Extraction of tumor regions keeping boundary shape information from chest X-ray CT images and benign/malignant discrimination

Yasushi Hirano[a], Jun-ichi Hasegawa[b], Jun-ichiro Toriwaki[a], Hironobu Ohmatsu[c] and Kenji Eguchi[d]

[a] Faculty of Engineering, Nagoya University,
Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8603 Japan

[b] School of Computer and Cognitive Sciences, Chukyo University,
101 Tokodachi, Kaizu-cho, Toyota, Aichi, 470-0393 Japan

[c] National Cancer Center East,
6-5-1 Kashiwanoha, Kashiwa, Chiba, 277-0882 Japan

[d] National Shikoku Cancer Center,
13 Horinouchi, Matsuyama, Ehime, 790-0007 Japan

Abstract : In this paper, we propose a new method to extract tumor regions from thin-slice chest X-ray CT images. The tumor regions which are used for benign/malignant discrimination must keep boundary shape information as exactly as possible, because the tumor regions specify the regions for which features are calculated. The proposed method is based on the thresholding of CT values and distance. Since the change of CT values suggest the existence of tumors, it is reasonable to use thresholding of CT values for extraction of tumor regions. We assume that the position of tumors are known beforehand, because the thin-slice chest X-ray CT images are taken with the view to be used in the close medical examination. We applied the method to seventy-eight practical CT images, and sixty-seven tumors were extracted properly. Furthermore, the benign/malignant discrimination was carried out suitably using the tumor regions extracted by the proposed method. The correct classification rate was improved from 0.84 to 0.91. This shows the usefulness of the method developed here.

Keywords : tumor extraction, boundary information, benign/malignant discrimination, chest X-ray CT image

1. INTRODUCTION

In this paper, we propose a new method to extract lung tumor regions in three dimensional (3D) chest X-ray CT images, keeping details of the shape in the boundary of tumors. Methods to extract lung tumor region have been reported already by several authors [1-3]. The goal of their methods was indication of the position and the rough size of tumors, because they intended to use the results for screening of lung cancer. Tumor regions which are used for benign/malignant discrimination in CAD(Computer Aided Diagnosis) system are required to have exact edge information, because tumor regions specify the regions for which features are calculated in order to discriminate for benign/malignant tumors and to quantify edge information itself. Many features for discrimination have been reported[4-6], but the accuracy of extracted figures were not enough for such classification.

The method proposed here is based on the thresholding against CT values and the distance transformation. Use of thresholding based on CT values is reasonable to extract a tumor region, because the significant difference in CT values is frequently observed with the existence of tumor.

2. METHOD

In the proposed method, the central part of the tumor region is first extracted roughly, and then the peripheral region of the tumor is added to it. The aim of the method is to be used for a close medical examination, so that we assume that the position of tumor is known, and also assume that at least one point (denoted by P_{sp}) near the center of the tumor region indicated beforehand.

The procedure is as follows, where α and β are the predetermined constants.

[Step 1] Thresholding of the CT image (threshold : T_{ct}) is performed. The voxels which have the CT values less than the threshold T_{ct} are set to 0, and the others are set to 1. Voxels outside the lung region are also set to 0. The resulting region is called the candidate region.

[Step 2] Euclidian distance transformation of the candidate region is performed. The distance at the point P_{sp} is denoted as R_{sp}.

[Step 3] Thresholding to the candidate region against the distance(threshold : $T_{dist} = R_{sp} - \beta$) is performed. Voxels which have the distance less than the threshold are set to 0.

[Step 4] A connected component which contains P_{sp} is called the central region 1 (C_1).

[Step 5] Inverse distance transformation of the boundary voxels of C_1 is applied after the constant α is added to the distance values of all voxels. The obtained region is called the expanded tumor region.

[Step 6] The set of the voxels of the candidate region which exist in the expanded tumor region and have the values more than T_{dist} is called the central region 2 (C_2).

[Step 7] If C_2 coincides with C_1, go to [Step 8]. If not, rename C_2 as C_1 and go to [Step5].

[Step 8] Logical AND of the expanded tumor region and the candidate region is calculated. Connected components which do not contain P_{sp} are deleted. The resulting region is called the rough tumor region.

[Step 9] The rough tumor region is eliminated from the set of candidate regions.

[Step 10] Connect components with the volume smaller than T_v is deleted. Remaining regions are called the detailed peripheral region.

[Step 11] Logical OR of the detailed peripheral region and the rough tumor region is calculated.

[Step 12] Connected components which do not contain the point P_{sp} are deleted. The resulting region is called the tumor region.

3. RESULTS

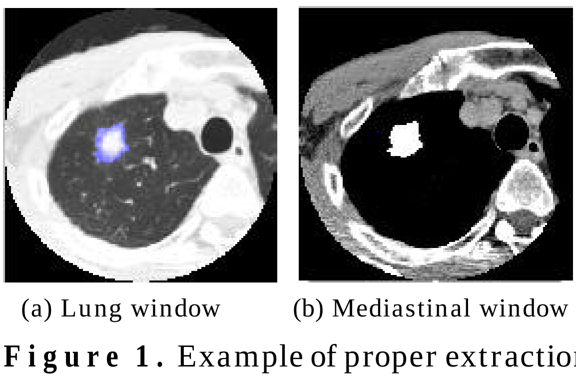
We applied the proposed method to seventy-eight chest X-ray CT images taken

Table 1. Specification of CT images

Slice size	512 x 512 [pixels]
Number of slices	43 ~ 64 [slices]
Reconstruction pitch	1.0 [mm]
Pixel size	0.29 ~ 0.42 [mm]
Slice thickness	2.0 [mm]

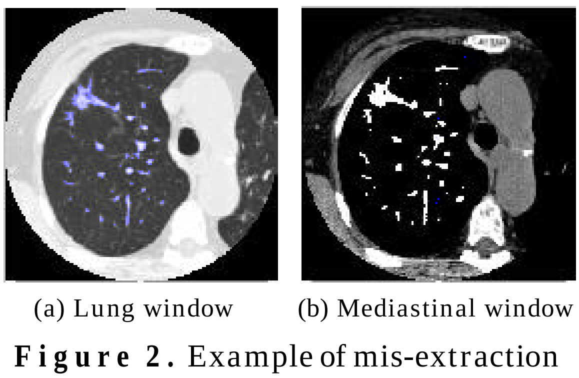
Table 2. Details of all cases

	Malignant	Benign	Total
Solid type	27	19	46
Middle type	8	3	11
Air-containing type	15	1	16
Unknown	1	4	5
Total	51	27	78



(a) Lung window (b) Mediastinal window

Figure 1. Example of proper extraction



(a) Lung window (b) Mediastinal window

Figure 2. Example of mis-extraction

by the Helical CT scanner(TCT-900s/Super Helix). The specification of images is shown in Table 1. Table 2 summarizes the details of cases which were used in this experiment. Figure 1 and Figure 2 present examples of the results. The extracted tumor regions are indicated as white regions in the lung area in (b) of each figure. Tumor regions in sixty-seven cases were extracted properly(Figure 1). The tumors which were not extracted connected with vessels which have almost the same radius as tumors(Figure 2). All tumors will be extracted if vessels are recognized beforehand. These sample images contained fifty-one malignant tumors and twenty-seven benign tumors. Forty-two malignant tumors and twenty-five benign tumors were extracted as tumor regions correctly.

Eighteen features for benign/malignant discrimination were calculated from tumor regions extracted. The details of cases used in the benign/malignant discrimination are given in Table 3. The best combination of features was selected from combinations of five or less features to obtain the best correct classification rate (= {(number of true positives) + (number of true negatives)} / (number of all cases)). The nearest neighbor method was used for the classification and the leave-one-out method was employed for performance evaluation. In the case using tumor regions extracted by the proposed method, the best feature set was the combination of the entropy of CT values in the tumor region, the volume of the set of inscribed

Table 3. Details of the cases used in the benign/malignant discrimination

	Malignant	Benign	Total
Solid type	25	18	43
Middle type	7	2	9
Air-containing type	9	1	10
Unknown	1	4	5
Total	42	25	67

Table 4. Result of benign/malignant discrimination without type classification

	Correct classification rate	Sensitivity	Specificity
Previous method	0.79	0.83(35/42)	0.72(18/25)
Proposed method	0.84	0.88(37/42)	0.76(19/25)

Table 5. Result of benign/malignant discrimination with type classification

	Correct classification rate	Sensitivity	Specificity
Solid type	0.90	0.90(28/30)	0.86(19/22)
Air-containing type	0.93	1.00(12/12)	0.67(2/3)
Total	0.91	0.95(40/42)	0.84(21/25)

spheres, and the standard deviation of CT values in the central region of tumor. The correct classification rate was 0.84. Using tumor regions extracted by the previous method, the combination of the standard deviation of CT values in the peripheral region of the tumor and the volume of the set of inscribed spheres was selected, and the correct classification rate was 0.79(Table 4).

Furthermore, tumors were discriminated between benign tumors and malignant tumors after classification into the solid type and the air-containing type. We also tried to classify the cases which belong to the middle type or the unknown into the above-mentioned two types. By introducing the type-classification process, the correct classification rate was improved from 0.84 to 0.91(Table 5). Two features were used in the type classification, and three features for the solid type and two features for the air-containing type in the benign/malignant discrimination. Here, we used again the nearest neighbor method and evaluated the performance by the leave-one-out method for both the type classification and benign/malignant discrimination, and combinations of five or less features were selected. The combinations of the features which was employed in each classification are given in Table 6.

4. DISCUSSION AND CONCLUSION

It was shown that the correct classification rate was improved using the tumors extracted by the proposed method. It is thought that this was achieved by the improved extraction of tumor region. In eleven out of seventy-eight cases, tumor regions were not extracted successfully. In these cases, tumors were connected to vessels which have almost the same radius. If the recognition of vessels or tumors is possible, tumors in these eleven cases will be extracted successfully.

By adding the process of the classification into the solid type and the air-con-

Table 6. Features which used for classification

Type classification	Mean of CT values in the tumor region SD of CT values in the tumor region
Benign/malignant discrimination	
for solid type	Volume of the tumor SD of CT values in the peripheral region of tumor Mean of CT values in the peripheral region of tumor
for air-containing type	Concentration index Mean of CT values in the peripheral region of tumor

SD = Standard deviation

taining type, the correct classification rate increased further.

The future work includes the improvement of tumor extraction method, development of better features for benign/malignant discrimination, and development of CAD system for the discrimination of lung tumor from chest X-ray CT images.

Acknowledgments

Authors thank to the members of Prof. Toriwaki's Laboratory of Nagoya University for useful discussions. The data conversion soft developed in Prof. N.Niki's laboratory (Tokushima University, Japan) was used to transform and transfer CT images. Parts of this research were supported by the Grant-In-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, the Grant-In-Aid for Private University High-Tech Research Center from Ministry of Education, Culture, Sports, Science and Technology and the Grant-In-Aid for Research from the Ministry of Health, Labor and Welfare, Japanese Government.

References

[1] J.Hasegawa, K.Mori, J.Toriwaki, et al. : Automated extraction of lung cancer lesions from multi-slice chest CT images by using three-dimensional image processing, Trans. of IEICE(D-II), J76-D-II, 8, pp.1587-1594, 1993. (in Japanese)

[2] S.Yamamoto, M.Matsumoto, Y.Tateno, et al. : Quoit Filter -A New Filter Based on Mathematical Morphology to Extract the Isolated Shadow, and Its Application to Automatic Detection of Lung Cancer in X-ray CT, Proc. of the 13th International Conference on Pattern Recognition, pp.3-7, 1996.

[3] T.Okumura, T.Miwa, J.Kako, et al. : Image processing for computer-aided diagnosis of lung cancer screening system by CT (LSCT), Proc. of SPIE's of International Symposium (SPIE'98), 1998.

[4] Y.Hirano, J.Hasegawa and J.Toriwaki : A feasibility study on computer classification of benign and malignant tumors in lung cancer diagnosis from chest X-ray CT images -The case of solid type of tumors, Proc. of CARS 2000 : pp.773-778, 2000.

[5] H.Shikata, H.Kitaoka, B.Keserci, et al. : Quantitative evaluation of the spatial distribution of vessels surrounding pulmonary nodules, Proc. of CARS 2000 : pp.761-766, 2000.

[6] Y.Kawata, N.Niki, H.Ohmatsu, et al. : Computer aided diagnosis of pulmonary nodules using topological and histogram feature, Proc. of CARS 2000 : p.1045, 2000.