Lymph Node Image Segmentation Based on Improved FCM Clustering and Multi-threshold

Yanling Zhang, Yuejia Zhang

School of Computer, Guangzhou University

Guangzhou, 510006, China

zyl@gzhu.edu.cn

Li Li

Sun Yat-sen University Cancer Center

Guangzhou, 510080, China

li2@mail.sysu.edu.cn

Abstract—**The pathological change of lymph node is an important basis of malignant tumor detection and judgment of metastasis of cancer (lung cancer, colorectal cancer, breast cancer, liver cancer, cervical cancer, etc.) An algorithm of lymph node image segmentation based on improved FCM clustering and multi-threshold is proposed to segment the lymph CT image with blurred edge. First, the improved FCM peak clustering is used to sharpen the fuzzy boundary of lymph CT image effectively. Then the multi-threshold algorithm based on image entropy change is introduced to segment enhanced images. The experiment shows that the above algorithm can obtain better segmentation results compared with the traditional FCM clustering method in the case of the fuzzy edge of the lymph node tissue.**

Keywords- fuzzy C-means (FCM) peak algorithm ; multi-threshold algorithm；lymph node image segmentation

I. INTRODUCTION

Lymph nodes are an important organ of the human body immune response. The pathological change of lymph node is an important basis for detection and judgment of tumor metastasis of cancer (lung cancer, colorectal cancer, breast cancer, liver cancer, cervical cancer, etc.) [1]. So the lymph image quality is very important for the doctor to do pathological analysis. However, because of the very low contrast of various organs and tissue of the body, the similarity between lesion area and the structure of the human organ and tissue, the obtained lymph image has the problem of blurry edge, low contrast, noise, etc. in some degree. Therefore, it is necessary for lymph node image to adopt some methods to obtain clear edge information to improve the accuracy of the doctor's diagnosis.

There are a lot of image segmentation algorithms, such as the classic edge extraction operator: Roberts, Sobel, Prewitt, LOG, Canny, etc. These operators are simpler, faster and more sensitive to noise interference [2][3].The lymph node segmentation method based on the color attributes and fuzzy C-means was proposed by David N. Olivieri in 2000 [4]. Zhang Junhua and Wang Yuanyuan introduced a segmentation algorithm based on edge flow of improved gradient vector flow (GVF) to segment lymph node ultrasound images [5]. Comprehensive Utilization of gray and texture features and regional restrictions from the lymph node characteristics improved segmentation result of ultrasound image with low contrast significantly. In 2007, Adam Kapelner, Peter P. Lee and Susan Holmes of Stanford University presented an interactive statistical image segmentation method to split the lymph nodes [6]. In October 2009, Liu Lu and Liu Wanyu adopted interactive segmentation to extract lymph nodes from CT images [7]. Zhang Lixi and Liu Binghan presented the image segmentation method based on texture spectrum characteristics mixed gray to isolate lymphoid nodules, lymphocytes effectively [8]

In this paper, the pelvic lymph CT image is studied as an example. CT images can better display soft tissue organs such as the lungs, liver, gallbladder, pancreas, and pelvic organs, and show lesions images under the background of good anatomical image. For the fuzziness and uncertainty of the edge, the shape and size of lymph nodes, the improved FCM clustering is used to sharpen blurry edges and the multi-threshold method is proposed to segment enhanced images.

II. FCM PEAK CLUSTERING

A. FCM clustering algorithm

Fuzzy C-Means clustering algorithm [9] (FCM, Fuzzy C-Means) has good adaptability to fuzzy boundaries, it can be used in image enhancement, but traditional FCM algorithm needs to determine the number n of cluster centers empirically, then to iterate it. When the objective function meets certain requirements, the algorithm terminates.

Fuzzy C-Means clustering algorithm based on the mini-cost function is achieved by the objective function of equation (1):

$$J(U,V) = \sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}^{m} \| x_j - v_i \|^2 \tag{1}$$

Where, U is the membership function from each sample to the cluster centers, V is a collection of the cluster center, c is the number of cluster centers, n is the number of samples, μ_{ij} is the membership from the j sample to the i class, v_i is a cluster center, $m \geq 0$ is the fuzzy weighted index.

After determining the cluster center, the equation (2) is used to recalculate the degree of membership μ_{ij} when the samples processed is not in the last position of cluster center:

$$\mu_{ij} = \left(\sum_{k=1}^{c} \left(\frac{\|x_j - v_i\|}{\|x_j - v_k\|} \right)^{\frac{1}{m-2}} \right)^{-1}, j = 1,2,3,...,c$$

(2)

The equation (3) is used to recalculate the degree of membership μ_{ij} when the samples processed are in the last position of cluster center:

$$\mu_{ij} = 1, \mu_{ik} = 0, k \neq j, i = 1,2,...,n$$

（3）

And then perform iterative formula (4), to recalculate the value of the cluster center:

$$v_i = \frac{\sum_{j=1}^{n} \mu_{ij}^m x_j}{\sum_{j=1}^{n} \mu_{ij}^m}, i = 1,2,\cdots,c$$

(4)

This algorithm is very sensitive to the selection of the initial value of the cluster center, the result of the algorithm is directly related to the initial position and the number of the cluster centers and weighted index, and is easy to fall into local minimum value.

B. FCM peak clustering

Peak method is the processing method based on the image histogram or frequency plot. the peak of image frequency plot or histogram is swept each time, then the both sides of peak are scanned till the nearest trough to do a trough symmetrical mapping, after processing scanned area, a new histogram is produced, and then the new a histogram is done the same processing, continuous iteration is performed until the peak is under a certain threshold value, or the number n of cluster centers is obtained. The advantage of such a method is that it can find pixel values with continuous peaks in the gradation region. Peak method is applied to the initialize cluster centers in this article. Assuming gray histogram as a function of $y = f(x)$, concrete steps are as follows:

Step 1: find the highest point of the gray histogram (x_i, y_i). When y_i is greater than the threshold and the labeled cluster center are less than n, x_i is marked as cluster centers. Otherwise, the algorithm ends.

Step 2: scan the left and right sides of x_i, find the nearest trough x_j, assuming $a = |x_i - x_j|$, then the value of y in the interval of $[x_i - a, x_i + a]$ in the gray histogram is set 0, the gray histogram is renewed, return step 1.
Initial cluster centers calculated by peak method are shown in Figure 1:
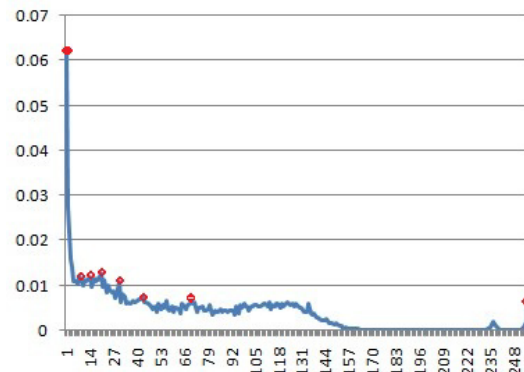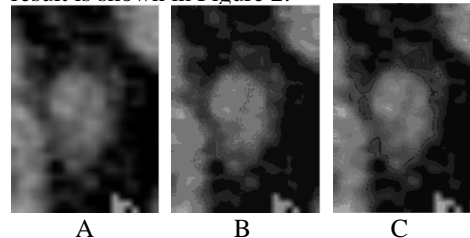

Figure 1 initial cluster centers obtained by peak

a number of initial cluster centers are identified by peak method (red dots in Figure 1), which are the highest points of the histogram. A threshold value is needed to set, the calculation is stopped when the highest points of the histogram is smaller than this threshold. The processed result is shown in Figure 2:

| A | B | C |

(A is the original image, B is the traditional FCM clustering, C is peak method FCM clustering)
Figure 2: FCM peak clustering
The representative and uniformly distributed initial cluster centers can be obtained by the peak method; Shortcomings caused by the cluster centers determined by the experience in the traditional FCM are removed. In Figure 2, image contrast is significantly enhanced by the FCM peak clustering method; the fuzzy edge of the lymph nodes is sharpened.

III. MULTI-THRESHOLD SEGMENTATION METHOD

Traditional maximum class variance (OTSU) algorithm divides images into two categories: one is the target, the other is the background, this will make the characters subdivided in two categories disappear and be not beneficial to extract features. Therefore, a new algorithm is needed to further subdivide these two classes. Multi-threshold method uses OTSU to divide images into two categories, and then split them again till obtain optimal threshold. In the multi-threshold iteration, entropy is used to describe the object and the background instead of the gray value because the entropy can better reflect the distribution of image gray uniformity [10].

Entropy is a measure of the uncertainty, the image is divided into two regions by the gray gate threshold t, i.e. the target and the background is divided into two source entropy respectively in the image processing:
Target source entropy:

$$H_A = -\sum_{i=0}^{t} p_i \log p_i \quad （5）$$

Background information source entropy:

$$H_B = -\sum_{i=t+1}^{L-1} p_i \log p_i \quad （6）$$

Entire entropy:

$$H_0 = -\sum_{i=0}^{L-1} p_i \log p_i \quad （7）$$

Where, p_i is the probability of image gradation level，L is gradation。

In this paper, the gray level and entropy change of the image is used to determine the local conditions of the recursive segmentation termination. Concrete methods are as follows:

a) Calculate the mean ω_0 and the entropy H_0 of the entire image;

b) calculate the first segmentation threshold t of the entire image using OTSU；
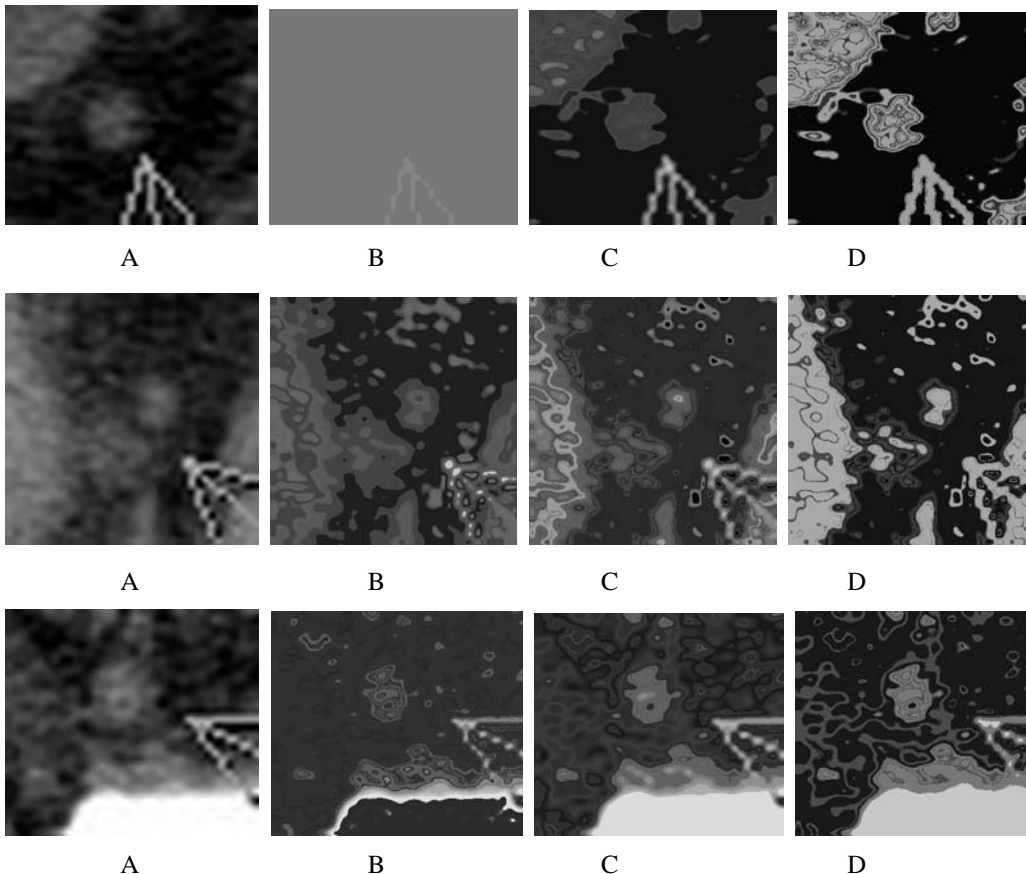
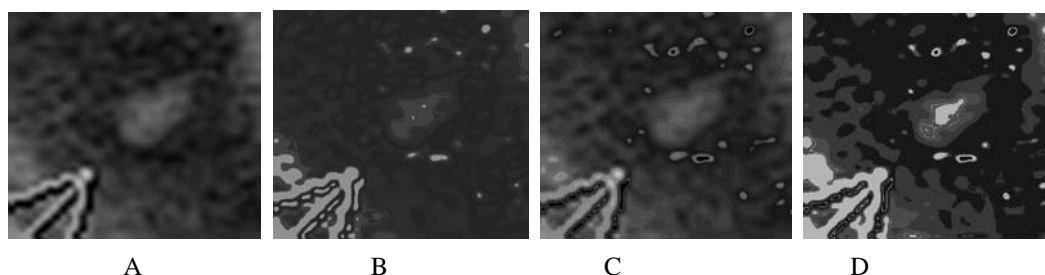c) According to t, calculate mean and entropy of those image pixels whose gray is greater than the threshold value ,

Assuming $\omega_1 =$ gray mean of those image pixels whose gray is greater than the threshold value, $H_1 =$ entropy of those image pixels whose gray is greater than the threshold value.

If $(\omega_1 - \omega_0)/\omega_0 > 0.2$ and $(H_1 - H_0)/H_0 > 0.1$, then recursion will go on, otherwise, the recursion will end. Local recursive segmentation uses the equation (8) to calculate the optimum threshold value.

The experimental results are shown in Figure 3:

$$t^* = Arg \underset{0 \le t \le L-1}{Max}[(H_A - H_0)^2 (H_B - H_0)^2 /((H_A - H_0)^2 + (H_B - H_0)^2)^2] \quad (8)$$

A   B   C   D

A   B   C   D

A   B   C   D

| A | B | C | D |

(A is the original image, B is the traditional FCM, C is FCM peak clustering, D is FCM peak clustering and multi-threshold)

Figure 3 Comparison of segmentation algorithms

The experiments show that the clearer boundary can be obtained by using FCM peak clustering and multi-threshold to segment lymph nodes image with blurry edge.

IV. CONCLUSION

In this paper, the method of combining FCM peak clustering with multi-threshold is proposed to segment lymph nodes image with blurry edge. FCM peak Clustering method can get the typical uniformly distributed initial cluster centers, and multi-threshold segmentation method based on image entropy change has better self-adaptability. The experimental results show that this method significantly enhances image contrast, sharpens fuzzy edge of the lymph nodes image, get a clearer image boundary.

Sometimes, the phenomenon that cluster centers value appears repeat ably after a number of iterations is found in the experiment of calculating FCM clustering centers using the peak method , but it does not affect the clustering effect. Multi-threshold segmentation method based on image entropy change may obtain undesirable results in the case of low signal-to-noise ratio of the image, and need to be improved later.

V. ACKNOWLEDGMENT

REFERENCES

[1] Zhou Yong-chang, Guo Wan-xue. Ultrasonic Iatrology[M]. Bei Jing, Publishing House of Technology Literature, 2003:   363-373

[2] Gonzalez R C, Woods R E. Image Processing,Second Edition[M].Ruan Qiu-qi,Ruan Yu-zhi.Beijing, Publishing House of Electronics Industry, 2003: 463-474.

[3] Canny J.A computational approach to edge detection[J]. IEEE PAMI, 1986, 8(6) : 679-698.

[4] David N. Olivieri，Francisco Vega. Image Prototype Similarity Matching for Lymph Node Hemopathology. Proceedings of the International Conference on Pattern Recognition (ICPR'00) [C]. 2000:1051-4651

[5] Zhang Jun-hua, WangYuan-yuan. Analysis and application for sonographic image of cervical lymph node[D]. [doctor dissertation]. Fu Dan University, 2007: 5-24.

[6] Adam Kapelner, Peter P. Lee, Susan Holmes. An Interactive Statistical Image Segmentation and Visualization System. International Conference on Medical Information Visualisation-Biomedical Visualisation (MediVis) [C]., 2007: 4-6

[7] Liu Lu, Liu Wan-yu. Classification of Tumid Lymph Nodes Metastases and Non-Metastases from Lung Cancer in CT   Image[J]. Journal of Electronics & Information Technology, 2009,31(10):2476-2482.

[8] Zhang Li-xi, Liu Bing-han. Auto-segmentation of the lymphoid tissue structure color pathological Images[J]. Journal of Ji   mei University(Natural Science), 2007,12(1):73-77.

[9] Liu Ruijie, Zhang jinbo. Fuzzy c-means clustering algorithm [J]. Journal of Chongqing Institute of Technology.2008.22(2): 139-141.

[10] Xiao Chaoyun, Zhu Weixing. Threshold Selection Algorithm for Image Segmentation Based on Otsu Rule and Image Entropy [J]. Computer Engineering. 2007, 33(14): 188-209.