is the value obtained from K-Means algorithm discussed in section-2.

Updation of Initial Estimates through EM algorithm

The initial estimates of μ_i, σ_i and α_i that are obtained from section – 4 are to be refined to obtain the final estimates. For this purpose EM algorithm is utilized. The EM algorithm consists of 2 steps E-step and M-Step. In the E-Step, the initial estimates obtained in section – 4 are taken as input and the final updated equations are obtained in the M-Step. The updated equations for the model parameters μ_i, σ_i and α_i are given below.

$$\mu^{(l+1)} =$$

$$y + \sigma^{2\,(l)} + \frac{1}{\int_{-\infty}^{\alpha^{(l)}\left(\frac{y-\mu^{(l)}}{\sigma^{(l)}}\right)} e^{-\frac{1}{2}\left(\frac{t-\mu^{(l)}}{\sigma^{(l)}}\right)^2} dt} +$$

$$\int_{-\infty}^{\alpha^{(l)}\left(\frac{y-\mu^{(l)}}{\sigma^{(l)}}\right)} (t-\mu^{(l)})\, e^{-\frac{1}{2}\left(\frac{t-\mu^{(l)}}{\sigma^{(l)}}\right)^2} dt -$$

$$\sigma^{(l)}\alpha^{(l)}\, e^{\frac{\left[\left(\alpha^{(l)}+\sigma^{(l)}\right)\mu^{(l)} - \alpha^{(l)}y\right]^2}{2\sigma^{4\,(l)}}} \tag{6}$$

$$\sigma^{(l+1)} =$$

$$1 \Big/ \frac{\left(y-\mu^{(l)}\right)^2}{\sigma^{3\,(l)}} + \frac{1}{\int_{-\infty}^{\alpha^{(l)}\left(\frac{y-\mu^{(l)}}{\sigma^{(l)}}\right)} e^{-\frac{1}{2}\left(\frac{t-\mu^{(l)}}{\sigma^{(l)}}\right)^2} dt} + \int_{-\infty}^{\alpha^{(l)}\left(\frac{y-\mu^{(l)}}{\sigma^{(l)}}\right)} \frac{\left(t-\mu^{(l)}\right)^2}{\sigma^{3\,(l)}}\, e^{-\frac{1}{2}\left(\frac{t-\mu^{(l)}}{\sigma^{(l)}}\right)^2} dt + \alpha^{(l)} \left(\frac{\mu^{(l)}-y}{\sigma^{2\,(l)}}\right) e^{\frac{\left[\left(\alpha^{(l)}+\sigma^{(l)}\right)\mu^{(l)}-\alpha^{(l)}y\right]^2}{2\sigma^{4\,(l)}}} \qquad (7)$$

$$\alpha^{(l+1)} =$$

$$\frac{\sqrt{2}\,\sigma^{2\,(l)}}{\mu^{(l)}-y} \left[\log\left(\int_{-\infty}^{\alpha^{(l)}\left(\frac{y-\mu^{(l)}}{\sigma^{(l)}}\right)} e^{-\frac{1}{2}\left(\frac{t-\mu^{(l)}}{\sigma^{(l)}}\right)^2} dt\right) - \log\left(\frac{y-\mu^{(l)}}{\sigma^{(l)}}\right)\right]^{\frac{1}{2}} - \frac{\sigma^{(l)}\mu^{(l)}}{\mu^{(l)}-y} \qquad (8)$$

Segmentation Algorithm

After obtaining the final estimates, the next step is image reconstruction by allocating the pixels to the segmentation. This operation is done by segmentation algorithm. This segmentation algorithm is given as follows:

Step-1: Obtain the pixel intensities of the gray image. Let they be represented by x_{ij}.

Step-2: Obtain the number of regions by k-means algorithm and divide the (image) pixel into regions.

Step-3: For each region obtain the initial estimates using moment methods of estimation for μ_i, σ_i. Let α_i=1/k be the initial estimate for α_i.

Step-4: Obtain the refined estimates of μ_i, σ_i, α_i for i=1….k using updated equations for the parameters derived by EM algorithm with step 3 estimates as initial estimates.

Step-5: Implement the segmentation and retrieval algorithm by considering maximum Likelihood estimate.

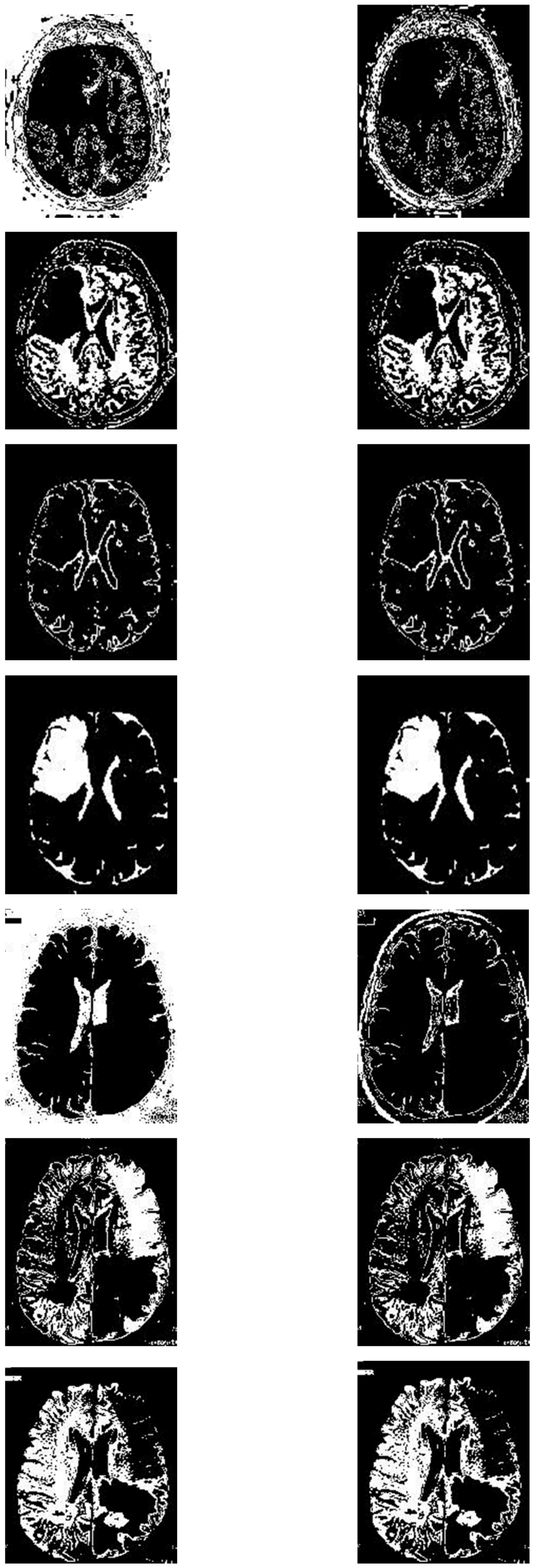
Step-6: With the step 5 obtain the image quality metric.

Step-7: The image segmentation is carried out by assigning each pixel into a proper region (Segment) according to maximum likelihood estimates of the j^{th} element L_j according to the following equation

$$L_j = = \mathrm{Max}_j \left\{ \sqrt{\frac{2}{\pi}}.\quad e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} \left[\int_{-\infty}^{\alpha\left(\frac{y-\mu}{\sigma}\right)} \frac{e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}}{\sqrt{2\pi}}\, dt\right]\right\}$$

Experimentation

INPUT IMAGE OUTPUT IMAGE



INPUT IMAGE OUTPUT IMAGE

Fig. 2-Brain MRI images

The above developed segmentation algorithm is applied on brain images obtained from web brain images. To evaluate our developed algorithm, both T_1 & T_2 weighted images were utilized. We have considered mainly 2 images on brain having deformities. The white matter and gray matter are segmented appropriately by the developed algorithm, where by helping out in the identification of damaged tissues.

Performance Evaluation

In order to evaluate our proposed model, we demonstrated our segmentation algorithm with Finite Skew Gaussian Mixture Model with K-Means algorithm and applied it to eight different images both of type T_1 and T_2. In T_1 weighted images the water is shown as darker and fat as brighter and in T_2 images fat is shown as darker and White matter is shown lighter. Among these images, T_1 images provide good gray matter and it highlights the fat decomposition. The input medical images are obtained from brain web images. We have assumed that the pixel intensities inside the brain images are non-symmetric and follow a Skew Gaussian distribution and the whole medical image is a mixture of Skew Gaussian distribution. The initialization of parameters for each segment is done using K-Means algorithm. The performance evaluation of the retrieved images can be done by subjective testing or objective testing. Objective testing is always preferred since they are based on numeric results. The performance of developed algorithm is evaluated by using quality metrics given by Eskicioglu et al[14]. The performance of the developed algorithm was compared with the medical Image Segmentation algorithm based on Finite Gaussian Mixture Model by using image quality metrics namely, Average Difference, Maximum Distance, Image Fidelity, Mean Squared Error, Signal to Noise Ratio, Jaccard index and Volume Similarity. The formulas for evaluating these metrics are given below in Table-1.

The developed method is compared with Gaussian Mixture Model. The results are shown in Table-2 and Table-3 and the corresponding graphs are shown in Graphs-1 and Graphs-2. From the Table – 2, Table – 3, Graph-1, Graph-2 and Figure – 2 it can be clearly observed that the developed algorithm performs much superior to the existing algorithm with respect to image quality metrics. This model is well suited in particular for medical image, where the shape of the image depends on the body structure.

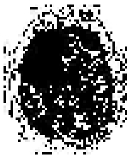
Table 1- Image Quality Metrics Formulae

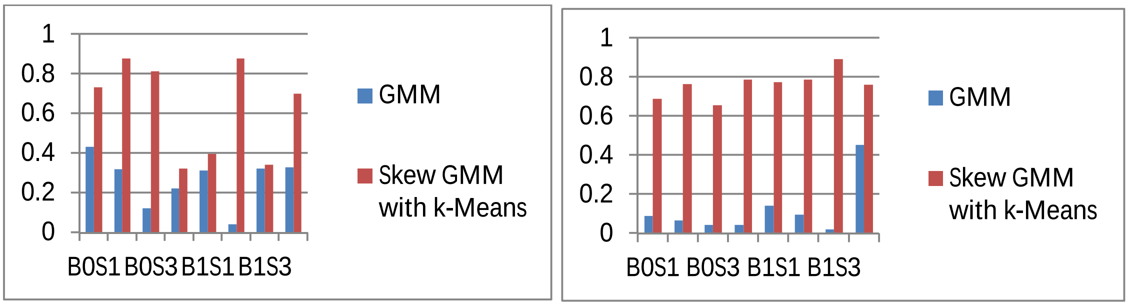
Quality metric	Formula to Evaluate
Average Difference	$\sum_{j=1}^{M}\sum_{k=1}^{N}[F(j,k) - \hat{F}(j,k)] \,/\, MN$ Where M,N are image matrix rows and colomns
Maximum Distance	Max{$\lvert F(j,k) - \hat{F}(j,k)\rvert$}
Image Fidelity	$1 - [\sum_{j=1}^{M}\sum_{k=1}^{N} [F(j,k) - \hat{F}(j,k)]^2 \,/\, \sum_{j=1}^{M}\sum_{k=1}^{N} [F(j,k))]^2]$ Where M,N are image matrix rows and colomns
Mean Squared error	$\frac{1}{MN}\sum_{j=1}^{M}\sum_{k=1}^{N} [O\{F(j,k)\} - O\{\hat{F}(j,k)\}]^2 \,/\, \sum_{j=1}^{M}\sum_{k=1}^{N}[O\{F(j,k)^2$ Where M,N are image matrix rows and colomns
Signal to noise ratio	$20.\log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right)$ Where, MAX_I is maximum possible pixel value of image, MSE is the Mean squared error
Jaccard quotient	$\frac{\lvert X \cap Y \rvert}{\lvert X \cup Y \rvert} = \frac{a}{a+b+c}$ Where, $a = \lvert X \cap Y \rvert$, $b = \left\lvert\frac{X}{Y}\right\rvert$, c= $\left\lvert\frac{Y}{X}\right\rvert$, d= $\lvert \overline{X \cup Y} \rvert$ and X, Y are input and output image intensities
Volume Similarity	$1 - \frac{\lvert\lvert X\rvert - \lvert Y\rvert\rvert}{\lvert X\rvert + \lvert Y\rvert} = 1 - \frac{\lvert b-c\rvert}{2a+b+c}$ Where, $a = \lvert X \cap Y \rvert$, $b = \left\lvert\frac{X}{Y}\right\rvert$, c= $\left\lvert\frac{Y}{X}\right\rvert$, d= $\lvert \overline{X \cup Y} \rvert$ and X, Y are input and output image intensities

Table 2- Segmentation Quality Metrics

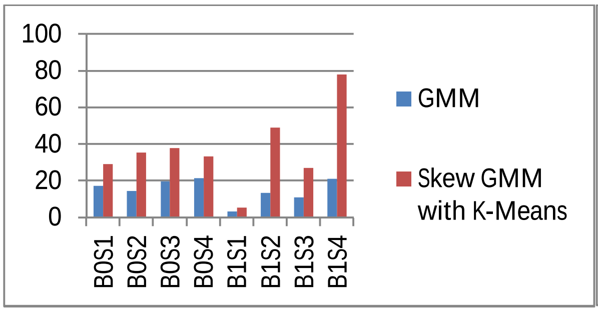
Image	Quality Metric	GMM	Skew GMM with k-Means	Standard Limits	Standard Criteria
B0S1	JC VS	0.089 0.432	0.689 0.733	0 to 1 0 to 1	Close to 1 Close to 1
B0S2	JC VS	0.0677 0.3212	0.7656 0.8767	0 to 1 0 to 1	Close to 1 Close to 1
B0S3	JC VS	0.0434 0.123	0.6567 0.812	0 to 1 0 to 1	Close to 1 Close to 1
B0S4	JC VS	0.0456 0.2233	0.7878 0.3232	0 to 1 0 to 1	Close to 1 Close to 1
B1S1	JC VS	0.141 0.313	0.776 0.397	0 to 1 0 to 1	Close to 1 Close to 1
B1S2	JC VS	0.098 0.04334	0.7892 0.878	0 to 1 0 to 1	Close to 1 Close to 1
B1S3	JC VS	0.0222 0.3223	0.8926 0.3429	0 to 1 0 to 1	Close to 1 Close to 1
B1S4	JC VS	0.455 0.329	0.762 0.7001	0 to 1 0 to 1	Close to 1 Close to 1

Table 3-Image quality metrics

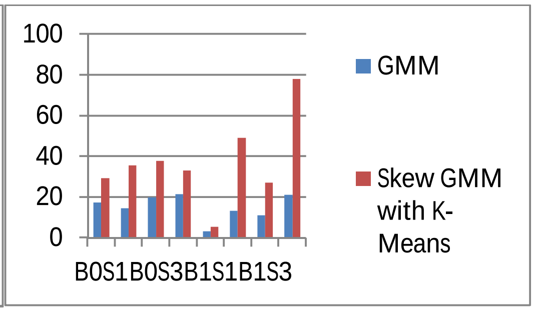
Image	**Quality Metric**	**GMM**	**Skew GMM with K-Means**	**Standard Limits**	**Standard Criteria**
	Average Difference Maximum Distance Image Fidelity Mean Squared error Signal to noise ratio	0.573 0.422 0.416 0.04 17.41	0.773 0.922 0.875 0.134 29.23	-1 to 1 -1 to 1 0 to 1 0 to 1 -∞ to ∞	Closer to 1 Closer to 1 Closer to 1 Closer to 0 As big as Possible
	Average Difference Maximum Distance Image Fidelity Mean Squared error Signal to noise ratio	0.37 0. 221 0.336 0 2404 14.45	0.876 0.897 0.876 0.211 35.65	-1 to 1 -1 to 1 0 to 1 0 to 1 -∞ to ∞	Closer to 1 Closer to 1 Closer to 1 Closer to 0 As big as Possible
	Average Difference Maximum Distance Image Fidelity Mean Squared error Signal to noise ratio	0.456 0.345 0.44 0.22 19.88	0.76 0.879 0.86 0.23 37.98	-1 to 1 -1 to 1 0 to 1 0 to 1 -∞ to ∞	Closer to 1 Closer to 1 Closer to 1 Closer to 0 As big as Possible
	Average Difference Maximum Distance Image Fidelity Mean Squared error Signal to noise ratio	0.231 0. 224 0.212 0.24 21.42	0.473 0.977 0.813 0.121 33.28	-1 to 1 -1 to 1 0 to 1 0 to 1 -∞ to ∞	Closer to 1 Closer to 1 Closer to 1 Closer to 0 As big as Possible
	Average Difference Maximum Distance Image Fidelity Mean Squared error Signal to noise ratio	0.342 0.317 0.391 0.2514 3.241	0.764 0.819 0.812 0.228 5.514	-1 to 1 -1 to 1 0 to 1 0 to 1 -∞ to ∞	Closer to 1 Closer to 1 Closer to 1 Closer to 0 As big as Possible
	Average Difference Maximum Distance Image Fidelity Mean Squared error Signal to noise ratio	0.21 0.21 0.2134 0.06 13.43	0.3653 0.892 0.787 0.145 49.22	-1 to 1 -1 to 1 0 to 1 0 to 1 -∞ to ∞	Closer to 1 Closer to 1 Closer to 1 Closer to 0 As big as Possible
	Average Difference Maximum Distance Image Fidelity Mean Squared error Signal to noise ratio	0.3232 0.123 0.233 0.01 11.11	0.322 0.212 0.897 0.4345 27.267	-1 to 1 -1 to 1 0 to 1 0 to 1 -∞ to ∞	Closer to 1 Closer to 1 Closer to 1 Closer to 0 As big as Possible
	Average Difference Maximum Distance Image Fidelity Mean Squared error Signal to noise ratio	0.314 0.241 0.293 0.18 21.214	0.338 0.249 0.683 0.197 78.19	-1 to 1 -1 to 1 0 to 1 0 to 1 -∞ to ∞	Closer to 1 Closer to 1 Closer to 1 Closer to 0 As big as Possible



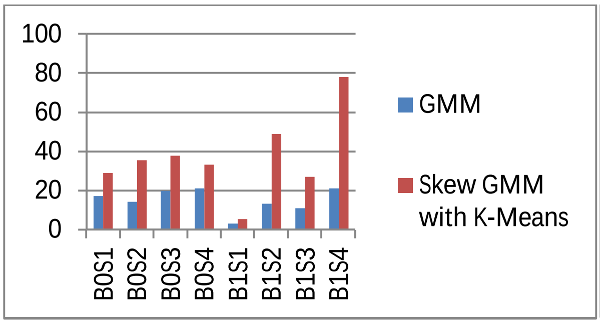
Graph – 1: Graphs for Jaccard Coefficient and Volume Similarity



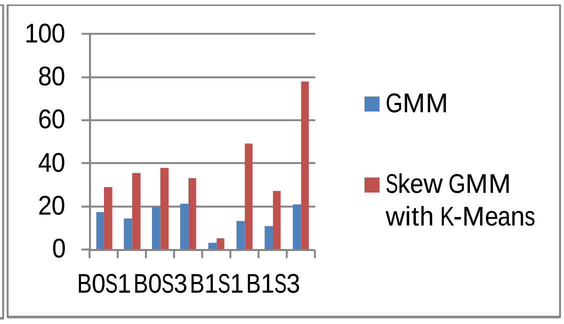
a) Average Difference



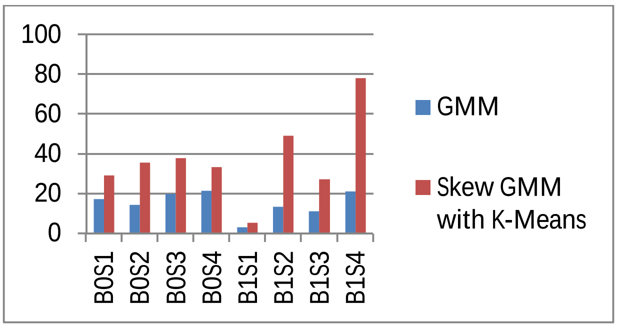
b) Maximum Distance



c) Image Fidelity



d) Mean Squared error



e) Peak Signal to noise ratio

Graph-2: Image quality metrics

Conclusion

In brain medical analysis, segmentation plays a vital role. In particular cases such as Acoustic neuroma, it is assumed that there is a possibility of hearing loss, dizziness and other symptoms related to brain. Some acoustic neuromas can be treated with surgery. Therefore, it is needed to segment the image more accurately, which helps to identify the damaged tissues to be repaired and can be corrected by surgery. Hence, in this paper, a new novel segmentation algorithm based on Skew Gaussian distribution is proposed which helps to identify the tissues more accurately. Due to the basic structure of Skew Gaussian distribution, it is well suited for both symmetric as well as asymmetric distribution. The performance evaluation is carried out by using quality metrics. The results show that, this developed algorithm outperforms the existing algorithm.

References

[1] Pham D. L., Xu C. Y. and Prince J. L. (2000) *Annu. Rev. Biomed.Eng.*, 2, 315–337.

[2] Van Leemput K., Maes F., Vandeurmeulen D. and Suetens P. (1999) *IEEE Trans. Med. Imag.*, 18(10), 897–908.

[3] Dugas-Phocion G., González Ballester M. Á., Malandain G., Lebrun C. and Ayache N. (2004) *Int. Conf. Med. Image Comput. Comput. Assist. Int. (MICCAI)*, 26–33.

[4] Van Leemput K., Maes F., Vandeurmeulen D. and Suetens P. (2003) *IEEE Trans. Med. Imag.*, 22 (1), 105–119.

[5] Prastawa M., Bullitt E., Ho S. and Gerig G. (2003) *Int. Conf. Med. Image Comput. Comput. Assist. Inter (MICCAI)*, *530–537.*

[6] Yamazaki T. (1998) *Introduction of EM algorithm into color Image Segmentation, Proceedings of ICIRS'98, 368-371.*

[7] Srinivas Yarramalle and Srinivas Rao K. (2007) *Current Science, 71 – 84, 2007.*

[8] Jayaram K. Udupa, Vicki R. LeBlanc, Ying Zhuge, Celina Imielinska, Hilary Schmidt, Leanne M. Currie, Bruce E. Hirsch, James Woodburn (2006) *Computerized Medical Imaging and Graphics, 30, 75 – 87.*

[9] Hayit Greenspan, Amit Ruf and Jacob Goldberger (2006) *IEEE Transactions on Medical Imaging, 25(9), 1233 – 1245.*

[10] Gajanayake GMNR., Yapa R.D., Hewawithana B. (2009) *ICIIS, pp. 301 – 305.*

[11] Bouix S. (2007) *Journal of NeuroImage 36,1207 – 1227.*

[12] Rodrigo de Luis-Garcia, Carl-Fredrik Westin, Carlos Alberola-Lopez (2011) *Journal of Computerized Medical Imaging and Graphics 35, 16-30.*

[13] Rahman Farnoosh & Behnam Zarpak (2008) *IUST International Journal of Engineering and Sciences, 19 (1–2), 29–32.*

[14] Eskicioglu, A.M. and Paul S.Fisher (1993) *IEEE Transaction. Commum.*, 43.