Automatic Lung Segmentation in CT Images using Watershed Transform

Rushin Shojaii [1]
[1] Department of Electrical Engineering
Ryerson University
Toronto, Canada
rshojaii@ee.ryerson.ca

Javad Alirezaie [1,2]
[2] Department of Systems Design Eng.
University of Waterloo
Waterloo, Canada
javad@rousseau.uwaterloo.ca

Paul Babyn
Department of Medical Imaging
University of Toronto
Toronto, Canada
Paul.babyn@sickkids.ca

***Abstract*— The preprocessing step of most Computer-Aided Diagnosis (CAD) systems for identifying the lung diseases is lung segmentation. We present a novel lung segmentation technique based on watershed transform, which is fast and accurate. Lung region is precisely marked with internal and external markers. The markers are combined with the gradient image of the original data and watershed transform is applied on the combined data to find the lung borders. Rolling ball filter is used to smooth the contour and fill the cavities while preserving the original borders. The proposed method eliminates the tasks of finding an optimal threshold and separating the attached left and right lungs, which are two common practices in most lung segmentation methods and require a significant amount of time. We have applied our new approach on several pulmonary CT images and the results reveal the speed, robustness and accuracy of this method.**

Keywords-medical image processing; lung segmentation; watershed transform; pulmonary CT image;

I. INTRODUCTION

CT scan of the thorax is widely used to diagnose and evaluate numerous lung diseases. These scans yield a large amount of image data. The expanding volume of thoracic CT studies along with the increase of image data, elucidate the need for computer-aided diagnosis (CAD) schemes to assist the radiologists. Lung segmentation is the preprocessing step of most CAD systems for identifying the lung disease.

Several research groups have proposed and developed techniques to segment the lung regions in pulmonary CT images. 3D region growing with specified seed points manually is proposed in [1]. Samuel *et al.* have used border tracking to find lung region boundaries [2]. In more recent studies by Hu *et al.* [3] and Brown *et al.* [4] dynamic programming has been used to find maximum cost path, which corresponds to the junction line between the left and right lung. Hu *et al.* have utilized optimal threshold to segment lung regions instead of fixed threshold. Smoothing of the segmented lung regions is then accomplished in three separate stages to deal with three different issues: filling indentation caused by pulmonary vessels, removing small airways and removing large airways [3]. However most of these approaches are not fully automated and require substantial time to accomplish the task.

In this paper, we present an automatic method for segmenting the lungs in CT images. The main transform used in this technique is 2D marker-based watershed transform. Since the separation of the left and right lungs and extraction of the borders are automatically done during the watershed transform and the optimal threshold is eliminated, the computation time is highly reduced. This technique has several main steps. First a gradient image of the CT image is created. In the next step, which is the most important, the internal marker is specified. The ability of thin edge segmentation is also embedded in this part to separate the left and right lung. Then the external marker is identified. By performing morphological operations on the internal marker the external marker is extracted and regional minima are imposed in the marked area. The next main step is to obtain the watershed lines by applying watershed transform on the marked area. Smoothing the segmented lung borders using a simple and accurate technique is the final step.

This paper is organized as follows. Section II gives a description of the methods used in this paper including marker-based watershed transform and the proposed method. Results are presented in section III for different parts of the thorax. Conclusion and the advantages of this method along with the related future works are described in section IV.

II. METHODS

The watershed transform can be classified as a region-based segmentation approach. Buecher and Lantu´ejoul [6] formalized the concept and later Vincent and Soille [7] turned it into an "immersion-based" algorithm. It finds "catchment basins" and "watershed ridge lines" in an image by treating it as a surface, where light pixels are high and dark pixels are low [8]. Several definitions of the watershed transform have been declared. Marker-based watershed algorithms on discrete images [8], [9] are considered in this paper.

Generally, the watershed transform is computed on the gradient image, where the boundaries of the catchment basins are located at high gradient points. The principle of

0-7803-9134-9/05/$20.00 ©2005 IEEE

the immersion-based watershed algorithm [7] can be illustrated by imagining the gradient image as a relief, with the 'height' variable being the grey-value for each pixel position. Imagine, water immersing from the bottom of the relief (the darkest parts), every time the water reaches a minimum, which corresponds to a region in the original image, a catchment basin is 'grown'. When two neighboring catchment basins eventually meet, a dam is created to avoid the water merging at that level of flooding and spilling from one basin into the other. We utilized this property to make the left and right lung separation during the transformation. When the water reaches the maximum gray-value, the union of all dams forms the watershed lines [8]. Watershed transform has several advantages: it is a simple intuitive method, which is fast and can be parallelelized [10], [11].

Over-segmentation is a well-known drawback in watershed segmentation. Since every regional minimum, even if tiny and insignificant, forms its own catchment basin, over-segmentation occurs. This phenomenon is illustrated in figure 1-c, which is the result of applying standard watershed transform on the gradient image of a pulmonary CT slice (figure 1-a). By using marker-based watershed transform, we can decrease the regional minima and bound them within the region of interest to prevent over-segmentation. There are several techniques to define markers and choosing a technique is highly dependent on the application.

The details of our approach are described in the following steps.

A. Gradient Image

Since watershed transform is applied on the gradient image, the first step is to obtain the gradient image. The Sobel masking operator is applied on the original pulmonary CT image in both horizontal and vertical directions to create the gradient image (figure 1-b).

B. Internal Marker

The key point of our approach is the internal marker. The internal markers are the connected components of the pixels with almost the same intensity values, whose external boundary pixel values are all greater than n. The value of n is a gray level value, which specifies the approximate gray level value for non-body pixels in CT images. Since in pulmonary CT images the air will appear with a mean intensity of approximately -1000 Hounsfield Units (HU), lung region will be in the range of -1000 to -400 HU and the chest wall, blood and bone will be much dense and well above -400 HU. In order to specify the internal markers the regions with pixel values lower than -400 HU are selected.

The background of the CT image is almost black (lower than -400), so, it makes an erroneous regional minimum. Background is removed by eliminating the objects, which are attached to the border of the binary image [2]. After eliminating the background, the small objects in the markers which are caused by the veins should be eliminated. The open morphological operator is

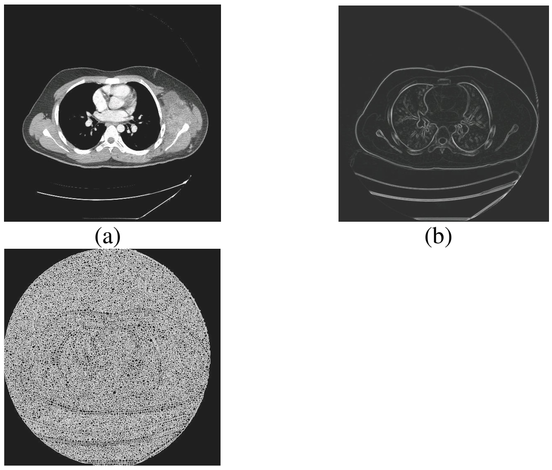


Figure 1: (a) Original pulmonary CT image. (b) The gradient image of figure 1-a using Sobel masking. (c) Result with oversegmentation obtained using standard watershed transform on the gradient image.

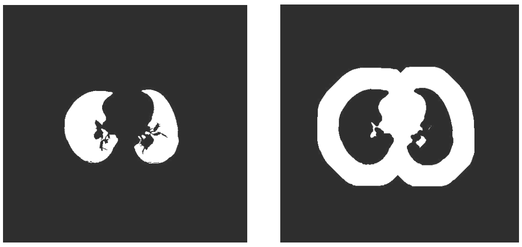
applied on the complement image to remove the objects with less than 200 pixels. Then to eliminate the remaining bright small separate objects, the objects with small number of pixels are removed to obtain the pure internal marker. The results are shown in Figure 2-a.

To remove trachea and other large airways from the internal marker binary image, region-growing algorithm is utilized [5]. The location of the trachea is identified by searching for the large, circular object near the center of the first few slices. Each slice provides the seed point for the next slice. When the size of the region on a slice highly increases, the region's growing procedure stops and the pixels, which correspond to trachea or main bronchi, are turned off in the internal marker binary image.

In CT images of certain disease conditions such as emphysema the left and right lung are attached to each other or separated with a very thin wall (with the intensity lower than -400 HU). A single large lung region in the internal marker binary image indicates that the left and right lung regions are merged. To extract the border in this case along with other tissues during the watershed transform, no regional minimum should be imposed on the area between the two lung regions. The pixels of the internal marker on this area, which is specified by a black strip near the center of the image, are turned off to avoid imposing regional minima on the area. A strip with a width of six pixels (twice the distance between the two markers) is selected to include the slightly curved border between the two lung regions.

C. External Marker

Lung border is located in the neighborhood area of the internal marker. To bind this area an external marker is



(a) (b)

Figure 2. (a) Internal Marker (b) External Marker

required. The internal marker is dilated with two circular structuring elements with different ratios. By subtracting the results, external marker is specified. It is a wide strip around the internal marker with a small distance. The width of the strip is selected at least 45 pixels to cover the whole minima in the neighborhood. The ratio of the smaller structuring element specifies the distance between the two markers. Figure 2-b shows the external marker for the lung CT image of figure 1.

D. Imposing Regional Minima

In this step the regional minima will be placed only in the marked area of the gradient image. This is required to compute the watershed lines in the region of interest. Minima imposition procedure, which utilizes morphological reconstruction, is used to place regional minima only within the area of the union of the two markers (figure 3-a).

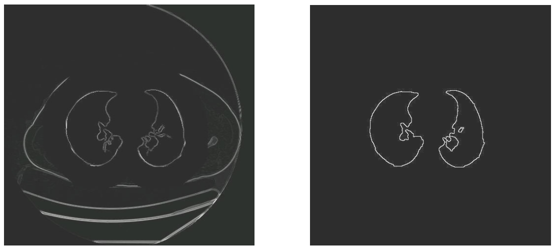
E. Watershed Transform

Watershed transform of the obtained image finds the watershed lines corresponding to the most significant edges between the markers. These watershed lines are the borders of the lung regions (figure 3-b). Since during this transform a dam is created between the two lung regions to prevent merging, the thin wall between the attached left and right lung with emphysema disease is also extracted.

F. Smoothing and Filling the Cavities

This is one of the most common practices in many lung segmentation schemes. In some CT images, the segmented lung region excludes dense structures, such as juxtapleural nodules and hilar vessels (figure 3-b).

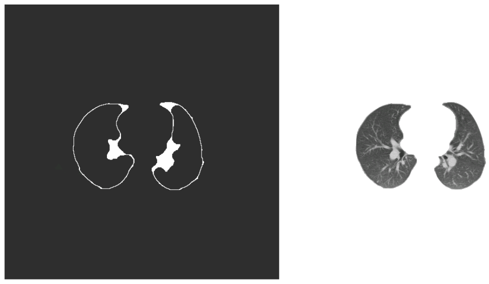
To include these structures the rolling ball filter is used [2], [8]. The segmented lung border is superimposed on the original image and the closed morphological operator is applied. A rolling ball with the ratio of 9 is selected as a 3D structuring element. Indeed, the ball rolls along the lung contours identifying the pixel along the ball's circumference with tangential slope that matches the tangential slope of the lung contour at the current contour pixel. The smoothed image is shown in figure 4-a. At this stage those parts of the border that need to be smoothed are



(a) (b)

Figure 3. (a) Imposed regional minima in the marked area.
(b) Segmented lung contour.

changed and other parts remain the same, which is the unique advantage of our technique. The segmented lung regions are shown in figure 4-b.



(a) (b)

Figure 4. (a) Smoothed lung contour using rolling ball filter
(b) Segmented lung regions

III. RESULTS

The pictorial lung segmentation results for 512 x 512 x 16 3D helical CT images of thorax with the slice thickness of 5 millimeters are given here. The slices are selected from different parts of the chest to show the accuracy of this technique. The results for the images of upper and lower parts of the chest with small regions of lung, which are successfully segmented, are shown in figure 5.

To smooth and retrieve the dense structures, which are adjunct to the lung border (figure 6-a), a rolling ball with the ratio of 9 is used. The images in figure 6-b and 6-c illustrate the segmented lung borders before and after the rolling ball filtering. Comparing the segmented lung regions (fig. 6-d) with the original image (figure 6-a) confirms the accuracy of our technique.

IV. CONCLUSIONS AND FUTURE WORK

The recent studies demonstrate that the lung segmentation method as a preprocessing step for CAD systems may affect CAD results. We have developed an automated lung segmentation method using watershed transform, which is accurate, fast and can be parallelized to make it less time consuming for the high volume dataset.

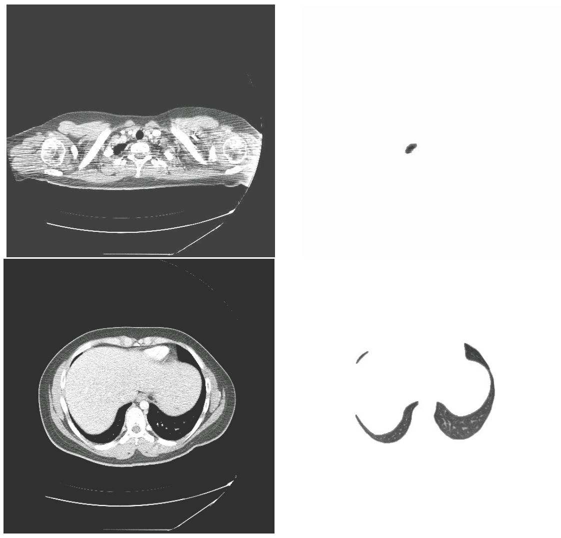


Figure 5: Right: Original CT images. Left: Segmented lung Regions.

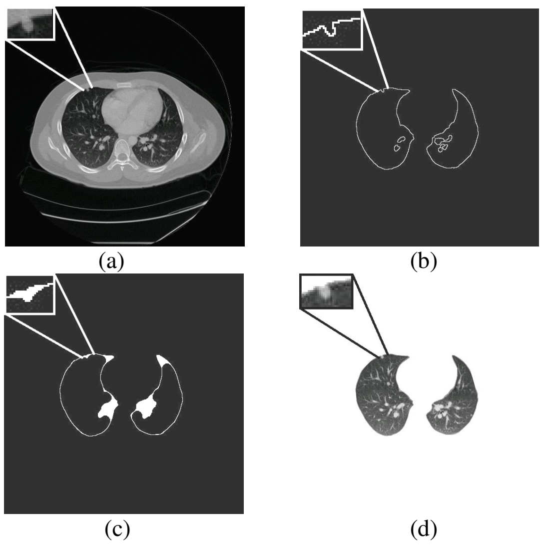


Figure 6: (a) Originl pulmonary CT image. (b) Segmented lung borders. (c) Smoothed lung borders. (d) Lung regions

Even if there are no strong edges between the markers, our technique always detects a significant edge in the desired area that demonstrates the robustness of this technique. We applied this technique on more than 180 CT images from different parts of the chest and the results are visually acceptable.

This method can be also utilized to detect pulmonary nodules in the segmented lung region. For further improvement we consider parallelization of this method to make it faster for huge dataset or very high-resolution images.

ACKNOWLEDGMENT

We would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for providing the research fund for this project.

REFERENCES

[1] W. A. Kalender, H. Fichte, W. Bautz and M. Skalej, “Semiautomatic evaluation procedures for quantitative CT of the lung.” J Comput Assist Tomogr Vol.: 15 pp: 248-255, 1991.

[2] S. G. Armato III and W. F. Sensakovic, “Automated Lung Segmentation for Thoracic CT: Impact on Computer-Aided Diagnosis”*Academic Radiology*: Vol.11, No. 9, pp. 1011-1021, September 2004.

[3] S. Hu, E. A. Hoffman and J. M. Reinhardt, "Automatic Lung Segmentation for Accurate Quantitation of Volumetric X-Ray CT Images", *IEEE Transactions on Medical Imaging* Vol 20, No 6, June 2001.

[4] M. S. Brown, M. F. Mcnittgray, N. J. Mankovich, J. G. Goldin, J. Hiller, L. S. Wilson and D.R. Aberle, “Method for segmenting chest CT image data using an anatomical model - preliminary results”, *IEEE Trans Med Imaging. Vol. 16, pp. 828-839, Dec. 1997.*

[5] Y. Masutani, K. Masamune and T. Dohi, “Region-growing based feature extraction algorithm for tree-like objects,” *in Lecture Notes in Computer Science, K. H. Hohne and R. Kikinis, Eds. Berlin, Germany: Springer-Verlag,* vol. 1131, pp. 161-171,1996.

[6] S. Buecher and C. Lantu´ejoul, “Use of watershed in contour detection.” In *Proc. Int. Workshop Image Processing, Real-Time Edge and Motion Detection/Estimation, Rennes, France*, pp. 17–21. September 1979.

[7] L. Vincent and P. Soille, “Watersheds in digital spaces: An efficient algorithm based on immersion simulations.” *IEEE Trans.Patt.Anal.Mach.Int.* 13(6), pp. 583–598, June 1991.

[8] R. C. Gonzalez, R. E. Woods, “Digital Image Processing”, *Prentice Hall Publishing Company*, 2002.

[9] V. Grau, A. U. J. Mewes, M. Alcañiz, R. Kikinis and S. K. Warfield, “Improved Watershed Transform for Medical Image Segmentation Using Prior Information”, *IEEE Transactions on Medical Imaging, Vol. 23, No. 4, April 2004.*

[10] A. N. Moga and M. Gabbouj, “Parallel marker-based image segmentation with watershed transformation”, *Journal of Parallel and Distributed Computing* 51, 27{45}, 1, 1998.

[11] J. B. T. M. Roerdink and A. Meijster, “The watershed transform: Definitions, algorithms and parallelization strategies,” *Fundamenta Informaticae*, vol. 41, pp. 187–228, 2001.