Image Segmentation and Detection of Tumor Objects in MR Brain Images Using FUZZY C-MEANS (FCM) Algorithm

M. Rakesh[1], T. Ravi[2]
[1]M. Tech Project Student, Dept. of ECE, K. L. University, A. P, India
[2]Assoc. Prof., Dept of ECE, K. L. University, A. P, India

Abstract- The brain is a highly specialized organ. It serves as the control center for functions of the body. Words, actions, thoughts, and feelings are centered in the brain. We do, however, know that each part of the brain has a specific, important function, often a profoundly important function, and each part contributes to the healthy functioning of our body. The location of tumors in the brain is one of the factors that determine how a brain tumor effects an individual's functioning and what symptoms the tumor causes. A color based segmentation method that uses the k-means clustering technique is to track the tumor objects in the Magnetic Resonance (MR) brain images. The key concept in color-based segmentation algorithm with K-means is to convert a given gray-level MR image into a color space image and then separate the position of tumor objects from other items of an MR image by using K-means clustering and histogram-clustering. In this paper, the FUZZY C-MEANS algorithm is used for image segmentation and detecting the tumor objects that are found in the MR brain image.

Keywords : segmentation, Magnetic Resonance (MR), Fuzzy C-Means (FCM), K-means Clustering, Histogram Clustering.

I. INTRODUCTION

Imaging is an essential aspect of medical science to visualize the anatomical structures of the human body [3, 4]. Several new complex multidimensional digital images of physiological structures can be processed and manipulated to help visualize hidden diagnostic features that are otherwise difficult or impossible to identify using planar imaging methods. Segmentation is an important process in most medical image analysis and classification for radiological evaluation or computer-aided [3] diagnosis. Basically, image segmentation methods can be classified into three categories: edge based methods, region-based methods [6], and pixel-based methods. K-means clustering is a key technique in pixel based methods.

Because pixel-based methods based on K-means clustering are simple and the computational complexity is relatively low compared with other region-based or edge-based methods, the application is more practicable. Furthermore, K-means clustering is suitable for biomedical image segmentation as the number of clusters is usually known for images of particular regions of the human anatomy. Many researchers have proposed related research into K-means clustering segmentation [1, 5]. The improvements achieved by [1, 5] have been remarkable, but more computational complexity and extra software functionality are required. In this paper, we carefully select the appropriate features from brain images as the clustering features to achieve good segmentation results while maintaining the low computation aspect of the segmentation algorithm. Because the color space transformation function in our proposed method is a fundamental operation for most image processing systems, the color space translation does not cause extra overhead in the proposed scheme. Therefore, by using color-based segmentation with K-means clustering to magnetic resonance (MR) brain tumors, the proposed image tracking method maintains efficiency. The experimental results also confirm that the proposed method helps pathologists distinguish exact lesion sizes and regions.

The rest of this paper is organized as follows. Section II introduces the basic theorem of histogram statistics segmentation and the K-means clustering method and also the segmentation scheme is presented. Section III introduces our proposed method. Experimental results are illustrated and discussed in Section IV. Finally, concluding remarks are given in Section V.

II. REVIEW OF HISTOGRAM CLUSTERING AND K-MEANS CLUSTERING

Two pixel-based segmentation methods are applied in our proposed method. One is histogram statistics and the other is K-means clustering. The histogram method defines single or multiple thresholds to classify an image pixel-by-pixel. A simple approach to determine the gray value threshold *T* is by analyzing the histogram for peak values and finding the lowest point, which is typically located between two consecutive peak values of the histogram. If a histogram is clearly bi-modal, the histogram statistics method can provide good results. By comparing the gray value of each pixel with the determined threshold *T*, a pixel can be classified into one or the two classes. An image *f(x,y)* can be segmented into two classes using a gray value threshold *T* so that

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) > T \\ 0 & \text{if } f(x, y) \le T \end{cases}$$

Where g(x, y) is the segmented image with two classes of binary values, "1" and "0", and *T* is the threshold assigned to the lowest point, which is located between two peak values of the histogram.

K-means is a widely used clustering algorithm to partition data into *k* clusters. Clustering is the process for grouping data points with similar feature vectors into a single cluster and for grouping data points with dissimilar feature vectors into different clusters. Let the feature vectors derived from *l* clustered data be $X=\{x_i \mid i=1,2\ldots, l\}$. The generalized algorithm initiates *k* cluster centroids $C=\{c_j \mid j=1,2,\ldots k\}$ by randomly selecting *k* feature vectors are grouped into *k* clusters using a selected distance measure such as Euclidean distance so that,

$$d=\|x_i-c_j\|$$

The next step is to recompute the cluster centroids based on their group members and then regroup the feature vector according to the new cluster centroids. The clustering procedure stops only when all cluster centroids tend to converge.

Basically, feature space selection is a key issue in K-means clustering segmentation. The original MR brain image is rendered as a gray-level image that is insufficient to support fine features. To obtain more useful feature and enhance the visual density, the proposed method applies pseudo-color transformation, a mapping function that maps a gray-level pixel to a color-level pixel by a lookup table in a predefined color map. An RGB color map contains R, G, and B values for each item. Each gray value maps to an RGB item. The proposed method has adopted the standard RGB color map, which gradually maps gray-level values 0 to 255 into blue-to-green-to-red color.

The flowchart of histogram clustering and the k-means clustering is as shown in the Fig 1.

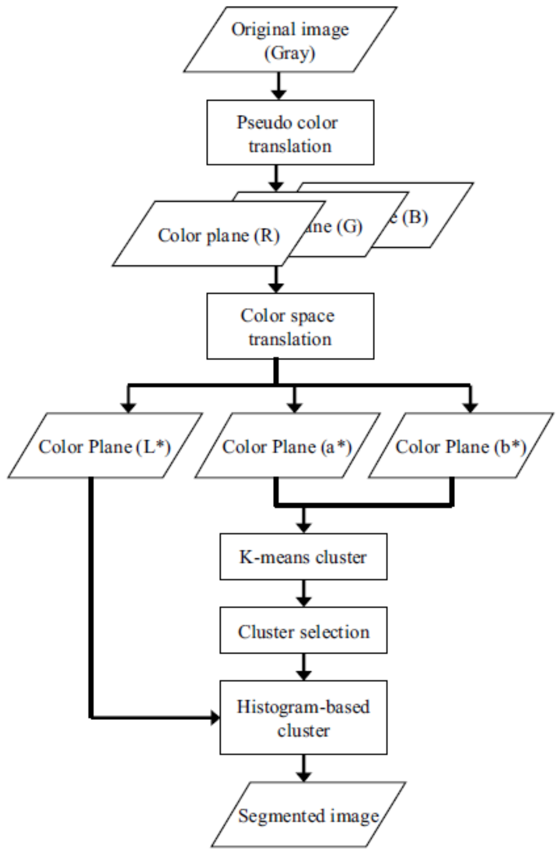

Fig 1: block diagram of histogram and k-means clustering

To retrieve important features to benefit the clustering process, the RGB color space is further converted to a CIELab color model (*L*a*b**) [2]. The *L*a*b** space consists of a luminosity layer *L**, a chromaticity-layer *a**, which indicates where color falls along the red-green axis, and a chromaticity-layer *b**, which indicates where the color falls along the blue-yellow axis. The translating formula calculates the tri-stimulus coefficients first as

$$W=0.4303R + 0.3416G + 0.1784B,$$
$$Y= 0.2219R + 0.7068G + 0.0713B,$$
$$Z= 0.0202R + 0.1296G + 0.9393B.$$

The CIELab color model is calculated as

$$L^* = 116(h(Y/Y_S))-16,$$

$$a^* = 500(h(W/W_s))-h(Y/Y_S)$$

$$b^* = 200(h(Y/Y_S)-h(Z/Z_S)),$$

$$h(q) = \sqrt{q} \qquad q > 0.008856$$

$$= 7.787q + 16/116 \qquad q \leq 0.008856,$$

Where Y_S, W_S, and Z_S are the standard stimulus coefficients.

III. PROPOSED METHOD

In this work, the FCM algorithm is implemented using the data compression technique without including the weight factor in the cluster center updation criterion which further speeds up the process besides yielding considerable segmentation efficiency. The modified FCM algorithm is used for clustering abnormal MR brain images from four tumor classes namely metastate, meningioma, glioma and astrocytoma. Textural features namely correlation, contrast and entropy are extracted from the images and used for the clustering algorithm. The segmented outputs are analyzed based on the segmentation efficiency and convergence rate. A comparative analysis is performed with the conventional FCM algorithm to show the superior nature in terms of convergence rate. Experimental results show promising results for the modified FCM algorithm.

Clustering is one of the widely used image segmentation techniques which classify patterns in such a way that samples of the same group are more similar to one another than samples belonging to different groups [7]. There has been considerable interest recently in the use of fuzzy clustering methods, which retain more information from the original image than hard clustering methods. Fuzzy C-means algorithm is widely allows pixels to belong to multiple classes with varying degrees of membership. But the major operational complaint is that the FCM technique is time consuming [8]. Several modifications have been done on the existing network to improve the performance.

This work consists of the following stages, viz. MR image database, feature extraction, FCM based segmentation and modified FCM based segmentation. The technique for MR brain tumor image segmentation is shown in Fig 2.

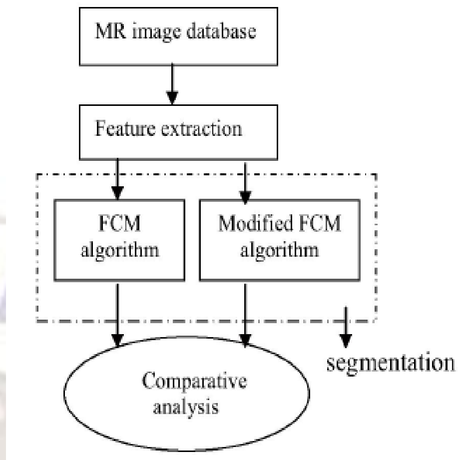

Fig 2 : Image Segmentation Method

A. MR Image Database

A Set of A set of MR brain tumor images comprising of the four tumor types are collected from radiologists. The images used are 256*256 gray level images with intensity value ranges from (0 to 255). Initially, these MRI images are normalized to gray level values from (0 to 1). Since normalization reduces the dynamic range of the intensity values, feature extraction is made much simpler. Some samples have been displayed in Fig 3.

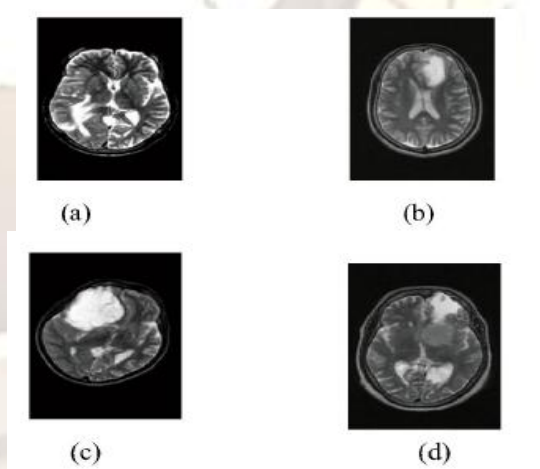

Fig 3: Sample data sat : (a) Metastase (b) Glioma (c) Astrocytoma (d) Meningioma

B. Feature extraction

The purpose of features extraction is to reduce the original data set by measuring certain properties, or features, that distinguish one input pattern from another pattern [9]. The extracted feature should provide the characteristics of the input type to the classifier by considering the description of relevant properties of the image into a feature space. Three textural features namely contrast, correlation and entropy based on the gray level co-occurrence (GLCM) have been used in this work.

Spatial gray level co-occurrence estimates image properties related to second-order statistics. Haarlick analysis [10] suggested the use of gray level co-occurrence matrices (GLCM) which have become one of the most well-known and widely used texture features. GLCM {P $_{(d, \Theta)}$ (i, j)} represents the probability of occurrence of a pair of gray-levels (i,j) separated by a given distance d at angle Θ. The commonly used unit pixel distances and the angles are O^0, 45°, 90° and 135°. A detailed algorithm of calculation of GLCM {P $_{(d, \Theta)}$ (i, j)} has been given in [11]. The features are calculated using the formulae given below.

Contrast:

$$S_c = \sum_i \sum_j (i-j)^2 P(i, j)$$

Correlation :

$$S_O = \frac{\sum_i \sum_j (ij) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$

Where $\mu_x,\mu_y,\sigma_x,\sigma_y$ are the means and standard deviations of p_x and p_y

Entropy :

$$S_E = -\sum p(i,j) \log\{p(i,j)\}$$

Each set of features are individually normalized to the range of 0 to 255. The features used in this paper are selected based on the previous works. These features work well especially for MRI brain tumor images.

C. Conventional FCM Technique

Fuzzy C-means (FCM) is a method of clustering which allows one pixel to belong to two or more clusters. The FCM algorithm attempts to partition a finite collection of pixels into a collection of "C" fuzzy clusters with respect to some given criterion. Depending on the data and the application, different types of similarity measures may be used to identify classes. Some examples of values that can be used as similarity measures include distance, connectivity, and intensity. In this work, the images are segmented into four clusters namely white matter, particular cluster which can be easily extracted. But grey matter, CSF and the abnormal tumor region based on the feature values.

Fuzzy c-means algorithm is based on minimization of following objective function:

$$J(U, c_1, c_2, \ldots, c_c) = \sum_{i=1}^{c} J_i = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{\ m} d_{ij}^{\ 2}$$

U_{ij} is between o and 1;
C_i is the centroids I;
d_{ij} is the Euclidian distance between i_{th} centroid (c_i) and the j^{th} data point.
m ε [1,∞] is a weighting exponent.

Fuzzy partitioning of the known data sample is carried out through an iteration optimization of the objective function shown in the above equation, with update of membership U_{ij} and the cluster centers C_i

D. Modified FCM Technique.

Clustering can also thought of as a form of data compression, where a large number of samples are converted into a small number of representative prototypes or clusters [15]. High dimensional feature space based image segmentation is time intensive than in one dimensional feature spaces. The modified FCM algorithm is based on the concept of compression where the dimensionality of the input is highly reduced. The data compression includes two steps: quantization and aggregation.

The quantization of the feature space is performed by masking the lower 'm' bits of the feature value. The quantized output will result in the common intensity values for more than one feature vector. In the process of aggregation, feature vectors which share common intensity values are grouped together. A representative feature vector is chosen from each group and they are given as input for the conventional FCM algorithm. Once the clustering is complete, the representative feature vector membership values are distributed identically to all members of the quantization level. Since the modified FCM algorithm uses a reduced dataset, the convergence rate is highly improved when compared with the conventional FCM.

IV. EXPERIMENTAL RESULTS

The experimental results of MR brain tumor images are shown as in the following figures.

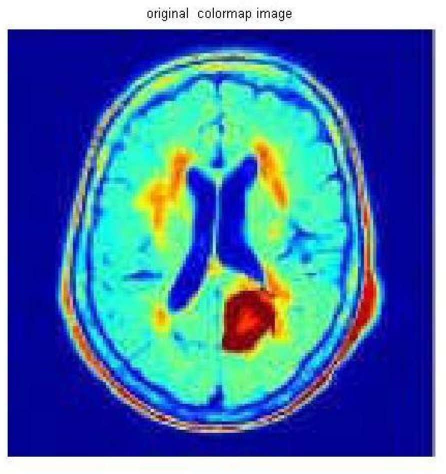

Fig 4 : Original Color Map Image

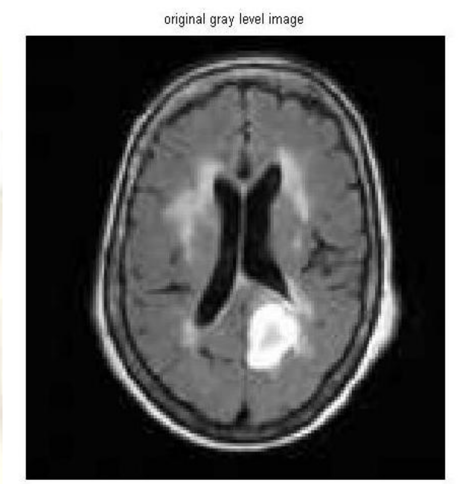

Fig 4: Original Gray level Image

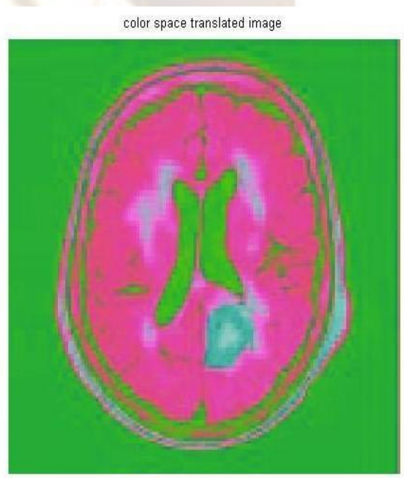

Fig 6 : Color Space Translated Image

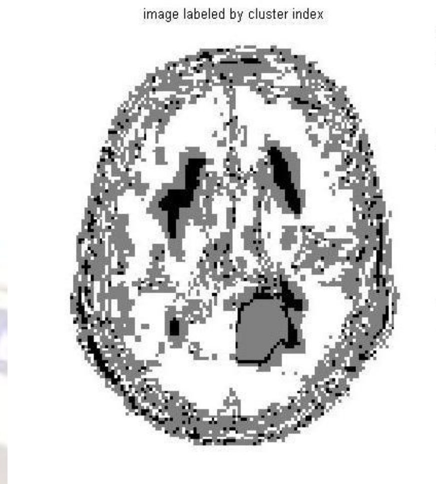

Fig 7 : Image Labeled by Cluster Index

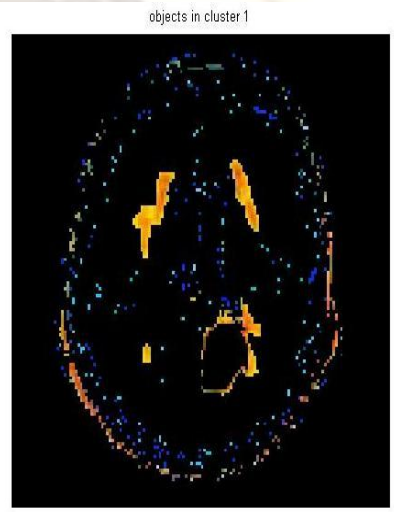

Fig 8 : Objects in cluster 1

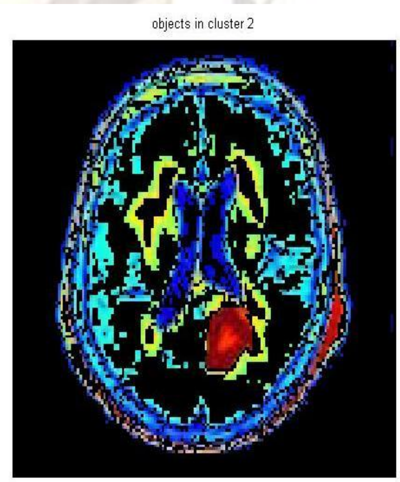

Fig 9 : Objects in Cluster 2

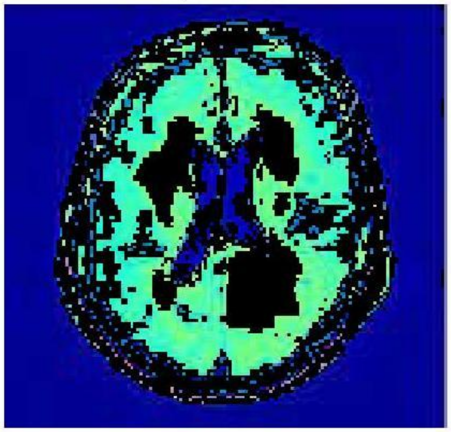

Fig 10. Objects in Cluster 3

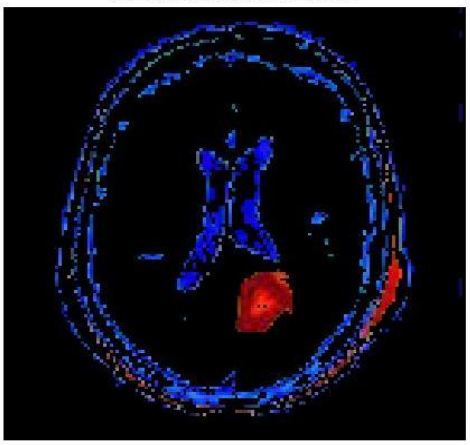

Fig 11 : Segmented Image using C-means

V. CONCLUSION

Average speed-ups of as much as 80 times a traditional implementation of FCM are obtained using the modified FCM algorithm, while yielding segmentation efficiency that are equivalent to those produced by the conventional technique. Thus, modified FCM algorithm is a fast alternative to the traditional FCM technique.

ACKNOWLEDGEMENT

The authors like to express their thanks to management and department of ECE K. L. University for their support and encouragement during this work.

REFERENCES

[1] Chen, C. W, Luo, J., Parker, K. J., "*Image Segmentation via Adaptive K-mean Clustering and Knowledge-based Morphological Operations with Biomedical Applications*," *IEEE Transactions on Image Processing*, Vol. 7, No, 12, Dec. 1998, pp. 1673-1683.

[2] Christine, C., Thomas, F., "*A Study of Efficiency and Accuracy in the Transformation from RGB to CIELAB Color Space*," *IEEE Transactions on Image Processing*, Vol. 6, No. 7, July 1997, pp. 1046-1048.

[3] Dhawan, A. P., "*A Review on Biomedical Image Processing and Future Trends*," Computer Methods and Programs in Biomedicine, Vol. 31, No.3-4, 1990, pp.141-183.

[4] Gonzalez, R. C.; Woods, R. E., *Digital Image Processing,2nd ed.*, Prentice-Hall, Englewood Cliffs, NJ, 2002.

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

[6] Tsai, C. S., Chang, C. C., "*An Improvement to Image Segment Based on Human Visual System for Object-based Coding*," *Fundamenta Informaticae*, Vol. 58, No. 2, 2004, pp. 167-178.

[7]. Wen, P, Zheng, L and Zhou, J, "*Spatial credibilistic clustering algorithm in noise image segmentation*", IEEE International Conference on Industrial Engineering and Engineering pp:309-315, 2005. Management, pp: 543 -547, 2007.

[8] Jie Yu, Peihuang Guo, Pinxiang Chen, Zhongshan Zhang and Wenbin Ruan, "*Remote sensing image classification based on improved fuzzy c -means*", Ceo-Spatial Information Science, vol.11, no.2, pp:90-94, 2008.

[9]. Anil.K.Jain, "Fundamentals of digital image processing" Prentice-Hall publications, ISBN:0-13-336165-9, 1989

[10] R.M Haarlick, "Statistical and structural approaches to texture", Proceedings of the IEEE, vol.67, pp.786-804, 1979.

[11]. C.H.Chen, L.F.Pau, P.S.P.Wang, "The handbook of pattern recognition and computer vision",2nd edition, World scientific publishing company, 1998.

AUTHOR'S BIBLOGRAPHY

M. Rakesh was born in 1987 at Guntur of Andhra Pradesh state, India. He Graduated in Electronics and Communication Engineering from KMCET, JNTU, Hyderabad. Presently he is pursuing his M. Tech in Communications and Radar Systems in K. L. University. His interested areas are Image Processing and Communications.

T. Ravi did his B. Tech at KSRMCE kadapa and later he did his M. E at Karunya University and M. Tech in JNTU Kakinada. He Worked as HOD in Universal College of Engg. and Tech. Guntur. And Presently he is working as a Assoc. Prof in K. L. University. His interested areas are Image Processing and Signal Processing.