Research Article

Brain Tumor Segmentation using K-Means Clustering Algorithm

Sanghamitra T. Kamble†* and M. R. Rathod†

†Department of Electronics and Communication, M.I.T, Aurangabad India

Abstract

Since image segmentation is a classic inverse problem which consist of achieving a compact region based description of the image scene by decomposing it into meaningful or spatially regions sharing similar attributes. Tumor is nothing but uncontrolled growth of tissues in any part of the body. Tumors are of different types and they have different treatments. The K-means algorithm is an iterative technique that is used to partition an image into K clusters. In developed countries most research show that the number of people who have brain tumors were died due to inaccurate detection of tumor .CT scan or MRI is directed into intracranial cavity produces a complete image of brain and this image is visually examined for detection.

Keywords: *Magnetic Resonance Imaging(MRI),Brain tumor, K-means, CT (Computerized tomoghraphy) scan, Thresholding, Image segmentation*

1. Introduction

Image segmentation is an important and, perhaps, the most difficult task in image processing. Segmentation refers to the grouping of image elements that exhibit similar characteristics, i.e. subdividing an image into its constituent regions or objects.

Segmentation is a fundamental process in digital image processing which has found extensive applications in areas such as medical image processing, compression, diagnosis arthritis from joint image, automatic text hand writing analysis, and remote sensing. The clustering methods can be used to segment any image into various clusters based on the similarity criteria like color or texture. K-means clustering algorithm divides the image into K clusters based on the similarity between the pixels in that cluster.

Brain tumor segmentation deals with the implementation of simple algorithm for detection of range and shape of tumor in brain MR image. Normally the anatomy of the brain can be viewed by the MRI scan or CT scan for diagnosis. There are different types of algorithm were developed for brain tumor segmentation. But they may have some drawback in detection and extraction and result people who have brain tumor died. K-mean clustering algorithm is used for the accurate detection of the brain tumor. The tumor may be primary or secondary. The brain tumor is primary when it is an origin and the brain tumor is secondary when the part of the tumor is spread to another place and grown.   propensity (Udroiu, 2002),

(Corneli, *et al*, 2004), (Seidu, 2008). It can be used to predict iron shrinkage tendency and help the foundry to control scrap.

2. Related work

The widely used algorithm in image segmentation system is K-mean clustering algorithm. There were Several attempts made by researchers to improve the effectiveness and efficiency of the *k*-means algorithm.
There are many researchers suggest initialized method of cancroids of k-means algorithm.
Stephen proposed a method for choosing *K* instances randomly from database as seeds. The drawback of this technique is the computational complexity; There are several iterations of the *k*-means algorithm are needed after that each instance is assigned, which in a large database is extremely burdensome.
Douglas and Michael suggested a method to select a good initial solution by partitioning dataset into blocks and applying *k*-means to each block. But the time complexity is slightly more. Though the above algorithms can help finding good initial centers for some extent, they are quite complex and some use the *k*-means algorithm as part of their algorithms, which still need to use the random method for cluster center initialization.
Haralick and Shapiro suggested that there is no full theory of clustering, and, therefore, no full theory of image segmentation. They have established the following qualitative guideline for good image segmentation.

*Corresponding author: **Sanghamitra T. Kamble**

- Segmented regions should be unified and homogeneous with respect to some characteristic such as gray level or texture.
- Region interiors should be simple and without many small holes.
- Adjacent segmented regions should have significantly different values with respect to the characteristic on which they are considered unified.
- Boundaries of each segment should be simple, not ragged, and must be spatially accurate.

That means the image segmentation techniques are generally ad hoc and differ on how they emphasize one or more of the desired properties. Therefore, the final implementation of each image segmentation algorithm depends very much on the end of the application. However, these differences usually center on the choices of parameters or methods of how to adapt certain parameters to the image.

3. System Design

In order to overcome the difficulties of K-means clustering in existing method the following design is carried out. Fig 1 consist of mainly four modules: preprocessing, segmentation, Feature extraction, and approximate reasoning. Pre processing is done by filtering. Segmentation is carried out by advanced Kmeans and Fuzzy C-means algorithms. Feature extraction is by thresholding and finally, Approximate reasoning method to recognize the tumor shape and position in MRI image using edge detection method. The proposed method is a combination of two algorithms.
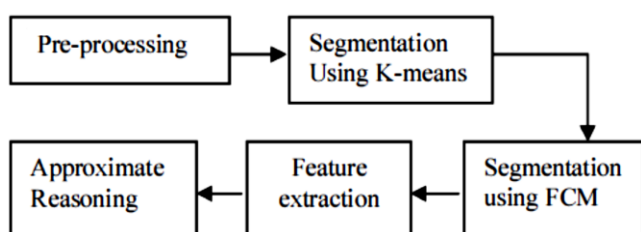
Fig.1 Block diagram of proposed method

3.1 Pre-Processing

It performs filtering of noise and other artifacts in the image and sharpening the edges in the image. In pre-processing RGB to grey conversion and Reshaping takes place .The main aim of this paper is to detect and segment the tumor cells. But for the complete system it needs the process of noise removal. For better understanding the function of median filter, we added the salt and pepper noise artificially and removing it using median filter. It includes median filter for noise removal. The possibilities of arrival of noise in modern MRI scan are very less. It may arrive due to the thermal effect.

3.2 K-Means Segmentation

A. K-means clustering

Clustering is one of the major data analysis methods widely used in many practical applications of emerging areas. Clustering is the process of finding groups of objects such that the objects in a group will be similar to one another and different from the objects in other groups. High quality clusters with high intra-cluster similarity and low inter-cluster similarity this is a good clustering method. 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.

B. Algorithm

1. Give the no of cluster value as k.

2. Randomly choose the k cluster centers

3. Calculate mean or center of the cluster

4. Calculate the distance b/w each pixel to each cluster center

5. If the distance is near to the center then move to that cluster.

6. Otherwise move to next cluster.

7. Re-estimate the center.

8. Repeat the process until the center doesn't move

C. Mathematical representation
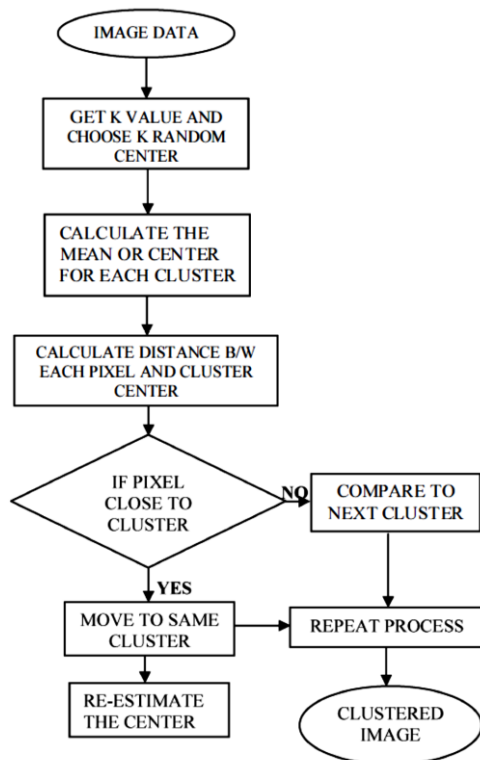
For a given image, compute the cluster means m

$$M = \frac{\sum_{nc(i)=k} Xi}{Nc} , k = 1 \dots K \tag{1}$$

Calculate the distance between the cluster center to each pixel

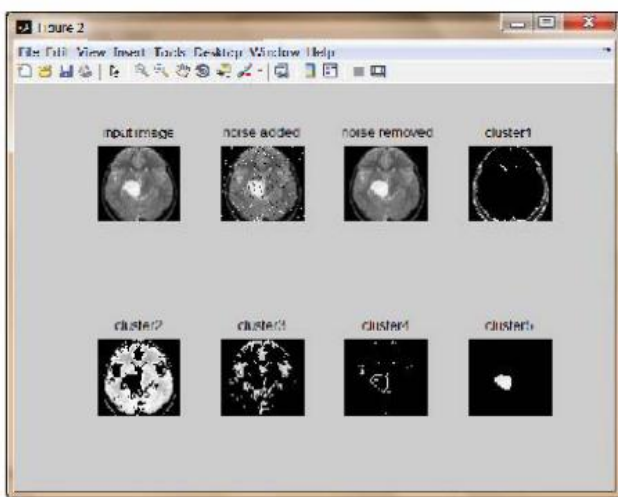$$D(i) = \arg\min ||Xi - Mk||^2 , i = 1 \dots N \tag{2}$$

Repeat the above two steps until mean value convergence.

D. Flowchart of k-means algorithm

The diagrammatic representation of the k-means algorithm and its flow

E. Screen shot or pre-processing and K-means

Fig. 2 Output image for pre-processing and k-means for k=5

4. Segmentation using fuzzy c-means

In fuzzy logic way to processing the data by giving the partial membership value to each pixel in the image. The membership value of the fuzzy set is ranges from 0 to 1. Member of one fuzzy set can also be member of other fuzzy sets in the same image. It is based on reducing the following function.

$$Ym=\sum_{i=1}^{N}\sum_{j=1}^{C} Mij ||Xi - Cj|| \qquad (3)$$

Where

m- any real number greater than 1,
Mij_ degree of membership of X; in the cluster j,
x,_ data measured in d-dimensional,
Rj - d-dimension center of the cluster,

Where

$$M_{ij} = \frac{1}{\sum_{k=1}^{C}(\frac{||X_i - C_j||}{||X_i - C_j||})^{\frac{2}{m-1}}} \qquad (4)$$

$$RR_j = \frac{\sum_{i=1}^{N} X_i . M_{ij}^m}{\sum_{i=1}^{N} M_{ij}^m} \qquad (5)$$

The above process ends when,

$$max_{ij}\{ |M_{ij}^{(k+1)} - M_{ij}^{(k)}| \} \angle \delta \qquad (6)$$

Where,

K= no of iteration steps.
C. The Fuz=y c-means Algorithm

5. Feature Extraction

In feature extraction image of tumor extracting the cluster which shows the predicted tumor area at FCM output . Whatever the extracted tumor giving to the thresholding process. That entire image applies the binary mask value then the dark pixel becomes dark and and white pixel becomes brighter.

6. Approximate reasoning

In the approximate reasoning binarization method is used to calculate tumor area. In binarization method image consist of two values either black or white. Black value assigned 0 and white value assigned to 1.

Image , I=$\sum_{w=0}^{255}\sum_{h=0}^{255}[f(0) + f(1)]$

Pixels = Width (W) X Height (H) = 256 X 256
f (0) = white pixel (digit 0)
f(l) = black pixel (digit 1)

No. of white pixel P=$\sum_{w=0}^{255}\sum_{h=0}^{255}[f(0)]$

Where,

P = number of white pixels (width*height)
1 Pixel = 0.264 mm
The area calculation formula is

Size of tumor, S=[($\sqrt{P}$)*0.264] mm²

P= no-of white pixels; W=width; H=height.

Conclusion

There are different types of tumors are available. They may be as mass in brain or malignant over the brain. Suppose if it is a mass then K- means algorithm is enough to extract it from the brain cells. If there is any noise are present in the MR image it is removed before the K-means process. The noise free image is given as a input to the k-means and tumor is extracted from the MRI image. For accurate segmentation of brain tumor K-means clustering algorithm is more efficient.

Acknowledgement

References

M.H. Fazel Zarandia, M. Zarinbala, M. Izadi b (2011), Systematic image processing for diagnosing brain tumors: A Type-II fuzzy expert system approach, Applied soft computing 11,285-294

Anil Z Chitade (2010),Colour based image segmentation using k-means clustering, International Journal of Engineering Science and Technology Vol. 2(10), 5319-5325

S.Mary Praveena ,Dr.Vennil (June 2010), Optimization Fusion Approach for [mage Segmentation Using K-Means Algorithm, International Journal of Computer Applications (0975 - 8887) Volume 2 , No.7.

A. Suman Tatiraju (July-2008), Image Segmentation using k-means clustering, EM and Nonnalized Cuts, Symposium of Discrete Algorithms. Tse-Wei Chen , Yi-Ling Chen , Shao-Yi Chien (2010), Fast Image

Segmentation Based on K-Means Clustering with Histograms in HSV Color Space, Journal of Scientific Research ISSN I452-2I6X Vol. 44 No.2, pp.337-35

M. Sahu and D. K. Parvathi (August 2013), Segmentation of Colour Data Base Image by Implementing K-means Clustering, International Journal of Innovations in Engineering and Technology, vol. 2, no. 4, pp. 229-234.

K. Alsabti, S. Ranka, and V. Singh (Mar. 1998), An Efficient k-means Clustering Algorithm, Proc. First Workshop High Performance Data Mining.

S. Zulaikha BeeviM, Mohamed Sathik (20I0). An Effective Approach for Segmentation of MRI Images:Combining Spatial Information with Fuzzy C-Means Clustering, European Journal of Scientific Research, Vol. 41 No.3 pp.437-45I

K.S. Ravichandran and B. Ananthi (2009), Color Skin Segmentation Using K-Means Cluster, International Journal of Computational and Applied Mathematics Volume 4 Number 2,pp. 153 -157

Douglas Steinley, Michael J. Brusco (2007), Initializing K-means Batch Clustering: A Critical Evaluation of Several Techniques, Journal of Classification DOI: 10.1007/s00357-007-0003-0, 24:99-121.

D. Van De Ville, M. Nachtegael, D. Van Der Weken, E.E. Kerre, W. Philips, I.Lemahieu (2003), Noise reduction by fuzzy image filtering, IEEE Transactions on Fuzzy Systems.

Segmentation (September 1991), Proceedings of the Pfefferkorn Conference on Signal and Image Processing in Microscopy and Microanalysis, pp. 299–314.

Malik, Khan (2009), Modified Watershed Algorithm for Segmentation of 2D Images, Journal of Information Science & Information Technology, 6, No.3, pp. 546-552.