A Novel Segmentation Algorithm for Pulmonary Nodule in Chest Radiograph

Erling Wei[1] Jiayong Yan[2] Mantao Xu[2] Jiwu Zhang[2]
[1]Department of Computer Science and Engineering, Fudan University, P.R. China
[2]Carestream Health Corporation R&D, China
[1]062021139@fudan.edu.cn
[2]jiayong.yan@carestreamhealth.com

Abstract

Segmentation of pulmonary nodules in chest radiographs is a difficult task due to heavy noise and superposition of ribs, vessels, and other anatomical structures in lung field. In this paper, a polynomial fitting based ray-casting algorithm is proposed for pulmonary nodule segmentation in chest radiographs. Instead of directly detecting nodule edge points, the nodule intensity profiles are first fitted by using polynomials. Then, the edge positions are identified through analyzing the local minimum of the fitted curves. The algorithm has been experimentally tested over an image database with 148 nodule cases in chest radiographs collected from a variety of digital radiograph modalities. Preliminary results have shown that the proposed segmentation algorithm has obtained a high rate of successful segmentations.

1. Introduction

Lung cancer is one of the most common cancers in the world. It is a leading cause of cancer death in men and women, which has approximately accounted for a death of 975,000 men and 376,000 women in the year 2007 [1]. Early detection of lung cancer can potentially save lives, and has been proved to increase the five-year survival rate from 12% to 70%. Early detection of lung cancers could benefit from a successful screening program, where projection radiographs often serve as a common screening technique. However, 30% of small pulmonary nodules in chest radiograph, which is an indication of early-stage lung cancers, are missed in diagnosis workflow. This is due to a limitation of image quality for small nodules as well as a superposition of their surrounding anatomical structures. Those overseen or missed cancer cases often go untreated for a couple of years and therefore their cancer survival rates are very low. In this sense, computer aided detection (CAD) can provide clinician a tool in avoidance of the oversight of small pulmonary nodules, which has received a considerable interest in diagnosis of early stage lung cancers. Naturally, the segmentation of pulmonary nodules plays an important role in automated detection of pulmonary nodules in chest radiograph. This is because all of region-based image features for the suspicious nodules are extracted according to the resulting contour by nodule segmentation. But only a few algorithms [2-4] have been developed so far for nodule segmentation due to a limited image quality of pulmonary nodules in chest radiograph. In other words, a suspicious nodule in chest radiograph is often overlapped or surrounded by anatomical noises, which poses great difficulty on determining right edges. In [2-3], a multi-threshold algorithm has been exploited in identification of the nodule regions in chest radiograph after the regions of interest were enhanced by using anatomical structure suppression filters. However, a prior contour or knowledge of the surrounding anatomical structures (e.g., ribs) must been known for the filters in advance. Another intuitive approach is the so-called ray-casting algorithm [4] by assuming that nodule edges correspond to the pixels with high value of gradients. One should remember calculation of gradient images is very sensitive to image noise. Multi-scale filters and edge focusing techniques can be applied for the sake of avoiding this sensitivity to image noise. But they are not applicable to the weak edges, which were often caused by the surrounding anatomical structures, and therefore result in a false detection of nodule contours. Motivated by those operational difficulties for pulmonary nodule segmentations, this paper presents a novel and robust algorithm by incorporating polynomial curve fittings into the ray-casting algorithm developed in [4]. Namely, the intensity profile is fitted into a polynomial curve sampled on each ray line via the nodule center, by which a possible edge point can

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

be identified by analyzing the inflection points of the resulting fitted curve. The initial experimental results have validated that the proposed segmentation algorithm is able to achieve a desirable rate of successful segmentations.

2. Polynomial fitting based ray-casting algorithm

2.1 Image enhancement

Pulmonary nodules are very hard to be segmented directly in original chest radiographs due to the occlusions of their surrounding anatomical structures (e.g., ribs) as well as the variation of image quality for chest radiograph. The suspicious region representing a nodule in chest radiograph must be enhanced before any segmentation algorithm is conducted. In [2-3], both a shape-matching filter and a ring average filter were applied separately to obtain a nodule-enhanced image and anatomical noise suppressed image. Then, the nodule segmentation was applied to a difference image, which was obtained by subtracting the suppressed image from the nodule-enhanced image. But in this paper, enhancement will be same as that of [4] of the chest radiograph so that a good performance comparison can be carried on with the previous gradient-based approach in [4]. It is a local normalization of chest radiograph

$$L_{LN} = (L - \widetilde{L}) / (\widetilde{L}^2 - (\widetilde{L})^2)^{1/2} \quad (1)$$

where L is the resized image (1024×1024 pixels) of an input chest radiograph and a tilde is a *Gaussian* filter with a kernel size of 25 pixels same as [4].

$$\sigma_{LN} = 25 \quad (2)$$

Put into words, the local deviation of the image intensity from the local average is normalized on the local standard deviation. Likewise in [4], a multi-scale blob detection algorithm is applied to select a number of initial candidates or seeds for suspicious nodules. The main purpose for blob detection is to identify all the suspicious regions indicating pulmonary nodules, in which the center of suspicious region and the scale size of the corresponding blob can be determined meanwhile. Once the resized image is locally normalized in terms of (1), the nodule edge was further strengthened by using zero-crossings of the resulting normalized image L_{LN} through

$$L' = \begin{cases} -\alpha_1 L_{LN} & L_{LN} < 0 \\ \alpha_2 L_{LN} & L_{LN} >= 0 \end{cases} \quad (3)$$

where L_{LN} denotes the pixel intensities in the LN image, α_1 and α_2 are predefined positive constants that are selected by 1 and 50 respectively in the experiments (same as [4]).

2.2 Polynomial fitting

For some cases, the gradient-based ray-casting approach in [4] undergoes a severe under-segmentation problem; namely, the resulting contour of nodule segmentation is much smaller than its actual boundary, which can be well observed through the two resulting examples in fig. 1.

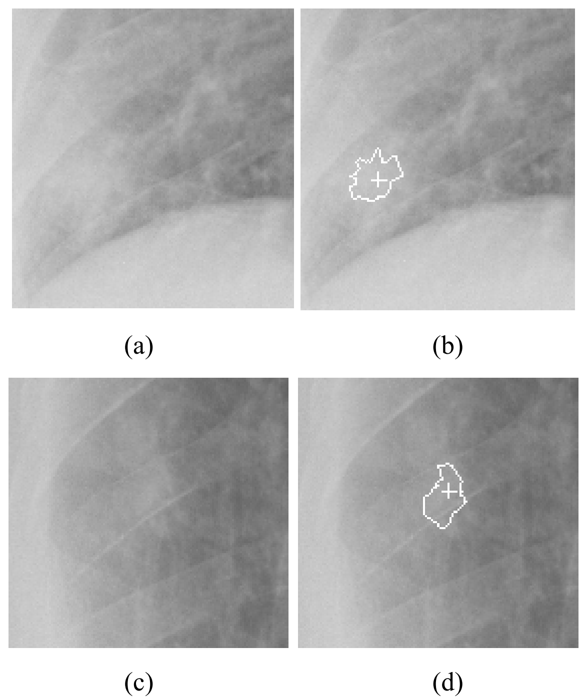


Fig. 1. Two exampled results of nodule segmentation obtained by using the gradient-based approach in [4]: (a) and (c) are the two original images for region of interest respectively whereas (b) and (d) are their segmentation results over (a) and (c) correspondingly.

The severe problem often occurs when the nodule region is highly overlapped by anatomical structures or noises, e.g., ribs. This is because calculation of gradient image only takes several neighborhood pixels into account, which makes itself very sensitive to the noisy pixels in the neighborhood window. The pixel with the local maximum gradient is obviously not on the practical nodule boundary if a rib edge or some isotropic noise caused by local normalization in (1) appears inside the nodule. The resulting edge point with local maximum gradient in [4] is offset by a significant displacement in the direction of pointing to the nodule center. As a matter of fact, many pixels whose gradients are local maximum are far from the actual edge of nodules since the surrounding anatomical noises are hard to suppress or even present as strong edges sometimes. This challenge remains almost in every previously developed algorithm, for instance, the anatomical structure suppression filter in [2-3]. To address these issues, this section has proposed a segmentation algorithm by using polynomial fittings based on the ray-casting algorithm in [4].

We cast 30 rays according to [4], in a region of interest with a homogeneous orientation distribution through the central pixel. The range of each cast ray line **z** in search for the edge point should be a line segment, [**c - 3r**, **c** + 3**r**], where **r** is the vector pointing from central pixel **c** to the blob boundary pixel, **r**, which can be identified in blob detection. An intensity profile of cast ray can be constructed by a sampling of image intensities over the ray segment **z**.

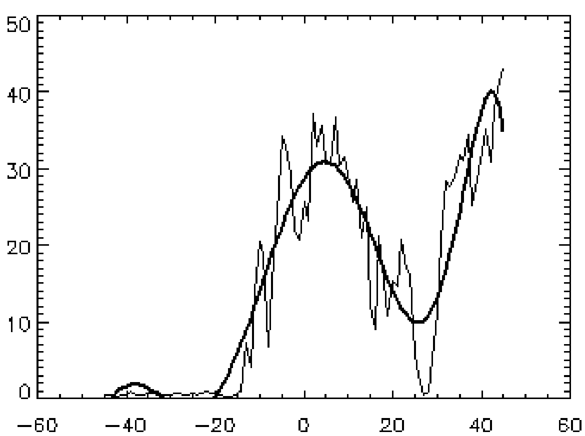
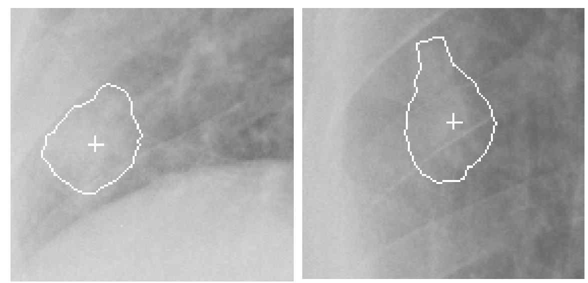


Fig. 2. Illustration of an intensity profile and its polynomial fitting curve: the thinner line is the intensity profile and the thicker line is the fitted curve.

Instead of computing the gradients, we fit the intensity profile of each cast ray by using a 7-order polynomial curve *C* (This parameter is acquired through experiment, which performs over other orders). The main advantage of using polynomial fitting hinges on its good estimation of change of image intensities across an object boundary since it is not sensitive to most of noises appearing nearby the boundary. But most of pulmonary nodules appear in a shape of blob, where the pixels inside blob are much brighter than those in the surrounding. We can assume that the detected blob should be located inside the actual boundary of pulmonary nodule. Thus, we can limit the search for edge point in two distinct subsets of [**c - 3r**, **c** + 3**r**], namely, $\mathbf{z}_L$ = [**c - r, c - 3r**] (i.e., the ray segment left to the central pixel) and $\mathbf{z}_R$ = [**c + r**, **c** + 3**r**] (i.e., the ray segment right to the central pixel).



(a) (b)

Fig. 3. Two segmentation results obtained by using the proposed polynomial fitting algorithm: (a) and (b) are the segmentation result over the same two testing images in fig. 1 respectively.

We first seek two pixels, $\mathbf{g}_L$ and $\mathbf{g}_R$, in the ray segments $\mathbf{z}_L$ and $\mathbf{z}_R$, such that they are the local minimum points of the fitted curve *C* but they are nearest to the central pixel **c** amongst all the minimum points respectively. However, in many cases, a significant fitting error may appear for the pixels nearby a isolated noise or a strong noise pixel. This inevitably leads to a strong noise in the resulting segmented contour accordingly. Hence the resulting two pixels, $\mathbf{g}_L$ and $\mathbf{g}_R$, can be further refined by using a smoothed profile using

$$I_{Smoothed} = I_{profile} + w*(I_{profile} - I_{fit}) \qquad (4)$$

where *w* is a weighting parameter, which is set to be 0.1 (acquired through experiment) and using

$$\begin{aligned} \mathbf{g}_L^* &= \arg\min_{\mathbf{x}\in[\mathbf{g}_L,\mathbf{c}-\mathbf{r}]} I_{profile}(\mathbf{x}) \\ \mathbf{g}_R^* &= \arg\min_{\mathbf{x}\in[\mathbf{g}_R,\mathbf{c}+\mathbf{r}]} I_{profile}(\mathbf{x}) \end{aligned} \qquad (5)$$

where $\mathbf{g}^*_L$ and $\mathbf{g}^*_R$ are the resulting edge pixels in left hand side and right hand side of central pixel **c** respectively so that they have the lowest *smoothed* intensity $I_{Smoothed}$ respectively in the two segments [$\mathbf{g}_L$, **c–r]** and [$\mathbf{g}_R$, **c+r]**. Then the list of resulting boundary points obtained by the ray-casting algorithm must be smoothed by a median filtering [4]. The thinner curve in fig. 2 is an example of the original intensity profile for cast ray while the thicker one is its corresponding polynomial fitted curve *C*.

3. Experimental results

We have evaluated the proposed polynomial fitting ray-casting algorithm over an image library with 148 cancer cases collected from several hospitals. As a comparison benchmark, we also tested the original ray-casting algorithm by using gradient features in [4]. The number of nodules per image ranges from 1 to 3 while the size of nodules is from 5mm to 30mm. A *successful segmentation* can be defined if the automated segmented region of nodule is highly overlapped with the ground truth manually drawn by a panel of radiologists (5 radiologists) whereas a *failed segmentation* can be defined if they are overlapped in very low level. The performance comparison of the original ray-casting segmentation algorithm in [4] and the proposed algorithm is demonstrated in fig. 3. It can be observed from the figure that the proposed algorithm significantly outperforms the original ray-casting algorithm in [4] by using gradient features.

In the experiment, the algorithm was tested only over the nodules inside the clear lung field. In polynomial fitting of each intensity profile, we used a 7th-order of polynomial curves. Amongst the 148 cases for testing, 21 nodules were missed by blob detections, which must be excluded in order to evaluate the segmentation performance for nodules. In other words, we investigated the original ray-casting algorithm and the proposed algorithm only on a subset of the original image database (i.e., with 127 cases), in each case of which all nodules could be found by blob detection. The *successful segmentation* rate obtained by using the proposed polynomial fitting is 83.46% while the original ray-casting algorithm has only achieved 69.29% *successful segmentations*. Please be reminded that our testing cases are from different digital radiograph modalities in several hospitals and therefore their image qualities may be in a large variation. Thus, it is evident that the proposed polynomial fitting based algorithm is very robust in achieving a desirable segmentation of pulmonary nodules in chest radiograph.

4. Discussion

In this paper, we proposed a polynomial fitting based ray-casting algorithm for pulmonary nodule segmentation in chest radiograph. The initial experiment results have shown that the polynomial fitting based algorithm is very robust and efficient in a comparison to the original ray casting algorithm based on gradient features. The proposed algorithm can be viewed as a significant pilot study for robust segmentations for pulmonary nodules in chest radiograph.

Even if the proposed algorithm has archived a high rate of *successful segmentation* in average, the best order of polynomial curve fitting is still unknown. Estimation of an appropriate order for polynomial curve fitting may make the segmentation very robust to noise and variation of image quality. Future work can be done both on adaptive selection of polynomial curve order and investigation of the correlations between adjacent cast rays.

References

1. Global Cancer Facts and Figures 2007-08. American Cancer Society. Http://www.cancer.org.

2. X. Xu, K. Doi, T. Kobayashi, et al., "Development of an improved CAD scheme for automated detection of lung nodules in digital chest images", Med. Phys., Vol.24, No.9, pp.1395-1403, 1997.

3. M. Aoyama, Q. Li, S. Katsuragawa, et al., "Automated computerized scheme for distinction between benign and malignant solitary pulmonary nodueles on chest images", Med. Phys., Vol.29, No.5, pp.701-708, 2002.

4. Arnold M.R. Schilham, Bram van Ginneken, Marco Loog, "A computer-aided diagnosis system for detection of lung nodules in chest radiographs with an evaluation on a public database", *Medical Image Analysis*, Vol. 10, No. 2, pp.247-258, 2006.