Volume 5, Issue 3, March 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Liver Segmentation using K-Means Algorithm and Mathematical Morphology

Vijendra P. Meshram, Rajesh.D. Thakare, Prashant Wanjari, Vishwas V. Balpande, Ishan A. Patil
Electronics Dept., DBACER, Nagpur, Maharashtra,
India

Abstract-- This study describes a new liver segmentation method for purpose of transplantation surgery as a treatment for liver tumors. Liver segmentation is not only the key process for volume computation but also fundamental for further processing to get more anatomy information for individual patient. Due to the low contrast, blurred edges, large variability in shape and complex context with clutter features surrounding the liver that characterize the liver CT images, it is a convoluted problem and still a challenge task to robustly and accurately segment the liver. In this paper, we overcome these difficulties with a novel variational model based on the idea of intensity probability distribution propagation and region appearance propagation with which we can focus on the target liver regardless of how complex the uninterested background is. In the segmentation process in order to divide the image, five separate regions are identified on the computerized tomography image frames. The merit of the proposed method lays in its potential to provide fast and accurate liver segmentation and 3-D rendering as well as in delineating tumor region(s), all with minimal user interaction.

Keywords-- Liver segmentation, CT images, MATLAB, K-means.

I. INTRODUCTION

Images are considered as one of the most important medium of conveying information. Understanding images and extracting the information from them such that the information can be used for other tasks is an important aspect of Machine learning. One of the first steps in direction of understanding images is to segment them and find out different objects in them. Thus image segmentation plays a vital role towards conveying information that is represented by an image and also assists in understanding the image. Image segmentation is the process of dividing the given image into regions homogenous with respect to certain features, and which hopefully correspond to real objects in the actual scene. Segmentation plays a vital role to extract information from an image to create homogenous regions by classifying pixels into groups thus forming regions of similarity. The homogenous regions formed as a result of segmentation indwell pixels having similarity in each region according to a particular selection criteria e.g. Intensity, color etc.Segmentation 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 segmentation of (textured) natural images [1]. After the selection and the extraction of the image features[usually based on color and or texture and computed on (possibly) overlapping small windows centered around the pixel to be classified],the feature samples, handled as vectors, are grouped together in compact but well-separated clusters corresponding to each class of the image. The set of connected pixels belonging to each estimated class thus defined the different regions of the scene. The method known as k-means [2] (or Lloyd's algorithm).The applications of Image segmentation are widely in many fields such as image compression, image retrieval, object detection, image enhancement etc.

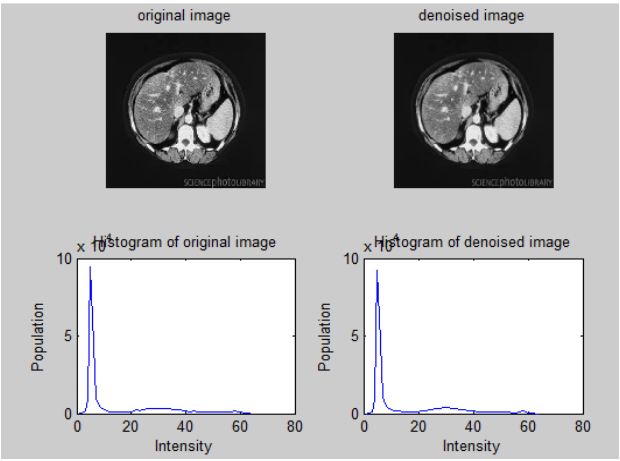


Fig. 1. Original and Denoised Image and its Histogram

II. LIVER SEGMENTATION

The main idea of the image segmentation is to group pixels in homogeneous regions and the usual approach to do this is by 'common feature. Features can be represented by the space of colour, texture and gray levels, each exploring similarities between pixels of a region. Segmentation [1] refers to the process of partitioning a digital image into multiple regions (sets of pixels). The goal of segmentation is to simplify and change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. The result of image segmentation is a set of regions that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region are similar with respect to some characteristic or computed property, such as color, intensity, or texture [3][4]. The segmentation is based on the measurements taken from the image and might be greylevel, colour, texture, depth or motion. Image segmentation techniques are categorized into three classes: Clustering, edge detection, region growing .Some popular clustering algorithms like k-means are often used in image segmentation [5] Adjacent regions are significantly different with respect to the same characteristic(s). Segmentation is mainly used in medical imaging,Face recognition,Fingerprint recognition,Traffic control systems, Brake light detection, and Machine vision

A. Histrogram

An image histogram is a type of histogram that acts as a graphical representation of the tonal distribution in a digital image. It plots the number of pixels for each tonal value. By looking at the histogram for a specific image a viewer will be able to judge the entire tonal distribution at a glance. Image histograms are present on many modern digital cameras. Photographers can use them as an aid to show the distribution of tones captured, and whether image detail has been lost to blown-out highlights or blacked-out shadows. The horizontal axis of the graph represents the tonal variations, while the vertical axis represents the number of pixels in that particular tone.]The left side of the horizontal axis represents the black and dark areas, the middle represents medium grey and the right hand side represents light and pure white areas. The vertical axis represents the size of the area that is captured in each one of these zones. Thus, the histogram for a very dark image will have the majority of its data points on the left side and center of the graph. Conversely, the histogram for a very bright image with few dark areas and/or shadows will have most of its data points on the right side and center of the graph distribution across the entire liver region in the CT scans.

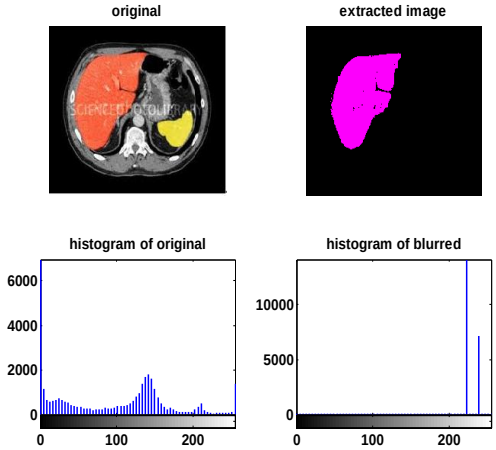


Fig. 2. Extraction Of Liver Image And Respective Histogram

B. Clustering

Clustering refers to the process of grouping samples so that the samples are similar within each group. The groups are called clusters. Clustering is a data mining technique used in statistical data analysis, data mining, pattern recognition, image analysis etc. Different clustering methods include hierarchical clustering which builds a hierarchy of clusters from individual elements. Because of its simplicity and efficiency, clustering approaches were one of the first techniques used for the segmentation of (textured) natural images [6].In partitional clustering; the goal is to create one set of clusters that partitions the data in to similar groups. Other methods of clustering are distance based according to which if two or more objects belonging to the same cluster are close according to a given distance, then it is called distance based clustering. In our work we have used K-means clustering approach for performing image segmentation using Matlab software. A good clustering method will produce high quality clusters with high intra-class similarity and low inter-class similarity. The quality of clustering result depends on both the similarity measure used by the method and its implementation. The quality of a clustering method is also measured by its ability to discover some or all of the hidden patterns. Image Segmentation is the basis of image analysis and understanding and a crucial part and an oldest and hardest problem of image processing. Clustering means classifying and distinguishing things that are provided with similar properties[17].Clustering techniques classifies the pixels with same characteristics into one cluster, thus forming different clusters according to coherence between pixels in a cluster. It is a method of unsupervised learning and a common technique for statistical data analysis used in many fields such as pattern recognition, image analysis and bioinformatics

C. Clustering techniques

Clustering is a data mining technique used in statistical data analysis, data mining, pattern recognition, image analysis etc. Different clustering methods include hierarchical clustering which builds a hierarchy of clusters from individual elements. Because of its simplicity and efficiency, clustering approaches were one of the first techniques used for the segmentation of (textured) natural images [4].In partitional clustering; the goal is to create one set of clusters that partitions the data in to similar groups. Other methods of clustering are distance based according to which if two or more objects belonging to the same cluster are close according to a given distance, then it called distance is based clustering. In our work we have used K-means clustering approach for performing image segmentation using Matlab software. A good clustering method will produce high quality clusters with high intra-class similarity and low inter-class similarity. The quality of clustering result depends on both the similarity measure used by the method and its implementation.The quality of a clustering method is also measured by its ability to discover some or all of the hidden patterns. Image Segmentation is the basis of image analysis and understanding and a crucial part and an oldest and hardest problem of image processing. Clustering means classifying and distinguishing things that are provided with similar properties [5]. Clustering techniques classifies the pixels with same characteristics into one cluster, thus forming different clusters according to coherence between pixels in a cluster.

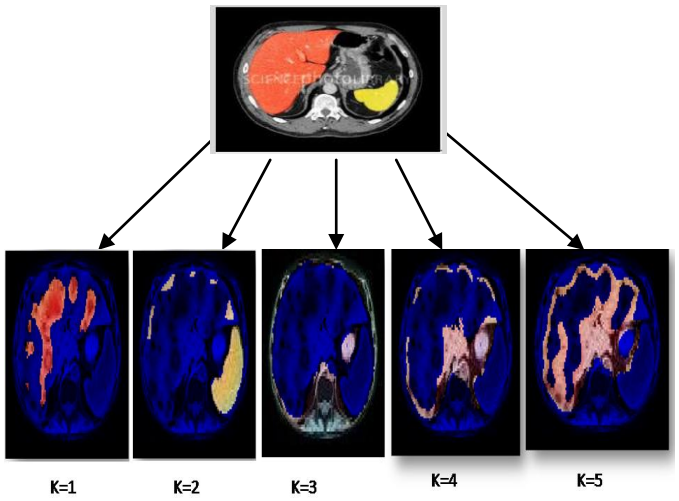
1) K-means clustering:

For the modified *k*-means approach adopted in the procedure segments the different regions of the CT slice around the user selected points. The user selected points act as the seeds for each of the five masks.The segmentation results using the *k*-means yield five masks namely M1 through M5 corresponding to the aforementioned five regions are depicted in Fig. 2. The masks liver (M1) and surrounding organs (M2) are logically ORed together to obtain the final mask (Mfinal). Based on empirical results, it was determined that the optimal mask would require a combination of the first two identified regions since in some cases the entire liver is not seen in maskM1 due to the inhomogeneous intensity. K-means is the clustering algorithm used to determine the natural spectral groupings present in a data set. K-Means algorithm is an unsupervised clustering algorithm that classifies the input data points into multiple classes based on their inherent distance from each other. The algorithm assumes that the data features form a vector space and tries to find natural clustering in them. The dataset is partitioned into K clusters and the data points are randomly assigned to the clusters resulting in clusters that have roughly the same number of data points. Clustering is a way to separate groups of objects. K-means clustering treats each object as having a location in space. It finds partitions such that objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible. K-means clustering requires that you specify the number of clusters to be partitioned and a distance metric to quantify how close two objects are to each other. K-means is used to cluster the objects into three clusters using the Euclidean distance metric. The most popular method for image segmentation is k-means clustering[6][7].

For each data point; the distance from the data point to each cluster is calculated. If the data point is closest to its own cluster, leave it where it is. If the data point is not closest to its own cluster, move it into the closest cluster. The above steps are repeated until a complete pass through all the data points results in no data point moving from one cluster to another. At this point the clusters are stable and the clustering process ends. The choice of initial partition can greatly affect the final clusters that result, in terms of inter-cluster and intra-cluster distances and cohesion. As K-means approach is iterative, it is computationally intensive and hence applied only to image subareas rather than to full scenes and can be treated as unsupervised training areas . The objective function

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

where $\left\| x_i^{(j)} - c_j \right\|^2$ 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 centres.



III. K-MEANS ALGORITHM

K-means is a clustering algorithm, which partitions a data set into clusters according to some defined distance measure. Images are considered as one of the most important medium of conveying information. Understanding images and extracting the information from them such that the information can be used for other tasks is an important aspect of Machine learning. An example of the same would be the use of images for navigation of robots. One of the first steps in direction of understanding images is to segment them and find out different objects in them. To do this, we look at the algorithm namely K-means clustering. It has been assumed 1that the number of segments in the image is known and hence can be passed to the algorithm. [19], [20]. K-Means algorithm is an unsupervised clustering algorithm that classifies the input data points into multiple classes based on their inherent distance from each other. The algorithm assumes that the data features form a vector space and tries to find natural clustering in them. The functions of k-means are as follows. IDX = kmeans(X,k) partitions the points in the n-by-p data matrix X into k clusters. This iterative partitioning minimizes the sum, over all clusters, of the within-cluster sums of point-to-cluster-centroid distances. Rows of X correspond to points, columns correspond to variables. Kmeans returns an n-by-1 vector IDX containing the cluster indices of each point. By default, kmeans uses squared Euclidean distances [8,9]. When X is a vector, kmeans treats it as an *n*-by-1 data matrix, regardless of its orientation.[IDX,C] = kmeans(X,k) returns the k cluster centroid locations in the k-by-p matrix C.

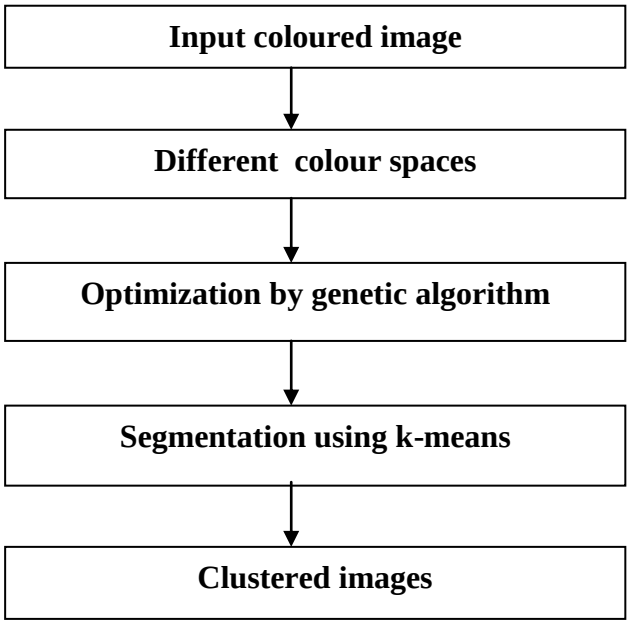


Fig 4. Flow Chart For Obtaining Clustered Image

IV. RESULT

Our segmentation method of liver was tested with ten abdomens (512x512 CT images) of ten patients. Each sample shows eight images that are the original image, the result of thresholding, application of median filter, the result of the detection of GCC, morphological closing, edge detection of the liver, the final result, manual segmentation of the expert. In that case of execution, the result obtained by our system seems to be medium by report has the manual segmentation of the expert. Apparently, it is impossible to obtain always a perfect result for the automatic segmentation. That's returns to various causes: in the images is different from a scanner to another, so the class of images is not the same and this poses a problem of automatic recognition as the nature of the histogram .

V. CONCLUSION

We have 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. Hence it is necessary to re-run the code several number of times for same and different values of k inoder to compare the quality of clusters obtained.

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].The advantage of K-Means algorithm is simple and quite efficient. 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, but does not ensure that the result has a global minimum of variance.

- The clarity in the segmented image is very good compared to other segmentation techniques.
- The clarity of the image also depends on the number of clusters used.

REFERENCES

[1] Alasdair Mc Andrew, "An introduction to digital image processing with MATLAB", school of computer science and mathematics,Victoria university of technology

[2] Gonzaled R. C., Woods R. E., Eddins S. L. "Digital Image Processing",3rd edition,Pearson Education (Singapore) Pvt. Ltd, 2009.

[3] Manjit Chitalpalii "Image segmentation by clustering using mahalanobis distance".

[4] Paoula campadelli, Elena Kaseraghi "Liver segmentation from computed tomography scans: A survey and a new algorithm", journal: Artificial intelligence in medicine, vol 45 issue 2-3 February 2009.

[5] Schmid, P. "Image segmentation by color clustering", http://www.schmid-saugeon.ch/publications.html, 2001.

[6] G. Milligan and M. Cooper, "An examination of procedures for determining the number of clusters in a data set," Psychometrika, vol. 5, pp.159–179, 1985.

[7] 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.

[8] T. Pappas "An adaptive clustering algorithm for image segmentation,"IEEE Transactions on Signal Processing, vol. 40, no. 4, pp. 901–914,1992.

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

