Lung Area Extraction from X-ray CT Images for Computer-aided Diagnosis of Pulmonary Nodules by using Active Contour Model

NORIYASU HOMMA
Tohoku University
Cyberscience Center
6-6-05 Aoba, Aoba-ku
Sendai 980-8579
JAPAN
homma@ieee.org

SATOSHI SHIMOYAMA
Tohoku University
Graduate School of Medicine
2-1 Seiryo-machi, Aoba-ku
Sendai 980-8575
JAPAN
simoyama@yoshizawa.ecei.tohoku.ac.jp

TADASHI ISHIBASHI
Tohoku University
Graduate School of Medicine
2-1 Seiryo-machi, Aoba-ku
Sendai 980-8575
JAPAN
tisibasi@mail.tains.tohoku.ac.jp

MAKOTO YOSHIZAWA
Tohoku University
Cyberscience Center
6-6-05 Aoba, Aoba-ku
Sendai 980-8579
JAPAN
yoshizawa@ieee.org

Abstract: In this paper, we develop a lung area extraction technique from X-ray computed tomography (CT) images for computer-aided diagnosis (CAD) systems. In lung cancer cases, pulmonary nodules are typical pathological changes and thus they are the target to be detected by CAD systems. The isolated nodules can be detected more easily by CAD systems developed previously, while previous CAD systems are often hard to detect non-isolated nodules. The extraction technique can then be used for transforming the non-isolated pulmonary nodules connected to the walls of the chest into isolated ones. The technique proposed here is based on an active contour model, but such model is often trapped into a local optimum solution. To avoid the local optimum solutions, an essential core of the proposed technique is to select an appropriate initial contour by using an anatomical feature of the lung shape in X-ray CT slices. Some experimental results demonstrate the usefulness of the proposed technique for assisting the CAD systems to detect non-isolated nodules more accurately.

Key–Words: Computer aided diagnosis, Active contour model, Pulmonary nodules, Anatomical feature, and X-ray CT images.

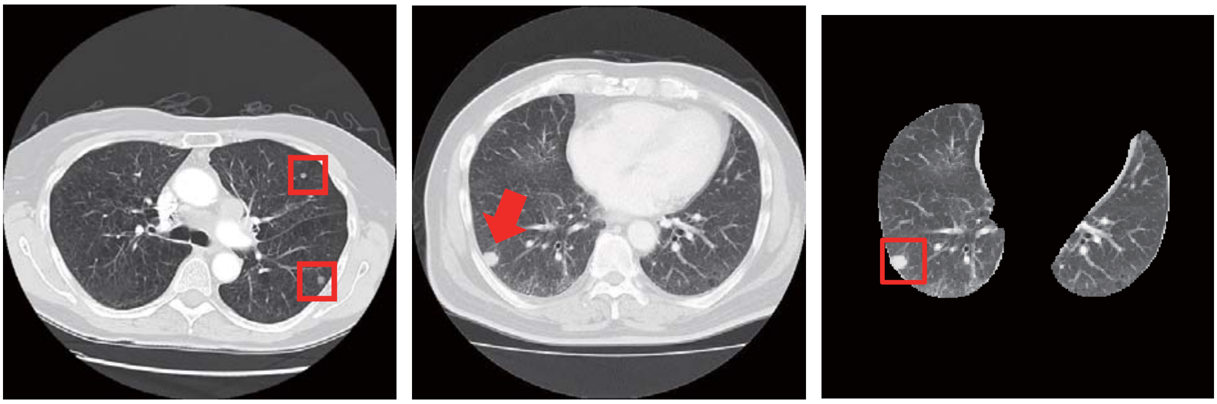
1 Introduction

An early stage detection of the lung cancer is extremely important for survival rate and quality of life (QOL) of patients [1]. Although a periodical group medical examination is widely conducted by diagnosing chest X-ray images, such group examination is not often good enough to detect the lung cancer accurately and there is a high possibility that the cancer at an early stage cannot be detected by using only the chest X-ray images. To improve the detection rate for the cancer at early stages, X-ray computed tomography (CT) has been used for a group medical examination as well [2, 3].

Using the X-ray CT, pulmonary nodules that are typical shadows of pathological changes of the lung cancer [4] can be detected more clearly compared to the chest X-ray examination even if they are at early stages. This is an advantage of the X-ray CT diagnosis. In fact, it has been reported that the survival rate of the later ten years can reach 90% after the detection at early stages using X-ray CT images [5].

On the other hand, using the X-ray CT may exhaust radiologists because the CT generates a large number of images (at least over 30 images per patient) and they must diagnose all of them. The radiologists' exhaustion and physical tiredness might cause a wrong diagnosis especially for a group medical examination where most of CT images are healthy and only very few images involve the pathological changes. Therefore, some computer-aided diagnosis (CAD) systems have been developed to help their diagnosis work [6, 7, 8, 12]. A core technique of CAD systems can be pattern recognition. Because of the fuzziness of the diagnosis target in the medical images, it is often different from a method for artificial targets. However, some extensions or modified versions of feature extraction techniques such as in [9, 10, 11] can be used for the CAD systems.

Miwa *et al.* [12] have developed a variable N-quoit filter to detect isolated pulmonary nodules and Homma *et al.* [13] have further improved the detec-



(a) Red squares show locations of isolated-nodules (b) Red arrow indicates a non-isolated nodule (c) Red square shows a converted isolated nodule from non-isolated one

Figure 1: (a) Isolated and (b) non-isolated nodules, and the conversion (c) from non-isolated into isolated one.

(a) Isolated target (b) Non-isolated target

Figure 2: A conceptual difference between isolated and non-isolated targets.

tion accuracy by discriminating between the isolated nodules and blood vessels those are both in a circle-like shape in CT images. The discrimination was achieved by developing a combination method of new features extraction techniques. These methods, however, aimed at only isolated circle-like shapes with the some morphological features, and thus non-isolated nodules (pathological changes) may not be detected by such methods. Indeed, it has been demonstrated that the conventional methods can detect isolated nodules as shown in Fig. 1 (a) [13], but cannot or hard to detect a non-isolated nodule as shown in Fig. 1 (b). A conceptual difference between isolated and non-isolated targets is depicted in Fig. 2.

Although non-isolated nodules are not very often seen in lung cancer observations, they can be a lung cancer with a high possibility and should not be missed from the viewpoint of the early stages detection of cancers [5].

In this paper, to improve the detection rate of such non-isolated nodules, we propose a technique transforming the non-isolated nodules connected to the walls of the chest into isolated ones that can be detected more easily by the conventional CAD systems aiming at isolated nodules. The transformation of Fig. 2 (a) into (b) can be achieved by extracting the lung area from the original CT image as shown in Fig. 1 (c).

The rest of this paper consists of as follows. In section 2, a fundamental theory and a local optimum problem of active contour models [14] that can be used for such extraction will be introduced. Then, by setting appropriate initial contours for solving the local optimum problem, a novel extraction technique based on the contour model will be developed in section 3. Experimental results with clinical data of X-ray CT images will be discussed to demonstrate the usefulness of the proposed method in section 4. Concluding remarks will be given in section 5.

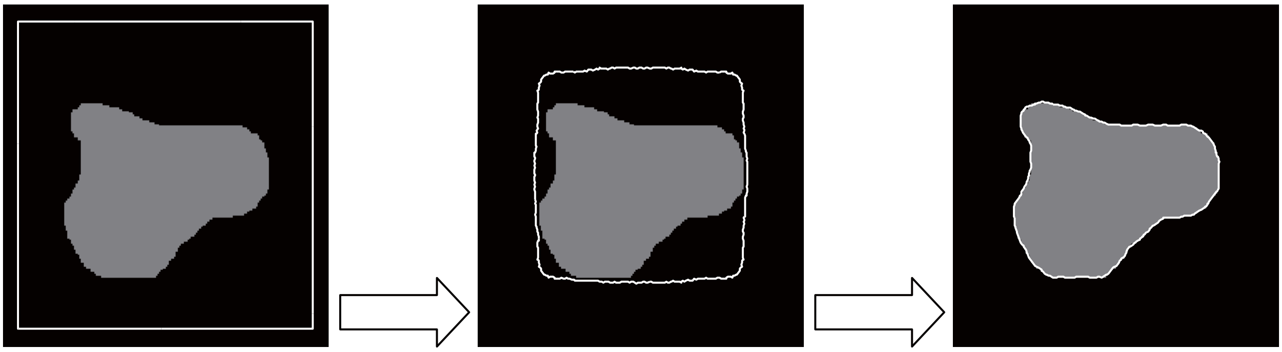
2 Active Contour Model

The active contour model proposed by Kass *et al.* [14] is a gradient decent-based optimal method. The optimality can be defined by an energy function E. The time evolution of the model is controlled by the following partial differential equation.

$$\frac{\partial v}{\partial t} = -\eta \frac{\partial E}{\partial v} \tag{1}$$

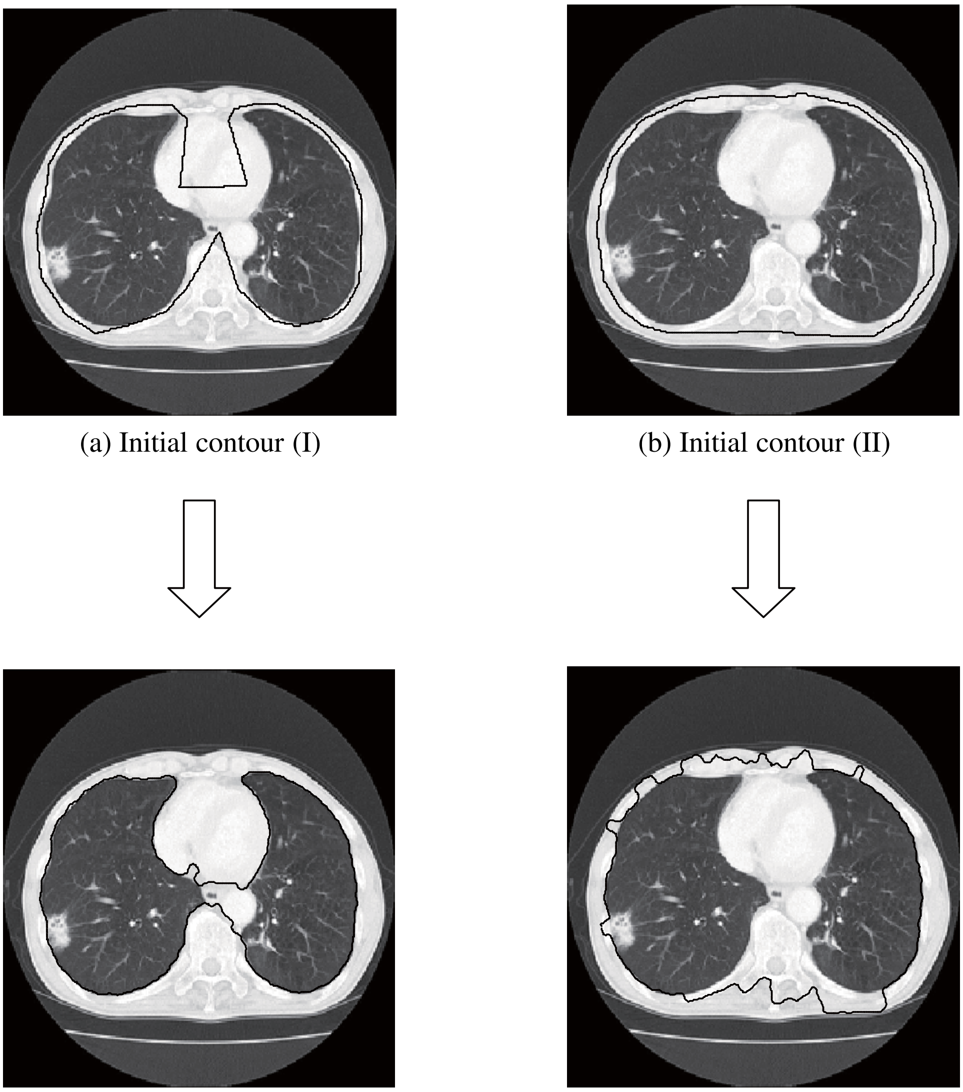
where $v(t, x, y)$ is a function of time t and coordinates x and y in the two dimensional space of the original image. $\eta > 0$ is a coefficient. The contour can be defined by a set of coordinates (x, y) satisfying a condition $v = L$ where L is a constant. Obviously, the final contour evolved by Eq. (1) is depended on the energy function E.

A well known and simple energy function is related to the edge of the original image and can be de-



(a) Initial contour (b) Contour in a halfway (c) Final contour

Figure 3: A sample time evolution of active contour. White lines show contours.



(a) Initial contour (I) (b) Initial contour (II)

(c) Final contour for the initial contour (I) (d) Final contour for the initial contour (II)

Figure 4: Effect of initial contours on the final results: Examples using the same lung X-ray CT image. Black lines near the walls on the CT images are contours.

fined as follows.

$$E = -\int_{\Omega} |\nabla I(x, y)|^2 dxdy \quad (2)$$

where $I(x, y)$ is a pixel value at the coordinates (x, y) and $\nabla I(x, y)$ is the spatial gradient of the pixel value. Ω is a domain of the coordinates (x, y) on the contour, i.e, $\Omega = \{(x, y)|v(x, y) = L\}$. By using the energy function E in Eq. (2), the final contour may be on an edge of the original image in which the gradient of the pixel value is the local maximum. Fig. 3 shows an example of the time evolution of the contour given by the active contour model where the energy function was defined by Eq. (2).

Since the active contour model is controlled by a gradient-decent evolution as mentioned above, the final result is also depended on the initial settings of the contour. In other words, such model can converge to a local optimal solution instead of the global optimal one. Thus, an appropriate setting of the initial contour is required to obtain the desired contour as well as the right design of the energy function. Fig. 4 shows examples illustrating results obtained from different initial contours for the same X-ray CT image. In fact, as is clear from this figure, the results are quite different from each other.

In addition, note that the result (I) in Fig. 4 (c) may be more desirable than the result (II) in Fig. 4 (d) because the result (I) seems more similar to the target contour inside the walls of the chest. This is because the initial contour (I) in Fig. 4 (a) is more similar and thus appropriate than the initial contour (II) in Fig. 4 (b). Consequently, if an initial contour as similar as possible to the desirable contour could be given, it may be expected that the final result is the most desirable one since the number of local optimal contours encountered during the time evolution can be the minimum compared to those for the other initial settings.

3 Proposed Method

As expected in the last paragraph of section 2, the local optimum problem can be avoided by starting from the appropriate initial contours. Note that a lung shape changes smoothly in axial direction as shown in Fig. 5 and recently the interval between X-ray CT slices next to each other is at most 10 [mm] in the direction. Then lung shapes in CT slices next to each other are almost the same or at least similar as shown in original CT images of Fig. 6. Thus a novel technique proposed here initializes the contour by using such anatomical feature of the lung shape. That is, the final contour obtained from the active contour model on the CT slice next to a target slice can be an appropriate candidate for the initial contour of the target CT

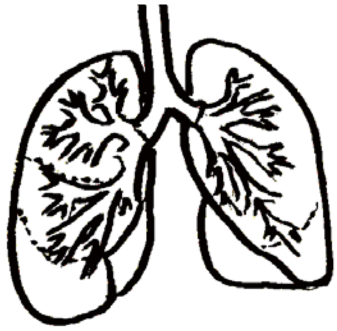


Figure 5: A schema of a human lung.

slice. This is a key idea of the proposed initialization. Let us define, in this paper, a lung area as inside the thorax that includes the center area of heart and aorta, and consider the walls of the chest that does not include the center area.

A flowchart of the proposed algorithm for extracting the lung area is shown in Fig. 7. In this algorithm, only the first CT slice is needed to be initialized in a specific way and called the *initial slice* of a series of the slices. Because of the specific initialization, steps (i) and (ii) in the flowchart for the initial slice are different from those of the other slices. In the followings, it is assumed, for simplicity, that the algorithm processes the series of CT slices from the head to the legs in the axial direction, but the algorithm is the same for the reverse direction.

(i) Selection of the target slice: If the current target is the initial slice of the series, select a slice without non-isolated nodules connected to the walls of the chest. Otherwise, select the slice below the previous target slice.

(ii) Initialization: There are many local optima during the time evolution of the model due to the edges created by the costae (bones) in the walls of the chest as shown in Fig. 4 (d). The final contour of the previous target slice can be a good candidate for the initial contour of the current slice as described above. The initialization except for the initial slice can thus be done easily by setting the candidate.

There is, however, no previous final contour for the initial slice. In this case, to remove such undesirable edges, an equalization of the pixel values that are larger than a threshold is conducted within the walls of the initial slice. The equalization can be given as follows.

$$I'(x, y) = \begin{cases} I_{\max}, & (I(x, y) > I_{Th}) \\ I(x, y), & (\text{otherwise}) \end{cases} \quad (3)$$

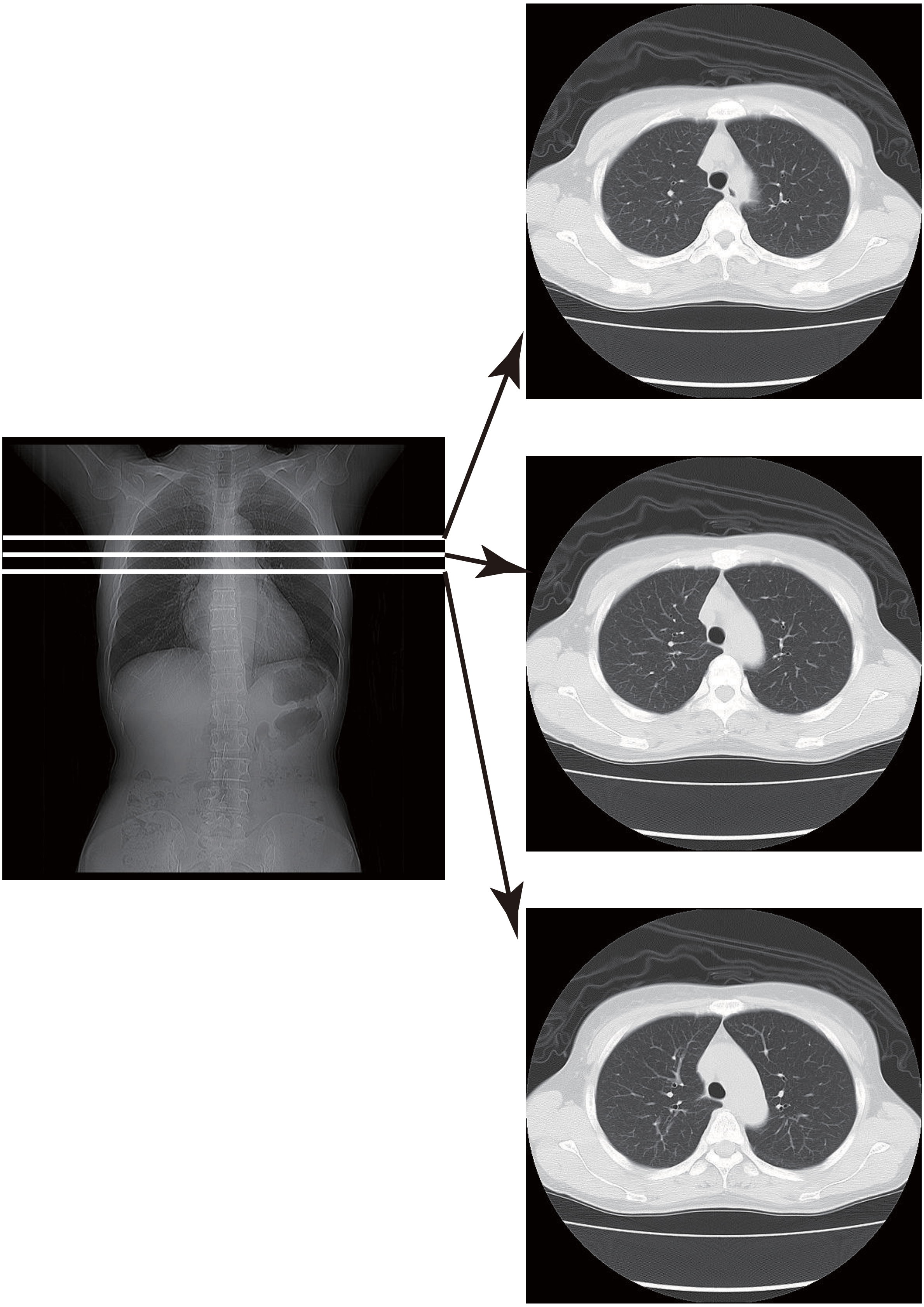


Figure 6: Similar lung shapes between CT slices next to each other.

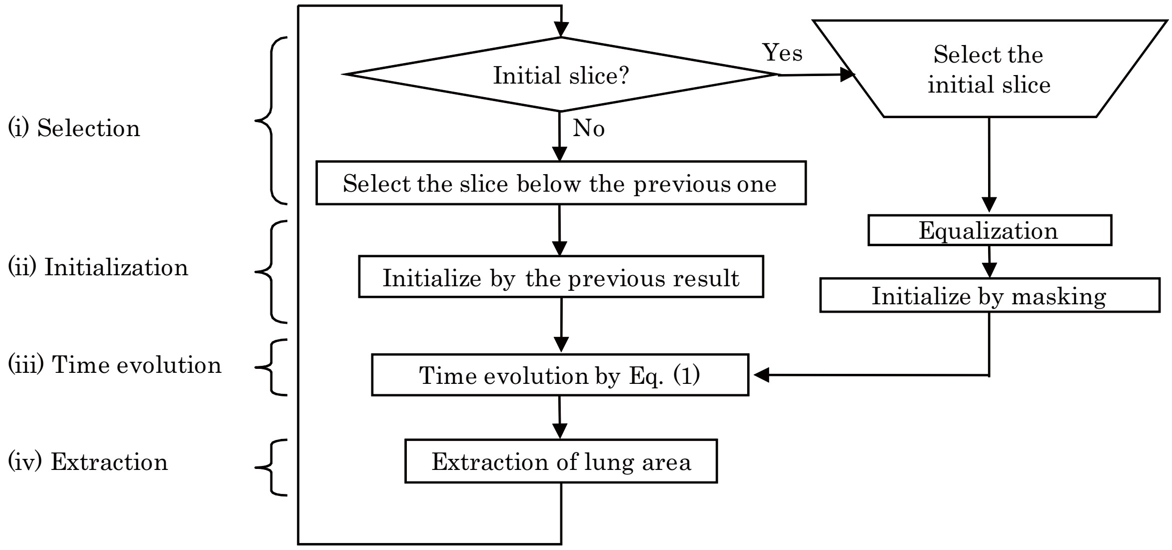


Figure 7: Flowchart of the proposed method.

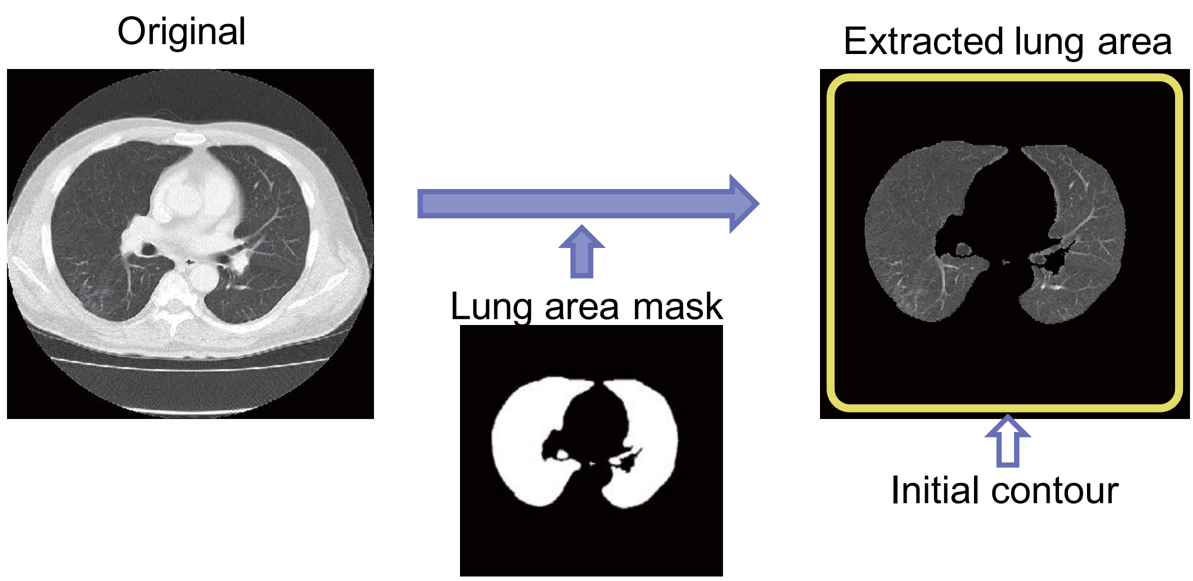


Figure 8: A mask processing to extract the lung area.

where $I'(x, y)$ denotes a new pixel value after the equalization, I_{Th} is the threshold, and $I_{\max}$ is the maximum pixel value that usually represents the white color.

As shown in Fig. 8, lung area of the initial slice can be extracted by using a mask processing. Then, a good result can be obtained from any contour outside the mask area. Note that lung area, however, could not often be extracted correctly if there is a non-isolated nodule connected to the walls of the chest as shown in Fig. 9. In this case, the non-isolated nodule that we want to detect is regarded as outside the lung area and thus cannot be detected by the mask processing. This is only the reason why we need to select the initial slice manually.

(iii) Time evolution: By using Eq. (1), the final result for the current target slice selected in step (i) can be obtained from the contour initialized in step (ii).

(iv) Extraction: The lung area for the current target is extracted as the inside the final contour obtained in step (iii).

Steps (i) - (iv) are repeatedly conducted until all lung areas in all CT slices are extracted.

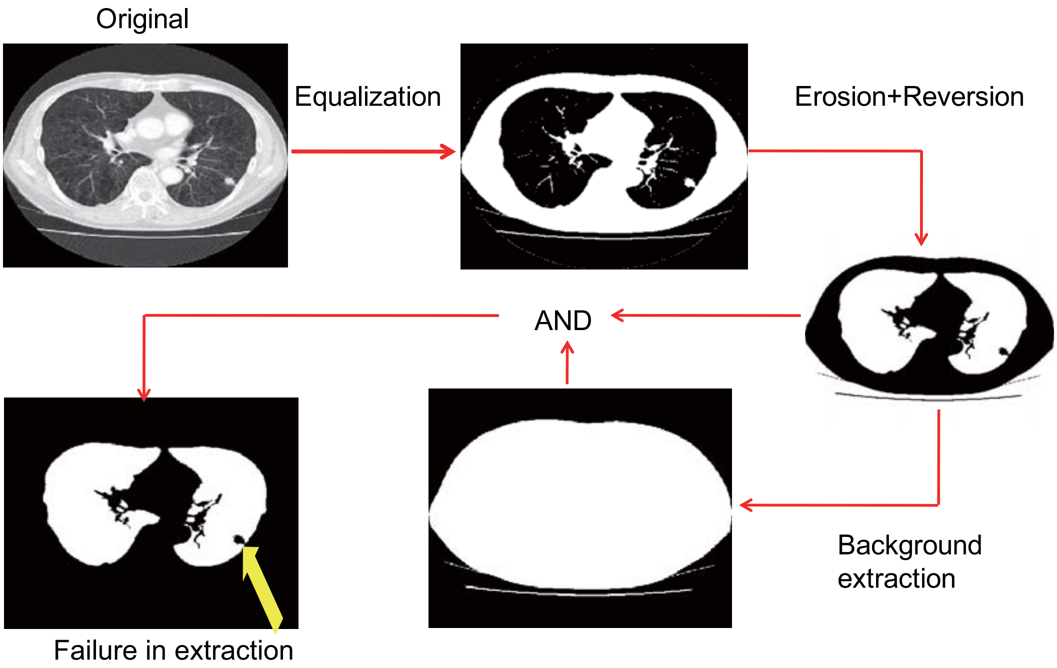


Figure 9: A failure case of the mask processing for a slice where there is a non-isolated nodule connected to the walls of the chest.

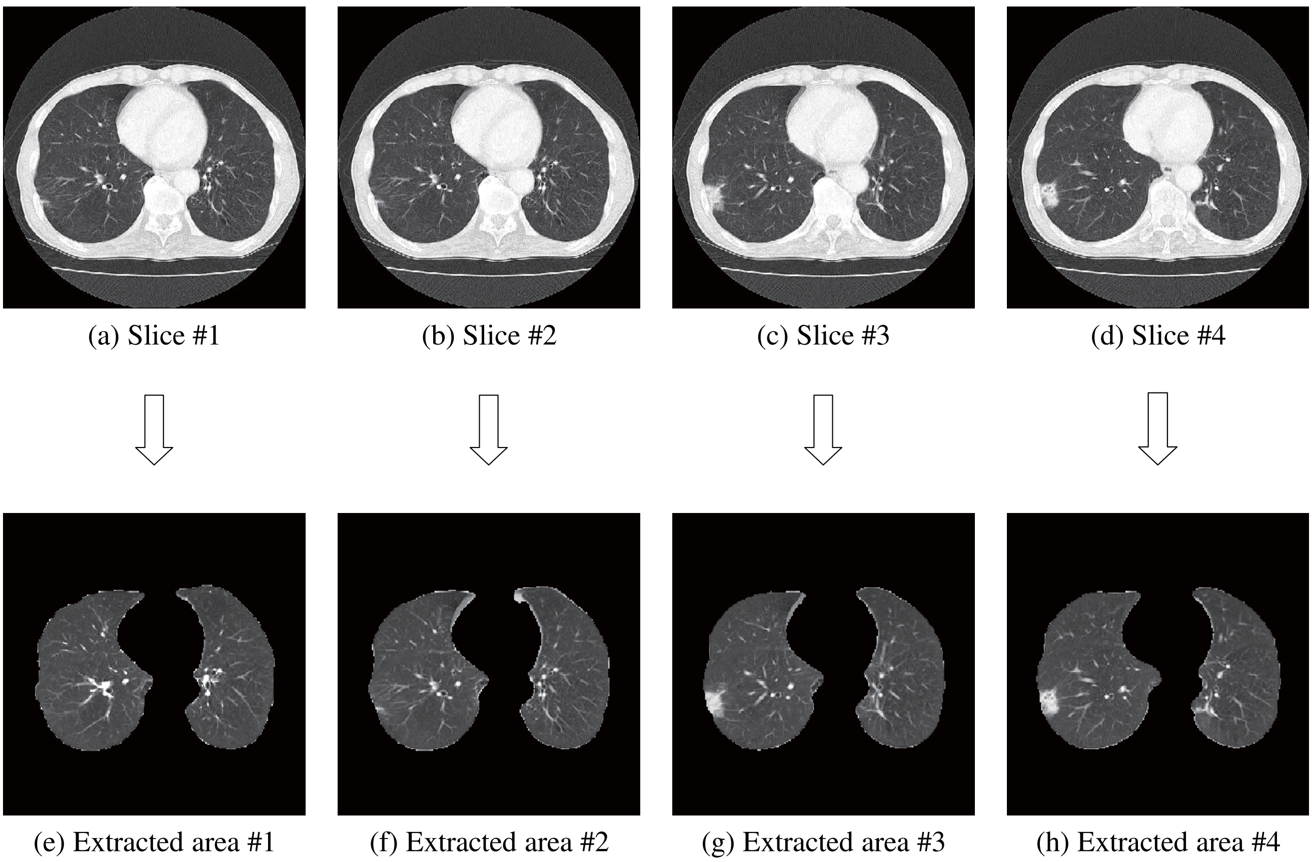


Figure 10: Extracted results for case 1 by the proposed active contour method. (a)-(d): Original CT images. (e)-(h): Extracted lung areas.

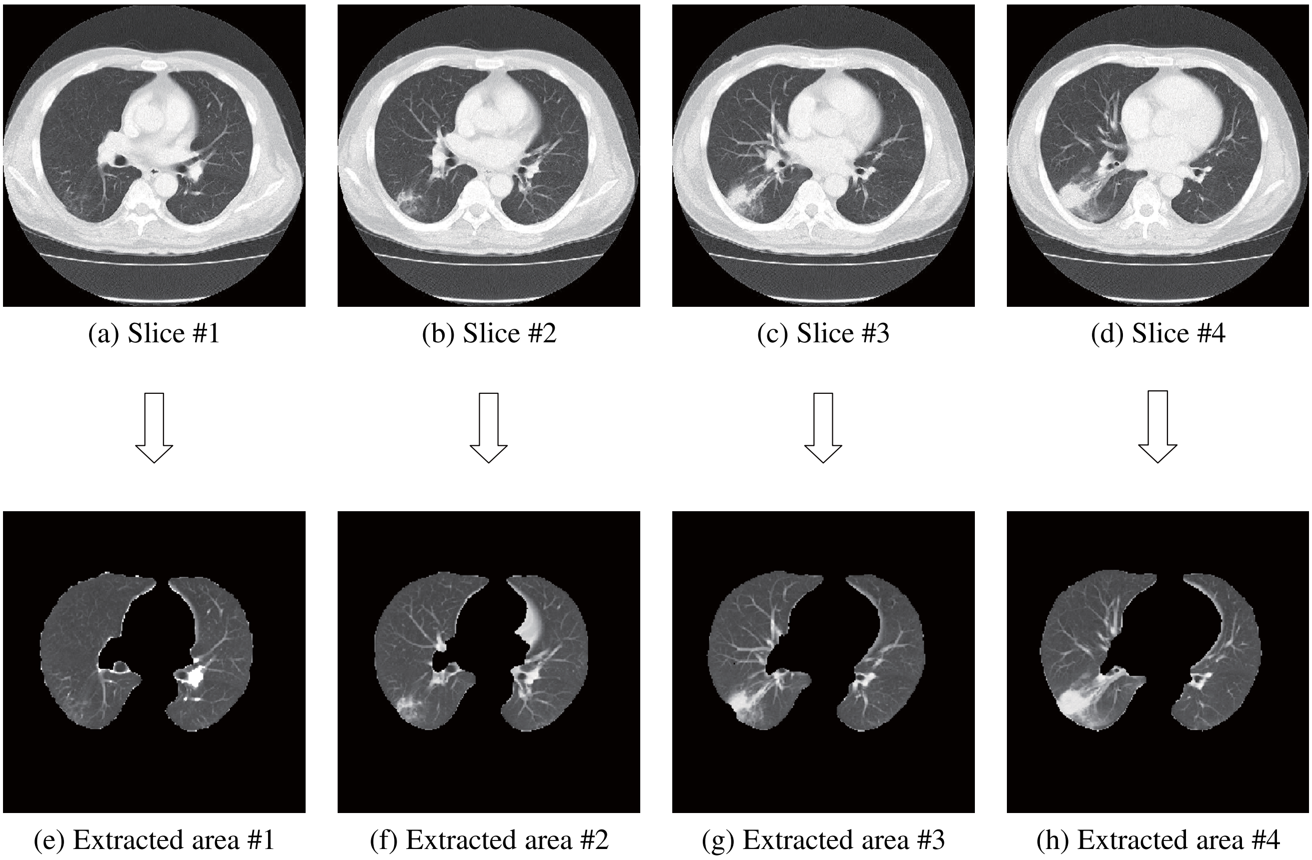


Figure 11: Extracted results for case 2 by the proposed active contour method. (a)-(d): Original CT images. (e)-(h): Extracted lung areas.

4 Results and Discussions

We have tested the proposed method using an extraction task in which the clinical CT images [15] including non-isolated nodules connected to the walls of the chest are used. Examples of the extraction results are shown in Figs. 10 and 11. It is clear that the proposed method can extract the lung area including the non-isolated nodules.

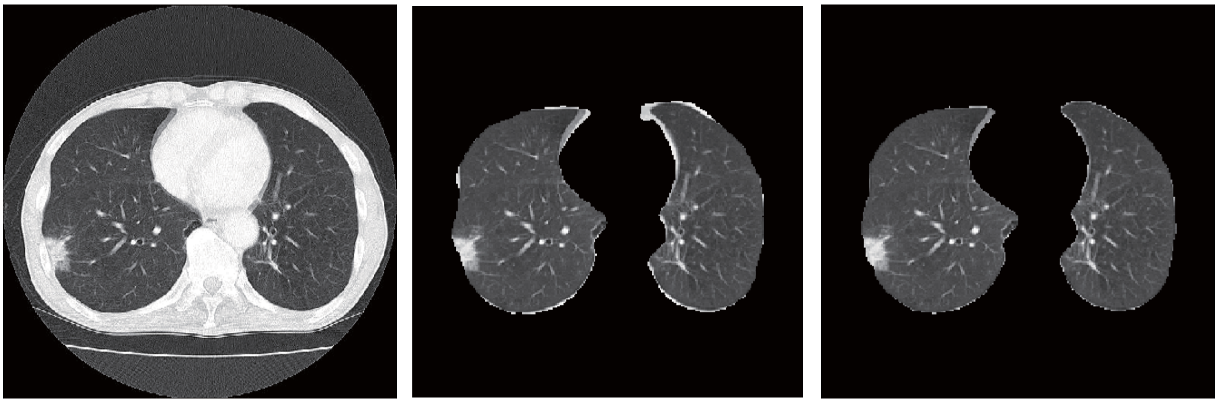
Extracted results by the initial and the final contours for the original slice shown in Fig. 10 (c) are shown in Fig. 12. Note that the initial contour that is the final contour obtained in the above slice in Fig. 10 (f) is similar enough and thus, the final result in Fig. 10 (g) is good enough.

On the other hand, there are a few examples in which non-isolated nodules were not extracted as the lung area, but regarded as within the walls. In such case, still non-isolated nodules cannot be detected by the conventional CAD systems for the isolated nodules. This problem may, however, be solved by designing a further appropriate energy function. That is, note that the contour curvature of the walls changes smoothly in general, but the curvature involving the connected nodules changes more sharply. Differences in the curvature may be incorporated into a new energy function to discriminate such non-isolated nodules from the walls of the chest.

Furthermore, the active contour model has an ability of making a smooth contour line even if the initial contour has a sharp corner with a high curvature. We can then select the initial slice in an automatic way, i.e., random selection, the first, middle, or last slice of the series, and so on. The masking problem with the initial slice including non-isolated nodules connected to the walls of the chest can be solved by applying the proposed algorithm with an appropriate parameters setting repeatedly to the same series. This direction of future works can be important for clinical use.

5 Conclusions

In this paper, we have taken into account non-isolated nodules connected to the walls of the chest that cannot be detected by the conventional CAD systems for lung cancer. To detect such nodules, we have proposed a technique to transform the non-isolated nodules into the isolated ones by using an active contour



(a) Original CT image (same as in Fig. 10 (c))
(b) Extracted area by the initial contour that is the final contour of the above slice
(c) Extracted area by the final contour (same as in Fig. 10 (g))

Figure 12: The appropriate initial contour and the final contour for the CT slice #3 in Fig. 10 (c).

model to extract the lung area from the original CT image. The promising results suggest that the detection accuracy of the CAD systems can be further improved by incorporating the proposed technique.

Acknowledgements: This work was partially supported by The Ministry of Education, Culture, Sports, Science and Technology under Grant-in-Aid for Scientific Research #19500413 and the Okawa Foundation.

References:

[1] T. Naruke, T. Goya, R. Tsuhiya, et al., "Prognosis and survival in resected lung carcinoma based on the new international staging system," J. Thorac Cardiovasc Surg, Vol. 96, pp. 440-447, 1988.

[2] Takeshi Iinuma, Yukio Tateno, Toru Matsumoto, Shinji Yamamoto, and Mitsuomi Matsumoto: "Preliminary Specification of X-ray CT for Lung Cancer Screening (LSCT) and its Evaluation on Risk-Cost-Effectiveness," Nippon Acta Radiologica, Vol. 52, pp. 182-190, 1992 (in Japanese).

[3] Shinji Yamamoto, Ippei Tanaka, Masahiro Senda, Yukio Tateno, Takeshi Iinuma, Toru Matsumoto, Mitsuomi Matsumoto: "Image Processing for Computer Aided Diagnosis in the Lung Cancer Screening System by CT(LSCT)," Trans. Institute of Electronics, Information and Communication Engineers, Vol. 76-D-2, pp. 250-260, 1993 (in Japanese).

[4] M. Prokop and M. Galanski, Spiral and Multislice Computed Tomography of the Body, Thieme Medical Publishers, Stuttgart, 2003.

[5] International Early Lung Cancer Action Program (I-ELCAP) : Survival of Patients with Stage I Lung Cancer Detected on CT Screening, NEJM, Vol. 355, No. 17, pp. 1763-1771, 2006.

[6] T. Okumura, T. Miwa, J. Kako, S. Yamamoto, M. Matsumoto, Y. Tateno, T. Iinuma and T. Matsumoto, "Variable-N-Quoit filter applied for automatic detection of lung cancer by X-ray CT," Proc. of Computer-Assisted Radiology, pp. 242-247, 1998 (in Japanese).

[7] Y. Lee, T. Hara, H. Fujita, S. Itoh and T. Ishigaki, "Nodule detection on chest helical CT scans by using a genetic algorithm," Proc. of IASTED International Conference on Intelligent Information Systems, pp. 67-70, 1997.

[8] Shinji Yamamoto, Masato Nakayama, Masahiro Senda, Mitsuomi Matsumoto, Yukio Tateno, Takeshi Iinuma, Tohru Matsumoto: "A Modified MIP Processing Method for Reducing the Lung Cancer X-ray CT Display Images," Medical Imaging Technology, Vol. 12, No. 6, 1994 (in Japanese).

[9] Cristian Molder, Mircea Boscoianu, Iulian C. Vizitiu, Mihai I. Stanciu: "Dicision Fusion for Improved Automatic License Plate Recognition," WSEAS Trans. Information Science and Applications, Vol. 6, Issue 2, pp. 291-300, 2009.

[10] Zhengmao Ye, Habib Mohamadian, Su-Seng Pang, Sitharama Iyengar: "Contrast Enhance-

ment and Clustering Segmentation of Gray Level Images with Quantitative Information Evaluation," WSEAS Trans. Information Science and Applications, Vol. 5, Issue 2, pp. 181-188, 2008.

[11] Santiago Akle, Maria-Elena Algorri, Ante Salcedo: "Use of Wavelet-based Basis Functions to Extract Rotation Invariant Features for Automatic Image Recognition," WSEAS Trans. Information Science and Applications, Vol. 5, Issue 5, pp. 664-673, 2008.

[12] Tomoko Miwa, Jun-ichi Kako, Shinji Yamamoto, Mitsuomi Matsumoto, Yukio Tateno, Takeshi Iinuma, Toru Matsumoto: "Automatic Detection of Lung Cancers in Chest CT Images by the Variable N-Quoit Filter," Trans. Institute of Electronics, Information and Communication Engineers, Vol. 82-D-II, pp.178-187, 1999 (in Japanese).

[13] N. Homma, K. Takei, and T. Ishibashi: "Combinatorial Effect of Various Features Extraction on Computer Aided Detection of Pulmonary Nodules in X-ray CT Images," WSEAS Trans. Information Science and Applications, Vol. 5, Issue 7, pp. 1127-1136, 2008.

[14] Michael Kass, Andrew Witkin, and Demetri Terzopoulos: "Snakes: Active Contour Models," International Journal of Computer Vision, pp. 321-331,1988.

[15] National Cancer Imaging Archive (NCIA), https://imaging.nci.nih.gov/ncia/faces/baseDef.tiles