http://dx.doi.org/10.6109/ijice.2011.9.6.661

An Efficient Extraction of Pulmonary Parenchyma in CT Images using Connected Component Labeling

Kiran Thapaliya, Il-Cheol Park, and Goo-Rak Kwon, *Member, KIICE*

***Abstract*— This paper presents the method for the extraction of the lungs part from the other parts for the diagnostic of the lungs part. The proposed method is based on the calculation of the connected component and the centroid of the image. Connected Component labeling is used to label the each objects in the binarized image. After the labeling is done, centroid value is calculated for each object. The filing operation is applied which helps to extract the lungs part from the image retaining all the parts of the original lungs image. The whole process is explained in the following steps and experimental results shows it's significant.**

***Index Terms*— Connected Component Labeling, Centroid, Standard Deviation, Mean, Thresholding.**

I. INTRODUCTION

IMAGE processing is a vast field which is growing rapidly. It has been used in various fields for many purposes. One of the most important fields of digital image processing is bio-medical. It is used for the extraction of various kinds of information from the images. Among the different parts of the body, lung is one of the most important organs of human body. So in the medical field the study of lung for different kinds of the diseases are very important. So the careful extraction of lung part for the diagnosis of diseases is most necessary. Among the other disease the frequently cause disease in the lungs part is the cancer. So early detection of cancer can save the life of human. At the present lung CT scanning is widely used to diagnosis the various pulmonary diseases including the lung cancer. CT images give a great opportunity for the detection of small long nodules due to their high resolution and high quality images. The doctor during the surgical planning has to view the detail parts of the lungs to find out the diseases. So the proper extraction of lungs can reduce the overload of doctor. The accurate of any lung irregularity can be determined from an efficient lung segmentation technique. During the segmentation, it is very vital to know that the technique is able to retain the complete part of the lung and not a small or single part which is present in the original image is eroded.

So many algorithms have been proposed for the segmentation of lungs. Samule et al. [2] applied border tracking to find region boundries. Brown et al. [3] adopted anatomical knowledge including expected volume, shape, relative positions and X-ray attenuation of organs to guide the segmentation process. Hu et al. [1] applied the optimal threshold to segment lung regions. In order to remove the work of finding the optimal threshold, Shojaii et al. [4] presented a lung segmentation technique based on the gradient technique and watershed transform to find the lung borders. Kim et al. [5] applied a novel lung segmentation technique based on anisotropic diffusion and morphological operation. Kanazawa et al. [6] applied an automatic extraction method based on grey level thresholding technique. El-Baz in [9] extracted the lung regions by eliminating the background followed by using an adaptive thresholding approach and the median filter. Julian in [7] used thresholding, followed by sequential morphological filtering, edge detection and thinning to extract the lung using TRACE algorithm. Chen et al. [8] used Otsu method for the first to binarize the image followed by the flood fill algorithm and morphological operation like erosion. However Chen et al. [8] has been success to extract the lungs part but has some defects. The author has applied the erosion operation to fill the holes in the lungs but erosion operation reduces the information from the lungs and the author has used the Otsu method as a first step to binarize an image. But Otsu method fails when the background is not uniform.

The proposed algorithm uses the mean and standard deviation as the first step to binarize an image and uses the connected label component to determine the centroid of each label component and used the flood fill algorithm. The proposed method can be easily implemented and the experimental results prove the proposed algorithm is better than the above algorithm.

Manuscript received September 22, 2011; revised October 11, 2011; accepted October 15, 2011.

Kiran Thapaliya, Il-Cheol Park, and Goo-Rak Kwon are with the Department of Information and Communication Engineering, Chosun university, Gwangju, 501-759, Korea (Email: grkwon@chosun.ac.kr)

II. PROPOSED ALGORITHM

A. Image Preprocessing

As we know that the medical images may contain lots of noises, dirt and dust which affects in the process of segmentation. So it's very important to remove the noises from the images which make doctor to examine the lungs part properly. So the first step of the image segmentation is the removing the noises which is present in the image. Median filter is used here because it preserves the sharp edges of the image. The original image and the filtered image are shown in Fig. 1.

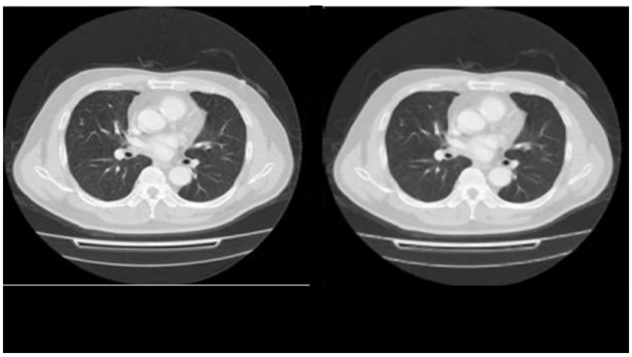

Fig. 1. The original CT image and the result after median filtering.

B. Binarization

The pulmonary parenchyma contains air whose mass density is lower in compare to the surrounding tissue. So by choosing the appropriate value for threshold, it can segment the lung region from the CT images. So to calculate the effective value of threshold we can find out the mean and standard deviation of the filtered image. The mean and standard deviation of the image is calculated by the following equation.

$$\mu = \frac{1}{X \times Y} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} f(x,y) \quad (1)$$

$$\sigma = \frac{1}{X \times Y} \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} (f(x,y) - \mu) \quad (2)$$

Where μ and σ are the mean and standard deviation of the image.

As the value of mean and standard deviation is calculated, the threshold value is determined by the following equation.

$$T = a * \sigma + b * \mu \quad (3)$$

Where a and b are constant variable, σ is the standard deviation of an image and μ is the mean of an image.

With this threshold value T the image is binarize which is shown in Fig. 2

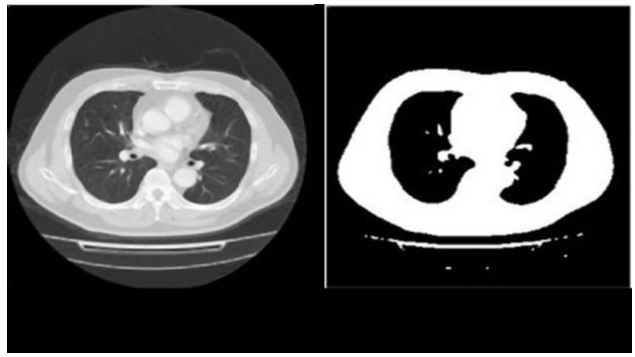

Fig. 2. The filtered image and the obtained binary image using threshold value *T*.

C. Connected Component Labeling

Connected Components labeling works based on scanning an image and group its pixels in to components based on pixel connectivity which means all pixels in a connected component share similar pixel intensities values and are connected with each other by any means. After determining these all the groups each pixel is labeled according to the component it was assigned to. Connected component labeling works on both binary or gray-level images and different level of connectivity is possible. But in this algorithm Connected component labeling is used in binary image with 4 label connectivity. The working of connected component is explained below.

The Connected components labeling operator scans the image by moving along until it comes to a point *p* where *p* denotes the pixel to be labeled at any stage encountered in the scanning process for which V=(1). When this is true, it examines the 2 neighbors (i.e. to the left and above it) of *p* which has been encountered in the scanning process. Based on this information, the labeling of *p* occurs as follows:

1) If all two neighbors are 0, assign a new label to *p*, else
2) If only one neighbor has V= (1), assigns its label to *p*, else
3) If more than one of the neighbors have V= (1), assign one of the labels to *p* and make a note of the equivalences.

After the scan is complete, the equivalent label pairs are sorted into equivalence classes and a unique label is assigned to each class. The final step is followed by the second scan through the image, during which each label is replaced by the label assigned to its equivalence.

The assigned labeled to the binary image is shown in Fig 3.

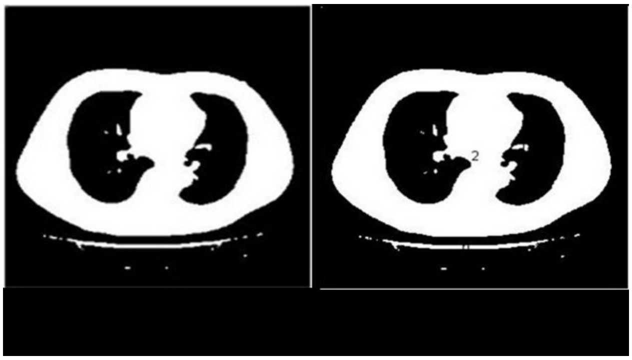

Fig. 3. Binary image and the labeled image done to the binary image.

D. Calculation of Centroid and applying the flood-fill algorithm

After the labeling of the each class, the centroid C of each class is calculated. This centroid value is used to remove the other unnecessary objects from the image and used to fill the holes in the image. The centroid value is again used as a threshold value to achieve the next level of binarize image.

If the value of centroid (C) is greater than the centroid value C=1.9 and less than 80 then we obtained the next level of binary image $B2$.

$$B2 = C > 1.9 \text{ } and \text{ } C < 80 \qquad (4)$$

After obtaining the image after applying the value of C, we used flood-fill algorithm.

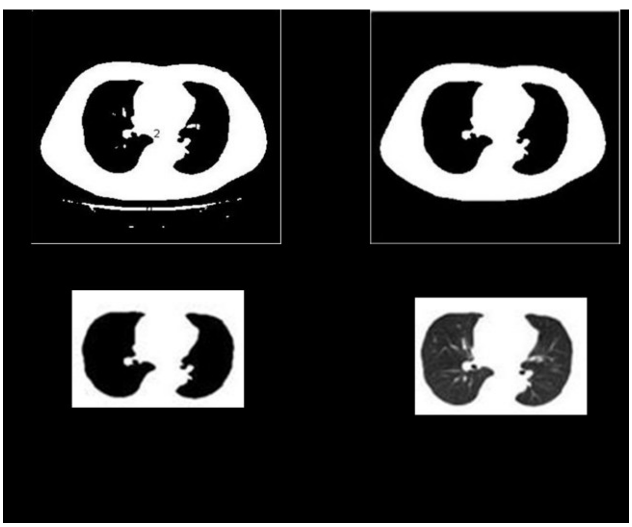

Fig. 4. The labeled image, the image obtained after applying the value L to the labeled image , the image obtained after applying the flood-fill algorithm and the finally extracted image.

Finally, the obtained binary image in Fig. 4. (c) as a mask to the original image, all the pixel corresponding to the white pixels in the mask image are kept unchanged and the other pixel values are set to the same values to the original CT image. Then the lung part is easily extracted which is shown in Fig. 4. (d).The obtained image after thresholding with centroid value C, applying the flood-fill algorithm and the final extracted image is shown in Fig. 4

III. FLOW CHART OF THE PROPOSED ALGORITHM

The Flow chart of the proposed algorithm for the extraction of the pulmonary parenchyma is shown in Fig. 5.

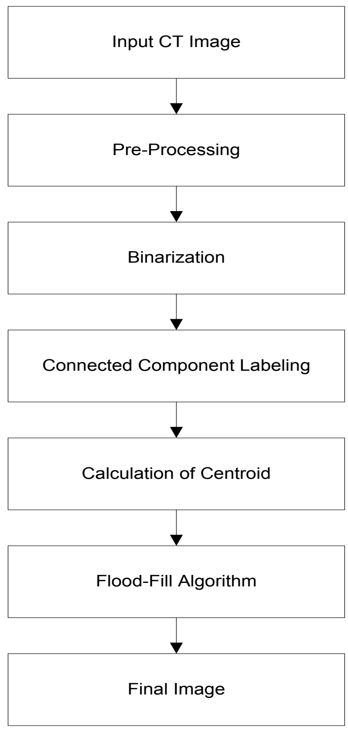

Fig. 5. The flowchart of the proposed algorithm.

IV. EXPERIMENTAL RESULTS

The proposed algorithm is tested in many different types of CT images. The proposed method gives the better result which is able to retain the all the parts of the lungs that are present in the original image. The proposed algorithm when applied to different CT images gives the result which is shown below. The value of C in equation (4) is determined experimentally. The given algorithm was tested using the MATLAB 2009. The experimental value found for the C =1.9 and C=80. Fig. 5. (a) is an original image and Fig. 5. (b) is the image obtained by applying the Chen method and Fig.5. (c) is the image obtained by output image extracted by using the proposed algorithm. From the experimental results it is clear that the proposed algorithm is better than the Chen method. Chen method is unable to retain all the parts of the lungs which are in the original image but the proposed algorithm retains all the parts of the lungs which are in the original image which is clear from the experimental results.

Fig. 6. The original image (a), the image (b) is obtained by Chen method and the final image (c) is obtained by proposed method.

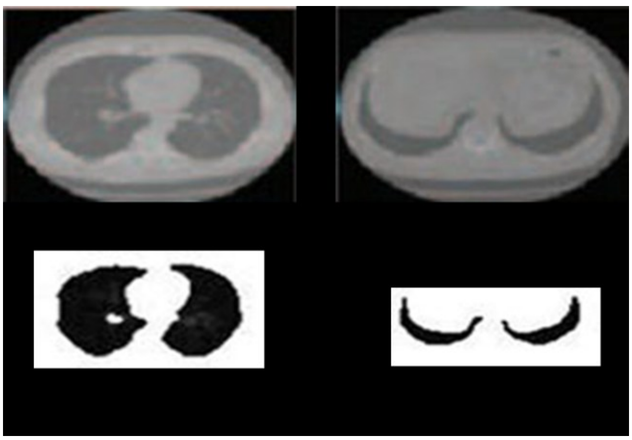

Fig. 7. Figure (a) and (b) are original image and figure (c) and (d) are finally extracted image respectively

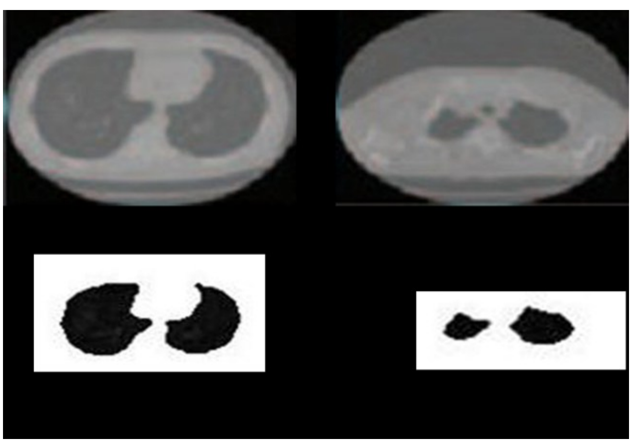

Fig. 8. The above figure (a) and (b) are the original image and figure (c) and (d) are the final image extracted respectively.

The proposed algorithm works to the image whose background is not uniform where as for the Chen method it has used Otsu method so fails when the background is not uniform. The other experimental results applied for the various lungs images are shown below.

V. CONCLUSIONS

The proposed algorithm is very efficient and easy to implement. Since it uses the mean and standard deviation as the first step, the image is properly binarized. As seen from the experimental results of different kinds of lungs image proposed algorithm can be used to apply different types of images to extract the lungs part effectively and since it does not use the erosion algorithm as used by Chen, our algorithm is able to retain the information of the lungs image which has been extracted. The result of proposed algorithm is better and able to retain tinier parts of the original image.

REFERENCES

[1] Hu S Y, Hoffman E A, and Reinhardt J M., "Automatic Lung Segmentation for Accurate Quantization of Volumetric X-ray CT Images," *IEEE Trans. Medical Imaging*, 20:490-498, 2001.

[2] S.G.Armato III and W.F. Sensakovic, "Automated Lung Segmentation for Thoracic CT: Impact on Computer –Aided Diagnosis," *Academic Radiology*, vol.11, pp. 1011-1021, 2004.

[3] M.S. Brown, M.F. Mcnittgray, N.J. Mankovich, J.G. Goldin, J. Hiller, L.S. Wilson, and Aberle, " Method for Segmentation Chest CT Image Data using an Anatomical Model-Preliminary Results," *IEEE Trans Medical Imaging*, vol.16, 828-839, 1997.

[4] Rushin Shojaii, Javed Alierezaie, and Paul Babyn, "Automatic Lung Segmentation in CT Images using Watershed Transform," *Proc. IEEE 8th International symposium on Computers and Communication*, vol.2, pp. II-1270-3, 2005.

[5] Hyoungseop Kim, Tooru Nakashime, Yoshinori Itai, Shinya Maeda, Joo Kooi Tan, and Seiji Ishikawa, "Automated Detection of Ground Glass Opacity from the Thoracic MDCT Images by using Density Features," *International Conference on Control, Automation and Systems*, pp. 1274-1277, 2007.

[6] K. Kanazawa, Y. Kawata, N. Niki, H. Satoh, H. Ohmatsu, R. kakinuma, M. Kaneko, K.Eguchi, and N.Moriyama, "Computer-Aided Diagnosis for Pulmonary Nodules based on Helical CT Images," *IEEE, Computerized medical image and graphics*, vol.22, no.2, pp. 157-167, 1998.

[7] A. El-Baz, A. A. Farag, R. Falk, and R.L. Rocco, "Detection, Visualization and Identification of Lung Abnormalities in Chest Spiral CT Scans: Phase I," *Information Conference on Biomedical Engineering*, Egypt 2002.

[8] Amal A. Farag, S.Y. Elhabian, S.A. Elshazly, and A.A. Farag, "Quantification of Nodule Detection in Chest CT: A Clinical Investigation Based on the ELCAP Study," *Proc. Of 2nd Int. Workshop on Pulmonary Image Processing in Conjunction with MICCAI-09*, pp. 149-160, Sept. 2009.

[9] Wei, Q., Hu, Y., Gelfand, G., and MacGregor, J.H., "Segmentation of Lung Lobes in High Resolution Isotropic CT Images," *IEEE Trans. Med. Imag*., vol. 56, no. 5, pp. 1-11, May, 2009.

[10] Anitha. S1 and Sridhar. S2, "Segmentation of Lung Lobes and Nodules in CT Images," *An International Journal SIPIJ* vol. 1, no. 1, Sept. 2010.

Kiran Thapaliya received the BIT (Bachelor in Information Technology) degree from College of Information Technology and Engineering, Purbanchal University, Nepal in 2009. He is currently pursuing the M.S. in Information and Communication Engineering, in Chosun University, South Korea. His research interests include bio-medical image processing, image enhancement and image segmentation.

Il-Cheol Park received his B.S. degree in information and communication engineering from the University of Chosun, Gwangju, South Korea, in 2010. Since March 2010, he has been the M.S. Student in information and communication engineering, University of Chosun, South Korea. His research interests include multimedia image processing, still image codec, video codec.

Goo-Rak Kwon received the M.S. degree from SungKyunKwan University, in Electronic Engineering, in 1999. He also received a Ph.D. degree from Korea University in Mechatronics in 2007. In 2008, he joined the Department of Information and Communication Engineering at Chosun University where he is currently an assistant professor. His interest research fields are A/V signal processing, multimedia communication, and applications. He is currently a Member in the IEEE, IEICE, and IS&T. He has been a Member in the Institute of Electronics Engineers of Korea (IEEK). And also he has served as an Editor for Journal of the Multimedia Society of Korea (KMMS).