Nodule Detection on Chest Helical CT Scans by Using a Genetic Algorithm

Yongbum Lee*, Takeshi Hara*, Hiroshi Fujita*
Shigeki Itoh** and Takeo Ishigaki***
* Department of Information Science, Faculty of Engineering, Gifu University
Yanagido 1-1, Gifu 501-11, Japan
** Department of Radiology, Nagoya Memorial Hospital,
Hirabari 4-305, Tenpaku-ku, Nagoya 468, Japan
*** Department of Radiology, Nagoya University School of Medicine,
Tsurumai-cho 65, Showa-ku, Nagoya 466, Japan

Abstract

The purpose of this study is to apply a genetic-algorithm (GA) template-matching method to detect lung nodules in chest helical x-ray CT (computed tomography) images. We combined GA and template matching to search the positions of nodules and to calculate adaptation scales of individuals on GA, respectively. We used four simulated nodules created by Gaussian distribution, whose sizes were different each other, as reference patterns in the GA template matching. The GA selected an adequate reference image from four ones and searched adequate positions to template matching. We used cross-correlation as similarity of tmplate matching and as adaptation scales of individuals on GA. It was possible to detect 23 nodules from 45 ones that did not touch lung walls without consideration of their sizes. It was also possible to detect all nodules that touched lung walls by using conventional template matching along lung walls. The total detection rate was approximately 67%. The number of false positives per one slice was over 10. To improve the detection performance and to decrease the number of false positives, we are now working on considering operators and their parameters of GA.

1. Introduction

Early detection of suspicious lesions is effective to reduce the number of death caused by lung cancer. Although a conventional chest radiogram has been used for the screening of lung cancer, mass screening by helical CT scan has started for the high risk groups of smokers in some facilities of Japan, and it shows good performance for the diagnosis. However numerous scans (approximately 30 scans per patient) have to be interpreted by a radiologist for each case. Therefore, development of computer-aided diagnosis (CAD) system which can assist the radiologists' decision is required. The purpose of this study is to develop a fundamental technique to detect lung nodules for the future CAD system in chest helical CT scans. We used an extended template matching technique for three-dimensional images based on a genetic algorithm (GA) in this study [1, 2].

2. Materials and methods

The detection technique is based on GA and template matching method (TM) in a three-dimensional space. GA is a probability search method based on the evolution principal of living things. TM is a simple image recognition method that can determine the position of object image in other images by comparing the similarity between reference and observed images. We combined GA and template matching to search the positions of nodules and to calculate adaptation scales of individuals on GA, respectively. We supposed that it was possible to detect a nodule to set the higher adaptation scales surrounding the nodules, so that the individuals could gather on the place for the environmental configuration. We used a similarity, which was defined by the difference of CT values between three-dimensional images, for adaptation scales of

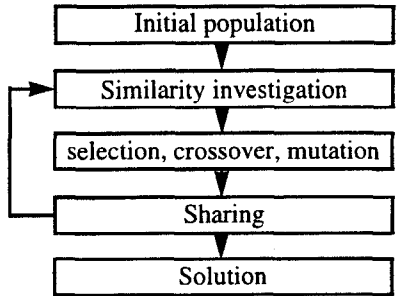


Fig.1 GA process

Proceedings of the 1997 IASTED International Conference on Intelligent Information Systems (IIS '97)
0-8186-8218-3/97 $10.00 © 1997 IEEE

individuals. The similarity is calculated by template matching between two images.

2.1 Setting GA

Figure 1 shows the GA process. Each individual of GA has a chromosome that includes information for locations in three-dimensional space of helical CT images and to select an adequate reference image. A chromosome is represented by a binary digit system and is applied genetic operators as selection, crossover and mutation. 50% individuals of recessive are selected and are replaced with crossovered ones of 50% of dominance. Here, we used one-point crossover and sharing to get some detection points [3]. The probability of mutation was 10% for all chromosomes. The number of individuals and generations were 124 per generation and 200 at maximum in each search area, respectively.

2.2 GA Template matching

Template matching is one of the most fundamental methods to detect an object within an image field. If the similarity between an unknown object (observed image) and the template (reference image) is sufficiently high, the unknown object can be considered as the detection object. We used three-dimensional images of the chest helical CT scans and simulated nodule images for observed and reference images, respectively. Each of the CT scans (observed image) has 512 ~512 pixels and consists of 27 scan images. Table 1 shows the specification of chest helical CT scans used in this study. We made four artificial simulated nodule images which includes a similar image to be detected (Fig.2). Here, lung nodule shadow generally shows a spherical shape and the CT values tend to have a Gaussian distribution, we used some simple sphere models with three-dimensional Gaussian distribution as simulated nodules of reference images. Four reference images ranged 3 slices and their diameter were 10, 20, 30 and 40 pixels. The observed image was cut off at a position which was determined by place information in each chromosome of GA. At the same time, each chromosome selected a reference image from four ones. The size of cutting image was as same as reference image. The similarity between a cut and a reference images was calculated to evaluate

Table 1 Specification of helical CT scans

Tube Voltage	120 kV
Tube Current	50 mA
Slice Thickness	10 mm
Scan Time	1 sec/round
Table Feed Rate	20 mm/sec

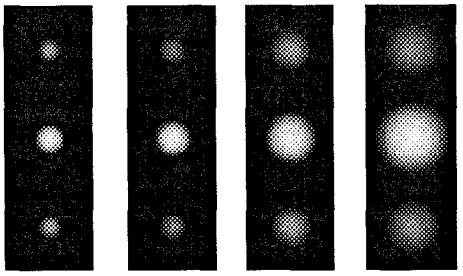


Fig.2 Reference images

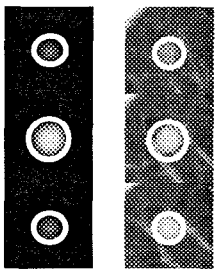
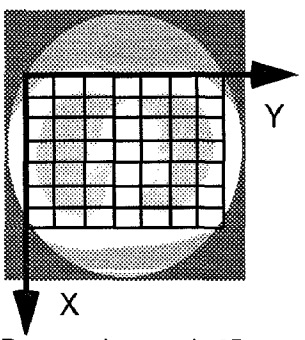
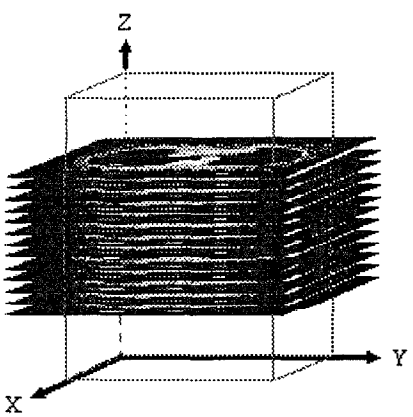


Fig.3 Region to be evaluated (Inside of circles)



(a) Rectangular area in 2D space



(b) Rectangular prism for determination of searching area in 3D space

Fig.4 Determination of searching area

Proceedings of the 1997 IASTED International Conference on Intelligent Information Systems (IIS '97)
0-8186-8218-3/97 $10.00 © 1997 IEEE

adaptation scale of the individual on GA. We used cross-correlation as the similarity. The similarity was only evaluated within the common area that contained only a part of nodules between a cut and a reference images. Therefore it was possible to evaluate the similarity without depending on the background patterns (Fig.3).

2.3 Automatic determination of search area

The number of searching iterations was decreased by determining a minimum area including whole lung fields. Therefore, the rectangular area including approximated lung fields on each slice image was determined by the pixel thresholding and the extraction of the largest rectangle in all slices [Fig.4(a)]. The region of interest for searching was indicated by an inner part of the rectangular prism as illustrated in Fig.4(b). Finally, the determinded region was segmented in 7 areas at each direction of x and y axis.

3. Result and discussion

Our method was applied to 11 clinical cases (4 patients, 291 slices) of the chest helical CT scans in which 67 nodules existed. Table 2 shows the detection rate in terms of the patient (A~D) and the diameter of the nodules. We used four Gaussian-distributed spheres as reference images, which sizes were 10, 20, 30 and 40 pixels in diameter. So we were able to detect 23 nodules from 67 ones without consideration of their sizes. All detected nodules in this detection step did not touch the lung walls (Fig.5). Table 3 shows the detection rate in terms of the patient (A~D) and the actual nodule locations. Nodules that touch lung walls were not detected (Fig.6). Therefore we used conventional template matching along lung walls as an effective technique to detect nodules that touch lung walls.

3.1 Conventional template matching along lung walls

Figure 7 shows the process of conventional template matching along lung walls. The rough lines of lung walls were determined by the pixel thresholding (Fig.8). Squre region as search area were determined along the rough line of lung walls (Fig.9). Conventional template matching was employed in determined squre region. We used eight reference images to detect nodules that touch lung walls. A reference image was selected by considering an angle of one position on the lung wall. Reference images were made from a circle model of two-dimensional Gaussian distribution, which diameter was approximately 10 pixels. The length of circle model was expanded to the size of 0.8 times, and was divided into two parts lengthwise. Then we

Table 2 Initial detection rates in terms of patient and size

Patient/Diameter	5~	10~	15~	20~	35~	Total
A (3scans)	0/0	0/0	0/0	0/0	1/3	1/3
B (2scans)	0/2	2/4	0/0	0/2	0/0	2/8
C (2scans)	0/14	0/10	0/2	0/0	2/2	2/28
D (4scans)	1/4	9/12	4/8	4/4	0/0	18/28
Total	1/18	11/28	4/10	4/6	3/5	23/67

Table 3 Initial detection rates in terms of patient and nodule location

Patient/Location	In the lung area	On the lung wall	Total
A (3scans)	0/0	1/3	1/3
B (2scans)	0/4	2/4	2/8
C (2scans)	0/14	2/14	2/28
D (4scans)	0/4	18/24	18/28
Total	0/22	23/45	23/67

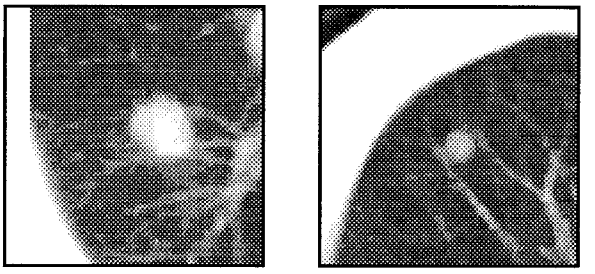


Fig.5 Nodules in the lung area

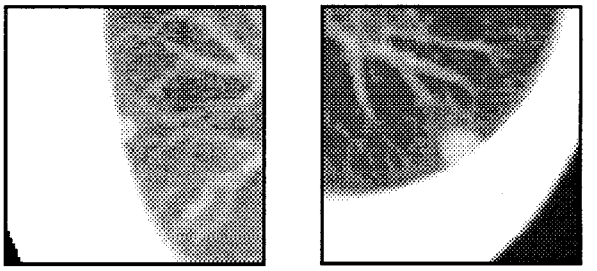


Fig.6 Nodules on the lung wall

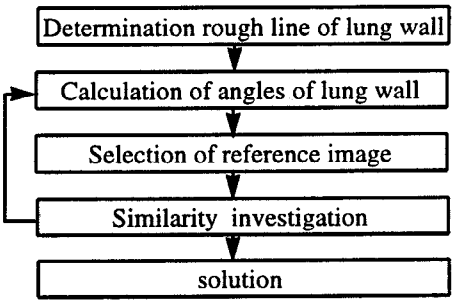


Fig.7 Process of template matching along the lung walls

Proceedings of the 1997 IASTED International Conference on Intelligent Information Systems (IIS '97)
0-8186-8218-3/97 $10.00 © 1997 IEEE

got a semicircular simulated nodule. Finally, we got eight reference images by rotating the semicircular simulated nodule 45 by 45 (Fig.10). Table 4 and Table 5 shows the detection rate when we used both GA template matching and conventional template matching along lung walls. It was possible to detect all nodules that touch lung walls by using conventional template matching along lung walls. But the number of false positives per one slice image was over 10. We supposed that the number of false positives was large because we expected to detect subtle nodules which diameter was less than 10 pixels.

3.2 Nodules which did not detect

It was not able to detect nodules that touch others except lung walls, example mediastinum, internal organs (Fig.11). We suppose it is posible to detect those nodules by determining exact lung region and improving of conventional template matching along lung walls.

4. Conclusion

It was possible to detect 23 nodules from 45 ones that did not touch the lung walls without consideration of their sizes by adding parameters of GA to select the fittest image in four reference ones whose sizes were different each other. It was possible to detect all nodules that touched lung walls by using conventional template matching along the lung walls. The total detection rate was approximately 67%. The number of false positives per one slice image was over 10. Now we are working on decreasing the number of false positives and considering operators and their parameters of GA to improve the detection performance.

References

[1]Takeshi Hara, Hiroshi Fujita : Template Matching of Gray-Scale Images Using a Genetic Algorithm : The Transactions of IEICE (Institute of Electronics, Information and Communication Engineers)D-II, Vol.J78-D-II, No.2, pp.385-388, 1995

[2]Takeshi Hara, Hiroshi Fujita, Hitoshi Yoshimura, Tsuneo Matsumoto : Automated Detection of Nodular Shadows in Chest Radiography : Application of Genetic Algorithm : Medical Imageing Technology, Vol.15, No.1, pp.73-81, 1997

[3]Hitoshi Sakano, Hideo Saito : Temptation and Nightmare of Genetic Algorithms in the Field of Pattern Recognition : The Transactions of IEICE, Vol.79, No.10, pp.961-966, 1996

Table 4 Improved detection rates in terms of patient and size

Patient/Diameter	5~	10~	15~	20~	35~	Total
A (3scans)	0/0	0/0	0/0	0/0	1/3	1/3
B (2scans)	2/2	4/4	0/0	0/2	0/0	6/8
C (2scans)	12/14	0/10	2/2	0/0	2/2	16/28
D (4scans)	1/4	9/12	8/8	4/4	0/0	22/28
Total	15/18	15/28	8/10	4/6	3/5	45/67

Table 5 Improved detection rates in terms of patient and nodule location

Patient/Location	In the lung area	On the lung wall	Total
A (3scans)	0/0	1/3	1/3
B (2scans)	4/4	2/4	6/8
C (2scans)	14/14	2/14	16/28
D (4scans)	4/4	18/24	22/28
Total	22/22	23/45	45/67

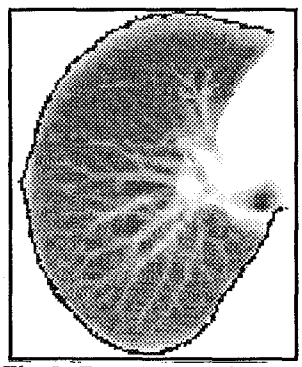


Fig.8 Rough line of lung walls

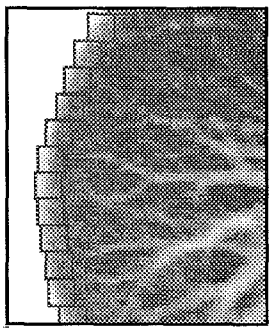


Fig.9 Square regions along lung walls

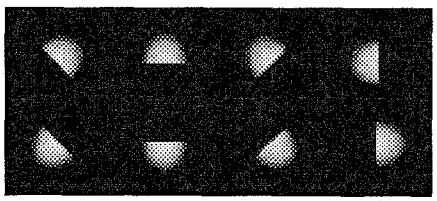
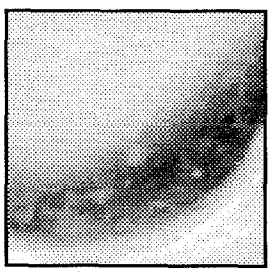


Fig.10 Reference images for conventional template matching along lung walls

(a) A nodule touching mediastinum



(b) A nodule touching internal organs

Fig.11 Missed nodules in detection

Proceedings of the 1997 IASTED International Conference on Intelligent Information Systems (IIS '97)
0-8186-8218-3/97 $10.00 © 1997 IEEE