AN ENHANCEDMEDICAL IMAGESEGMENTATIONUSING SKEW GAUSSIAN MIXTURE MODEL AND HIERARCHICAL CLUSTERING ALGORITHM

Nagesh Vadaparthi¹, Srinivas Yarramalle², Suresh Varma.P.³

¹ Dept. of I.T., M.V.G.R. College of Engineering, Vizianagaram, India. itsnageshv@gmail.com
² Dept. of IT, GITAM University, Visakhapatnam. India. sriteja.y@gmail.com
³ Aadikavi Nannaya University, Rajahmundry, India. vermaps@yahoo.com

ABSTRACT

Image segmentation is an important process in the field of medical imaging. Recently, much work has been reported in medical image segmentation. Among these techniques, finite Gaussian mixture models are considered to be more recent and accurate. However, in this approach, a number of segments that an image can be divided are taken through apriori and if these segments are not initiated properly it leads to misclassification. Hence, to overcome this disadvantage, we proposed an algorithm for Medical Image Segmentation using Hierarchical Clustering and Skew Gaussian Mixture. The experimentation is done with two different brain images and the results obtained are evaluated using both Image Quality metrics and Segmentation Quality Metrics.

Keywords

Segmentation, finite Gaussian mixture model, Skew Gaussian distribution, Classification, image quality metrics.

1. INTRODUCTION

The research in medical field has geared up tremendously in the recent years. This may be due to the fact that many new diseases related to brain were emerging which needs rigorous research to trigger these diseases. As the research in this direction is spreading, many new models are available in literature to identify these diseases. Among these, model based MRI brain segmentation has gained popularity [1],[2],[3],[4],[5]The brain is surrounded with many other tissues apart from White Matter (WM), Gray Matter (GM) and Cerebro Spinal Fluid (CSF), in addition some noise gets embedded into the system as a default at the time of acquisition and inhomogeneity in magnetic fields also aid to considerable changes in medical images.

Some diseases that are related with brain are Parkinson's, Acoustic neuroma, which leads to memory loss and hearing loss. Effective segmentation techniques help to identify these diseases there by driving towards effective treatment. Segmentation is a process of converting inhomogeneous data into homogeneous data. There are many segmentation algorithms in

DOI: 10.5121/ijma.2012.4207

literature consisting of both parametric and non-parametric models. Among these models, parametric modeling, in particular, model based on finite Gaussian mixture model has gained popularity [2], [3]. This is due to the fact that, of the basic assumption that every image considered in reality follow a bell shaped distribution [6]. But the brain structures are non-rigid, complex in shape and vary from person to person[7]. Segmentingbrain images is a challenging task and Gaussian mixture model models are not well suited due to the different structures of the brain. Hence, it is necessary to develop new algorithms which help in segmenting brain images more efficiently and effectively.

Hence, in this paper Skew Gaussian mixture model is proposed to cater the non-uniform patterns of the brain structure. It also includes Gaussian Mixture model as a limiting case. This paper is organized as follows: In section -2, Hierarchical clustering is briefed and Section -3 describes about the developed model. In section -4, the initialization parameters are given and in section -5, the segmentation algorithm is presented. Section -6 deals with the experimentation that is carried out by using four types of brain images and the performance evaluation that is carried out is discussed in section -7 and in section -8 conclusions are presented.

2. HIERARCHICAL CLUSTERING

Clustering aims at partitioning the data without using the training data, hence, they are called unsupervised models. Clustering is defined as a technique where the objects of interest with similarity along the dimension of interest are kept close and the other objects are apart. The dimension of interest depends on the application [8].

A Hierarchical Clustering goes one step further by collecting similar clusters at different levels into a single cluster by forming a tree which gives better selection of clusters for further exploration and hence, in this method Hierarchical Clustering is utilized.

Given a set of N items to be segmented and an $M \times N$ distance (or similarity) matrix, the basic process of hierarchical segmenting is as follows.

- (1) First, assign each item to a segment, so that if we have N items, it implies that we have N segments, each containing just one item. Let the distances (similarities) between the segments be the same as those (similarities) between the items they contain.
- (2) Find the closest (most similar) pair of segments and merge them into a single segment, i.e. we will now have one segment less.
- (3) Compute distances (similarities) between the new segment and each of the old segments.
- (4) Repeat steps 2 and 3 until all items are segmented into a single segment of size N.

Step 3 can be done using single-linkage method. In single-linkage segmenting (also called the connectedness or minimum method), we consider the between one segment and another to be equal to the shortest distance from any member of one segment to any member of the other segment. If the data consist of similarities, we consider the similarity between one segment and another to be equal to the greatest similarity from any member of one segment to any member of the other segment. The M × N proximity matrix is D = [d(i, j)]. The segmenting is assigned sequence numbers 0, 1..., (n – 1) and L(k) is the level of the kth segmenting. A segment with sequence number m is denoted as (m) and the proximity between segments (r) and (s) is denoted as d [(r), (s)]. The algorithm is composed of the following steps:

- (1) Start with the disjoint segments having level L(0) = 0 and sequence number m = 0.
- (2) Find the least dissimilar pair of segments in the current s, say pair (r), (s), where the minimum is over all pairs of segments in the current segmenting.

- (3) 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)].
- (4) 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 the old segment (k) is defined as d[(k), (r, s)] = min (d[(k), (r)], d[(k), (s)]).
- (5) If all objects are in one segment, stop. Else, go to step.2

3. 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,
$$\Phi(\propto z) = \int_{-\infty}^{\infty z} \phi(t) dt$$
. (2)

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

Let,
$$y = \mu + \sigma z$$
.
 $z = \frac{y - \mu}{\sigma}$ (4)

Substituting equations 2, 3 and 4 in equation 1,



Fig. 1.Skew Normal Distributions

4. 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 Hierarchical Clustering algorithm as proposed in section-2. 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 Hierarchical Clustering algorithm discussed in section-2.

5. SEGMENTATION ALGORITHM

After obtaining the initial 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_{ii}.
- Step 2: Obtain the number of regions by Hierarchical Clustering 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 $\alpha_i=1/k$ is the initial estimate for α_i .
- Step 4: Implement the segmentation and retrieval algorithm by considering maximum likelihood estimate.
- Step 5: With the step 4 obtain the image quality metric.

6. EXPERIMENTATION

The above developed method is applied on 8 sub-images of 2 different brain images obtained from brain web images. The input images and their corresponding histograms are presented in Figure-2a and Figure-2b. To evaluate our algorithm, we have used both T_1 and T_2 type images. In order to initialize the Hierarchical Clustering, we have used the histogram of the image as input, and basing on the peaks we have obtained the initial estimates of k. The experimentation is carried out in MATLAB environment by considering the images of fixed size. We have used both T_1 and T_2 images, where the fat decomposition varies in both these images. The segmentation process is carried out by using the segmentation algorithm presented in Section-5.The outputs images obtained after performing the segmentation are presented in figurea-3a and figure-3b. After segmenting, the performances of the segmentation results were evaluated by using both Segmentation quality metrics and Image quality metrics.



The International Journal of Multimedia & Its Applications (IJMA) Vol.4, No.2, April 2012

7. **PERFORMANCE EVALUATION**

In order to evaluate the proposed algorithm, we have used both T_1 weighted images where water is shown as darker and fat as brighter, T_2 images where fat is shown as darker and Whiter 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 Hierarchical Clustering algorithm. The experimentation is carried out using segmentation algorithm given in section -5 and the results obtained are evaluated using Segmentation quality metrics such as Jaccard quotient (JC), Volume Similarity (VS), Global Consistency Error (GCE), Probabilistic Rand Index (PRI) and Variation of Information (VOI). The formulas for computing the above quality metrics are as follows

Jaccord Coefficient (JC) =
$$\frac{|X \cap Y|}{|X \cup Y|} = \frac{a}{a+b+c}$$

(6)
Volume Similarity (*VS*) = $1 - \frac{||X| - |Y||}{|X| + |Y|} = 1 - \frac{|b-c|}{2a+b+c}$
(7)

Where, $a = |X \cap Y|$, $b = \left|\frac{X}{Y}\right|$, $c = \left|\frac{Y}{X}\right|$, $d = |\overline{X \cup Y}|$ and X, Y are input and output image intensities

$$GCE(S, S') = \frac{1}{N} \min\{\sum LRE(S, S', x_i), \sum LRE(S', S, x_i)\}$$
(8)

Where, LRE = $\frac{|C(S,x_i) \setminus C(S',x_i)|}{|C(S,x_i)|}$, S and S' are segment classes and x_i is the pixel.

VOI
$$(X,Y)$$
= H (X) = H (Y) – 2I $(X;Y)$
(9)

Where, X and Y are two clusters

$$PRI(S_t, \{S\}) = \frac{1}{\binom{N}{2}} \sum_{i,j,i < j} \left[I\left(l_i^{S_t} = l_j^{S_t}\right) p_j + I\left(l_i^{S_t} \neq l_j^{S_t}\right) (1 - p_j) \right]$$
(10)

Where, $p_j = P(l_i = l_j) = \frac{1}{K} \sum_{k=1}^{K} I(l_i^k = l_j^k)$ and the values range from 0 to 1. 1 denotes the segments are identical.

The results of the quality metrics are tabulated in Table-1.

	Tuoro Trooghionitation Trootion					
Image	Quality Metric	GMM	Skew GMM with k- Means- EM	Skew GMM with Hierar. Clustering- EM	Standard Limits	Standard Critrial
	JC	0.089	0.689	0.703	0 to 1	Close to 1
	VS	0.432	0.733	0.8799	0 to 1	Close to 1
B0S1	VOI	2.3665	5.3173	5.142	$-\infty$ to ∞	As big as Possible
B0S2	GCE	0.2802	0.5964	0.561	0 to 1	Close to 1
	PRI	0.504	0.6396	0.619	0 to 1	Close to 1
	JC	0.0677	0.7656	0.7921	0 to 1	Close to 1
	VS	0.3212	0.8767	0.8814	0 to 1	Close to 1
B0S2	VOI	1.9724	3.924	4.35	$-\infty$ to ∞	As big as Possible
	GCE	0.2443	0.4741	0.419	0 to 1	Close to 1
	PRI	0.416	0.5016	0.514	0 to 1	Close to 1
	JC	0.0434	0.6567	0.7143	0 to 1	Close to 1
	VS	0.123	0.812	0.916	0 to 1	Close to 1
B0S3	VOI	0.7684	0.2916	1.659	$-\infty$ to ∞	As big as Possible
	GCE	0.089	0.031	0.107	0 to 1	Close to 1
	PRI	0.576	0.5853	0.632	0 to 1	Close to 1
	JC	0.0456	0.7878	0.874	0 to 1	Close to 1
	VS	0.2233	0.3232	0.54	0 to 1	Close to 1
B0S4	VOI	1.268	1.569	3.354	$-\infty$ to ∞	As big as Possible
	GCE	0.056	0.091	0.157	0 to 1	Close to 1
	PRI	0.189	0.191	0.496	0 to 1	Close to 1
	JC	0.141	0.776	0.791	0 to 1	Close to 1
	VS	0.313	0.397	0.784	0 to 1	Close to 1
B1S1	VOI	1.6499	4.0874	3.951	$-\infty$ to ∞	As big as Possible
	GCE	0.1874	0.4487	0.418	0 to 1	Close to 1
	PRI	0.9256	0.6678	0.6258	0 to 1	Close to 1

Table - 1: Segmentation Metrics

B1S2	JC	0.098	0.7892	0.877	0 to 1	Close to 1
	VS	0.04334	0.878	0.881	0 to 1	Close to 1
	VOI	2.3215	2.8047	3.91	$-\infty$ to ∞	As big as Possible
	GCE	0.2838	0.3407	0.339	0 to 1	Close to 1
	PRI	0.3807	0.369	0.485	0 to 1	Close to 1
	JC	0.0222	0.8926	0.9124	0 to 1	Close to 1
	VS	0.3223	0.3429	0.3543	0 to 1	Close to 1
B1S3	VOI	1.2411	0.9988	2.665	$-\infty$ to ∞	As big as Possible
	GCE	0.1466	0.1157	0.398	0 to 1	Close to 1
	PRI	0.9576	0.9662	0.652	0 to 1	Close to 1
B1S4	JC	0.455	0.762	0.815	0 to 1	Close to 1
	VS	0.329	0.7001	0.7158	0 to 1	Close to 1
	VOI	-8.8e-16	0	0.19	$-\infty$ to ∞	As big as Possible
	GCE	0	0	0.212	0 to 1	Close to 1
	PRI	1	1	0.27	0 to 1	Close to 1

The International Journal of Multimedia & Its Applications (IJMA) Vol.4, No.2, April 2012





(C) Variation of Information

(D) Global Consistency Error



Graph – 1: Comparison of Image Segmentation Techniques

The reconstruction is carried out by ascribing each pixel to its appropriate position by using a method of random number generation. 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; hence the performance of the retrieved image is carried out by using image quality metrics such as Average difference, Maximum distance, Image Fidelity, Mean Squared Error and Signal – to – Noise Ratio. The formulas for evaluating the image quality metrics are tabulated below in Table-1.

Table – 1: Formulas for Image Quality Metrics						
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 columns					
Maximum Distance	$\max\{ F(j,k) - \hat{F}(j,k) \}$					
Image Fidelity	$1 - \left[\sum_{j=1}^{M} \sum_{k=1}^{N} \left[F(j,k) - \hat{F}(j,k)\right]^2 / \sum_{j=1}^{M} \sum_{k=1}^{N} \left[F(j,k)\right]^2\right]$ Where M,N are image matrix rows and columns					
Mean Squared error	$\frac{1}{MN}\sum_{j=1}^{M}\sum_{k=1}^{N}\left[O\{F(j,k)\} - O\{\widehat{F}(j,k)\}\right]^2 / \sum_{j=1}^{M}\sum_{k=1}^{N}\left[O\{F(j,k)^2\right]^2$ Where M,N are image matrix rows and columns					
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					

The developed algorithm is compared with K-Means algorithm and the results obtained are tabulated and presented in Table -2.

Image	Quality Metric	GMM	Skew GMM with K-Means	Skew GMM with hierarchical clustering	Standard Limits	Standard Criteria
	Average Difference	0.573	0.773	0.812	-1 to 1	Closer to 1
	Maximum Distance	0.422	0.922	0.9325	-1 to 1	Closer to 1
	Image Fidelity	0.416	0.875	0.923	0 to 1	Closer to 1
	Mean Squared error	0.04	0.134	0.094	0 to 1	Closer to 0
	Signal to noise ratio	17.41	29.23	33.89	$-\infty$ to ∞	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	0.749 0.912 0.859 0.2019 39.85	$\begin{array}{c} -1 & \text{to } 1 \\ -1 & \text{to } 1 \\ 0 & \text{to } 1 \\ 0 & \text{to } 1 \\ -\infty & \text{to } \infty \end{array}$	Closer to 1 Closer to 1 Closer to 1 Closer to 0 As big as Possible
	Average Difference	0.456	0.76	0.81	-1 to 1	Closer to 1
	Maximum Distance	0.345	0.879	0.807	-1 to 1	Closer to 1
	Image Fidelity	0.44	0.86	0.917	0 to 1	Closer to 1
	Mean Squared error	0.22	0.23	0.2123	0 to 1	Closer to 0
	Signal to noise ratio	19.88	37.98	39.71	-∞ to ∞	As big as Possible
	Average Difference	0.231	0.473	0.4991	-1 to 1	Closer to 1
	Maximum Distance	0.224	0.977	0.971	-1 to 1	Closer to 1
	Image Fidelity	0.212	0.813	0.892	0 to 1	Closer to 1
	Mean Squared error	0.24	0.121	0.1192	0 to 1	Closer to 0
	Signal to noise ratio	21.42	33.28	37.41	$-\infty$ to ∞	As big as Possible

Table – 2: Evaluated Image Quality Metrics

Average Difference	0.342	0.764	0.7015	-1 to 1	Closer to 1
Maximum Distance	0.317	0.819	0.854	-1 to 1	Closer to 1
Image Fidelity	0.391	0.812	0.876	0 to 1	Closer to 1
Mean Squared error	0.2514	0.228	0.1759	0 to 1	Closer to 0
Signal to noise ratio	3.241	5.514	5.68	$-\infty$ to ∞	As big as Possible
Average Difference	0.21	0.3653	0.232	-1 to 1	Closer to 1
Maximum Distance	0.21	0.892	0.912	-1 to 1	Closer to 1
Image Fidelity	0.2134	0.787	0.791	0 to 1	Closer to 1
Mean Squared error	0.06	0.145	0.594	0 to 1	Closer to 0
Signal to noise ratio	13.43	49.22	20.39	$-\infty$ to ∞	As big as Possible
Average Difference	0.3232	0.322	0.4592	-1 to 1	Closer to 1
Maximum Distance	0.123	0.212	0.456	-1 to 1	Closer to 1
Image Fidelity	0.233	0.897	0.923	0 to 1	Closer to 1
Mean Squared error	0.01	0.4345	0.119	0 to 1	Closer to 0
Signal to noise ratio	11.11	27.267	29.86	-∞ to ∞	As big as Possible
Average Difference	0.314	0.338	0.497	-1 to 1	Closer to 1
Maximum Distance	0.241	0.249	0.317	-1 to 1	Closer to 1
Image Fidelity	0.293	0.683	0.791	0 to 1	Closer to 1
Mean Squared error	0.18	0.197	0.213	0 to 1	Closer to 0
Signal to noise ratio	21.214	78.19	99	-∞ to ∞	As big as Possible

The International Journal of Multimedia & Its Applications (IJMA) Vol.4, No.2, April 2012





Figure C: Image Fidelity

Figure D: Mean Squared Error

The International Journal of Multimedia & Its Applications (IJMA) Vol.4, No.2, April 2012



Figure E: Signal – to – Noise Ratio Graphs – 2: Comparison of Techniques

From the above table-1, table-2, Graphs-1, Graphs-2, figures-2 and figures-3 it can be easily seen that the developed methods outperforms the existing method based on Medical image Segmentation based on Gaussian Mixture model and Skew Gaussian Model based on K-Means algorithm. The performance of the developed method outperforms the Gaussian method due to the fact that the segmentation process in Gaussian Mixture model always assume that the intensities inside the medical images are bell shaped. But the shape of the images is defined on the body structure. Hence, it is necessary to go for better models. Hence, this paper mainly focuses on this disadvantage.

8. 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. This model is well suited in particular for medical image, where the shape of the image depends on the body structure. The performance evaluation is carried out by using quality metrics. The results show that, this developed algorithm outperforms the existing algorithm.

REFERENCES:

- D. L. Pham, C. Y. Xu, and J. L. Prince: A survey of current methods in medical image segmentation. Annu. Rev. Biomed. Eng. vol. 2, pp. 315–337 (2000).
- [2] K. Van Leemput, F. Maes, D. Vandeurmeulen, and P. Suetens: Automated model-based tissue classification of MR images of the brain.IEEE Trans. Med. Imag. vol. 18, no. 10, pp. 897–908 (Oct. 1999).
- [3] G. Dugas-Phocion, M. Á. González Ballester, G. Malandain, C. Lebrun, and N. Ayache: Improved EM-based tissue segmentation and partial volume effect quantification in multi-sequence brain MRI.Int. Conf. Med. Image Comput. Comput. Assist. Int. (MICCAI). pp. 26–33, (2004)
- [4] K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens: A unifyingframework for partial volume segmentation of brainMRimages.IEEE Trans. Med. Imag. vol. 22, no. 1. pp. 105–119 (Jan. 2003).

- [5] M. Prastawa, E. Bullitt, S. Ho, and G. Gerig: Robust estimation for brain tumor segmentation. Int. Conf. Med. Image Comput. Comput. Assist. Inter (MICCAI). pp. 530–537 (2003).
- [6] Yamazaki T.: Introduction of EM algorithm into color Image Segmentation.Proceedings of ICIRS'98, pp. 368-371. (1998)
- [7] Srinivas Yarramalle, K. Srinivas Rao: Unsupervised image segmentation using finite doubly truncated Gaussian mixture model and hierarchical clustering. CURRENT SCIENCE, pp. 71 84. (2007).
- [8] Sanjuo.K.Bhatia, "Hierarchical clustering for image databases", 2005, IEEE Explorer.
- [9] GMNR. Gajanayake, et. al., "Comparison of standard image segmentation methods for segmentation of brain tumors from 2D MR Images", *ICIIS 2009, pp. 301 305.*
- [10] S Bouix, et al., "On evaluating brain tissue classifiers without a ground truth", Journal of NeuroImaging 36(2007), pp. 1207 – 1227. (Elsevier).
- [11] Eskicioglu, A.M, et al "Image Quality measures and their performance", IEEE Transaction. Commum. 1993,43.
- [12] Karandeep Singh Chawla and P.K.Bora, "PMM based segmentation of Gray scale images" IEEE, 2009
- [13] Carey E.Priebe, Michael I Miller and J.T.Ratnanather, "Segmenting magnetic resonance images via hierarchical mixture modeling", Comput. Stat. Data Anal., 2006-50(2), pp.551-567.