Automatic Lung Segmentation in HRCT Images

Huaqing Chen, Ramakrishnan Mukundan
Dept. Computer Science and Software Engg.
University of Canterbury
Christchurch, New Zealand
hch60@uclive.ac.nz, mukundan@canterbury.ac.nz

Anthony Butler
Dept. Radiology
University of Otago
Christchurch, New Zealand
anthony.butler@otago.ac.nz

Abstract—**This paper proposes a fully automatic algorithm for accurately segmenting lung regions in High-Resolution Computed Tomography (HRCT) images. A region based intensity characterization allows the image to be considered as consisting primarily of three regions: the CT background, the lungs, and the thorax region surrounding the lungs. A novel method based on flood fill algorithm is used to effectively identify the surrounding region. This step facilitates the use of another fast method for the removal of the CT background using linear scans originating from border pixels. Connected components that represent parts of the trachea are removed by noting the separation of the mean and standard deviation of intensity values between the trachea and the lungs. The segmented lung images are further enhanced to restore the intensity values of the pixels on the bronchi and the lung boundary. The proposed technique is not only computational inexpensive, but also robust and accurate in detecting the lung boundary. This paper presents the complete framework including examples and experimental results.**

Keywords-Computer-Aided Diagnosis; HRCT lung images; Image analysis; Segmentation; flood-fill method; Image dilation;

I. INTRODUCTION

High-resolution computed tomography (HRCT) imaging is used as a non-invasive tool for diagnosing and evaluating lung diseases. The HRCT lung images show much more information than plain radiography x-ray images do, but requires additional skills for the interpretation and extraction of pathological information. In order to improve the efficiency and accuracy of detecting lung diseases, Computer Aided Diagnosis (CAD) systems help radiologists to characterize the distribution of the disease patterns found in the HRCT images [1]. Lung segmentation is the key part of such CAD systems. Lung segmentation is also required in the preprocessing phase of visualization and analysis tools [2].

This paper proposes a novel method for fast identification of the thorax regions surrounding the lung, as well as the CT background using a flood-fill method, where the seed point is automatically determined. Image regions containing the lung and the bronchi are further processed to remove the main bronchi, and to enhance the lung boundary. The paper is organized as follows. Section II outlines background research related to the area of lung segmentation. Section III describes the proposed methodology for lung segmentation. Section IV presents the results and outlines the advantages and the limitations of the method. Section V provides a summary of the work and possible future research directions.

II. BACKGROUND

There are several computer-assisted methods proposed and developed to segment the lung regions in chest CT images [2]-[7]. In [3], 3-D region growing method was used to find the left and right lung boundary, with seed point manually allocated. In [4], a 2-D edge tracking method was used to find the lung boundaries. The graph-cut method was implemented to obtain globally optimal object segmentation in N-dimension [5]-[6]. The methods presented in [2]-[7] all require manual initialization of seed points. In [7], the anterior and posterior junction lines are manually generated to separate the left and right lungs.

However, there have also been some other fully automatic methods proposed in recent years. In [8], Brown et al. adopted anatomical knowledge including expected volume, shape, relative position and X-ray attenuation of organs to guide the segmentation process. Shojaji et al. [9] presented a lung segmentation method based on the gradient technique, and the watershed transform to find the lung boundaries.

In this paper, we base our automatic segmentation method on well defined intensity characteristics of different regions of HRCT lung images, and use flood fill, linear scan and morphological operations to identify the exact lung region. The method also guarantees that every internal pixel of the processed and extracted lung region has the same intensity value as the original image.

III. PROPOSED METHOD

This section describes the proposed method for segmenting lung regions in HRCT images. Fig. 1 shows a flow chart consisting of seven main steps used in the method. An outline of these steps is given below.

- Using a partitioning of the HRCT lung image into 16x16 boxes, a seed point position for the flood fill algorithm is automatically found in the body region.

- The flood fill process is initiated with a pre-specified threshold value.

- The flood fill method visits all pixels in a connected component of the body region.

- A test is made to see if the flood fill has completely visited the surrounding tissue region. If the surrounding tissue is not completely processed, a new flood fill seed point is automatically allocated in the

unprocessed surrounding tissue region and the flood fill process will start again.

- The region corresponding to the CT background is now removed using an iterative algorithm.

- Non-lung connected components are removed from the extracted coarse lung region base on their pixel intensity mean and deviation.

- Lung region and the airway region can be distinguished by their pixel intensity difference. A 3×3 horizontal and vertical dilation mask is applied to recover the lung information lost during the flood fill process.
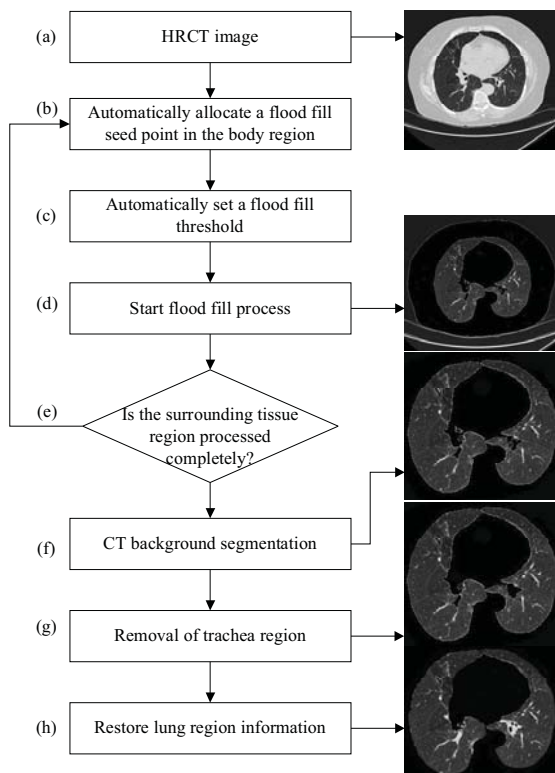
Figure 1. The flow chart of the proposed method.

A. Region Separation

In an average sense, the intensity values around a pixel in an HRCT image can be roughly categorized as belonging to one of the three regions: (i) the CT background, (ii) the body tissues surrounding the lungs and (iii) the lung region. Our method takes advantage of this separation of intensity values, and uses a flood fill method to first identify the pixels corresponding to tissues in the thorax surrounding the lung. Once this process is completed, the region outside this segment, corresponding to the CT background can be easily identified using a fast linear scan from each border pixel. Fig. 2 shows the pixel intensity mean difference between the lung region, surrounding tissue region, and the CT background, computed from several HRCT image slices. The pixel intensity values range from 0 to 65535. Fig. 2 also suggests that a threshold of $T_1 = 15000$ would be adequate to initiate a recursive flood-fill in the body region. The main task here is to

identify a seed pixel in the region, and also to make sure that all possibly disjoint, sections of this region are visited by the algorithm.
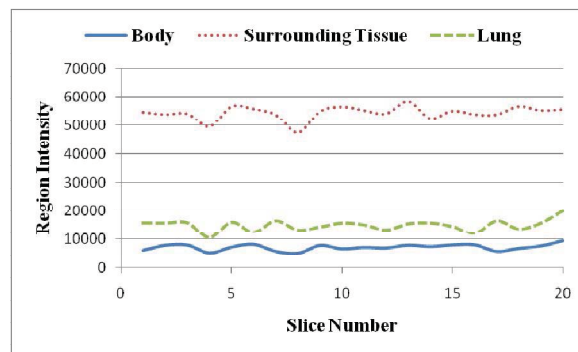

Figure 2. Intensity difference of the three main regions in an HRCT lung image.

B. Automatic Seed Point Selection

As seen in the previous section, the surrounding thorax tissue region is located in between the CT background and the lung region in HRCT lung images. Implementing the flood fill method in the surrounding tissue region can therefore separate the CT background and the lung region. The seed point is defined as the point with the maximum intensity within a neighbourhood of 16x16 pixels having intensity values in the upper region of Fig. 2. A 16x16 window is found to adequately characterize the nearly uniform distribution of intensities in a neighbourhood, and it also provides the average intensity within the window. The HRCT image is therefore divided into 16×16 cells R_{ij}. where $i, j = 1,...,32$, and the average intensity r_{ij} in each cell is computed. Cells with average intenstiy greater than $T_2 = 45000$ contain only pixels belonging to the thorax region. We select the cell G with the maximum average intensity $g = \max_{i,j}(r_{ij})$, and define the seed point as the pixel with the highest intensity in G. The previous steps are guaranteed to provide a cell in the thorax region surrounding the lungs, and a seed point in that region (Fig. 3(a)). A recursive 4-connected flood fill algorithm using the threshold T_1 visits an entire connected component within this region (Fig. 3(b)). All visited pixels and cells are marked, and the whole process is repeated if there exits another component of un-marked cells with average intensity greater than T_2. In Fig. 3(a), such disjoint components exist because of the anterior and posterior junctions (indicated by arrows). So, checking for the completeness of the flood fill process is an essential step for lung segmentation. Please note that parts of the bronchi also have similar intensity values as the surrounding tissues, and may get removed by the flood fill operation (see Figs. 3(b), 3(c)). These regions will be recovered at a later stage.
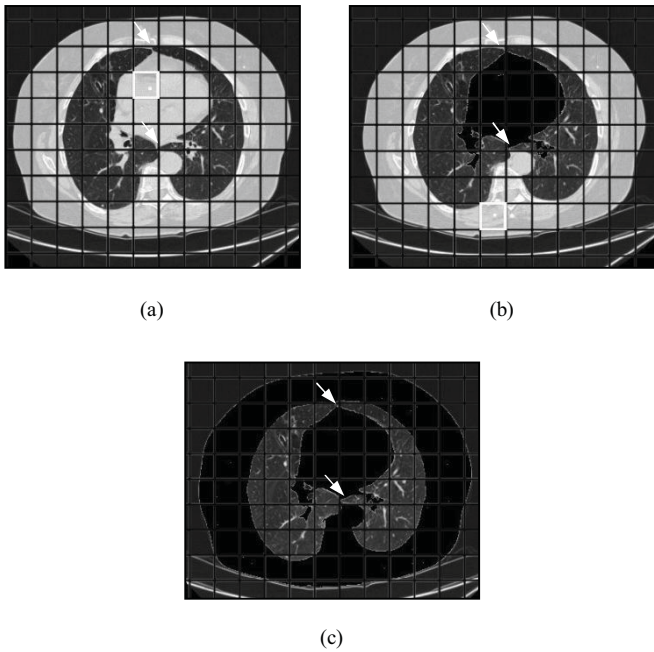
(a)   (b)

(c)

Figure 3. (a) First seed point allocation. (b) Second seed point allocation. (c) Identification of the complete surrounding tissue region using flood fill.

C. CT Background Segmentation

Having marked the entire body region surrounding the lungs, the next task is to identify the exterior of this region, which is the CT background. This is done by initiating a linear scan from each border pixel towards its pair on the opposite border, marking off each pixel and stopping the scan when a pixel belonging to the surrounding region is encountered. Since, in the most general case, the surrounding region has a non-convex shape, it is necessary to perform both horizontal scan (from left and right border pixels) and vertical scan (from top and bottom pixels) to completely remove all external pixels belonging to the CT background. The process is illustrated in Fig. 4.
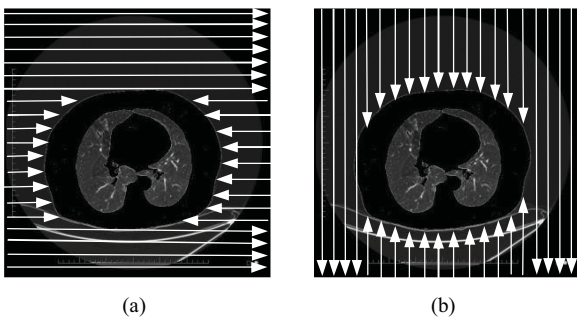

(a)   (b)

Figure 4. (a) Horizontal scan. (b) Vertical scan.

We will now only need to focus on the coarse chest region located inside of the marked body region. All other information will be removed. Fig. 5 shows the coarse chest region, which contains airways (Trachea and bronchi) and lungs.
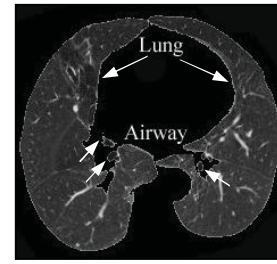

Figure 5. Coarse lung region

D. Removal of trachea region

A number of authors have developed techniques for detecting lung airways from CT lung images [10]-[12]. In [11], Hu et al. used slice-by-slice region growing method to remove airways and used the airways location on current CT slice to estimate the airway's location on the next slice. However, the slice-by-slice region growing method is not capable of segmenting the airways on some transverse slices [13].

In this paper, we consider specific characteristics of the intensity distribution within the lung region to identify and remove image segments that correspond to the airways. The images shown in Fig. 6 have two things in common. Firstly, the trachea region has lower pixel intensity than the lung region. Secondly, the presence of veins in the lung region makes the lung region vary in pixel intensity. So the lung region has higher pixel intensity mean and deviation than the air region. Fig. 7 shows the pixel intensity mean and deviation difference between the two regions. This separation in the intensity space can be effectively used of for identifying connected components that represent parts of the trachea.
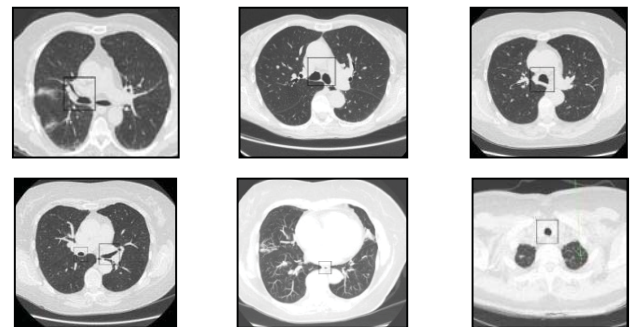

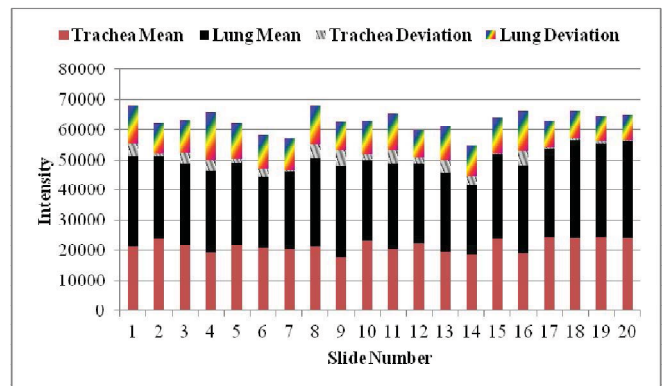Figure 6. Air filled regions vary in size, location and shape.

Figure 7. The statics of lung and trachea.

Each connected component in Fig. 5 is labeled with a positive number, which is further processed to determine its shape and size. In our case, we will also need to calculate the pixel intensity mean and deviation for each connected component. If a specific connected component belongs to the trachea region according to its pixel intensity mean and deviation, this region will be discarded.

E. Restoring lung region information

Bronchial walls, veins and nodules have generally higher intensity compared to the average lung region, and if they are near the boundary of the lung regions, they will most likely get removed by the flood fill process. In Fig. 8 (b), the lung nodule near the boundary is lost after extracting the coarse lung region. Restoring the lost pixel information within the lung boundary is important for proper segmentation. We propose a method below for accurate reconstruction of the lung image region.
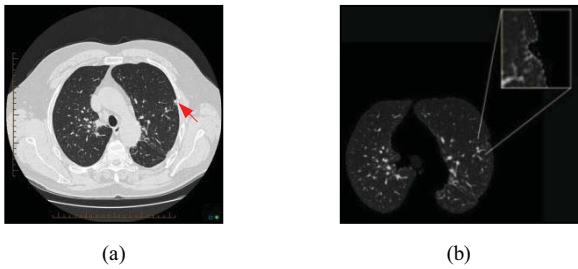

(a)   (b)

Figure 8. (a) Lung nodule on the side of the lung boundary. (b) Pixels on the lung nodule affected by the flood fill process.

1) Morphological dilation: In grayscale morphology, the operation of dilation is used for region growing. This operation also closes holes and gaps within the region. Dilation makes an object larger by adding pixels around its edges [14]. Fig. 9 shows a 3×3 horizontal and vertical dilation mask.

$$1\ 1\ 1$$
$$1\ 1\ 1$$
$$1\ 1\ 1$$

Figure 9. Horizontal and vertical mask [14]

In order to close the lung gaps completely, we convolve the image with the dilation mask four times. An example of this application is showin in Fig. 10(a), (b). The operation increases the intensity of all pixels within the region, and closes gaps that resulted from the flood fill operation. The border also expands by four pixels. We now replace all non-zero intensity pixels in this image with the corresponding pixels from the original lung image. This operation yields a result as shown in Fig. 10 (c), where the bronchi regions have been restored. This image also has a four pixel wide border as a result of the region growing by dilation, and copying of the corresponding pixels from the surrounding region. Fig. 11 shows the result of the above sequence of operations when the image is dilated only once, twice, thrice and four times. A four-step dilation is used primarily to close the gaps representing the bronchi. The next operation restores the original lung boundary.
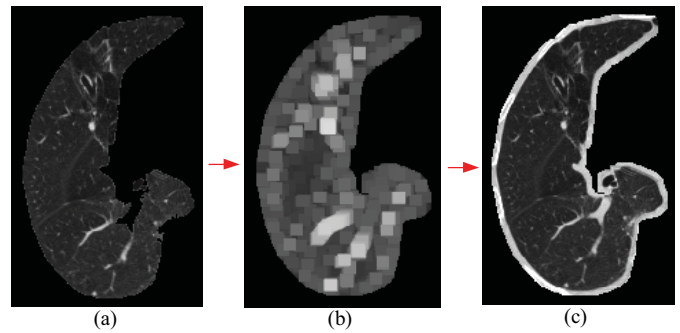

(a)   (b)   (c)

Figure 10. (a) Lung region after flood fill and CT background removal operations. (b) Result of applying the dilation mask four times (c) Dilated lung region shown by original pixel intensity
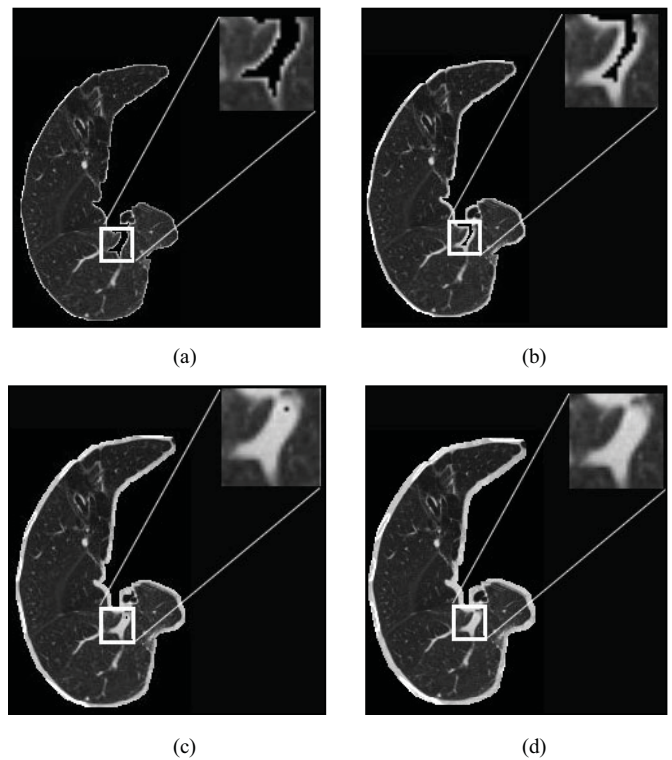

(a)   (b)

(c)   (d)

Figure 11. (a) dilation×1. (b) dilation×2. (c) dilation×3. (d) dilation×4.

2) Lung boundary enhancement: In Fig. 12, the grey cells in the middle of the enlarged window represent the pixels from the surrounding tissue region that were copied from the original image to the lung boundary. The cells marked with numbers are pixels in the lung region, and the black cells correspond to the background pixels. A scan-line algorithm is used to find the edge pixels which are adjacent to the black background pixels. Every edge edge pixel has a background pixel included in its 8 neighbouring pixels and is also not part of the lung region. Each scan line strips off edge pixels found on it by setting their pixel intensity to zero. Repeating the process four times removes the 4-pixel wide border introduced by the dilation operation.
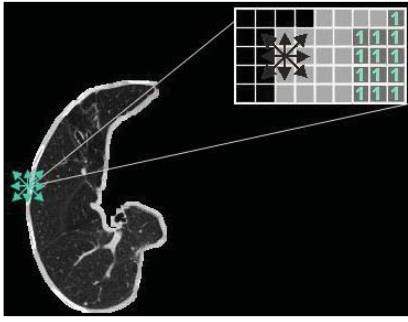
Figure 12. Optimize lung boundary.

The results of the edge peeling process for our sample image are shown Fig. 13 (a) to Fig. 13 (d). In the Fig. 13 (d), the border containing pixels from surrounding tissues is completely eliminated. Note also the interior of the lung region now contains the original pixel intensities.
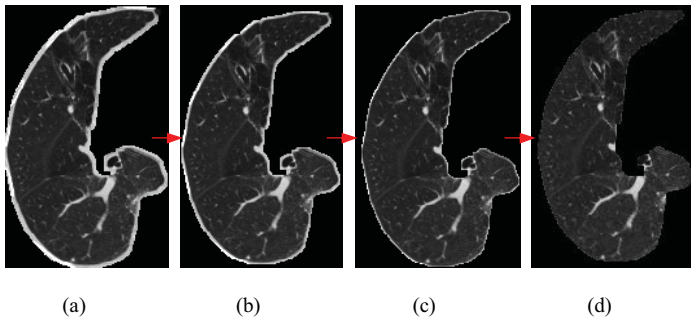

Figure 13. (a) Edge pixel peel ×1. (c) Edge pixel peel ×2. (d) Edge pixel peel ×3. (e) Edge pixel peel ×4.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A set of HRCT lung images was used to evaluate the performance of the proposed method in lung segmentation. Each image was stored and processed in RAW format, and had a size of 512x512 pixels, with a bit depth of 16 bits. Fig. 14 shows the results after lung segmentation for a set containing distinctly different images and features. In all the cases, the lung region has been successfully segmented. In order to show the robustness and accuracy of the proposed method, the lung regions in the test images have varying shape, size and location. As seen from the results, airways regions of different sizes and shapes have also been detected and removed by the proposed method.
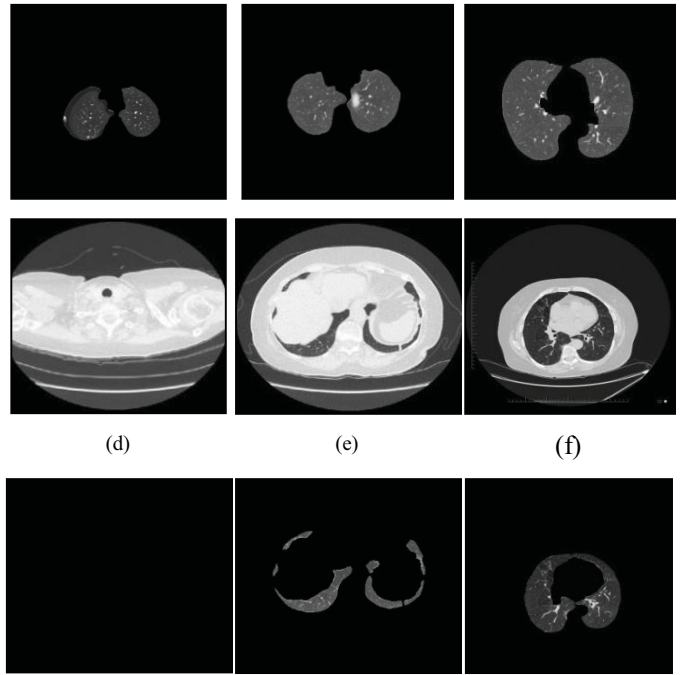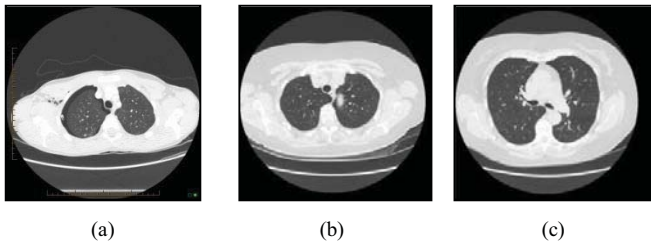

Figure 14. Selected HRCT images with corresponding segmentation results

A. Advantages of the proposed method

- The method is fully automatic; it does not require manual input of seed points or interactive selection of regions.

- The method provides accurate segmentation of the lung region, and effectively removes the trachea regions and the surrounding region of the thorax.

- The pixel intensity values within the interior of the lung region after segmentation are the same as the original image.

B. Limitations of the proposed method

- The method involves several stages of processing.

In order to evaluate the performance of the proposed method, we considered many factors that affect the accuracy of lung segmentation. In Fig. 15, these factors are called small trachea region (STR), irregular trachea shape (ITS), small Lung region (SLR), irregular lung shape (ILS), close lung junctions (CLJ) and close lung and trachea distance (CLTD). Let a be the number of HRCT images with one specific factor, b be the number of HRCT images that are segmented accurately. Then b/a represents the segmentation accuracy ratio. In our case, we set the value of a to 30 for each factor listed above. Fig. 15 shows the comparison of segmentation accuracy ratio.
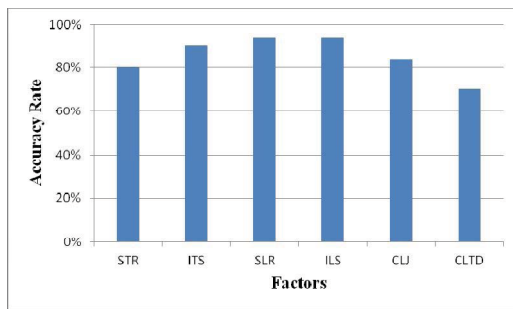
Figure 15. Comparison of segmentation accuracy ratio

As seen in Fig. 15, the images with irregular trachea shape (ITS), small lung region (SLR), irregular lung shape (ILS) and close lung junctions (CLJ) are segmented well with accuracy rate over 80%. For images with small trachea region (STR) and close lung and trachea distance (CLTD), accuracy ratio is found to be lower than 80%.

From the results, it is concluded that the airway region and the lung region can be effectively separated by using pixel intensity mean and deviation values.

V. CONCLUSION AND FUTURE WORK

This paper presented an automated lung segmentation method based on flood fill and morphological operations for effectively segmenting the lung regions of varying size, shape and location from HRCT images. Using the statistical characteristics of the intensity distribution, it is possible to easily distinguish, the lung and the airway. Experimental results obtained using a large collection of HRCT images show that the proposed method is highly effective.

For future improvement we plan to optimize the system by improving the flood fill algorithm for identifying the thorax region, minimizing the number of comparisons used in CT background segmentation, and eliminating the repeated operations of the morphological operations. The proposed method could be fasted enhanced by making of a stack of HRCT images with 3D morphological operators [15].

ACKNOWLEDGMENT

REFERENCES

[1] N. A. Memon, A. M. Mirza and S.A.M. Gilani, "Segmentation of Lungs from CT Scan Images for Eatly Diagnosis of Lung Cancer," Proceedings of World Academy of Science, Engineering and Technology, vol 14, August 2006.

[2] J. S. Silva, "A fast approach to lung segmentation in X-ray CT images", Proc. Portuguese Conference on Pattern Recognition, 2000, Pp. 415-418.

[3] L. W. Hedlund, R. F. Anderson, P. L. Goulding, J. W. Beck, E. L. Effmann, and C. E. Putman, "Two methods for isolating the lung area of a CT scan for density information," Radiology, vol. 144, pp. 353-357, 1982.

[4] J. M. Keller, F. M. Edwards, and R. Rundle, "Automatic outlining of regions on CT scans," J. Comput. Assist. Tomogr., vol. 5, no. 2, pp. 240-245, 1981.

[5] Y. Y. Boykov and MP. Jolly, "Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images," Int Conf Comput Vision (ICCV'01), 1: 105-112, 2001.

[6] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient N-D image segmentation," Int J Comput Vision, 70(2): 109-131, 2006.

[7] W. A. Kalender, H. Fichte, W. Bautz, and M. Skalej, "Semiautomatic evaluation procedures for quantitative CT of the lung," J. Comput. Assit. Tomogr., vol. 15, no. 2, pp. 248-255, 1991.

[8] M. S. Brown, M. F. McNitt-Gary, 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. Imag., vol. 16, pp. 828-839, Dec. 1997.

[9] R. Shojaji, J. Alirezaie and P. Babyn, "Automatic lung segmentation in ct images using watershed transform," In Processings of the 8th IEEE International symposium on Computers and Communication, vol. 2, pp. II-1270-3, 2005.

[10] M. Rudrapatna, A. Sowmya and P. Wilson, "Automatic detection of hilum and parenchymal bands on HRCT lung images," IEEE Conference. Cybernetics and Intelligent Systems., vol. 2, pp. 769-774, Jul. 2005.

[11] S. Hu, E. A. Hoffman, and J. M. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric x-ray CT images," IEEE Transaction on Medical Imaging, vol. 20, no.6, pp. 490-498, June. 2001.

[12] Q. X. Gao, S. J. Wang, D. Z. Zhao and J. Liu, "Accurate Lung Segmentation For X-ray CT Images," Third International Conference on Natural computation, 2007. ICNC 2007, vol. 2, pp. 257-261, Aug. 2007.

[13] J. Tschirren, Segmentation, anatomical labeling, ranchpoint matching, and quantitative analysis of human airway trees in volumetric CT images. PhD thesis, The University of Iowa, 2003.

[14] D. Phillips, Image Processing in C, 2nd ed., R&D Publications: Kansas, April 2000, pp. 156-160.

[15] J. Lerdsinmongkol, K. Chaisaowong, S. Roongruangsorakarn, T. Kraus and T. Aach, "Application of 3D Morphological Operations in the Framework of a Computer-Assisted Diagnonsis System to Construct Thorax Mask and Remove Trachea," The 3rd International Symposium on Biomedical Engineering. 2008. ThaiBME/IEEJ, pp. 260-265, Feb. 2009.