Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 3, March 2014, pg.431 – 437

SURVEY ARTICLE

A SURVEY ON CONTENT RETRIEVAL SCHEMES FOR MEDICAL IMAGES

Meeshika Arora

Dept. of CSE, Banasthali University, India

meeshika.arora@gmail.com

Abstract: Segmentation is the process which subdivides an image into its constituent regions. This technique can also identify region of interest in image or annotate the data. It is observed as the most critical function of image processing and analysis. Image segmentation plays a crucial role in many medical imaging applications by automating delineation of anatomical structures and other regions of interest. This paper makes a survey on the current segmentation techniques used in the retrieval of medical images. Numerous segmentation algorithms and techniques have been defined for image segmentation. This paper presents a detailed review of several segmentation techniques with their principal ideas, types, advantages and disadvantages.

Keywords: Image segmentation, Thresholding segmentation, Region growing, Clustering, Neural network, Classifier

I. INRODUCTION

For more than a decade content-based image retrieval has been an active research area .The foremost aim is to retrieve digital images based not only on textual annotations but on features derived directly from the images data which are stored alongside the image and serve as an index. One way to perform retrieval is often performed in a query by example fashion where a query image is provided by the user. The image database application is then searching through all images Content-based image retrieval (CBIR) has been an active research area for more than a decade . CBIR techniques have also been applied to various medical image databases successfully. Content-based retrieval of medical images has been shown to be useful for various applications.[1]

MEDICAL images play a crucial role in patient therapy, diagnosis, surgical planning, medical training and reference. The development of systems for diagnosing, screening, archiving, and annotating based on automatic analysis of medical images are recurring research topics. With the advent of digital imaging modalities such as single-photon emission computerized tomography (SPECT), computed tomography (CT), magnetic resonance imaging (MRI) as well as images digitized from conventional devices such as histological slides and X-rays, collections of medical images is increasingly being held in digital

form. The challenging issue with this research is to effectively maintain the medical database and used of data in an effective manner. [2]

The necessity of computers in facilitating their processing and analysis has become mandatory with the expansion in the number and size of medical images. In certain, computer algorithms for the delineation of anatomical structures and other regions of interest area key component in assisting and automating specific radiological tasks. Basically the image segmentation algorithms play an essential role in numerous biomedical imaging applications such as the diagnosis, quantification of tissue volumes, and localization of pathology, study of anatomical structure, and partial volume correction of functional imaging data, treatment planning and computer-integrated surgery

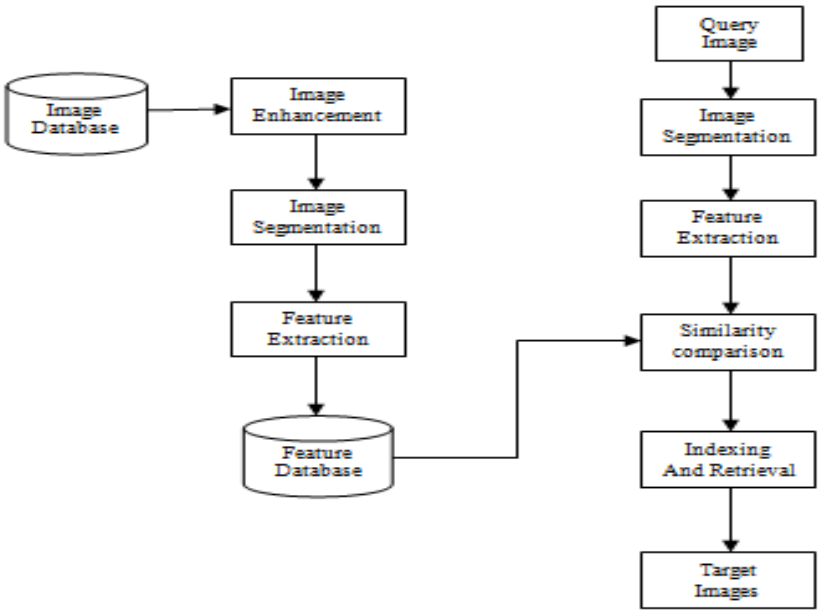

Fig. 1 Basic framework of CBIR

II. LITERATURE SURVEY

Soo Beom Park, Jae Won Lee, Sang Kyoon Kim proposed a method of content-based image classification using a neural network. In this classification process, the background regions are removed from an image in an attempt to minimize the case misclassified. And by using the wavelet transform and the sliding window-based feature extraction, the more feature information from images is acquired In addition; a neural network classifier is created using the learning pattern of the texture feature to reflect a shape of an object. A higher classification rate by removing the back-ground is achieved by results. And the diagonal moment shows the highest classification rate among the various texture features.

Our work is applicable to improving the performance of content-based image indexing or retrieval systems. And our system shows the potential that an automatic classification system could retrieve images more efficiently from the great many images of the Internet. [3]

Jianping Fan, Yau Elmagarmid & Aref's [2] paper presents an automatic image segmentation method using thresholding technique. This is based on the assumption that adjacent pixels whose value lies within a definite range belong to the same class and hence, good segmentation of images that include only two opposite components can be obtained. Jaskirat Kaur & Renu Vig.'s paper presented edge detection and thresholding being one of the most important aspects of image segmentation comes prior to feature extraction and image recognition system to examine the images. It helps in deriving the basic shape of an image, overlooking the unnecessary details. In this paper using image segmentation (thresholding and edge detection) techniques different medical images, geosatellite images and architectural images are analyzed. To determine the regularity of our results error measure is used [8]. Zhen Ma et al. proposed that for a concrete medical

image segmentation task, researchers should combine the application background and practical requirements to design proper algorithms. Accuracy, complexity, efficiency and interactivity of a segmentation algorithm should all be the considered factors.

III. SEGMENTATION APPROACHES

The necessity of computers in facilitating their processing and analysis has become mandatory with the expansion in the number and size of medical images. In certain, computer algorithms for the delineation of anatomical structures and other regions of interest area key component in assisting and automating specific radiological tasks. Basically the image segmentation algorithms, play an essential role in numerous biomedical imaging applications such as the diagnosis, quantification of tissue volumes, localization of pathology , study of anatomical structure, treatment planning, partial volume correction of functional imaging data and computer-integrated surgery.

The segmentation & edge detection approaches were studied under 5 categories. These are as follows-

1) Thresholding approaches
2) Region growing approaches
3) Clustering approaches
4) Neural network approaches
5) Classifier approaches

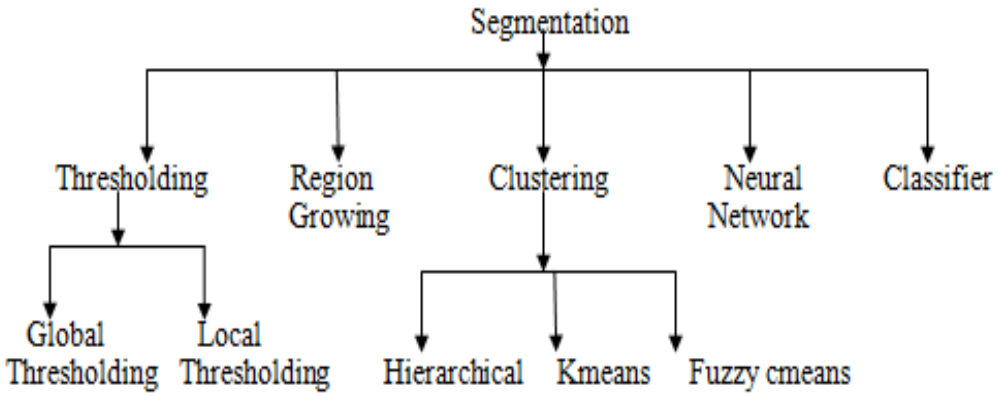

Fig. 2 Types of Segmentation

IV. SEGMENTATION TECHNIQUES

A. Threshold Segmentation

This method is based on a threshold value which will convert gray scale image into a binary image format i.e. it assigns the value of 0(background) or 1(objects) to each pixel of an image based on a comparison with some defined threshold value. The global thresholding appears when the value of T is constant. The main logic behind this is the selection of a threshold value. A thresholding procedure attempts to determine an intensity value, called the threshold, which separates the desired classes Threshold selection is typically done iteratively however it is possible to derive automatic threshold selection algorithm [4]. The segmentation procedure of segmentation tends to be achieved when all pixels with intensity greater than threshold appear into one class, and all other pixels into another class.

Let I(i,j) be an image,

$$I(i,j) = \begin{cases} 0, & p(i,j) < T \\ 1, & p(i,j) \geq T \end{cases}$$

Where p(i, j) refer to the pixel value at the position (i, j).
Types of threshold segmentation:
i) Global Thresholding: When a Threshold value T depends only on gray level values f(x,y) and the value of T solely relates to this character of the pixels, this thresholding techniques is called as global thresholding techniques[7]. A number of global thresholding techniques such as: minimum thresholding, Otsu, optimal thresholding iterative thresholding and so on.
ii) Local thresholding: If threshold value T depend on both f(x,y) and p(x,y) this thresholding is called local threshold[7].Local thresholding techniques are simple statistical thresholding, histogram transformation thresholding etc[5]
The advantage of using Threshold-based algorithms is that they do not need complex operations and are computationally efficient. Disadvantage of these algorithms is that they are sensitive to noise, hard to be applied to multichannel images and difficult to combine with spatial information. The segmentation results of threshold-based algorithms are usually far from satisfaction as medical images are usually noisy and suffer from intensity in homogeneity. [1]

B. Region Growing

Region growing is a technique for extracting a region of the image that is connected based on some pre defined criteria. This criterion can be based on edges in the image or on intensity information. Basically, region growing requires a seed point that is manually selected by an operator, and obtains all pixels connected to the initial seed with the same intensity value. [6]
The region growing approach is a threshold based algorithm which is having the same advantage that they do not need complex operations. But its primary disadvantage is that it requires manual interaction to obtain the seed point. Therefore, a seed must be planted for each region that needs to be extracted. Region growing can cause extracted regions to have holes or even become disconnected due to noise sensitivity. [6] In real images, it becomes rare to obtain regions from contours directly and vice versa. Regions are important for the interpretation of images because they may correspond to objects in a scene.

C. Clustering

In this technique data set is replaced by cluster pixels may belong together because of the some colour , texture etc. K-means, C-means and hierarchical clustering are the classes of cluster based methods. These are iterative technique that is used to partition an image into clusters that minimizes the variance between the pixel and the cluster centre. This algorithm is guaranteed to converge. The quality of solution depends on the initial set of clusters.[4]
Clustering techniques can be classified into supervised and unsupervised schemes. The unsupervised classification of data is known as clustering or exploratory data analysis, no labeled data are available .The goal of clustering is to separate a finite unlabeled data set into a finite and discrete set of "natural," hidden data structures.[7].
Clustering based method is divided into three classes
1) Hierarchical clustering
2) Overlapping (fuzzy c-means) clustering
3) Exclusive (k-means) clustering [4]

1. Hierarchical Clustering

This algorithm is used for image retrieval .In Hierarchical clustering we create a hierarchy of clusters which may be represented in a tree structure as a dendrogram. Here a single cluster containing all observations at the root, and the leaves correspond to individual observations. Usually the algorithms for hierarchical clustering are generally either agglomerative starts at the leaves and successively merges clusters together; or divisive in which one starts at the root and recursively splits the clusters.[8]

2. K-means Clustering

K-means is an algorithm to group objects based on features into k number of groups. The grouping (clustering) is done by minimizing the Euclidean distance between data and the corresponding cluster centroid. The k-means clustering algorithm has been used to classify image pixels, which calculates the initial means and their position in clusters. Also k-means algorithm which has been used in this study in order to differentiate the different texture regions of the image. The K-means clustering algorithm to perform the segmentation because it is simple and has low computational complexity [9]
Algorithmic steps for K-Means clustering ref
1) Initially set k to choose a desired no. of clusters.
2) To choose k initial starting points which are used as initial estimates of the cluster centroids. They are taken as the initial starting values.

3) In Classification scheme each point in the data set is examined and assigned to the cluster whose centroid is nearest to it.
4) Then the centroid is calculated on the basis of the point in the data set assigned to a cluster, also it is needed to recalculate the new k centroids.
5) The (iii) and (iv) steps are required to be repeated until no point changes its cluster assignment or until the centroids no longer move. This is the convergence criteria.

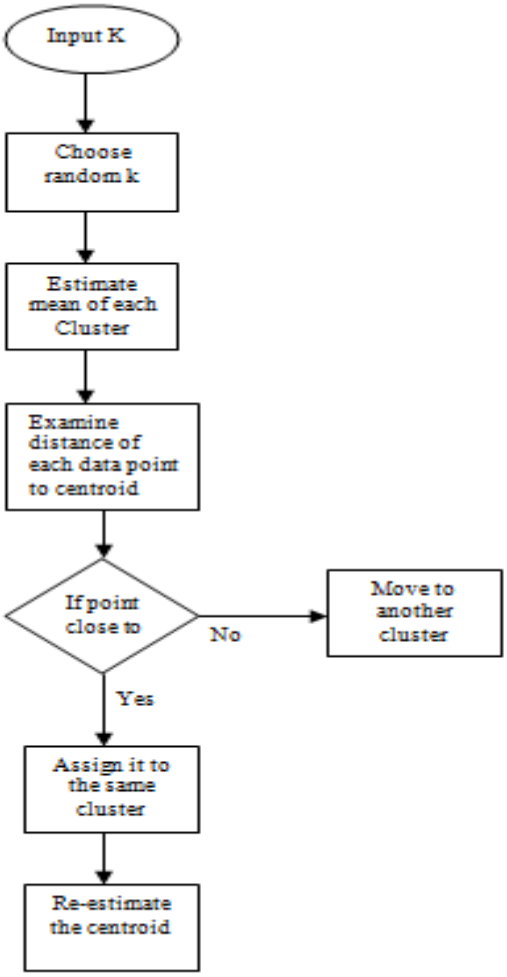

Fig.3 Flowchart of Kmeans

3. Fuzzy C-means

Fuzzy C-means segmentation could retain much more information from the original image because they are over hard segmentation.[4] Fuzzy C-means (FCM) is one of the method of clustering that allows one piece of data to belong to two or more than two clusters. It works by assigning membership to each data point corresponding to each cluster center on the basis of distance between the cluster center and data point. Membership towards the cluster center is more if more the data is near to the cluster center. Clearly summing up membership of each data point should be equal to one after each iteration membership and clusters are updated according to formula

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{ik}}\right)^{\left(\frac{2}{m}-1\right)}}$$

$$v_j = \frac{\sum_{i=1}^{n} \mu_{ij}^{m} x_i}{\sum_{i=1}^{n} \mu_{ij}^{m}} \qquad \forall = 1, 2, 3 \ldots\ldots\ldots C.$$

n- Number of data points.

$v_{j-\ j^{th}}$ Cluster center.c – number of cluster centers.

μ_{ij-} Euclidean distance between $i^{th}\ and\ j^{th}$ cluster center.

m= fuzziness index.

A common method is to run the algorithm several times regain the best clustering found. Advantage of using Clustering algorithms is that they essentially perform the same function as classifier methods with-out the use of training data. Thus, they are termed unsupervised methods. In order to compensate for the lack of training data, clustering methods iterate between segmenting the image and characterizing the properties of the each class. In a sense, clustering methods train themselves using the available data. Clustering algorithms do not directly incorporate spatial modeling and can therefore be sensitive to noise and intensity in homogeneities. This lack of spatial modeling, however, can provide significant advantages for fast computation.

D. Neural Network

An artificial neural network (ANN) is an information processing system which contains a large number of highly interconnected processing neurons. These neurons work together in a distributed manner to learn from the input information, to coordinate internal processing, and to optimize its final output Neural networks covering medical image registration, segmentation and edge detection for medical image content analysis, computer- aided detection Other applications of ANN include data compression, image enhancement and noise suppression and disease prediction etc. More recently, application of ANN for functional magnetic resonance imaging (MRI) simulation becomes a new research hotspot, where certain structured ANNs are employed to simulate the functional connectivity of brain networks . Due to the similar nature of ANN and human neurons, ANN has been proved to be a very useful for this new task .

E. Classifier

Classifier methods are pattern recognition techniques that seek to partition a feature space derived from the image using data with known labels. Classifier are known as supervised methods since they require training data that are manually segmented and then used as references for automatically segmenting new data. There are a number of ways in which training data can be applied in Classifier methods. A simple classifier is the nearest-neighbor classifier, where each pixel or voxel is classified in the same class as the training datum with the closest intensity. K-nearest neighbor is one of the simplest pattern recognition classification techniques. The disadvantage is the requirement of manual interaction for obtaining training data. Training sets can be acquired for each image that requires segmenting, but this can be time consuming and laborious. On the other hand, use of the same training set for a large number of scans can lead to biased results which do not take into account anatomical and physiological variability between different subjects.[1]

V. CONCLUSION AND FUTURE DIRECTIONS

The overall aim of this survey to investigate and discuss different traditional and popular image segmentation techniques. Here the basic fundamental properties different techniques have been highlighted. The merits and demerits of methods discussed in short. Although various techniques are available, each technique works on specific concept hence it is important which image segmentation techniques should be used as per application domain. According to this survey we concluded that segmentation technique works on specific concept. Hence it becomes important to use right technique according to their application. Also it has been seen that there is no single algorithm that works well for all types of images, but some work better than others only when we combine these techniques or we can combine them with any of the optimization technique. In future

work I will implement k means clustering technique combine with an optimization technique. So that more effective segmentation can be done.

REFRENCES

[1] Roman Starosoiski, Gerald Schaefer, MarekMoleda and Shao, Ying Zhu "Content-based retrieval of compressed medical images "Proceedings of the 5th International Conference on Information Technology and Application in Biomedicine, in conjunction with The 2nd International Symposium & Summer School on Biomedical and Health Engineering Shenzhen, China, pp 157-159, 2008

[2] H. Lilian Tang, Rudolf Hanka, Member, IEEE, and Horace H. S. Ip, Member, IEEE "Histological Image Retrieval Based on Semantic Content Analysis" IEEE Transactions on Information Technology in Biomedicine, pp 26-36, Mar. 2003.

[3] SooBeom Park, Jae Won Lee, Sang Kyoon Kim "Content-based image classification using a neural network"

[4] Mamta Bharadwaj, "Comparative Research on Fuzzy C-Means and K-Means Clustering Segmentation", International Journal of Computer Applications & Information Technology, Aug-September 2013.

[5] RashmiSaini*, MaitreyeeDutta and Robin Kumar,"A Comparative study of Several Segmentation Techniques", Journal of Information and Operations Management, pp-21-24, 2012.

[6] Dzung L. Phamé?ê, Chenyang Xué, Jerry L. Princeé "A Survey of Current Methods in Medical Image Segmentation", pp 1-26, 1998.

[7] Mrs. Monika Jain, Dr. S.K. Singh "A Survey On: Content Based Image Retrieval Systems Using Clustering Techniques for Large Data sets" International Journal of Managing Information Technology (IJMIT), pp 23-39, Nov. 2011

[8] Stuti Mehta and Ashok Kajal, "Analysis of Different Clustering Algorithms on Image Databases", 5thIEEE International Conference on Advanced Computing & Communication Technologies [ICACCT-2011] ISBN 81-87885-03-3.

[9] T. Kanungo and D. M. Mount, "An Efficient K-means Clustering Algorithm: Analysis and Implementation", Pattern Analysis and Machine Intelligence, IEEE Transactions on Pattern Analysis and Machine Intelligence., 2002 .

[10] Kailash Sinha, G.R.Sinha "Segmentation of Brain MRI Images for Tumor extraction by combining C-means clustering and Watershed algorithm with Genetic Algorithm" *International Journal of Digital Application & Contemporary research, pp 1-6, August 2012.*

[11] Shally HR, Chitharanjan K "Review of Brain Tumor Detection Techniques from MRI Images" Asian Journal of Computer Science and Information Technology ISSN 2449-5126.

[12] Kimmi Verma, Aru Mehrotra, Vijayeta Pandey, Shardendu Singh "Image Processing techniques for the Enhancement of Brain Tumor Patterns" International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, pp 1611-1615, Apr. 2013.

[13] Abduljawad A. Amory, Rachid Sammouda, Hassan Mathkour, Rami Mohammad Jomaa, "A Content Based Image Retrieval Using K-means Algorithm", IEEE, pp.221-225, 2012.

[14] Sagar Soman, Mitali Ghorpade, Vrushali Sonone, Satish Chavan, "Content Based Image Retrieval Using Advanced Color and Texture Features" International Conference in Computational Intelligence (ICCIA) Proceedings published in International Journal of Computer Applications® (IJCA), pp.10-14, 2011.

