Automated Skull Stripping Method using Clustering and Histogram Analysis for MRI Human Head Scans

[I]K.Somasundaram, [II]R.Siva Shankar
[I,II]Image Processing Lab, Dept. of Computer Science and Applications, Gandhigram Rural Institute-Deemed University, Gandhigram-624302, Tamil Nadu, India.

Abstract
Magnetic Resonance Images (MRI) are used to analyse the human organs without surgery. The analysis requires medical image segmentation techniques. Brain image segmentation is one of the most important tools and hence need by the clinicians. However, an accurate segmentation is a significant task and crucial for exact diagnosis. In this paper, we have proposed automated method for segmenting brain from T1 weighted MR images. Initially, Otsu thresholding technique is used to find the threshold value in order to eliminate low intensity pixels such as air and CSF from the image. K-Mean clustering technique is used to classify the image into three parts such as brain tissues, non-brain tissues and background. To eliminate non-brain pixels, we have analysed histogram of the image and finally Largest Connected Component is used to segment the brain.

Keywords
MRI Segmentation, OTSU Thresholding, K-Mean Clustering, T1 Weighted MRI, Histogram

I. Introduction

Magnetic Resonance Imaging and other medical images contain complicated anatomical structures that require precise and most accurate segmentation for clinical diagnosis [1]. Brain image segmentation from MRI images providing precise and accurate diagnostic information about spatial relationships between critical anatomical structures such as cortical areas. It is essential for pre-processing step in neuro-imaging applications. Many semi-automatic [2] - [4] and automatic brain MRI segmentation techniques [9] and [13] are available to replace the time consuming manual segmentation. Semi-automatic segmentation method includes region-based and boundary- based techniques. Region-based techniques [5] provide fast segmentation by assigning memberships to voxels according to homogeneity values, but sometimes the in-homogeneity voxels can result inaccurate segmentation. Boundary – based techniques [6] - [8] works well for low noise images. but these techniques are unreliable since image noise and low contrast edges between brain structures can result false or non-existent boundaries causing under or over segmentation[6] [7]. Fully automatic segmentation techniques require no human intervention and provide completely reproducible results for same data. Atkins et al.,[10] proposed a full automatic segmentation of the brain in MRI. In this method a rough brain is produced by an anisotropic operation and thresholding. Then it uses an active contour algorithm to produce accurate result. Sometimes it failed on dataset containing abnormal anatomic structures, extremely high noise or poor contrast. Another fully automatic segmentation of brain from T1-weighted MRI using bridge burner algorithm is proposed by Artem Mikheev et al., [11]. This algorithm is based on thresholding, connectivity, surface detection and a new operator of constraint growing. Automated brain segmentation algorithms provide acceptable accuracy. Such an automated segmentation algorithm [12] is proposed by Zu .Shan et al. This algorithm is based on Histogram analysis of the T1 weighted 3D head volume MR images. It consists of thresholding and simple morphological operations such as dilation, erosion and connectivity process. Histogram scale- space analysis and mathematical morphology techniques jointly used in T1-weigted brain MRI by J.F.Mangin et al.,[15]. They have analysed the crossings in scale-space of trajectories of extrema of different derivative orders follow regular topological properties. These properties allow designing a new structural representation of a 1D signal. Then they have proposed a heuristics using this representation to infer statistics on grey and white matter grey level values from the histogram. These statistics are used by an improved morphological process combining two opening sizes to segment the brain.

Unsupervised segmentation techniques are pixel-based methods as well. K-Mean Clustering [19]-[22] plays an important role in medical image segmentation and becomes a vital technique in pixel-based classification method. H.P.Ng et al., have make use of K-Mean clustering to produce primary segmentation of the input image, in their improved watershed segmentation algorithm [16] in order to achieve the objective of over segmenting in MR Head images. Many thresholding techniques [27]-[30] used the criterion-based concept to select the most suitable gray scale as the threshold value and it classify the image regions into foreground and background. One of the popular methods is Otsu's thresholding method [17] that employs discriminate analysis to find the maximum separability of classes. For every possibility of threshold value, Otsu evaluated the goodness of this value if used as the threshold. This evaluation incorporates the heterogeneity of both classes and the homogeneity of every class. Thresholded images results binary images. In such images, background being labelled b zero values and region represented by one value. The labelling of connected components [18] [19] in binary image is essential to many automated image analysis applications. This technique is very constructive to extract the Region of Interest (ROI) from MRI.

In this paper we have proposed a brain extraction method using clustering and histogram analysis. In section 2, we have given the basic principles used in our method. Our proposed work is presented in detail in section 3. Results generated from the proposed method are discussed in section 4. Finally, section 5 draws the conclusion.

Vol. 2, Issue 3 (July - Sept. 2014)

ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

II. Basic Principles used in Brain Segmentation

(i) Otsu Thresholding

Otsu's thresholding method is used to find the threshold value T which minimizes the intra-class variance (within class variance) or maximizes the inter-class variance (between – class variance) to separate the input image into two classes to produce the binary image. The intra-class variance is defined as the weighted sum of variances of each class.

$$\sigma^2_{within}(T) = n_B(T)\sigma^2_B(T) + n_O(T)\sigma^2_O(T) \quad (1)$$

where, class probability of background

$$n_B(T) = \sum_{i=0}^{T-1} p(i)$$

class probability of object

$$n_O(T) = \sum_{i=T}^{N-1} p(i)$$

$\sigma^2_B(T)$ = the variance of pixels in the background separated by the threshold T.

$\sigma^2_O(T)$ = the variance of the pixels in the object separated by the threshold T.

Subtracting the with-in class variance from the total variance, we can obtain between-class variance:

$$\sigma^2_{Between}(T) = \sigma^2 - \sigma^2_{Within}(T) = n_B(T)n_O(T)[\mu_1(T) - \mu_2(T)]^2 \quad (2)$$

Where σ^2 is the combined variance, where μ_1 and μ_2 are class means.

(ii) K-Mean Clustering

The purpose of k-means algorithm is to cluster the data. K-means algorithm is one of the simplest partitions clustering method.

K-Means is the one of the unsupervised learning algorithm for clusters. Clustering the image is grouping the pixels according to the some characteristics. In the k-means algorithm initially we have to define the number of clusters k. Then k-cluster center are chosen randomly. The distance between the each pixel to each cluster centers are calculated. The distance may be of simple Euclidean function. Single pixel is compared to all cluster centers using the distance formula. The pixel is moved to particular cluster which has shortest distance among all. Then the centroid is re-estimated. Again each pixel is compared to all centroids. The process continuous until the center converges. Let X_i be the data set, where i = 1, 2,...,n.

1. Set the cluster size as k.
2. Initialize the centroid of each cluster Ci=0, i=1,2...k.
3. Process the observations and assign to a cluster.
4. Find the centroids of each cluster.

The centroid of a cluster is found by minimizing the objective function

$$J = \sum_{j=1}^{k} \sum_{i-1}^{n} \left\| x_i^{(j)} - c_j \right\|^2 \quad (3)$$

Where $\left\| x_i^{(j)} - c_j \right\|^2$ is a measure of intensity distance between a data point xi and the cluster center cj. For simplicity, the Euclidean distance is used as the dissimilarity measure.

(iii) Labelling

To identify regions in the image, each region is labelled with a unique number. Such methods are called as labelling. Region labelling using run length encoded data is simple and fast. The algorithm has two passes. In first pass, the runs are analysed for neighbourhood and labelled row-wise.

The label collision issue in the first pass is solved by second pass.

First Pass: Use a new label for each continuous run in the first image row that is not part of the background.

For the second and subsequent rows, compare the positions of runs.

a. If a run in a row does not neighbour (in the 4 -or- 8 sense) any run in the previous row, assign a new label.
b. If a run neighbours precisely one run in the previous row, assign its label to the new run.
c. If the new run neighbours more than one run in the previous row, a label collision has occurred.

Collision information is stored in an equivalence table, and the new run is labelled using the label of any one of its neighbours.

Second Pass: Search the image row by row and re-labelled the image according to the equivalence table.

(iv) Largest Connected Component (LCC)

The run length identification scheme for region labelling described by Sonka et al. [24] is used to find the LCC among the regions as:

$$R_{LCC} = R(arg\ max\ R_A(i)) \quad (4)$$

where, the area $R_A(i)$ of ith region R(i) is the total number of pixels in that region.

III. Proposed Method

A. Proposed Method

The proposed algorithm consists of four steps. The first step is removing low-intensity pixels by applying an image intensity threshold to separate the background noise and foreground head tissues. The second step is to cluster the pixels into brain, non-brain and background using K-mean method. The third step is to eliminate non-brain regions using histogram analysis. The fourth step is to find largest connected component in order to obtain brain image separately.

Step 1: Foreground and Background Separation

We have used Otsu's thresholding technique to find the threshold value T of the input image. Any pixels with intensity lower than Totsu have been removed from the images. All the pixels which satisfy the condition in eqn. (5) were set to 0. In MRI brain image, these pixels included air and low-intensity components such as CSF.

$$I_T = \begin{cases} 0 & if\ I(x,y) \le T \\ I(x,y) & otherwise \end{cases} \quad (5)$$

Step 2: Cluster Brain, Non-brain and Background Pixels

The K-mean algorithm clusters the image according to similar intensity values. We classify the image by clustering intensities into three groups with k=3.The resultant image I_K consists of similar regions such as brain, non-brain and background pixels.

Step 3: Removing Non-brain regions using Histogram analysis.

The brain is connected to the skull and sub head/ neck tissues. However, these connections have lower intensity than GM, some portions posses a similar intensity to GM. In order to disconnect the brain region from these non-brain regions, we have analysed histogram of the image (Fig 1). The first peak shows the intensity range of GM. This value is taken as a rough estimate of multiplication of Otsu's threshold value T by two. A binary image I_B is obtained after applying the eqn. (6) on the image I_K.

$$I_B = \begin{cases} 0 & \text{if } I_K \leq T * 2 \\ 1 & otherwise \end{cases} \quad (6)$$

Pixels of I_K with intensities greater than or equal to T*2, were set to 0, otherwise set to 1.

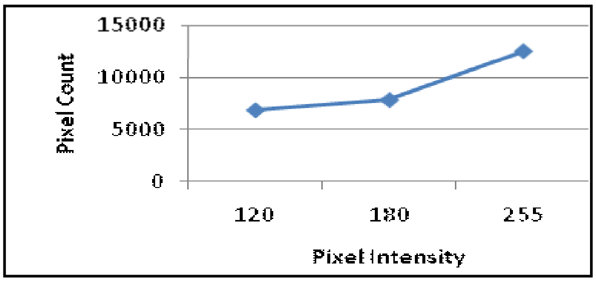


Fig.1: Histogram of a sample slice

Step 4: Segmenting Brain mask
Assuming brain tissue is the largest connected component in head MR image volume, we have used largest connected component analysis. Prior, we have labelled the components in I_B using region labelling and run length identification scheme is used to find LCC in the Image.

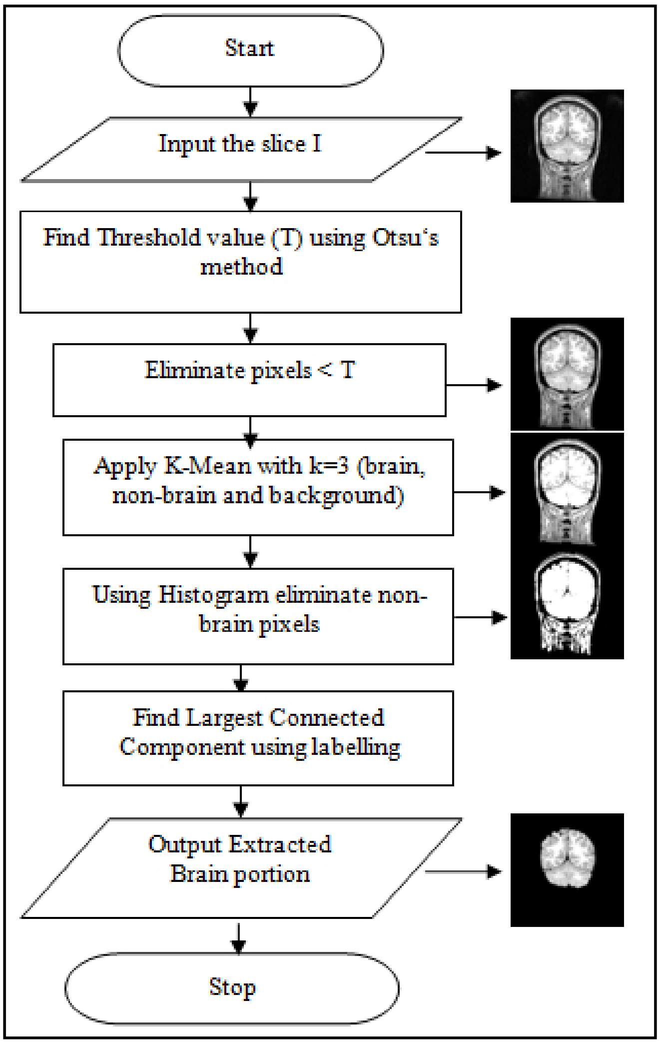


Fig. 2 : Flow Chart of Proposed Method

B. Materials Used

We have used twenty volumes of MRI T1 coronal datasets obtained from IBSR [23] website developed by Centre for Morphometric Analysis (CMA) at Massachusetts General Hospital, for the proposed method. The thickness of each slice is 3.0mm and its pixel size is 256X256. The manually segmented masks, ground truth or gold standard are also available at IBSR for these twenty volumes.

IV. Results and Discussions

We carried out the experiments by applying our method on twenty T1 weighted coronal MRI of head scans. To evaluate the performance of our method, we computed Jaccard(J) and Dice(D).

The Jaccard coefficient [25] is given by:

$$J(A,B) = \frac{A \cap B}{A \cup B} \quad (7)$$

The Dice coefficient [26] is given by:

$$D(A,B) = \frac{A \cap B}{A \cup B} \quad (8)$$

where A and B are two data sets. The value J as well as D varies from 0 for complete disagreement to 1 for complete agreement, between A and B. The Computed values of Jaccard and Dice Co-efficient for BET, BSE and our method are given below in the tables.

Table 1: Comparison of Jaccard co-efficients.

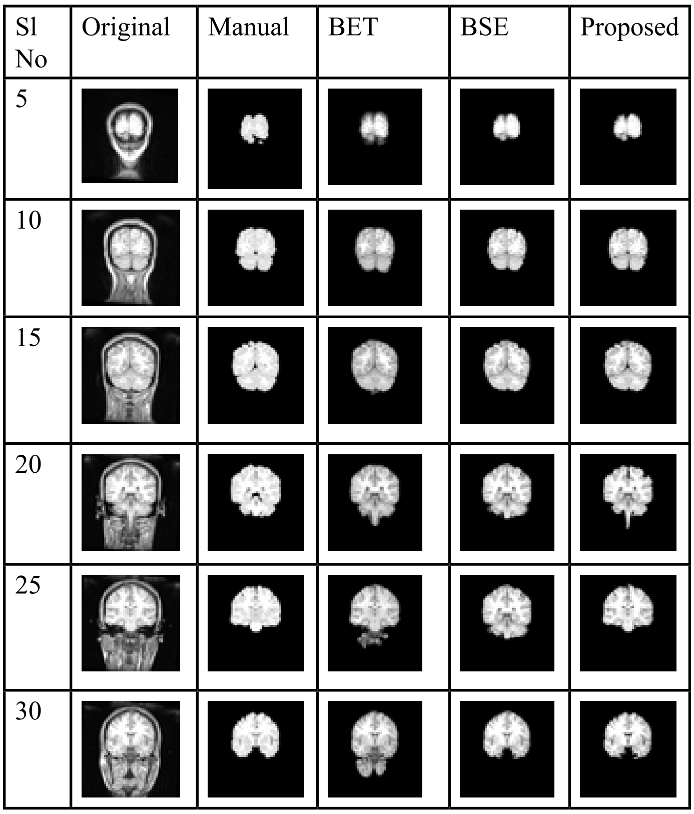
	Jaccard		
Volumes	Proposed	BET	BSE
1_24	0.991385	0.7581	0.8999
2_4	0.988119	0.8078	0.9204
4_8	0.977805	0.8281	0.9053
5_8	0.96558	0.7716	0.7122
6_10	0.96172	0.6666	0.7183
7_8	0.989786	0.6829	0.8912
8_4	0.98731	0.7778	0.9142
11_3	0.986909	0.8527	0.9141
12_3	0.982093	0.7878	0.87
13_3	0.984551	0.8566	0.8885
15_3	0.944404	0.6839	0.9089
16_3	0.987158	0.5401	0.8935
17_3	0.992036	0.5775	0.9084
100_23	0.989539	0.827	0.9175
110_3	0.988642	0.7772	0.9126
111_2	0.989213	0.855	0.915
112_2	0.991184	0.7791	0.9056
191_3	0.992661	0.8455	0.9273
202_3	0.938696	0.8582	0.9186
205_3	0.936754	0.7107	0.9256
Avg.	**0.978277**	**0.76221**	**0.888355**

ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

Table 2: Comparison of Dice co-efficients.

	Dice		
Vol-umes	Proposed	BET	BSE
1_24	0.995672	0.8624	0.9473
2_4	0.994015	0.8937	0.9585
4_8	0.988768	0.9059	0.9503
5_8	0.982388	0.871	0.8319
6_10	0.980249	0.7999	0.836
7_8	0.994861	0.8116	0.9424
8_4	0.993609	0.875	0.9552
11_3	0.993399	0.9205	0.9551
12_3	0.99095	0.8813	0.9305
13_3	0.9922	0.9227	0.9409
15_3	0.971295	0.8123	0.9523
16_3	0.993533	0.7013	0.9437
17_3	0.995999	0.7322	0.952
100_23	0.99473	0.9053	0.9569
110_3	0.994281	0.8746	0.9543
111_2	0.994571	0.9218	0.9556
112_2	0.995567	0.8758	0.9504
191_3	0.996315	0.9162	0.9622
202_3	0.967911	0.9237	0.9576
205_3	0.96642	0.8309	0.9613
Avg.	**0.988837**	**0.861905**	**0.93972**

For visual comparision the normal brain images and the extracted brain portion by BET , BSE and our proposed method are shown in Fig.4.1.



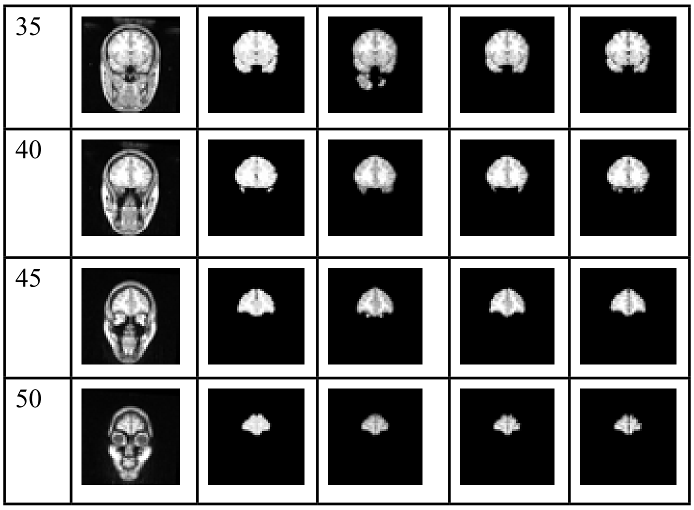


Fig 4.1 Original slices and manually segmented slices are shown in first Column and second column. The extracted brain images by BET, BSE and the proposed method are shown in third column, fourth column and Fifth column respectively.

V. Conclusions

This work proposed an automated brain segmentation method for T1-weighted MR Images. This method is based on clustering and histogram analysis. The quantitative results shows that our method has been produced the segmented results with higher accuracy than the popular methods BET and BSE. The segmented results were confirmed by the neurologist as well.

VI. AcknowledgEment

The authors would like to acknowledge Dr.Karunakaran., MBBS., DMRD., DNB(RD) Radiologist, Department of Imaging and Interventional Radiology, Meenakshi Mission Hospital and Research Centre, Madurai-and Dr.S.P.Balachandran., M.D.,D.M.,(NEURO) Neurologist., Dindigul Neuro Centre, Dindigul, Tamil Nadu, India., for giving suggestions and evaluating our segmented results.

References

[1] *Hall LO, Bensaid AM, Clarke LP, Velthuizen RP, Silbiger MS and Bezdek J," A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain", IEEE Transaction on Neural Networks,Vol. 3, No.5, pp.672–682, 1992.*

[2] *DubeyR.B., Hanmandlu M.,Gupta S.K., and Gupta S.K.,"Semi-automatic Segmentation of MRI Brain Tumor", ICGST-GVIP Journal, ISSN: 1687-398X, Vol.9, Issue 4, pp.33-40, 2009.*

[3] *Huy-Nam Doan,Greg Slabaugh,Gozde Unal and Tong Fang, "Semi-Automatic 3-D Segmentation Of Anatomical Structures of Brain MRI Volumes using Graph Cuts", International Conference on Image Processing (ICIP),pp. 1913-1916,2006.*

[4] *Kevin Karsch, Qing He and Ye Duan, "A Fast, Semi-Automatic Brain Structure Segmentation Algorithm for Magnetic Resonance Imaging", IEEE International Conference on Bioinformatics and Biomedicine, pp.297-302, 2009.*

[5] *Zhu S.C., Lee T.S., and Yuille A.L., "Region competition:Unifying snakes, region growing,and bayes/mdl for multiband image segmentation", International Conference on Computer Vision, pp. 416–423,1995.*

[6] *Kass M., Witkin A., and Terzopoulos D., "Snakes: Active*

contour models", Intl Journal of Computer Vision, Vol. 1, pp.321–331, 1987.

[7] Staib L.H. and Duncan J.S., "Boundary finding with parametrically deformable models", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, No.11, pp. 1061–1075, 1992.

[8] Mcinerney T. and Terzopoulos D., "Topology adaptive deformable surfaces for medical image volume segmentation", IEEE Transactions on Medical Imaging, Vol. 18, pp. 840–850, 1999.

[9] Mohammad Mehdi Khalighi, Hamid Soltanian-Zadeh and Caro Lucas, "Unsupervised MRI segmentation with spatial connectivity", SPIE International Symposium on Medical Imaging, San Diego, CA, pp.1-9, 2002.

[10] Atkins S.M. and Mackiewich B.T., "Fully Automatic Segmentation of the Bain in MRI", IEEE Transactions on Medical Imaging, Vol.17, No.1, pp.98-107, 1998.

[11] Artem Mikheev, Gregory Nevsky, Siddharth Govindan, Robert Grossman and Henry Rusinek, "Fully Automatic Segmentation of the Brain From T1-Weighted MRI Using Bridge Burner Algorithm ", Journal of Magnetic Resonance Imaging, Vol.27, pp.1235-1241, 2008.

[12] Zu Y,Shan, Guang H.ue and Jing Z.Liu, " Automated Histogram –Based Brain Segmentation in T1-Weighted Three-Dimensional Magnetic Resonance Head Images", NeuroImage, Vol.17, pp.1587-1598, 2002.

[13] Bouchaib Cherradi, Omar Bouattane, Mohamed Youssfi and Abdelhadi Raihani, "Fully Automatic Method for 3D T1-Weighted Brain Magnetic Resonance Images Segmentation", International Journal of Image Processing, Vol.5, Issue 2, pp.220-235, 2011.

[14] Lemieux L., Hagemann G., Krakow K. and Woermann F.G., "Fast, accurate, and reproducible automatic segmentation of the Brain in T1-weighted MRI data", Magnetic Resononance . Med.Vol. 42 , pp.127–135, 1999.

[15] Mangin J.F. , Coulon O. and Frouin V.," Robust Brain Segmentation using Histogram Scale – Space analysis and Mathematical Morphology", MICCAI'98, Springer-VerlogBerlin,pp.1230 – 1241,1998.

[16] H.P. Ng, S.H. Ong, K.W.C. Foong, P.S. Goh and W.L. Nowinski, "Medical Image Segmentation using K-Means Clustering and Improved Watershed Algorithm", pp.61-65,Southwest Symposium on Image Analysis and Interpretation, USA.,2006.

[17] Otsu N, "A threshold selection method from gray-level histograms", IEEE Transaction on Systems, Man, and Cybernetics, Vol. SMC- 9, No. 1, pp. 62-66, 1979.

[18] Tina Kapur, Eric W., Grimson L., William M.Wells III and Ron Kikinis, "Segmentation of Brain Tissue from Magnetic Resonance Images", Medical Image Analysis, Vol.1, No.2, pp.109-127, 1996.

[19] Mehdi Jafari and Shohreh Kasaei, "Automatic Brain Tissue Detection in MR Images Using Seeded Region Growing Segmentation and Neural Network Classification", Australian Journal of Basic and Applied Sciences, Vol.5, No.8, pp.1066-1079, 2011.

[20] Mankikar S.S.,"A Novel Hybrid Approach Using Kmeans Clustering and Threshold filter for Brain Tumor Detection",International Journal of Computer Trends and Technology, Vol.4,Issue 3,pp.206-209,2013.

[21] Ali Salem Bin Samma and Rosalina Abdul Salam, "Adaptation of K-Means Algorithm for Image Segmentation ", World Academy of Science, Engineering and Technology, Vol.50, pp.58-62, 2009.

[22] Dhanalakshmi P. and Kanimozhi T., "Automatic Segmentation of Brain Tumor using K-Means Clustering and its Area Calculation",International Journal of Advanced Electrical and Electronics Engineering, Vol.2, Issue 2, pp.130-134, 2013.

[23] International Brain Segmentation Repository, Center for Morphometric Analysis Massachusetts General Hospital, CNY-6, Building 149, 13th Street, Charlestown, MA, 02129-USA. http://www.cma.mgh.harvard.edu/ibsr/ibsr_data/sec3.sub2.html.

[24] Milan Sonka, Vaclav Hlavac and Roger Boyle, "Image Processing, Analysis and Machine Vision", Thomson Learning Inc., Second Edition, 2007.

[25] Jaccard P., "The distribution of the flora in the alpine zone", New Phytol. 11 (2), pp. 37–50, 1912.

[26] Dice L., "Measures of the amount of ecologic association between species", Ecology, Vol. 26, pp: 297- 302, 1945.

[27] Evelin Sujji G.,Lakshmi Y.V.S. and Wiselin Jiji G., "MRI Brain Image Segmentation based on Thresholding", International Journal of Advanced Computer Research, Vol.3, No.8,pp.97-101, 2013.

[28] Matei Mancas, Bernard Gosselin, Benoit Macq, "Segmentation using a region-growing thresholding", Image Processing: Algorithms and Systems, pp.388-398, 2006.

[29] Anam Mustaqeem,Ali Jave and Tehseen Fatima, "An Efficient Brain Tumor Detection Algorithm Using Watershed and Thresholding Based Segmentation",I.J. Image, Graphics and Signal rocessing, Vol.10, pp.34-39, 2012.

[30] Sudipta Roy, Ayan Dey, Kingshuk Chatterjee and Samir K. Bandyopadhyay, "A New Efficient Binarization Method for MRI of Brain Image", Signal & Image Processing : An International Journal (SIPIJ) Vol.3, No.6, pp.35-51, 2012.

Author's Profile and Image

Dr.Somasundaram was born in 1953. He did M.Sc in 1976 in Madras University. He obtained Ph.D in 1984 in Indian Institute of Science (IISC), Bangalore. He got Post graduate Diploma in Computer methods in 1989 from Madurai Kamaraj University, Madurai. He was a visiting research fellow in 1987 in International Centre for Theoretical Physics, Trieste, Italy. He is having 38 years of teaching experience and 30 years of research experience. He is a Member of IEEE Computer Society-USA, ACM-USA, Indian Society for Technical Education, New Delhi, World Enformatika Society and life member of Telemedicine society of India. He has written 4 books in Java. He has published over 130 International journals and conferences. His current areas of research are Image Compression, Image Enhancement and MRI Processing.

ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

Siva Shankar Ramasamy was born in 1984. He did Master of Computer Applications in 2008 in Gandhigram Rural University, Dindigul, Tamil Nadu, India. He is doing research in the field of Medical image segmentation and also working as Assistant Professor in the Department of Computer Science and Applications in the same Institute.