3D GENERAL LESION SEGMENTATION IN CT

Marie-Pierre Jolly and Leo Grady

Siemens Corporate Research
Imaging and Visualization Department
Princeton, NJ, USA
marie-pierre.jolly,leo.grady@siemens.com

ABSTRACT

This paper describes a general purpose algorithm to segment any kind of lesions in CT images. The algorithm expects a click or a stroke inside the lesion from the user and learns gray level properties on the fly. It then uses the random walker algorithm and combines multiple 2D segmentation results to produce the final 3D segmentation of the lesion. Quantitative evaluation on 293 lesions demonstrates that the method is ready for clinical use.

Index Terms— Cancer Imaging, Lesion Segmentation

1. INTRODUCTION

Cancer is the second leading cause of death in the United States as half a million people die each year. For this reason, it is important to diagnose the disease as early as possible and monitor it carefully as it is being treated. In addition to primary tumors, physicians are also interested in secondary tumors that might have metastasized through the rest of the body. The main areas of concern are the liver (because many forms of cancer generate liver metastases) and lymph nodes (because cancers spread through the lymphatic system).

The current standards to measure lesions for cancer monitoring are the WHO and RECIST criteria. The first criterion was proposed in 1979 by the World Health Organization [1] and measures a tumor in two dimensions by its maximum diameter in an axial image multiplied by its largest perpendicular diameter in the same image. In 2000, the Response Evaluation Criteria in Solid Tumor (RECIST) [2] stated that the maximum diameter in an axial image alone could be used to quantify tumor sizes. It has been shown however that volume measurements provide more accurate estimates of the lesion sizes than 1- and 2-D criteria [3]. Unfortunately, there are no good tools available to measure lesion volumes and it is very time consuming to manually outline them in 3D. In this paper, we propose a semi-automatic segmentation algorithm which, from a single click inside or a stroke through the lesion, can produce its 3D segmentation.

Most of the work on medical analysis for cancer screening and treatment has been in the context of lung cancer, mammography, and colon cancer. There has been very little interest in liver lesion segmentation [4, 5]. The work has focused on relatively simple hypodense lesions which are nicely contrasted against the parenchyma. Papers have presented simple image processing techniques which have been tested on few examples. Of note is the work in [6] which uses AdaBoost to learn 1-D profiles on 30 examples and segment 30 lesions in 2D. Unfortunately, no validation is presented. Lymph node segmentation has received more attention in recent years (see [7, 8] and references herein) with more mature techniques that have been tested on larger number of examples. The main disadvantage with techniques published so far is that they focus on one type of lesions and are not general. In clinical cases however, radiologists might want to look at primary tumors, some liver metastases, and the lymphatic system. Therefore, it is very important to provide a single tool capable of segmenting any type of lesion. This segmentation task presents various difficulties however. First, sizes and shapes vary significantly among lesions. In addition, the lesion itself can be highly heterogeneous and contain calcifications or necrotic regions. Finally, lesions and in particular lymph nodes are adjacent to iso-intensity soft tissues and high contrast structures such as blood vessels. In this work, we focus primarily on liver lesions (tumors and metastases) and lymph nodes, but have also applied the technique to a large number of other lesions.

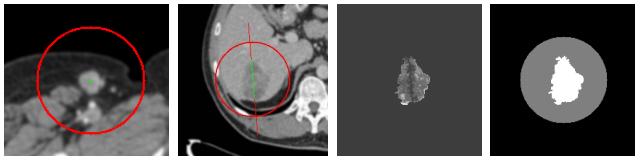
2. SEGMENTATION ALGORITHM

Our algorithm is divided into three steps. First, the system learns a rough gray level distribution for the lesion using the clicked point or the stroke. Second, the lesion is segmented on the 2D plane on which the user interacted, as well as on other 2D planes orthogonal to it. After each 2D segmentation, the gray level distributions for the lesion and the background are updated. Third, the random walker algorithm is used to produce the 3D segmentation.

2.1. Gray Level Distribution Estimation

Since lesions can be of any size or shape, we decided to rely solely on an intensity model for the lesion. However, lesions

are all different, so it is crucial to learn an intensity model for the lesion currently being segmented. The fuzzy connectedness algorithm [9] computes the cost of the best path between a set of seeds and every pixel in the image. In a sense, it provides a measure of how strongly a pixel belongs to the regions represented by the seeds. The region seeds are defined by the click point or the pixels on the stroke while the background seeds are placed along a circle and in the prolongation of the stroke (see Fig. 1(a)-(b)).



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

Fig. 1. Multiseeded fuzzy connectedness: (a) Click point seeds (green for region, red for background); (b) Stroke seeds (green for region, red for background); (c) Lesion cost $C_L(x,y)$; (d) Labels $L(x,y)$.

We bring two modifications to the traditional multi-region fuzzy connectedness segmentation. First, we use the multiseeded technique proposed by Herman and Carvalho [10] where the regions compete as the paths are being built. We also use a different cost function. In the traditional fuzzy connectedness approach, one might use the mean and variance of the region gray levels to prevent a slow varying path to pass through regions with different statistics. In our situation, these statistics are unknown and cannot be assumed to be Gaussian. So, instead, the cost function is the range of gray level intensities along the path. We implement fuzzy connectedness using Dijkstra's algorithm with a heap to keep track of the current best path. At every pixel (x,y), we store the minimum intensity $F^m(x,y)$, the maximum intensity $F^M(x,y)$ and the cost $C(x,y)$ along the current best path to that pixel. The information at its neighbor (i,j) is updated by:

$$F^m(i,j) = \min\{F^m(x,y), f(i,j)\},$$
$$F^M(i,j) = \max\{F^M(x,y), f(i,j)\},$$
$$C(i,j) = F^M(i,j) - F^m(i,j)$$

where $f(i,j)$ is the gray level at pixel (i,j). This cost function favors paths that do not vary too much in gray level and therefore stay within one homogeneous region. Heterogeneous lesions can still be recovered when each seed point on the stroke builds its own homogeneous region which, all put together, form a larger heterogeneous region.

The multiseeded algorithm from [10] outputs the cost image $C_L(x,y)$ which is the cost of reaching pixel (x,y) from a lesion seed, as well as a label image $L(x,y) = 1$ if pixel (x,y) belongs to the lesion and 0 otherwise (see Fig. 1(c)-(d)). The lesion histogram H_L is built by adding the gray value of every lesion labeled pixel a number of times inversely proportional to the cost at that pixel (i.e.: $100/C_L(x,y)$).

This method is not infallible and it could happen that an unknown region (belonging neither to the lesion nor to the background) gets labeled as lesion (see the rib in Fig. 2). To eliminate the wrong elements that might have been introduced in the lesion histogram, all modes in the histogram that are not sufficiently represented in the stroke are deleted.

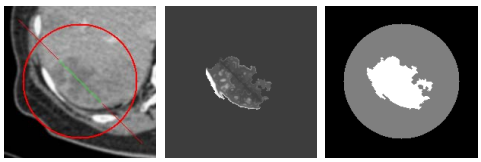


Fig. 2. Mislabeled pixels from the multiseeded fuzzy connectedness.

2.2. 2D Segmentation

Using the lesion statistics, the response from the histogram $H_L(f(x,y))$ is evaluated at every pixel and combined with the cost function $C_L(x,y)$ from the fuzzy connectedness to compute the response image $g(x,y)$:

$$g(x,y) = \begin{cases} H_L(f(x,y))\left(1 - \frac{C_L(x,y)}{C_L^{\max}}\right) & \text{if } L(x,y) = 1, \\ H_L(f(x,y))/2 & \text{if } L(x,y) = 0. \end{cases}$$

$g(x,y)$ emphasizes the difference between the pixels inside the lesion (which respond well to the lesion histogram and have low cost) and the pixels outside the lesion. The image is converted to log-polar coordinates, its gradient is computed (using a 3×21 horizontal mask), and the 2D segmentation is recovered using a shortest path algorithm.

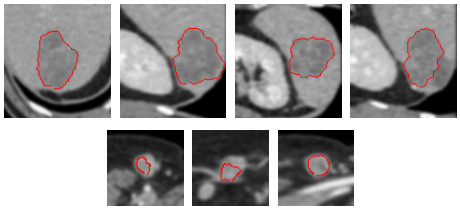


Fig. 3. Segmentations in the various 2D planes (4 planes in the case of a stroke and 3 planes in the case of a click point).

Once the lesion has been segmented in the interaction plane, pixels inside the lesion are used to update the lesion histogram H_L and pixels within 10 pixels outside the lesion are used to generate the background histogram H_B.

Additional 2D contours are recovered on other planes as follows. In the case of a clicked point, the two other main MPRs are processed. In the case of a stroke, we also process the plane that is perpendicular to the stroke plane and also contains the stroke. Fig. 3 shows the results of the 2D segmentations. Since background statistics have also been collected, the response image is now defined as:

$$g'(x,y) = \begin{cases} g(x,y) & \text{if } H_L(f(x,y)) > H_B(f(x,y)) \\ & \text{else} \\ \frac{3H_L(f(x,y))}{4}\left(1 - \frac{C_L(x,y)}{C_L^{\max}}\right) & \text{if } L(x,y) = 1 \\ H_L(f(x,y))/4 & \text{if } L(x,y) = 0 \end{cases}$$

2.3. 3D Segmentation

The final step of the algorithm consists of applying the random walker segmentation algorithm proposed by Grady [11]. The seeds for the random walker are generated directly from the 2D segmentations by simply choosing all the pixels that are a certain distance away from the segmentation boundaries (see Fig. 4). The algorithm computes the probability that a random walker initiating its walk at each voxel first arrives at a foreground seed before arriving at a background seed. It was shown in [11] that these probabilities can be efficiently computed by solving a sparse system of linear equations.

Typically, the random walker probabilities would be thresholded at 0.5 to produce the final binary segmentation. We propose the following size preserving smoothing algorithm. First, the size of the lesion is determined if the probabilities were to be thresholded at 0.5. Then, the probabilities are smoothed using a standard linearly separable Gaussian filter. Finally, the smoothed probabilities are thresholded at the level that leads to the same lesion size. The final segmentation result is shown in the bottom part of Fig. 4.

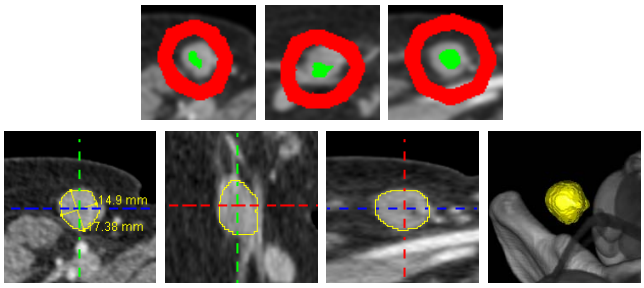


Fig. 4. Random walker segmentation: seeds in 2D planes (green for region, red for background) and resulting 3D segmentation (with volume rendering).

3. EXPERIMENTS AND RESULTS

We collected 115 CT datasets from patients with various types of cancer. The datasets were acquired at 10 different clinical sites with different protocols and slice thicknesses ranging from 1 to 5 mm. Radiologists at the different hospitals identified and manually segmented a total of 293 lesions in these datasets. Among these lesions, we had 159 liver lesions, 98 lymph nodes, and 36 other lesions in lungs, pancreas, colon, ovaries, rectum, etc. The lesions ranged in size from 0.1 to 700 mL. We also asked the radiologists to segment the lesions using the semi-automatic algorithm.

Fig. 5 shows different segmentation results. The first 7 are liver lesions (the 7th is a hyperdense lesion and the 5th contains a necrotic core). The next 3 images contain many lymph nodes. The 11th is an ovarian lesion, the 12th is a thyroid lesion, the 13th is a pancreas lesion, the 14th is a bladder lesion, and the 15th is a kidney lesion. The algorithm produces very good results on all these lesions. Fig. 6 shows examples where the segmentation results are bad. In all three cases, the algorithm has leaked into neighboring structures that are very similar in gray level.

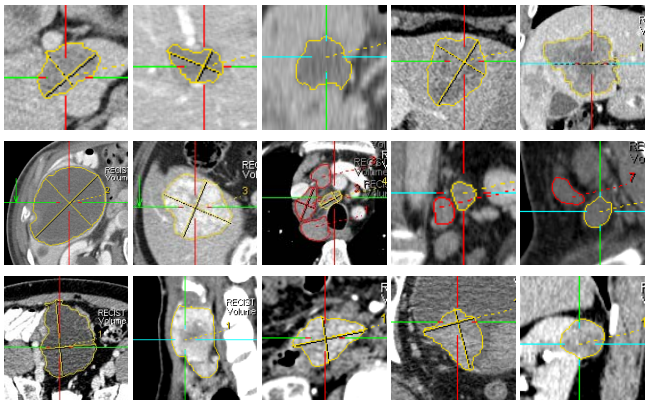


Fig. 5. Segmentation results on various lesions.

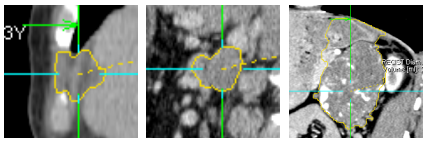


Fig. 6. Bad segmentation results.

To evaluate the performance of the algorithm quantitatively, we computed the following measures. The normalized volume difference is defined as $V_d = \frac{|V_A - V_{GT}|}{V_{GT}}$ where V_A is the automatically segmented volume, V_{GT} is the ground truth volume. The volume overlap reflects the relative position of the two objects better and is defined as $V_o = \frac{V_A \cap V_{GT}}{V_A \cup V_{GT}}$.

Fig. 7 shows the cumulative histogram of normalized volume differences over all lesions (bold curve), liver lesions, lymph nodes, and other lesions. This curve can be thought of as an ROC curve. It basically shows that in 50% of the cases, the normalized volume difference is below 25% (18.5% to be exact) and in 80% of the cases, it is below 50%. It can be seen that the type of lesion does not influence the behavior of the algorithm. Other lesions are segmented better probably because 20 of them are lung lesions which have high contrast and are relatively easier to segment.

Fig. 8 shows the cumulative histogram of volume non overlap ($1 - V_o$) over all lesions (bold curve), liver lesions, lymph nodes, and other lesions. Again, it can be seen that the algorithm perform roughly the same on all lesions, slightly better on other lesions. From visual inspection of segmented objects, radiologists agree that a volume non overlap of 30% is good, 30 to 60% is acceptable, and 60% is poor. It can be seen that 90% of the segmented lesions fall in the good to acceptable zone and 50% of the lesions are well segmented. For the lesions that are not correctly segmented, we provide editing tools that are beyond the scope of this paper [12].

From the binary 3D segmentation, the system automatically computes the volume, RECIST and WHO diameters

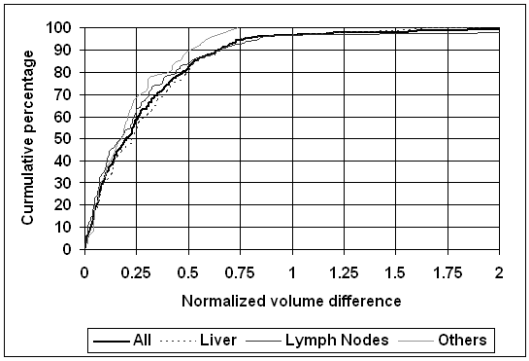


Fig. 7. Cumulative histogram of the normalized volume difference.

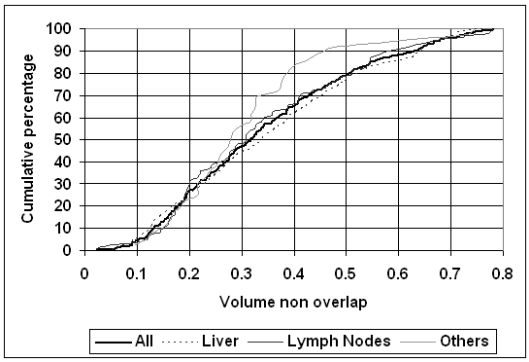


Fig. 8. Cumulative histogram of the volume non overlap.

and z-extent of the lesion. To evaluate the accuracy of the measurements, we scanned a physical phantom containing 31 lesions (spherical and gelcap shaped) of different known sizes, with a slice thickness of 2mm. We segmented each of the lesions using 10 different initialization strokes. We then computed the normalized difference in volume, RECIST diameter, WHO diameter and z-extent between the known measurements and the calculated measurements. The cumulative histogram for all 10 runs on all 31 lesions is shown in Fig. 9. It can be seen that in 90% of the cases, the WHO diameter is within 10% of the true diameter, the RECIST diameter within 15% of the true RECIST diameter, and the volume within 25% of the true volume.

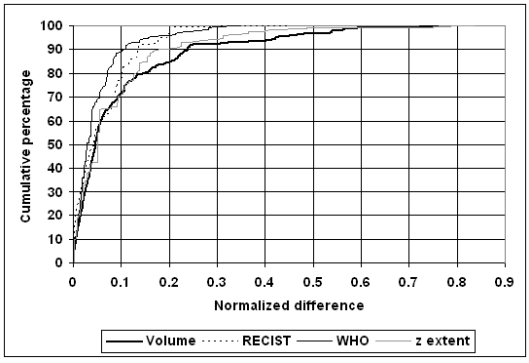


Fig. 9. Cumulative histogram of the normalized measure differences.

4. CONCLUSIONS

We have presented a general algorithm to segment any type of lesion in 3D CT. We have shown that the algorithm is successful in many cases and actually ready for clinical use. However, it still has some problems with some cases. In the future, we will focus more on very heterogeneous lesions.

5. ACKNOWLEDGMENTS

We would like to thanks M. Suehling, C. Leidecker, and M. Milite at Siemens Med-CT for helping establish the clinical collaborations and collect the data. We are also grateful to our collaborators X. Deng, H. Huang, L. Zhu, and G. Du at Siemens CT-MIA for providing data, ground truth, and measurement tools.

6. REFERENCES

[1] World Health Organization, "WHO handbook for reporting results of cancer treatment," 1979.

[2] P. Therasse, S. G. Arbuck, E. A. Eisenhauer, J. Wanders, R. S. Kaplan, L. Rubinstein, J. Verweij, M. van glabbeke, A. T. van Oosterom, M. C. Christian, and S. G. Gwyther, "New guidelines to evaluate the response to treatment in solid tumors," *Journal of the National Cancer Institute*, vol. 92, no. 3, pp. 205–216, 2000.

[3] L. van Hoe, E. van Cutsem, I. Vergote, A. L. Baert, E. Bellon, P. Dupont, and G. Marchal, "Size quantification of liver metastases in patients undergoing cancer treatment: reproducibility of one-, two-, and three-dimensional measurements determined with spiral CT," *Radiology*, vol. 202, no. 3, pp. 671–675, 1997.

[4] P. J. Yim and D. J. Foran, "Volume of hepatic metastases in computed tomography using the watershed and active contour algorithms," in *Proc. IEEE Symposium on Computer-Based Medical Systems*, New York, NY, 2003, pp. 329–335.

[5] M. Bilello, S. B. Gokturk, T. Desser, S. Napel, R. B. Jeffrey Jr., and C. F. Beaulieu, "Automatic detection and classification of hypdense hepatic lesions on contrast-enhanced venous-phase CT," *Medical Physics*, vol. 31, no. 9, pp. 2584–2593, 2004.

[6] Y. Li, S. Hara, and K. Shimura, "A machine learning approach for locating boundaries of liver tumors in CT images," in *Proc. ICPR*, 2006, vol. 1, pp. 400–403.

[7] J. Dornheim, H. Seim, B. Preim, I. Hertel, and G. Strauss, "Segmentation of neck lymph nodes in CT datasets with stable 3D mass-spring models," in *Proc. MICCAI*, 2006, pp. 904–911.

[8] G. Unal, G. Slabaugh, A. Ess, A. Yezzi, T. Fang, J Tyan, M. Requardt, R. Krieg, R. Seethamraju, M. Harisinghani, and R. Weissleder, "Semiautomatic lymph node segmentation in LN-MRI," in *Proc. Int. Conf. Image Processing*, Atlanta, GA, 2006, pp. 77–80.

[9] J. K. Udupa and S. Samarasekera, "Fuzzy connectedness and object definition: Theory, algorithms, and applications in image segmentation," *Graphical Models and Image Processing*, vol. 58, no. 3, pp. 246–261, 1996.

[10] G. T. Herman and B. M. Carvalho, "Multiseeded segmentation using fuzzy connectedness," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 5, pp. 460–474, 2001.

[11] L. Grady, "Random walks for image segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, no. 11, pp. 1768–1783, 2006.

[12] L. Grady and G. Funka-Lea, "An energy minimization approach to the data driven editing of presegmented images/volumes," in *Proc. MICCAI*, 2006, pp. 888–895.