A NOVEL METHOD FOR MEDICAL ULTRASOUND IMAGE SEGMENTATION

Mehmet K. Kocamaz and Daniel Freedman

Rennsselaer Polytechnic Institute

ABSTRACT

Segmentation of ultrasound images is difficult due to discontinuity and uncertainty of segmentation boundaries caused by speckle noise. Most of the ultrasound image segmentation algorithms require high computational time and do not give good results. In this paper, a novel algorithm for the ultrasound image segmentation is proposed. Our approach is based on semi-automatic classical region growing segmentation algorithm. Given a seed point by user, one can automatically and quickly segment the object from the image. We developed a mechanism to add groups of pixels to the segmented object. Our growing mechanism uses the distances between the distributions as the criteria to solve the speckle noise problem in the ultrasound images. Minimizing the distances between the distributions allow us to find the borders. The performance and the results of the proposed method are tested in different set of ultrasound images.

Index Terms— Ultrasound image segmentation, medical image segmentation, region growing, matching distributions.

1. INTRODUCTION

Ultrasound image segmentation is very important, difficult and useful in the medical area. It is important and useful because of its usage in the diagnosis and the treatment applications. For example, to use radiotherapy techniques for a thyroid cancer patient, the first step is to measure the volume of the thyroid from the given medical image. Then, according to the size of the thyroid, the amount of the radiation is going to be determined and applied. On the other hand, ultrasound image segmentation is difficult because of its some characteristic features. In the computer vision area, each of these features can be called as a challenge and some of them are followings:

- Challenge 1: Ultrasound images contain speckle noise. The inside or outside of the interested objects fairly consist of homogenous regions. There are big intensity differences among neighbor pixels. They do not have similar intensity profiles.

- Challenge 2: There are not sharp intensity changes in the boundaries. The speckle distribution changes separate both the outside and the inside of the interested object.

- Challenge 3: The boundary pixels are often not connected to each other with small intensity differences. It causes the leaks on the boundaries.

These challenges make the most of the image segmentation methods unreliable for the ultrasound images. For example, active contour model based algorithms fail to segment ultrasound images due to challenge 1 and 3. The most famous active contour segmentation method is called snakes [1]. Snakes algorithm requires the initialization of the first contour points. Then, it finds the boundary of the region. Due to speckle noise on the boundaries, snake algorithm fails for the ultrasound images. Region growing algorithm is another segmentation method. It first takes a seed point from the user. It measures intensity differences between two neighbor pixels, if the difference is smaller than a constant; it includes that pixel to the segmented region. High intensity differences among neighbor pixels cripple the regular region growing algorithm. Other segmentation algorithms, such as live-wire [2], random-walker [3] and level-sets [4] and automatic algorithms such as, k-means and graph-cuts [5] methods fail in the ultrasound images, mainly because of the problems caused by the speckle noise.

We propose a novel semi-automatic segmentation algorithm by combining region growing and matching distributions methods [6]. Our algorithm starts by taking one initial seed point from the user. It draws a small patch around this seed point and retrieve intensity distribution map of this patch. By taking this patch's distribution map as the initial model, the segmented region tries to be grown. In each iteration, a point on the border is chosen and a certain number of its neighbor pixels are taken as a test model from outside of the segmented region. The intensity distribution maps of our initial model, the map of the test model and the model which consists of the neighbor pixels to the test model on the border are compared. The most similar test model which lies near the border is added to the segmented object. Similarity between two models is measured according to intensity distribution maps. This procedure is repeated until no similar test models to the segmented region exist. Finally, the desired segmentation region that contains our initial seed point is found. We test our algorithm with

some ultrasound images and the results indicate superior performance of our method.

The remainder of this paper is organized as follows; we review similarity calculation techniques of the distribution maps in Section 2. Section 3 is the explanation of the technical details of our algorithm. Our experimental results of some ultrasound images are discussed in Section 4. Finally, conclusion remarks are in Section 5.

2. DISCUSSION ON THE DISTANCE FUNCTIONS BETWEEN TWO DISTRIBUTION MAPS

Given a model M, the distribution of the intensities inside of M can be defined as a function

$$P_i^M = \frac{\int\limits_{x \in M} F(I(x),i)dx}{\int\limits_{x \in M} dx} \qquad (1)$$

where $F(I(x),i)$ is 1, if $I(x) = i$ and 0 otherwise [6]. The numerator of this equation gives us the frequency of the each intensity in the model M. The denominator, which is the total pixels in the model M, normalizes the equation.

If there are two models model $M1$ and model $M2$ to measure the dissimilarity or distance between $M1$ and $M2$, K is the value of the distance between two models; we can construct a new function D which is defined as

$$K = D(P^{M1}, P^{M2}). \qquad (2)$$

In our segmentation algorithm, new pixels are added to segmented region as groups, so the distances between the models need to be measured. To calculate K in equation (2), there are several proposed mathematical techniques. Kullback-Leibler divergence

$$D(M1, M2) = \sum_{i=1}^{n} M1_i \log \frac{M1_i}{M2_i} \qquad (3)$$

is one of them. If we assume that intensity values range between 1 and 256 in gray level images, so n is 256 in the equation (3). Another function is derived from Bhattacharyya coefficient and it is formulated as

$$D(M1, M2) = 1 - \sum_{i=1}^{n} \sqrt{M1_i M2_i}. \qquad (4)$$

Both of these two distance functions have some problems to measure the distances between the intensity distributions of the ultrasound images. For example, let's say we have three gray level images $I1$, $I2$ and $I3$. $I1$ consists of all white pixels, $I1(i)$ is all 255, $I2$ consists of the pixels where each pixel has a value of 254, $I2(i)$ is all 254, and the last image $I3$ is black, $I3(i)$ is all 0. According to Kullback-Leibler divergence, which is stated in equation (3), the distance between images $I1$ and $I2$; and between the images $I1$ and $I3$ is $D(P^{I1}, P^{I2}) = D(P^{I1}, P^{I3}) = \infty$. However,

image $I2$ is very similar to image $I1$ in real sense and image $I1$ is very different than image $I3$. As seen in this example, Kullback-Leibler and Bhattacharyya formulations are not enough to solve the distance problem between the distributions in some cases. Since the ultrasound images contain highly speckle noise, we need another formulation to overcome this problem.

Cumulative Distribution Functions (cdfs) [7] is another distance measure method which takes into account the values of the random variables. For given two distribution models $M1$ and $M2$. Let's say, $M1_i = \sum_{j=1}^{i} M1_j$ and

$M2_i = \sum_{j=1}^{i} M2_j$ .Then the cdfs distance between model

$M1$ and $M2$ is as below:

$$D(M1, M2) = \sum_{i=1}^{n} |M1_i - M2_i| \qquad (5)$$

The cdfs distance calculation is more effective for our ultrasound image models. To see it, let's return back to our example three images $I1$, $I2$, and $I3$. By using cdfs, the distance between white image $I1$ and almost white image $I2$ is $D(P^{I1}, P^{I2}) = 1$, the distance between white and black image is $D(P^{I1}, P^{I3}) = 255$, and the distance between almost white image and black image is $D(P^{I2}, P^{I3}) = 254$. Our visual sense affirms that the cdfs distances on these three example images are true. Instead of using Kullback-Leibler or Bhattacharyya distance, because of its more reliable measurements, we use cdfs distance for the calculation of similarities between two models in our segmentation algorithm.

3. EXPOSITION OF OUR ALGORITHM

This section describes three main concepts of our algorithm. They are: Seed point selection and the block region growing, best test model selection function, and the segmentation algorithm.

3.1. Seed Point Selection and Group Region Growing

All region growing methods start from the given inital point by the user, and continue by adding new pixels to the segmented region if the distance between two pixels is smaller than a constant. In our approach, groups of the pixels are occupied from outside of the segmented region by comparing the distances between distribution maps. Adding group of pixels allows us the elimination of the speckle noise and weak boundary problems in the ultrasound images. Our initial model is a patch that includes the user given seed point. It might be chosen as a rectangle or a circle with a

constant number of pixels. We call the group of pixels which is going to be joined to the segmented region as test model. Its main attributes are first, its all pixels are connected to each other, and second, they are all neighbors of the segmented region. Particularly, they just lie near the border of the growing region.

Comparing the distances only between the initial model and test models does not help us in some circumstances. For example, if there is continuous increase in distances between the initial and the test models while the segmented region is growing and moving away from the seed point, the growing algorithm stops before it reaches the border. There are two ways to solve this problem. First way is to update the initial model distribution map each time when a test model is added to the segmented region. However, this approach causes high computational cost. It requires to re-run the algorithm for the border marked pixels after each update of the initial distribution map. Second approach is to introduce a new model for the distance measurement. We call our new model as the neighbor model. Neighbor model consists of the pixels which lie on the border and each pixel is neighbor of the test model. In this approach, the distance between the test and neighbor model is taking into account as well as the distance between the initial and test model. We choose to use the second way, because of the low computational cost comparing to first approach. Figure 1 illustrates the initial, test and neighbor models.
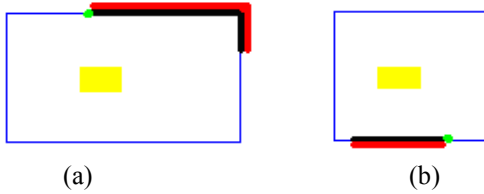

(a)   (b)

Fig.1. Examples of the initial, test, and neighbor models. Blue rectangles are segmented region borders. Yellow pixels are corresponds to the initial model. The green dot indicates a test pixel on the border. Red pixels are test model for the green pixel. They lie outside of the border. The black pixels are the neighbor model. The black pixels lie on the border of the segmented region. The size of the test and neighbor models differ in (a) and (b).

3.2. Best Test Model Selection Function

Regular region growing algorithm uses the method of comparison intensities between two pixels, which is stated as

$$|I(x) - I(y)| < \theta \qquad (6)$$

where $I(x)$ is the intensity of a point on the border of the growing object and $I(y)$ is the intensity of a neighbor point of x, and θ is a constant for the testing function. The pixel y belongs to the object, if the equation (6) holds. This function works well for one pixel growing algorithms. However, our algorithm is based on the block level growing, so we cannot use equation (6). Instead of regular approach, we develop a new block selection function and it is:

$$F(T,M,N) = \partial(D(T,M)) + (1-\partial)(D(T,N)) . \quad (7)$$

$D(T,M)$ is the distance between the initial model M, and the test model T. $D(T,N)$ is the distance between test model T, and neighbor model N. Fig.1 shows us the detail descriptions of the models. The distances are calculated using the formula (5). Lastly, ∂ is the weight parameter to stabilize the growing mechanism on the border of the object. ∂'s range changes between 1 and 0.

In (7), $D(T,M)$ allows us to see the dissimilarity between the user selected initial model and the test model. $D(T,N)$ indicates the dissimilarity just near the border of the object. The combination of these two distances predicts for us whether or not a block of pixels belongs to an object.

Finally, to decide whether or not the test model T is inside of the object, regular approach in the region growing algorithms is used. The equation

$$F(T,M,N) < \theta \qquad (8)$$

is tested. θ is a constant for the test and $F(T,M,N)$ is the function as defined in (7). Formula (8) indicates us that if $F(T,M,N)$ is smaller than θ, then the test model T belongs to the segmented object.

3.3. The Segmentation Algorithm

The detailed description of our segmentation algorithm is in Figure 2.

M, T and N are the models, s is a seed point, and b is a point on the border of the growing region. *MAP* is a map data structure to hold the given models with a mapped value.

Step 1. Obtain a seed point s from the user, take an initial model M around s, and calculate its intensity distribution map using equation (5).

Step 2. For each unvisited point b on the growing region border,

    Step 2.a.1. Take k number of points that are connected to each other and to the point b from the outside of the segmented region. Call these point as model T.

    Step 2.a.2 Find model T's neighbors on the border. Construct model N from these neighbor points.

    Step 2.a.3 Calculate $F(T,M,N)$ as indicated in the equation (7). If equation (8) holds, put test model T with its corresponding $F(T,M,N)$ distance value to the *MAP*.

    Step 2.a.4. Decrease the value of k and do **Step 2.a.1 - 2.a.3** again.

    Step 2.b. Choose the best model T with its smallest $F(T,M,N)$ value from the *MAP*. If *MAP* is empty, label b as visited and continue on **Step 2**.

    Step 2.c. Add model T to the segmented region and make the *MAP* empty.

Fig.2. The Segmentation Algorithm

As it can be seen in Step2.a in our algorithm, for each border point we are trying to find best similar test model. The aim of this step is that to minimize equation (7) for each border point and to add best similar block of pixels to the segmented region. The algorithm stops when all border points are labeled as visited.
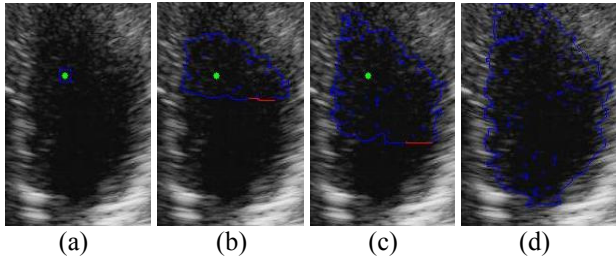

| (a) | (b) | (c) | (d) |

Fig.3. Illustration of our algorithm on the left ventricle ultrasound image while the segmented region is growing. User selected seed point is shown as green. Red pixels in the images indicate the new regions that are added to the object space in the each iteration.

4. ALGORITHMIC RESULTS

We present our experimental results on three image sets, ventricle, thyroid and follicle ultrasound images, displayed in Figures 3 and 4. Figure 3 illustrates the growing of the segmented region on the left ventricle while our algorithm is running. Initial model is taken around the seed point as a rectangle in all test images. Our test results might contain small areas which usually consist of just one pixel inside of and do not belong to the segmented regions. This is because of the best test model selection in our algorithm. There might be one pixel on the border of the segmented region which has very unusual intensity that maximizes the distance between the models, so that unusual pixel is not taken to the segmented region.
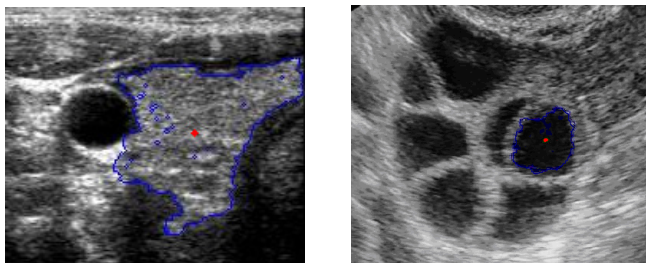

Fig3: Examples of the segmentations of the thyroid and a follicle. (a) is a thyroid of a patient and (b) is a follicle. Initial seeds are red points in both images.

We used a Pentium4 2ghz CPU and 512mb memory to test our segmentation algorithm. The implementation of the algorithm is done in Matlab. Even though we do not use a pre-compiled language to implement our algorithm, it is fast enough to segment the ultrasound images which contain highly speckle noise. Left ventricle segmentation in figure 3 takes just 15 seconds. Thyroid and follicle segmentation in figure 3 took 38 seconds and 18 seconds, respectively. Two constants are $\partial = 0.10$ in equation (7) and $\theta = 30$ in equation (8) in all test cases.

5. CONCLUSION

We have proposed a novel segmentation algorithm for the ultrasound images based on a user selected seed point. Our algorithm starts with initial small model and tries to expand this model until it reaches the border of the object. While growing, it allocates best similar pixels from image to the object space. Instead of one pixel growing as in regular region growing algorithms, our algorithm grows by adding block of pixels to the segmented region. Distance measurements between the distributions help us to identify the blocks that are similar to the object.The proposed algorithm in this paper depends on two free parameters (∂ and θ) defined in equation (7) and (8). We have demonstrated our approach on real ultrasound images and shown that it provides quality, robust and fast segmentation results.

6. REFERENCES

[1] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active Contour Models. International Journal of Computer Vision, pages 321–331, 1988.

[2] W. A. Barrett and E. N. Mortensen. Interactive live-wire boundary extraction. MedicalImage Analysis, 1(4):331–341, 1997.

[3] Leo Grady and Gareth Funka-Lea, "Multi-Label Image Segmentation for Medical Applications Based on Graph-Theoretic Electrical Potentials", in Proceedings of the 8th ECCV04, Workshop on Computer Vision Approaches to Medical Image Analysis and Mathematical Methods in Biomedical Image Analysis, p. 230-245, May 15th, 2004, Prague, Czech Republic, Springer-Verlag.

[4] M. Rumpf and A. Telea. A continuous skeletonization method based on level sets. In Joint EUROGRAPHICS-IEEE TCVG Symposium on Visualization, pages 151-157, 2002.

[5] Y. Boykov and M.-P. Jolly. Interactive organ segmentation using graph cuts. In Medical Image Computing and Computer-Assisted Intervention, pages 276–286, 2000.

[6] D.Freedman, R. J. Radke, T. Zhang, Y. Jeong, D. M. Lovelock, and G. T. Y. Chen. Model-Based segmentation of medical imagery by matching distributions.

[7] Yossi Rubner, Carlo Tomasi, and Leonidas J. Guibas. A Metric for Distributions with Applications to Image Databases. Proceedings of the 1998 IEEE International Conference on Computer Vision, Bombay, India, January 1998, pages 59-66