Detection of Pulmonary Nodules on CT and Volumetric Assessment of Change over Time

Margrit Betke[1] and Jane P. Ko[2]

[1] Computer Science Department
Boston College, Chestnut Hill MA 02167, USA
betke@oak.bc.edu
http://oak.bc.edu/~betke
[2] Department of Radiology, Massachusetts General Hospital
Harvard Medical School, Boston, MA 02114, USA

Abstract. We present a computer vision system that automatically detects pulmonary nodules in computed tomography (CT) scans of oncology patients, performs size analysis and assesses for change in volume over time.

Thresholding, backtracking, and smoothing algorithms have been developed to recognize the thorax and trace the lung border. The regions within the lung that potentially contain nodules are evaluated for their shape, size, and position. These candidate regions are then characterized as nodules versus other structures by comparing consecutive CT slices in the same study. A preliminary system for the registration of studies has also been developed. It estimates nodule volume in each study and evaluates the volumetric change over time.

Our system has been tested on initial and follow-up studies of four patients. Preliminary results were detection of 284 nodules ranging between 1 and 32 mm at various locations and assessment of their volumetric change over time.

Our techniques have future applications for determining disease progression, remission, and stability in oncologic patients in addition to coregistration of different modalities within the thorax.

1 Introduction

Chest computer tomography (CT) is used to diagnose pulmonary metastasis and evaluate for progression in cancer patients. CTs are obtained at intervals determined by treatment regimens, protocols, and acute clinical issues. The radiologist searches the CT images for pulmonary nodules and compares the results with neighboring images in the same study and corresponding images on preceding studies. The radiologist then determines if a given nodule is stable, increased, decreased or resolved from prior studies.

The radiologist typically uses bidimensional measurements in the axial plane to estimate volumetric change in nodule size. However, in general, volumetric change is not accurately quantified.

The process of systematically evaluating several CT studies of the same patient is very time consuming for the radiologist. If a study contains $n = 30$ images, there are $3(n - 2) = 84$ triplet images that are compared with each other, the image and its neighbors above and below. If two studies are evaluated, each image in one study is compared to approximately three images in the second study to register corresponding nodules, so a total of $9(n - 2) = 252$ comparisons are made.

Our work addresses both issues. We developed a computer vision system that detects nodules in CT images, analyzes the size of the nodules and then assesses for any change on subsequent studies. The system quantitatively evaluates whether a nodule has changed in size. This estimate helps evaluate for disease response or progression.

Kanazawa et al. [4] and Giger et al. [2] have described preliminary diagnostic systems which have high sensitivities for detecting nodules of diameters > 4 mm and 3 – 18 mm, respectively. Work has also been done in nodule detection on chest radiographs [1,5]. Our work complements the previous work in nodule detection and focuses on volumetric quantization and comparison of consecutive studies.

2   Algorithms

Our computer vision system processes each CT image separately to obtain a set of regions in the lung that contain potential nodules. The properties of candidate regions on sequential slices are then compared. The results of the comparisons are used to classify candidate regions as nodules and to estimate their volumes.

2.1   Analysis of Individual CT Images

Thorax and Lung Border Detection. The patient's thorax is detected in the CT image by analyzing vertical and horizontal profiles of pixel grey levels. Since the surrounding soft tissues and bones are denser than the air-filled lung, they are visible as bright grey values, while the lung is dark. To find the border of the lung, the CT image is thresholded to create a binary image that is then analyzed to find the lung border. We developed a recursive backtracking algorithm that is similar to well-known maze algorithms [6] to trace along the border. Backtracking is used to detect narrow channels that could be due to vessels or artifacts. Since such channels are not considered to be border-forming structures, the tracing algorithm bridges them. The cropped center of a typical chest CT image and its traced lung border is shown in Figure 1 d.

Lung Border Correction and Lung Interior Detection. Our system addresses the problem of misinterpreting nodules that abut the border. Nodules in this location can be falsely excluded from the lung. Therefore the computed curve along the lung border is corrected. The locations of such nodules are detected by comparing the slopes of the points on the curve. A rapid change in

slope indicates a nodule, large vessel, or bronchus. Pixels that are enclosed within the adjusted border curve are then labeled as belonging to the lung parenchyma. Figure 1 shows how the border curve is adjusted so that two large nodules are included in the lung.

The "main centroid" and the centroid of each individual lung are calculated using these pixels. The "main centroid" is computed using pixels of both lungs. Figures 1 d and 2 show the centroids as white crosses.
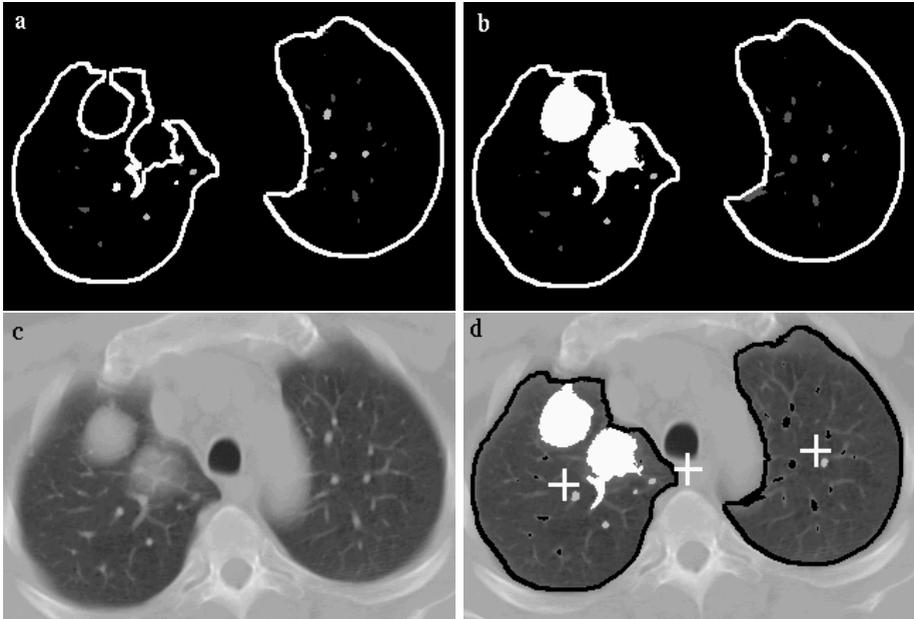
Fig. 1. The border curve in **1 a** is adjusted so that the two large nodules are included in the lung (**1 b**). **1 c** and **d** show a crop of the original CT image and the result of the processing. Detected nodules and centroids are shown in white, potential nodules in grey.

Determining Candidate Regions. Once the set of pixels belonging to the lung parenchyma has been determined, pixels can be divided into lighter and darker subsets. The lung parenchyma is dark on a CT image, except where there are normal vessels and bronchi, or nodules and other pathology. These would be lighter regions. The grey value at a pixel is a density average that describes the underlying volume (voxel). Therefore, lighter structures that fill the whole slice thickness appear bright. Structures that are smaller than the slice thickness are averaged with their surrounding air and appear fainter. Therefore, we use several grey-level thresholds to create binary images of the lung parenchyma that contain

different candidate regions (Fig. 2). We then use the sequential labeling algorithm
described by Horn [3] to distinguish and label candidate regions uniquely.
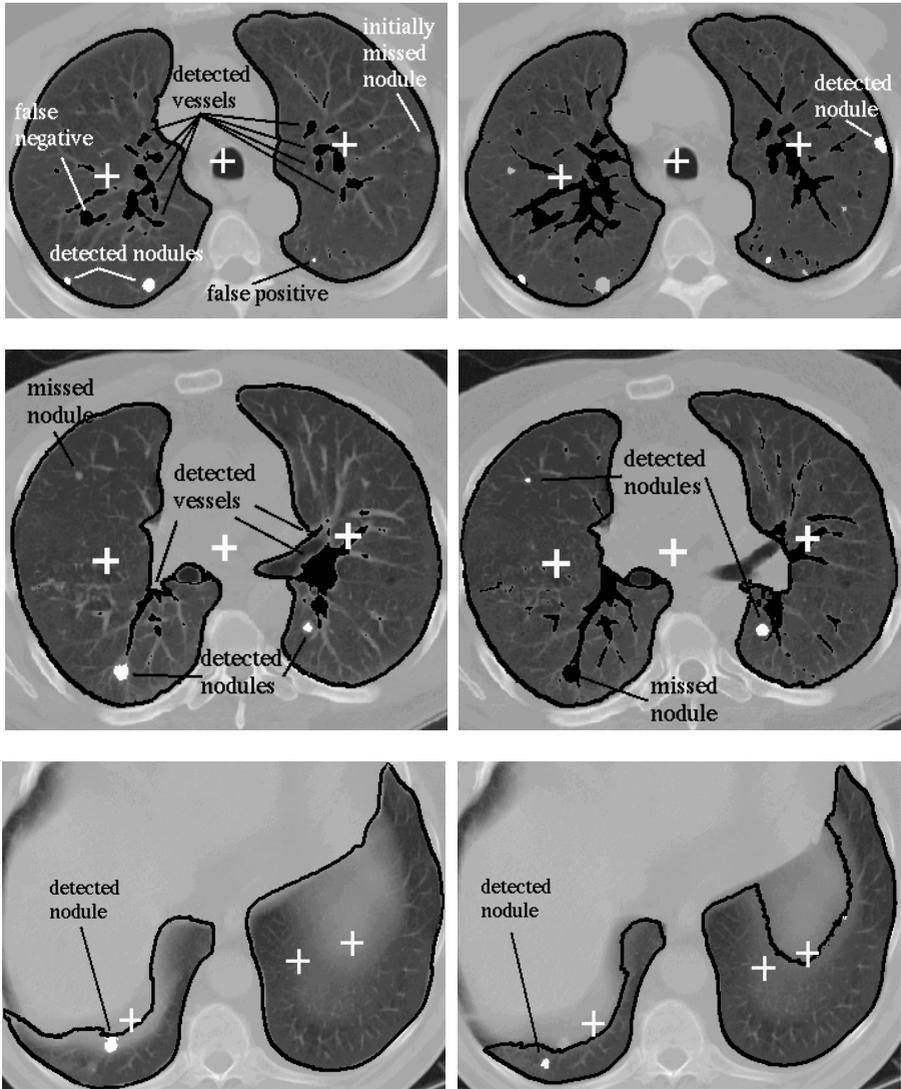
Fig. 2. On the left, each row shows the detection results for a CT image processed
with a high grey-level threshold. On the right, the same CT image is shown
processed with a lower threshold. Nodules initially missed using a high threshold
can be detected using a low threshold and vice versa.

Computing Properties of Candidate Regions. For each candidate region, the number of pixels within the region is computed as a first estimate of the *2D axial area* of the region. This estimate is then used to compute the region's centroid and its *distance* to the lung centroids and to the lung border. Large candidate regions close to the border of the lung are likely to be nodules.

The *shape* of a candidate region is important for classification of the region. Nodules are typically spherical. Therefore circular regions on 2D CT images may be spherical nodules, but they also could be vessels in cross section. Elongated regions indicate vessels. A combination of two methods is used to determine shape.

First, for each candidate region, we find the line of pixels for which the sum E_{min} of the square of the distance to points in the region is a minimum. This line is the axis of least second moment or the "major axis" of the region. Maximizing the sum E_{max} of the square of the distance to points in the region yields the minor axis of the region. The ratio E_{min}/E_{max} is zero for a straight-line region and one for a circular region. The ratio is therefore an important measure for how elongated or round the region is. Our second measure of shape is the percentage P_{csc} of pixels within the circle circumscribing the candidate region.

2.2   Analysis of Consecutive CT Slices

To compare consecutive CT slices, we first register global positions which include individual lung centroids, main centroid and lung borders in both images. We then compare local nodule positions in both images and determine correspondences.

Nodules that appear in only one CT image have a diameter that is smaller than the slice thickness, so their diameter can be upper bounded by this thickness. They can be lower bounded by the 2D diameter of the disk visible in the CT image.

Similarly, we can estimate the diameter of a nodule that appears in more than one CT image by bounding its diameter from above by the number of slices times their thickness. We then determine the 3D center of the nodule. We use the axial image on which the nodule appears brightest along with the 2D center of the nodule on this image. We use the diameter of the nodule cross section as a lower bound on the nodule's diameter.

2.3   Time Analysis

Identifying corresponding nodules in separate studies is considerably more difficult than identifying corresponding nodules in consecutive images within the same study. Differences in patient position and inspiration complicates the registration between studies.

Given a CT image, we first need to find the closest CT image on a subsequent study. In our preliminary system, we identify CT slices by hand in both studies

that show the apex of the lung. With the initial registration of apex images, we simplify the search for corresponding CT images in the rest of the lung. For each image in the earlier study we identify a possible matching image in the subsequent study and examine the two surrounding slices. Once corresponding CT images are identified, the images are registered similar to the 3D registration described above, and the volumes of corresponding nodules are compared.

3   Experiments and Results

We evaluated an initial and a follow-up study for four patients. The patients were identified through a database of reports as having multiple pulmonary nodules. The patients had melanoma, renal carcinoma, and rectal cancer. The studies were obtained 1, 2 or 6 months apart. The studies were chosen to reflect a wide variety of nodule number, size, distribution, and change over time.

The helical thoracic CT scans were obtained at 10 mm increments from the apices to the adrenal glands with 5 mm increments through the hila. On average 30 images were processed per CT study, which gives a total of 240 images. Each image contains a 512×512 matrix, each pixel is quantized using 16 bits.

A thoracic radiologist evaluated the scans using a clinical IMPAX PACS workstation (AGFA) with a 12-bit display and the follow-up scans were evaluated in comparison with the original studies. The radiologist interpreted sequential studies together and made comparisons for change in size.

The system is implemented in C on a Pentium II PC running Linux 5.2. We define nodules to be *large*, *medium*, and *small* if their diameter is greater than 1 cm, between 1 cm and 5 mm, and smaller than 5 mm, respectively. Our data contains a total of 323 nodules, 52 large, 108 medium, and 163 small nodules. The volume V of a nodule can be easily computed from the obtained diameter d assuming a spherical shape, $V = 4/3\pi(d/2)^3$.

The overall sensitivity of the system was 88% for detecting all nodules in eight studies. The individual study sensitivities are 84% (A1), 84% (A2), 100% (B1), 100% (B2), 85% (C1), 100% (C2), 84% (D1), and 92% (D2). Table 1 summarizes our nodule detection results.

Nodule sizes range between 1 mm – 4 cm, 2.5 mm – 1.6 cm, 2 mm – 1 cm, and 1.9 mm – 2.5 cm for patients A, B, C, and D, respectively. Size analysis results are shown in Table 2. Of the 39 missed nodules, 29 are small, 8 are medium, and 2 are large. The missed nodules either abut a lung border (20) or vessel (13), or are too small and faint (6).

The radiologist detected a total of 156 nodules changed and 44 remained stable, while the computer demonstrated change in 175 nodules with 25 nodules remaining stable (Table 3).

4   Conclusions

We presented a computer vision system for pulmonary nodule detection, automated quantification of volume, and assessment for change in volume over

Table 1. Nodule Detection: Class 1 = Detected Extremely Likely Nodule, Class 2 = Detected Likely Nodule, Class 3 = Missed Nodule. Total = Total Number of Nodules Identified by Radiologist

Class	Study 1				Study 2			
	1	2	3	total	1	2	3	total
Patient								
A	26	36	12	74	26	38	12	76
B	6	4	0	10	7	3	0	10
C	10	1	2	13	9	3	0	12
D	16	16	6	38	59	24	7	90

Table 2. Volumetric Analysis of Detected Nodules: Detected nodules are separated into size categories. Large nodules: > 1cm, medium-sized nodules: between 5mm and 1cm, and small nodules: < 5mm.

		Number of Nodules and Average Sizes in mm							
		Study 1				Study 2			
		l	m	s	total	l	m	s	total
Patient	Nodule								
A	number	5	13	44	62	5	12	47	64
A	av. size	24.7	6.3	3.3	5.6	25.1	7.0	3.1	5.1
B	number	1	4	5	10	3	3	4	10
B	av. size	15	7.3	3.1	6.0	14.4	7.8	3.5	8.1
C	number	1	8	2	11	1	8	3	12
C	av. size	10.0	6.4	3.3	6.2	10.5	6.7	3.5	6.2
D	number	5	18	9	32	29	34	20	83
D	av. size	13.7	7.1	3.7	7.2	13.8	6.6	3.5	8.4

Table 3. Change in Volume: Disease progression: decrease (d), no change ($=$) and increase (i) in nodule size.

Studies Compared	Change in Size of Nodules					
	Radiologist			Computer		
	d	$=$	i	d	$=$	i
A1 − A2	37	20	13	38	9	23
B1 − B2	2	4	6	3	3	6
C1 − C2	1	12	0	1	11	1
D1 − D2	15	8	82	17	2	86

time. The system detects nodules of a wide range in diameter (1 – 32 mm), and location, for example, abutting the lung border or a vessel. The system uses thresholding, backtracking, and smoothing techniques and 3D registration techniques. Registration of thoracic CT between studies is challenging because of differences in patient position and respiration. We presented a preliminary registration technique for comparing the same nodules on sequential studies. These techniques have future applications for determining disease progression, remission, and stability in oncologic patients in addition to coregistration of different modalities within the thorax.

Acknowledgements

References

1. Carreira, M., Diego, C. Penendo, M. Mosquera, A.: Computer-Aided Diagnosis: Automatic Detection of Lung Nodules. Med. Phys. **25**:10 (1998) 1998–2006.   246
2. Giger, M., Bae, K., MacMahon, H.: Computerized Detection of Pulmonary Nodules in Computed Tomography Images. Invest Radiol. **28** (1994) 459–465.   246
3. Horn, B. K. P.: Robot Vision. MIT Press, 1986.   248
4. Kanazawa, K., Kawata, Y., Niki, N., Satoh, H., Ohmatsu, H., Kakinuma, R., Kaneko, M., Eguchi, K., Moriyama, N.: Computer-Aided Diagnostic System for Pulmonary Nodules Using CT Images. Proc. Medical Image Computing and Computer-Assisted Intervention, 1998, 449–456.   246
5. Matsumoto, T., Yoshimura, H., Giger, M. Doi, K., MacMahon H. Montner, S. Nakanishi, T.: Potential Usefulness of Computerized Nodule Detection in Screening Programs for Lung Cancer. Invest Radiol. **27** (1992) 472–475.   246
6. Wirth, N. Algorithmen und Datenstrukturen. B. G. Teubner, 1983.   246