STUDY AND ANALYSIS OF SEGMENTATION FOR MEDICAL IMAGES BASED ON K-MEANS AND FUZZY C-MEANS CLUSTERING

Pammi kumari[1] Ravindra Kumar Gautam[2] Hari Kumar Singh[3]
Mridul Saran[4] Santosh Kumar[5]

[1,5]Department of Electronics and Communication Engineering
Institute Of Engineering and Technology Invertis University, Bareilly (India)
[2,3]Department of Electronics and Communication Engineering
Institute of Engineering And Technology M.J.P. Rohilkhand University, Bareilly (India)
[4]Department of Computer Science Engineering
Birla Institute of Applied Sciences Bhimtal, Nainital (U.K.)

ABSTRACT

Image segmentation have been the two major areas of research in the medical imaging community for decades and still are. In the context of segmentation and clustering methods are widely used for target structure definition such as prostate or head and back sole node areas. In the past two years, 45% of all articles published in the most important medical imaging journals and conferences have presented either segmentation or registration methods. This article reviews the literature on image segmentation methods by introducing a novel taxonomy based on the amount of shape knowledge being incorporated in the segmentation process .It presents a new approach for image segmentation by applying k-means clustering and fuzzy c-means clustering. In image segmentation, clustering techniques are very important as they are intuitive and are also easy to implement. In This paper proposes a color based segmentation method that uses K-means clustering and fuzzy c-means clustering technique The k-means clustering is an instinct technique used to partition an image into k clusters. It produces accurate segmentation results only when applied to images defined by homogenous regions with respect to texture and colour since no local constraints are applied to impose spatial continuity for medical images.

Keywords: Segmentation ,K-Means Clustering, Fuzzy C-Means Clustering, Survey Paper

I. INTRODUCTION

Medical Image segmentation and clustering have been important research topics over the last two to three decades and several state of the art surveys exist for segmentation[1,2,3,4,5,6].Image segmentation plays very important role in diagnosis of medical images. The objective of image segmentation on the basis of clustering is to partition an image into non overlapping, and homogeneous parts with respect to variation in intensity level and texture of that medical image. India is the second most popular country of the world and has changing socio-politic demographic and morbidity patterns that have been drawing global attention in recent years. About 75% of health infrastructure, medical man power and other health resources are concentrated in urban areas where 27% of the population lives. Contagious, infectious and waterborne diseases such as diarrhea, amoebiasis, typhoid, infectious hepatitis, worm infestations, measles malaria ,tuberculosis, whooping cough, respiratory infections, pneumonia and reproductive tract infections dominate the morbidity pattern, especially in rural areas. For problem identification of any medical images or for extraction of information from affected parts clustering approach is used. Detection of dead tissues based on clustering plays an important role in image understanding, image analysis and image processing. Because of its simplicity and efficiency, clustering approaches were one of the first techniques used for the textured natural images [1].

II. MEDICAL IMAGING

Medical imaging is the technique, process and art of creating visual representations of the interior of a body for clinical analysis and medical intervention. Medical imaging seeks to reveal internal structures hidden by the skin and bones, as well as to diagnose and treat disease. This rapidly evolving field of medicine originated in the first decade of the 19th century, when Wilhelm Rontgen, a professor of physics at Wurzburg University in Germany, discovered electromagnetic radiation. After the World War Two, the development of computer technology has triggered an amazing revolution in medical imaging techniques. There is a continuous drive not only to improve the diagnostic yield of medical imaging techniques for clinical use, but also the management of the huge amount of digital information available to medical imaging departments[2].

III. CLASSIFICATION OF SEGMENTATION TECHNIQUES

There are three main characteristics which influence the segmentation of an object in an image: object boundaries, object homogeneity and object shape. Object boundaries and object homogeneity are image or signal based characteristics. Therefore, they are affected by image specific disturbances like noise or reconstruction artifacts. Furthermore, they are modality dependent. An object's shape is image independent and in most imaging modalities apart from small deviations like perspective mapping distortions also independent from the acquisition technique. The concepts of object boundary, object homogeneity and object shape have a strong influence on the development of segmentation methods. Segmentation techniques try to detect boundaries and homogeneous regions in the images and incorporate shape information to restrict the shape of the resulting segmentation.

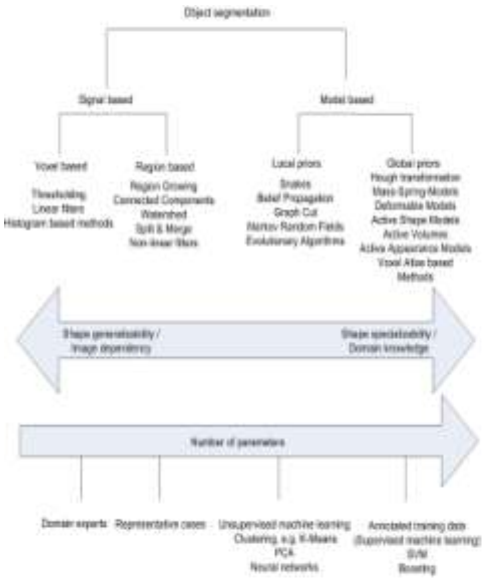

Figure:1 Classification scheme of method for segmenting objects for medical images

A large number of medical image segmentation approaches have been proposed in the literature working either on two-dimensional or three-dimensional data. In each classification group, the most important methods for medical image segmentation are described. Often complex segmentation methods consist of a whole pipeline using algorithms from different classification groups.

IV. SEGMENTATION METHOD

4.1 Otsu's Method

However in other cases the absolute image intensities of the classes are not known but it can be assumed that they are different from all other objects' intensities. Here, methods based on the intensity histogram can be used to automatically determine the thresholds. The probably most popular method for finding a single threshold separating two classes is Otsu's method [13] which has been extended to multiple classes [14]. It finds the global maximum of the between-class Variance.

$$\sigma^2_{between\,Modified} = \sum_{j=1}^{k} W_j \mu^2_j$$

where k is the number of object classes J is the mean intensity of the class and W is the probability of each class given by the histogram.

4.2 Region Based Methods

The second classification group for image segmentation according to the taxonomy presented in Figure 1 consists of region based methods. Region based methods are mainly based on image signal information, but they incorporate local relationships between samples, for example for building contiguous regions. One of the most prominent region based methods in 2d and 3d is the region growing algorithm. Here, segmentation grows from initially placed points called seeds by aggregating neighboring pixels or regions according to some similarity criterion. Region growing is often used to segment homogeneous regions like vessels trees which vary in shape but share a similar intensity. The main limitation of region growing clearly is its tendency to leak into neighboring objects sharing similar intensities with the structure to be segmented.

4.3 Geometric Model Based Segmentation

In geometric model based segmentation, shape is represented by geometric objects like point clouds and polygonal surfaces [2,3,5,6], simplex meshes [8], representations, level set representations [7,7,10,12], geometric grids [10] or finite element triangulations [14].

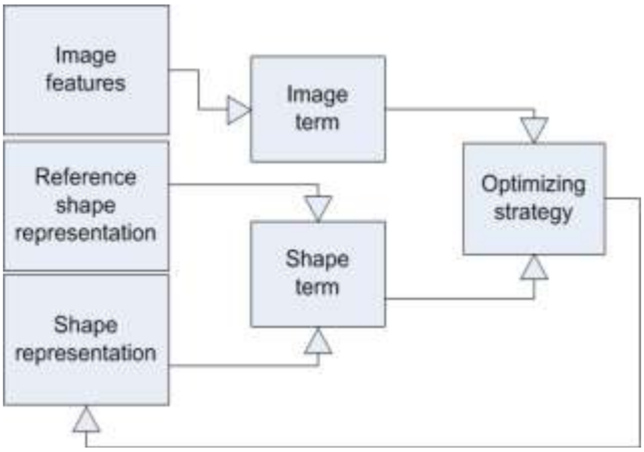

Figure: 2 Shows the Principle of Geometric Model Based Segmentation

4.4 K-Means Clustering

K-means is one of the simplest unsupervised learning algorithms. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. The main idea is to define k centroids, one for each cluster. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest centroid. When no point is pending, the first step is completed and an early group age is done. At this point we need to re-calculate k new centroids as bary centers of the clusters resulting from the previous step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest new centroid. A loop has been generated.. In other words centroids do not move anymore.Finally, this algorithm aims at minimizing an *objective function*, in this case a squared error function. The objective function

$$j = \sum_{j=1}^{k}\sum_{i=1}^{k} \left\| X^{(j)}{}_{i} - C_{j} \right\|^{2},$$

$\left\| X^{(j)}{}_{i} - C_{J} \right\|^{2}$ where is a chosen distance measure between a data point x_i^{j} and the cluster centre c_j is an indicator of the distance of the *n* data points from their respective cluster centers. In statistics and data mining, k-means clustering is a method of cluster analysis which aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean.

4.5 Fuzzy C-Means Clustering

The well-known Fuzzy C-Means (FCM) clustering algorithm was originally introduced by Dunn [11] and later it is enhanced by Bezdek [4]. The FCM algorithm is mainly an iterative clustering method, which results an optimal c partition by minimizing the weighted within group sum of squared error objective function O (U,C)In present scenario, one of the most difficult task in image analysis & computer vision is to classify the pixel in an image correctly, when there is no crisp boundaries between objects in an image thus in order to resolve this difficult, fuzzy clustering techniques are used. Fuzzy clustering divides the input pixels into clusters or groups on the basis of some similarity criterion. Fuzzy c-means is a method of clustering which allows one piece of data to belong to two or more clusters. It is based on minimization of the following objective function:

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} U^{m}{}_{ij} \left\| X_i - C_j \right\|^{2}, \quad 1 \le m < \infty$$

where *m* is any real number greater than 1, u_{ij} is the degree of membership of x_i in the cluster *j*, x_i is the *i*th of d-dimensional measured data, c_j is the d-dimension center of the cluster, and ||*|| is any norm expressing the similarity between any measured data and the center.

Fuzzy partitioning is carried out through an iterative optimization of the objective function shown above, with the update of membership u_{ij} and the cluster centers c_j by:

$$U_{ij} = \frac{1}{\sum_{k=1}^{c}\left[\frac{\left\| X_i - C_j \right\|}{\left\| X_J - C_K \right\|}\right]^{\frac{2}{m-1}}} \quad , \quad C_j = \frac{\sum_{i=1}^{N} U_{ij}.X_j}{\sum_{i=1}^{N} U^{m}{}_{ij}}$$

This iteration will stop when $\max_{ij}\left\{\left| U_{ij}{}^{(k+1)} - U_{ij}{}^{(k)} \right|\right\} < \varepsilon$ where ε is a termination criterion between 0 and 1, whereas *k* are the iteration steps. This procedure converges to a local minimum or a saddle point of J_m. The FCM is the most accepted method since it can preserve much more information than other approaches.

V. MATERIALS AND METHODS

The following image is taken under consideration

Fig:3 Sample image for training

This image is collected from Dr Somesh Malhotra. It shows the disease named back sole .A common allergic reaction that looks like wounds, hives are often itchy, and sometimes stinging or burning. Back sole vary in size and may join together to form larger areas. . Exaggerated surface markings in the overlying skin (Wickham's striae) may also be evident but are difficult to appreciate. The forearms, the middle of the back, and the anterior surfaces of the lower extremities are other common locations.

6.1 Experimental Results

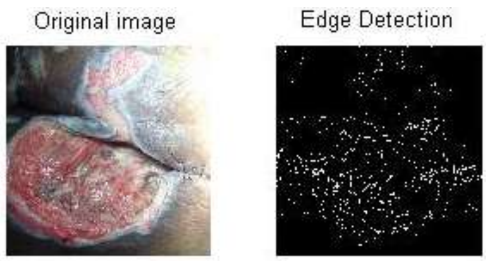

Fig:4 Image obtained by applying k means clustering method

In K-means algorithm, we firstly initiate cluster centers and then decide the number of iteration by a lot of tries to get the good quality of segmentation. The Figure shows the segmented result using K-means clustering.

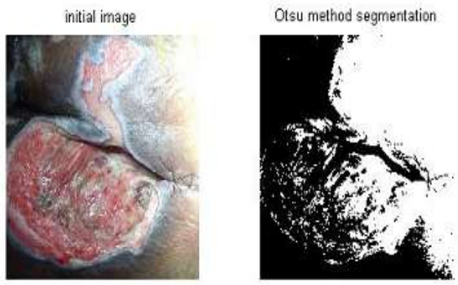

Fig:5 Image obtained by applying otsu method

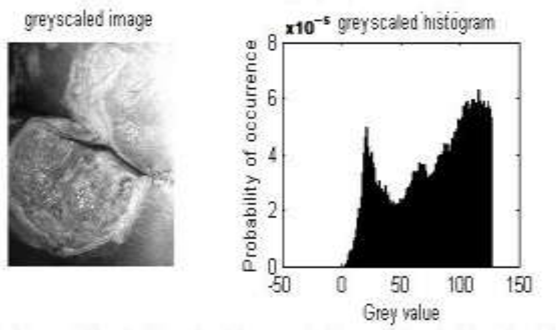

Fig:6 Image obtained by applying histogram thresholding

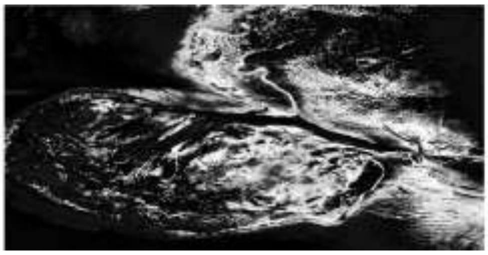

Fig:7 Image obtained by applying fuzzy c-means clustering

VI. CONCLUSION

After successfully implemented k-means clustering algorithm. for smaller values of k the algorithms give good results. for larger values of k, the segmentation is very coarse; many clusters appear in the images at discrete places .this is because euclidean distance is not a very good metric for segmentation processes. different initial partitions can result in different final clusters.the result aims at developing an accurate and more reliable image which can be used in locating tumors, measure tissue volume, face recognition, finger print recognition and in locating an object clearly from a satellite image and in more[9]. it works well when clusters are not well separated from each other this could be happen in web images. we proposed a framework of unsupervised clustering of images based on the colour feature of the image. it minimizes intra-cluster variance.

The work is done on the images that do not have very sharp variation in RGB values. when various segmentation techniques are applied to such images the results were not satisfactory, the proper visualization of the dead tissues of the images was not possible. the use of such system gave 60-70 % improvement in results. images were obtained showing the proper detection of dead tissues including performance, regression plot. By using clustering on medical images each area, of live and dead tissue was clearly visible. it gave almost 85% improved results .This paper represents various methods of segmentation and clustering which can be helpful for medical image segmentation.

REFERENCES

[1] Heimann T, Meinzer HP, Wolf I. A statistical deformable model for the segmentation of liver ct volumes. MICCAI Workshop on 3d Segmentation in the Clinic: A Grand Challenge (2007), 161-166. http://mbi.dkfz-heidelberg.de/mbi/TR/Papers/P19-07.pdf .

[2] Ling H, Zhou S, Zheng Y et al. Hierarchical, learning-based automatic liver segmentation. CVPR (2008), 1-8. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4587393.

[3] Zhen Y, Barbu A, Georgescu B et al. Fast automatic heart chamber segmentation from 3d ct data using marginal space learning and steerable features. Conf. on Computer Vision (ICCV) (2007), 1-8.

[4] Ecabert O, Peters J, Schramm H et al. Automatic model-based segmentation of the heart in ct images. IEEE Trans. Med. Imaging. 2008;27(9):1189-1201.

[5] S. P. Lloyd, ―[1] Least squares quantization in PCM," IEEE Trans. Inf.Theory, vol. IT-28, no. 2, pp. 129–136, Mar.1982.

[6] Pham, Dzung L.; Xu, Chenyang; Prince, Jerry L. (2000). "Current Methods in Medical Image Segmentation". Annual Review of Biomedical Engineering 2: 315–337.

[7] M. Tabb and N. Ahuja, Unsupervised multiscale image segmentation by integrated edge and region detection, IEEE Transactions on Image Processing, Vol. 6, No. 5, 642–655, 1997

[8] Novák, V., Perfilieva, I. and Močkoř, J. (1999) Mathematical principles of fuzzy logic Dodrecht: Kluwer Academic. ISBN 0-7923-8595-0

[9] "Fuzzy Logic". Stanford Encyclopedia of Philosophy. Stanford University. 2006-07-23. Retrieved 2008-09-30

[10] J.A Hartigan "Clustering Algorithms", New York Wiley 1975.

[11] Quek, C., & Zhou, R. W. (1999). "POPFNN-AAR(S): a pseudo outer-product based fuzzy neural network." IEEE Transactions on Systems, Man and Cybernetics, Part B, 29(6), 859-870.

[12] Abraham A., "Adaptation of Fuzzy Inference System Using Neural Learning, Fuzzy System Engineering: Theory and Practice", Nadia Nedjah et al. (Eds.), Studies in Fuzziness and Soft Computing, Springer Verlag Germany, ISBN 3-540-25322-X, Chapter 3, pp. 53–83, 2005. information on publisher's site.

[13] Bernd Jähne (2002). Digital Image Processing. Springer. ISBN 978-3-540-67754-3.

[14] Lin, C.-T., & Lee, C. S. G. (1996). Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems. Upper Saddle River, NJ: Prentice Hall.

[15] Wilhelm Burger and Mark J. Burge (2007). Digital Image Processing: An Algorithmic Approach Using Java. Springer. ISBN 978-1-84628-379-6.

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