

A novel lung nodules detection scheme based on vessel segmentation on CT images

Tong Jia^{a,b,*}, Hao Zhang^a and Haixiu Meng^a

^a College of Information Science and Engineering, Northeastern University, Shenyang 110819, China

^b Key Laboratory of Medical Image Computing, Ministry of Education Shenyang 110819, China

Abstract. Lung vessels often interfere with the detection of lung nodules. In this paper, a novel computer-aided lung nodule detection scheme on vessel segmentation is proposed. This paper describes an active contour model which can combine image region mean gray value and image edge energy. It is used to segment and remove lung vessels. A selective shape filter based on Hessian Matrix is used to detect suspicious nodules and remove omitted lung vessels. This paper extracts density, shape and position features of suspicious nodules, and uses a Rule-Based Classification (RBC) method to identify true positive nodules. In the experiment results, the detection sensitivity is about 90% and FP is 1/scan.

Keywords: Lung vessel segmentation, lung nodule detection, active contour model, CT image

1. Introduction

Lung nodule is important clinical manifestation of lung cancer in Computed Tomography (CT) image. Its two-dimensional shape is circular structure and three-dimensional shape is spherical structure. The lung CT image contains blood vessels, trachea, pleura, bone, and soft tissues etc, which can cause interference on lung nodules detection. Especially because lung vessel density, location and shape are all similar to nodules, radiologists usually distinguish them by evaluating the relevance of a series of CT images. However, it is very difficult for lung cancer CAD system.

Most of lung nodule detection methods are roughly classified into two categories: some methods are based on the image intensity, and the other methods are based on the model. The former often includes multiple threshold [1], clustering [2] and mathematical morphology, etc. [3]. The latter often includes multilevel binomial logistic prediction [4], object-based deformation [5] etc.

Lung nodule computer-aided detection is an ongoing research topic [6]. One major difficulty is that the lung nodules are usually located at the edge of vessels. Because they have similar CT values and the lung vessels tree's topology structure is very complex, the disturbance of lung vessel usually leads to the detection of false positive nodules. For this problem, the proposed method is summarized as follows.

- The lung vessel segmentation scheme is proposed based on an active contour model.

*Corresponding author: Jia Tong, College of Information Science and Engineering, Northeastern University, Shenyang 110819, China. Tel.: +86 2483688326; Fax: +86 2423891272; E-mail: jiatong@ise.neu.edu.cn.

- The suspicious nodule detection algorithm is proposed based on a selective shape filter, and the omitted vessel areas in the last step can also be filtered based on the algorithm.
- Several important nodule features are extracted and a simple and effective feature classifier is used.

2. Methods

2.1. Lung vessel segmentation

The lung parenchyma segmentation scheme has been developed and implemented by the authors in [7]. Because the topological structure of lung vessel tree is complex, an active contour model is proposed and the level-set method is used to implement it. Existing geometric active contour models include only image region or edge energy, such as Chan-Vese model [8], CURVES model [9] etc. Our method combines them together.

2.1.1. Region mean energy

Firstly, the two-dimensional Chan-Vese model is simplified. Secondly, the level set method is used to solve energy minimization problems, and the curve curvature calculation is changed to compute surface curvature. Finally, with optimizing the evolution formula, the final image segmentation contour C is computed.

Suppose the domain Ω of image $I(x,y)$ is divided into the object and background regions. They are two homogeneous regions. The energy function of Chan-Vese model is defined in Eq. (1).

$$F(c^+, c^-, C) = \mu \text{Length}(C) + \lambda^+ \int_{\text{inside}(c)} |I(x, y) - c^+|^2 dx dy + \lambda^- \int_{\text{outside}(c)} |I(x, y) - c^-|^2 dx dy \quad (1)$$

Where c^+ and c^- are the mean gray value of the object and background regions separately, $\mu > 0, \lambda^+, \lambda^- > 0$ is the weigh coefficient.

The original image $I(x,y)$ is divided into the object and background regions by contour C . C_1 and C_2 are mean gray value of the object and background region, respectively. Eq. (1) is simplified into Eq. (2).

$$E_{REG} = \alpha \left[\int_{\text{inside}(c)} |I(x, y) - c_1|^2 dx dy + \int_{\text{outside}(c)} |I(x, y) - c_2|^2 dx dy \right] + r \int_c ds \quad (2)$$

The first entry is energy fitting of image region, and the second one is geometric constraint. Let curve C as short as possible. α , and γ is constant, ds represents arc-length micro element. Therefore, the ∇E 's variation problem is converted into computing the corresponding surface evolution. The evolution equation is:

$$C_t = \alpha \left[(I - c_1)^2 - (I - c_2)^2 \right] \vec{n} - r H_t \vec{n} \quad (3)$$

where H_I represents the surface curvature. When the curve evolution stops, $\nabla E = 0$, thereby the optimal segmentation results are obtained.

2.1.2. Edge energy

I_x and I_y represents the first derivative in the horizontal and vertical direction respectively, $I(x,y)$ represents the original image. The image gradient direction vector is defined as:

$$\vec{\xi}(x, y) = \frac{\nabla I}{|\nabla I|} = \frac{\{I_x, I_y\}}{\sqrt{I_x^2 + I_y^2}}$$

The vertical vector is defined as:

$$\vec{\eta}(x, y) = \frac{\bar{\nabla} I}{|\nabla I|} = \frac{\{-I_y, I_x\}}{\sqrt{I_x^2 + I_y^2}}, \text{ therefore } \langle \vec{\xi}, \vec{\eta} \rangle = 0.$$

Let $I_{\eta\eta}$ be the vertical vector, $I_{\varepsilon\varepsilon}$ be the second derivative in the image gradient direction. The energy function is defined as:

$$\int_C \langle \nabla I, \vec{n} \rangle ds - \iint_{\Omega_c} k_I |\nabla I| dx dy \tag{4}$$

Where \vec{n} represents the unit direction vector, k_I represents the curvature and Ω_c represents the area inside C . When $I_{\varepsilon\varepsilon} \vec{n} = 0$, the Euler-Lagrange extremum is solved based on the variation method.

Because Eq. (4) is two-dimensional curve evolution function, it must be extended to surface evolution for 3D medical images segmentation, shown as Eq. (5).

$$E_{EDGE} = \iiint_{\Omega_s} H_I |\nabla I| dx dy dz - \iint_S \langle \nabla I, \vec{n} \rangle da \tag{5}$$

where H_I represents surface curvature, S represents surface, da represents surface area, \vec{n} represents surface unit direction vector.

2.1.3. Total energy calculation

The user can select a seed point in the vessel region, and then the initial surface can be obtained by the 3D region growing method. But when the initial surface is away from the actual object, it is difficult to locate the edges of all the objects accurately. Therefore, combining the object's inside and out

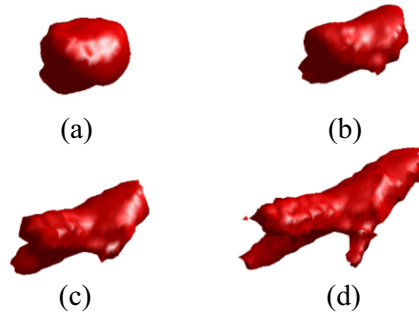


Fig. 1. The lung vessel local segmentation deformation results. (a)-(d) represent different iteration segmentation results respectively, which shows the surface deformation process.

side region is a better solution. The total energy function is defined in Eq. (6), where α is the parameter of edge energy, α is set at 1 in the experiment.

$$E = \alpha E_{EDGE} + E_{REG} \quad (6)$$

The total energy equation is defined as follows.

$$E = \alpha \left[\int_{inside(c)} |I(x, y) - c_1|^2 dx dy + \int_{outside(c)} |I(x, y) - c_2|^2 dx dy \right] - \left(\iint_S \langle \nabla I, \vec{n} \rangle da - \iiint_{\Omega_s} H_I |\nabla I| dx dy dz \right) + r \int_c ds \quad (7)$$

Three-dimensional local vessel segmentation results are shown in Figure 1.

2.2. Suspicious nodule detection

Because of the lung vessel tree's topology complexity (from the lung artery to artery branches nearly 17 parts, in diameter ranging from 20 mm to a few millimeters), it is inevitable for some small blood vessel branches to be ignored. To solve the problem, the shape filter is developed based on the Hessian matrix. Lung nodule is shown as a circular structure in two-dimension, and approximately spherical in three-dimension, while the lung vessel's three-dimensional shape is approximately tubular. The developed method can suppress the "tube" shape of vessel, but enhance the "spherical" shape of nodule.

The method constructs nodule and vessel's math model based on Gaussian function.

$$\begin{cases} d(x, y, z) = \exp\{-(x^2 + y^2 + z^2)/2\delta^2\} \\ l(x, y, z) = \exp\{-(x^2 + y^2)/2\delta^2\} \end{cases} \quad (8)$$

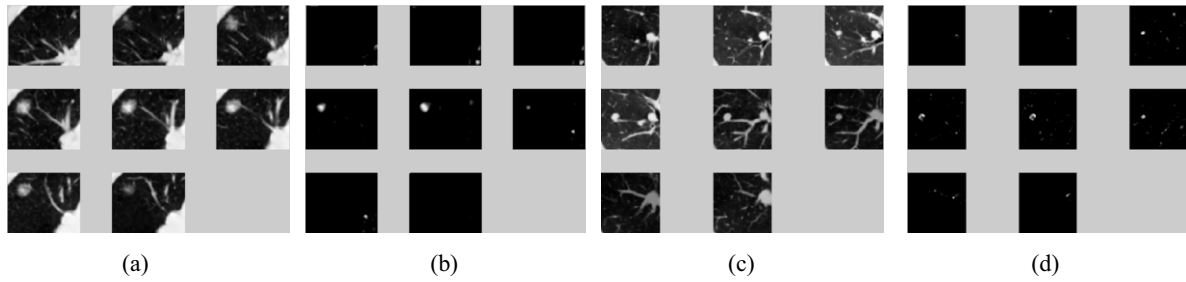


Fig. 2. The suspicious nodule detection results. (a) and (c) are two different original CT images, (b) and (d) are their corresponding nodule detection results, respectively.

where $d(x,y,z)$ represents a sphericity and $l(x,y,z)$ represents a tube in the form of a 3D Gaussian function, σ represents the scale of a sphericity or a tube. Let $f(x,y,z)$ represent the 3D original image, the Hessian matrix is defined as following.

$$H = \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix} \tag{9}$$

For the standard sphere, H must be a real symmetrical matrix, $f_{xy}, f_{xz}, f_{yz}, f_{yx}, f_{zx}, f_{zy}$ represents the six second derivatives respectively, and they are equal to zero. The three eigen values of H are represented as λ_1, λ_2 and λ_3 , respectively. The shape filter Z_{dot} is defined in Eq. (10).

$$\begin{aligned} Z_{dot}(\lambda_1, \lambda_2, \lambda_3) &= |\lambda_3|^2 / |\lambda_1| \quad \lambda_1 < 0, \lambda_2 < 0, \lambda_3 < 0 \\ Z_{dot}(\lambda_1, \lambda_2, \lambda_3) &= 0 \quad otherwise \end{aligned} \tag{10}$$

The suspicious nodule detection results are showed in Figure 2.

2.3. Feature extraction and classification

2.3.1. Feature extraction

This paper focuses on density, shape and position of suspicious nodules, and extracts the following features, including spatial distribution density, sphericity, maximum diameter, volume, CT mean value, CT max value and CT min value, etc.

2.3.2. Feature classification

A Rule-Based Classification (RBC) method is used to identify true positive nodules [10,11]. The theory is described in Eq. (11). For example, for the spatial distribution density, [0.3, 0.6] interval is more sensitive to true-positive nodule. For the sphericity, [2.5, 5] interval is more sensitive to false-positive nodule.

$$\text{If Character Satisfy then Classify} \tag{11}$$

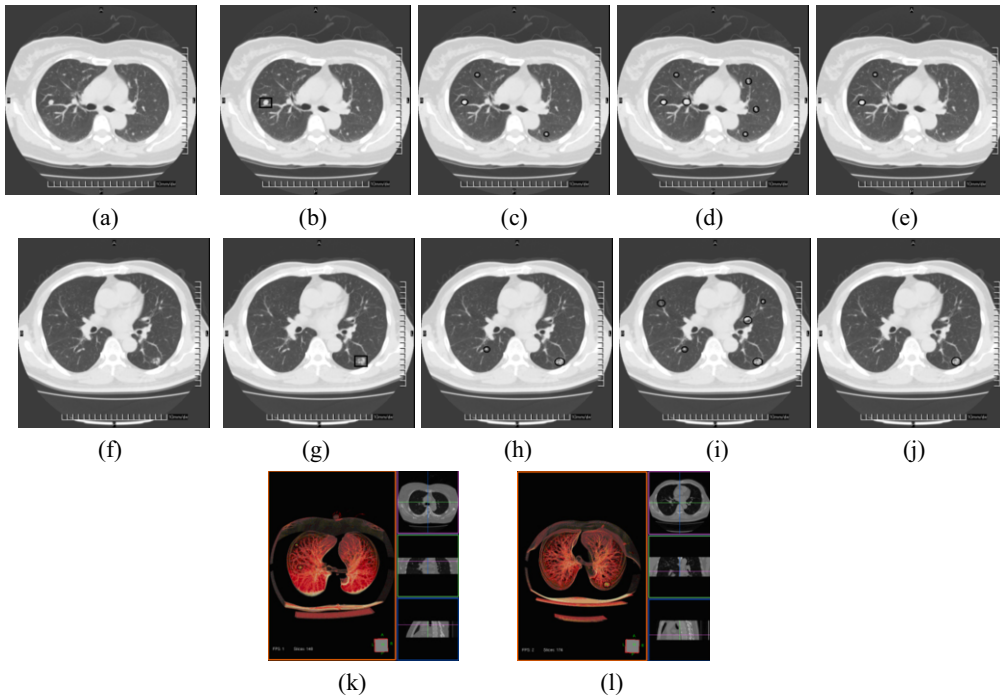


Fig. 3. The final nodule detection results. (a) and (f) are two different CT original images; (b) and (g) are radiologist's diagnosis results (black boxes mark); (c) and (h) are the detection results of literature [12] (black circle lines mark); (d) and (i) are the detection results of literature [13] (black circle lines mark); (e) and (j) are this paper's detection results (black circle lines mark); (k) and (l) are 3D visualization results (black circle lines mark).

Our scheme's detection results are compared with the detection results in [12] and [13], which are shown in Figure 3. It can be found that the detection results of our scheme are more accurate than others, and false-positive nodules are significantly reduced. While the other two schemes, the nodule detection results are interfered by lung vessels etc., which is especially apparent in [13]. Figures 3(k) and 3(l) are the 3D visualization results of detected nodules (marked by black circle).

3. Experimental results and analysis

The proposed scheme has been trained on a clinical dataset of 90 CT scans. The detection results contain 232 true-positive nodules. The data are randomly split into training and testing datasets, with the training datasets containing 45 scans (TP=124), and the testing datasets containing 45 scans (TP=108). Each CT image is 512×512 pixels, slice thickness is 1-2 mm, and pixel pitch is 0.7 mm, the X-ray tube current ranges from 30 to 250 mA.

To evaluate the CAD systems, this paper computes the detection sensitivity and FPs/scan, which is defined in Eq. (12).

$$\begin{cases} Sen = TP / (TP + FN) \\ FPs / scan = FP / N \end{cases} \quad (12)$$

Table 1
The detail test results

Sample Number	Nodule Number	Detection results			SEN	FPs/scan
		TP	FP	FN		
33%	55%	97	45	11	90%	1

where N represents the number of scans in testing datasets, TP represents the number of true-positive nodules, FP represents the number of false-positive nodules, and FN represents the number of false negative nodules. The detail test results are shown in Table 1.

4. Conclusion

Because the lung vessels are removed firstly, the proposed scheme improves the detection sensitivity. The detection rate is about 90% and FP is 1/scan. According to the test results, the detection sensitivity is high and the FP is low. Most of the tissues (blood vessels, pleural, scar, etc.) can be excluded. The scheme's high performance and fast computation time have been applied for clinical diagnosis of lung cancer.

Acknowledgement

This work was supported in part by the National Natural Science Foundation of China under Grant 61273078, and in part by the Doctoral Foundation of Ministry of Education of China under Grant 20110042120030, and in part supported by the Fundamental Research Funds for the Central Universities of China 130404012.

References

- [1] J. Zhang, C. Ya and C. Chui, Fast segmentation of bone in CT images using 3D adaptive thresholding, *Computers in Biology and Medicine* **40** (2010), 231–236.
- [2] B. Bouraoui, C. Ronse, J. Baruthio and N. Passat, 3D segmentation of coronary arteries based on advanced mathematical morphology techniques, *Comput. Med. Imaging Graph.* **34** (2010), 377–387.
- [3] T. Kubota, A.K. Jerebko, M. Dewan and M. Salganicoff, Segmentation of pulmonary nodules of various densities with morphological approaches and convexity models, *Medical Image Analysis* **15** (2011), 133–154.
- [4] Xiuhua Guo, Zhongwei Jia, Hongkai Li, Zhigang Liang, Kuncheng Li and Qian He, Multilevel binomial logistic prediction model for malignant pulmonary nodules based on texture features of CT image, *European Journal of Radiology* **74** (2010), 124–129.
- [5] Chunming Li, Chenyang Xu, Changfeng Gui and Martin D. Fox, Distance regularized level set evolution and its application to image segmentation, *IEEE Transactions on Image Processing* **19** (2010), 3243–3254.
- [6] J.D. Silverman and N.S. Paul, Investigation of lung nodule detect-ability in low-does 320-slice computed tomography, *Medical Physics* **36** (2009), 1700–1710.
- [7] Tong Jia, Lu Meng, Dazhe Zhao and Xu Wang, Automatic lung parenchyma segmentation on CT image, *Journal of Northeastern University (Natural Science)* **29** (2010), 965–975.
- [8] T. Chan and L. Vese, Active contours without edges, *IEEE Trans. of Image Processing* **10** (2001), 266–277.
- [9] L. Lorigo, O. Faugeras, W. Grimson and R. Keriven, Codimension-two geodesic active contours for the segmentation of tubular structures, *Proceeding of Computer Vision and Pattern Recognition* **1** (2000), 444–451.

- [10] Xujiong Ye, Xinyu Lin and Jamshid Dehmeshki, Shape-based computer-aided detection of lung nodules in thoracic CT images, *IEEE Transactions on Biomedical Engineering* **56** (2009), 1810–1820.
- [11] Qiang Li, Recent progress in computer-aided diagnosis of lung nodules on thin-section CT, *Computerized Medical Imaging and Graphics* **31** (2007), 248–257.
- [12] S.G. Armato III, F. Li, M.L. Giger, H. MacMahon, S. Sone and K. Doi, Lung cancer: performance of automated lung nodule detection applied to cancers missed in a CT screening program, *Radiology* **225** (2002), 685–692.
- [13] S. Yamamoto, H. Takizawa, H. Jiang, T. Nakagawa, T. Matsumoto, Y. Tateno and T. Iinuma, A CAD system for lung cancer screening test by X-ray CT, *Proceedings of the International Conference on Computer Assisted Radiography and Surgery (CARS)* **1** (2001), 605–610.