Automated Nodule Location and Size Estimation Using a Multi-scale Laplacian of Gaussian Filtering Approach

Artit C. Jirapatnakul, Sergei V. Fotin, Anthony P. Reeves, Alberto M. Biancardi,
David F. Yankelevitz, and Claudia I. Henschke

Abstract— Estimation of nodule location and size is an important pre-processing step in some nodule segmentation algorithms to determine the size and location of the region of interest. Ideally, such estimation methods will consistently find the same nodule location irregardless of where the the seed point (provided either manually or by a nodule detection algorithm) is placed relative to the "true" center of the nodule, and the size should be a reasonable estimate of the true nodule size. We developed a method that estimates nodule location and size using multi-scale Laplacian of Gaussian (LoG) filtering. Nodule candidates near a given seed point are found by searching for blob-like regions with high filter response. The candidates are then pruned according to filter response and location, and the remaining candidates are sorted by size and the largest candidate selected. This method was compared to a previously published template-based method. The methods were evaluated on the basis of stability of the estimated nodule location to changes in the initial seed point and how well the size estimates agreed with volumes determined by a semi-automated nodule segmentation method. The LoG method exhibited better stability to changes in the seed point, with 93% of nodules having the same estimated location even when the seed point was altered, compared to only 52% of nodules for the template-based method. Both methods also showed good agreement with sizes determined by a nodule segmentation method, with an average relative size difference of 5% and -5% for the LoG and template-based methods respectively.

Index Terms— pulmonary nodule, size estimation, Laplacian of Gaussian

I. INTRODUCTION

Often, small pulmonary nodules detected on high-resolution computed tomography scans are too small to be diagnosed through conventional means such as by biopsy, so the growth rate of a nodule is increasingly becoming an important indicator of malignancy. In previous studies, nodule growth rate was found to be a good discriminating feature for distinguishing benign from malignant nodules [1]. Recent studies using automated volumetric measurement methods have also found growth rate to be a useful feature [2], though some studies have observed malignant nodules with very slow growth rates [3], [4]. Studies have suggested that automated methods of nodule volume measurement tend to have less variability than manual methods [5] and may be more useful for the assessment of treatment response [6].

Most automated methods require the specification of a seed point within the nodule, while some methods also require the manual specification of the region of interest containing the nodule [7].

One of the goals of a measurement method is to have high reproducibility. This can be achieved in automated methods by consistently using the same set of parameters. In the case of methods that require the manual specification of a seed point or the region of interest, using an algorithm to quickly and consistently estimate the same nodule location and size irregardless of where the initial seed point is placed within the nodule would likely improve the consistency of the method. Estimating the size of the nodule allows smaller image regions to be analyzed, potentially reducing processing time. Locating the nodule involves finding the approximate center of the nodule, given a seed point anywhere inside the nodule. This is important to allow the algorithm to be robust to the position of a seed point and may possibly reduce variability caused by an observer.

Gaussian fitting has been used to locate and size pulmonary nodules [8], and other studies have shown that the characteristic scale of a Laplacian of Gaussian (LoG) agreed well with radiologists' estimates of nodule size [9], [10]. In this work, we present a fully automated method for estimating the size and location of a nodule using a multi-scale LoG filtering approach. We compared this method to a previously published method using template functions [11] on a dataset of 29 nodules imaged prior to biopsy. Both methods were evaluated for the stability of the estimated nodule location to changes in the seed point and for how well the nodule's size is approximated.

II. METHOD AND MATERIALS

We developed a multi-scale Laplacian of Gaussian (LoG) filtering method for the task of nodule location and size estimation. Given a seed point to identify the nodule in the CT scan, a volume of interest was extracted from the image around the nodule; for this study, it was fixed in size to be 60 mm^3. The volume of interest was resampled into isotropic space with a voxel size of 0.5x0.5x0.5 mm in order to reduce errors due to quantization, as suggested by Kostis et al [12]. Finally, LoG filtering was performed on the resampled volume of interest. The method was compared to a template-based method that was used as an initial step in a nodule segmentation method published by Reeves et al. [11] on the same dataset. The same set of pre-processing methods were used to prepare images for both methods.
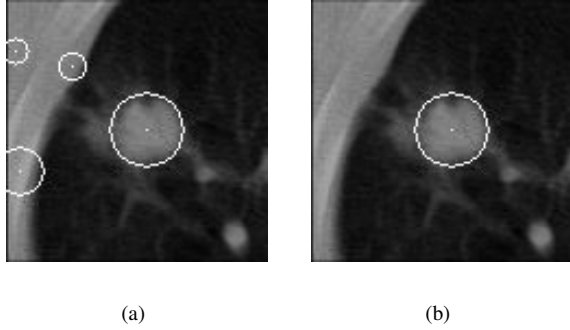
<center>(a)   (b)</center>

Fig. 1. Single slice of a volume of interest showing a) nodule candidates generated by the LoG filter and b) the selected candidate after pruning.

A. Scale-normalized LoG filtering method

We propose a multi-scale method using a Laplacian of Gaussian (LoG) filter for estimation of nodule size and location. This approach previously produced good results for nodule candidate generation in pulmonary nodule detection task [10]. For nodule location and size estimation, the algorithm generates nodule candidates in a small region of interest around the nodule, and the best candidate is selected through a rule-based method.

Nodule candidate space is identified by local minima in four-dimensional multi-scale image feature space $L(x, y, z, \sigma^2)$, created by filtering the original image performed as the series of convolutions with scale-normalized LoG kernels $\sigma^2 \nabla^2 G(x, y, z, \sigma^2)$ of different size:

$$L(x, y, z, \sigma^2) = I(x, y, z) * \sigma^2 \nabla^2 G(x, y, z, \sigma^2),$$

where $G(x, y, z, \sigma^2)$ is a three-dimensional Gaussian with variance σ^2. These convolutions are computed efficiently by using the fact that LoG can be approximated by a difference of Gaussians and can be performed by multiplication in the Fourier domain.

Each local minimum in this space identifies the spatial location and the best matching kernel size for a candidate nodule in that region of space:

$$< x, y, z, \sigma^2 >_i = \text{argminlocal} \left\{ L(x, y, z, \sigma^2) \right\}.$$

The size parameter of the kernel is directly related to the diameter of a candidate nodule as given by this expression $d^2 = 12\sigma^2$. For the task of identifying possible candidates, 150 normalized LoG kernels of incrementally increasing size corresponding to the diameter range from 3.0 to 30.0 mm were used.

The output of the LoG filtering stage is a list of candidate locations, sizes, and filter responses, which is pruned according to several rules. A threshold was applied to remove candidates with a low values of response $L(x, y, z, \sigma^2)$. Candidates with more than half of their volume outside the region of interest were removed, as were candidates that were located far from the manually specified seed point. Finally, the list of remaining candidates was sorted and the largest candidate was chosen as the best estimate for the nodule size

and location. A slice of a nodule with all the candidates in the slice is shown in Figure 1a, and the final selected candidate in Figure 1b.

B. Template-based method

The template-based method was used as a pre-processing step in the nodule segmentation algorithm developed by Reeves et al [11] and will be summarized here. The method was developed based on a two-level model where the nodule is modeled as a high-intensity spherical object surrounded by low-intensity lung parenchyma. Using this information, two template functions for the location, $L_{P,r}(x, y, z)$, and size, $S_{P,r}(x, y, z)$ are defined, where P indicates the nodule center point and r is the estimated radius of the nodule. Correlation between the template function and the image is used as the response of the template. A 3-D Gaussian was used as the sizing template because it fits the model of the nodule, while the negative of the Laplacian of Gaussian was used for the locating template because it has a negative weight further from the center, which is useful for locating juxtapleural nodules. The algorithm uses an iterative greedy search in four-dimensional space to find the best values of P and r. This method can be used by either specifying only a seed point, or by specifying two points on opposite sides of the nodule. In the first case, the method uses the seed point as the initial location estimate and assumes a small initial radius. For the second case, the initial location is estimated as the center of the line segment connecting the two points and the initial radius is estimated as half the distance between the two points. The method using only a seed point works well for spherically shaped nodules in uncomplicated situations, but with complex attachments, the method may estimate a nodule size and location over part of the nodule due to finding a local extremum. This can be addressed by using the two-point method, but this requires additional input from the radiologist and increases the subjectivity of the estimate. For this study, the initial location and an estimate of the radius were provided to the method.

C. Experiments

A method for nodule size and location estimation should be robust to changes in the position of the initial seed point. This was evaluated by comparing the estimated nodule location using 1) a seed point located at the approximate center of the nodule and 2) a seed point near the nodule periphery. The peripheral seed point was placed to represent seed point variability that might be expected to occur in clinical use; the center seed point was shifted by 50% of the radius in-plane and by 1 slice to obtain the peripheral seed point. The radius of the nodule was computed from the volume estimated by a previously published algorithm [11] for 22 nodules; the remaining seven nodules had the radius estimated manually. For example, a nodule with a diameter of 10 mm would have a peripheral seed point shifted 2.5 mm along the x and y axes and 1 slice from the center point. The Euclidean distance between the estimated location of each nodule using the center seed point and peripheral seed points

	Mean dist. (mm)	SD (mm)	% nodules no change
LoG	1.27	4.56	93% (27/29)
Template	3.23	5.14	52% (15/29)

TABLE II

RELATIVE DIFFERENCE IN DIAMETER ESTIMATES COMPARED TO
VOLUMETRIC SEGMENTATION-BASED METHOD.

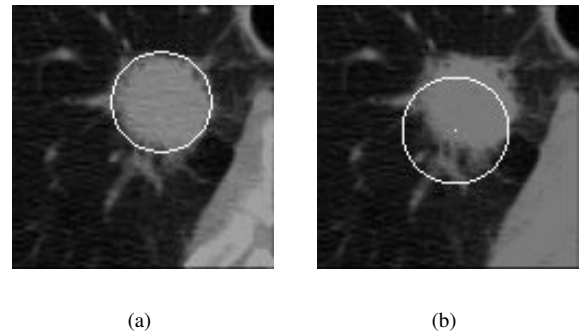| Method	Mean (%)	SD (%)
LoG	4.6	6.0
Template	-5.3	8.8

(a)   (b)

Fig. 2. Center slice of a nodule with the estimated nodule location and size indicated by a circle for the a) LoG method and b) template-based method. Both results were obtained for a peripheral seed point.
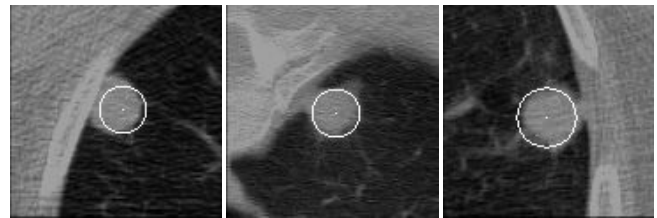

Fig. 3. Center slice of a juxtapleural nodules successfully located and sized by the LoG method

was computed, and in the ideal case, the distance between the estimated locations will be zero.

In addition to estimating the nodule location, these methods also estimate the nodule size in terms of diameter. For nodules which had automated volumetric measurements available, the diameters estimated by both methods were compared to the diameter equivalent of the volumetric measurements. The relative difference was computed and the mean and standard deviation were computed for each method.

D. Data

A dataset of 29 nodules with scans obtained during biopsy procedures was used to evaluate the methods. The nodules ranged in size from approximately 7 mm to 30 mm. The nodules were imaged on scans with a whole-lung field of view with slice thicknesses of 1.25 mm (20), 2.5 mm (7), or 5.0 mm (2). Scans were acquired using GE LightSpeed QX/i, LightSpeed Pro 16, or LightSpeed Ultra scanners using 120 kVp and a current in the range of 40 - 250 mA. Of these 29 nodules, 22 had size measurements from an automated segmentation method available.

III. RESULTS

The stability results of both methods are summarized in Table I as the average distance between the estimated nodule location using the center and peripheral seed points. The number of nodules with no change in the estimated nodule location is also given. The mean and standard deviation of the relative differences in size compared to a volumetric segmentation-based method are given in Table II for both methods for the 22 nodules with available segmentations. The nodule sizes were estimated based on the central seed point. Differences were computed from the volumetric method, so that a negative mean relative difference indicates a method that tends to overestimate, while a positive mean indicates a method that underestimates.

IV. DISCUSSION

The stability of both methods was relatively good, with the LoG method exhibiting better stability than the template-based method; the LoG method was only provided with the

seed point, while the template-based method required the seed point and an estimate of the nodule size. The LoG method had an average distance between nodule locations that was less than half the distance of the template-based method. Additionally, the LoG method found the same seed point in 93% (27/29) of nodules despite altering the seed point, while the template-based method only found the same seed point in 52% (15/29) of nodules. The LoG method had large errors in the two cases where it failed to find the same nodule location, whereas the template-based method had a wide range of errors. Depending on the application, catastrophic errors may be preferable because they can be detected and reported to the user.

The LoG method worked in many difficult cases. In Figure 2, the LoG method correctly identified the location and size of the nodule with many spiculations despite using a peripheral seed point, whereas the template-based method found an incorrect nodule location. The template-method selected the wrong location because of its use of a LoG template for locating the nodule; moving the location in any direction would have lowered the template response. The LoG method was also successful for juxtapleural nodules, as shown in Figure 3.

The LoG method was designed for spherical nodules with attachments; nodules that significantly deviated from this model resulted large variations of the seed point location. An example of a nodule where the LoG method incorrectly estimated the nodule size and location, with a large difference in the estimated location when the seed point was altered, is shown in Figure 4; the estimated nodule location from the
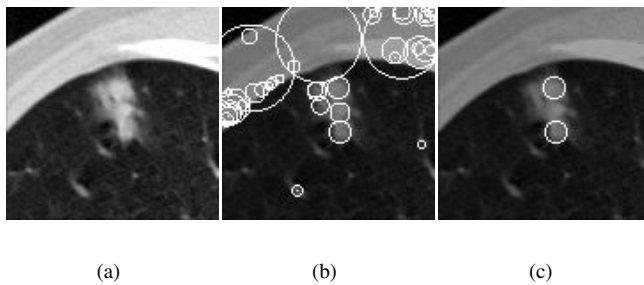
<div align="center">(a)   (b)   (c)</div>

Fig. 4. a) Single slice of a flat elongated nodule showing b) candidates generated by the LoG method and c) the final selected nodule locations. Different candidates are selected depending on the location of the seed point; the upper circle is the estimate from the central seed point, while the lower circle is the from the peripheral seed point. Images are windowed at different levels to improve contrast.

center seed point is in the upper left, while the estimated location using the peripheral seed point is the lower right. In this case, due to the variations in density within the nodule, no candidates for the entire nodule were generated, and as the peripheral seed point was closer to a different candidate, it was chosen over the candidate from the center seed point.

The size estimates provided by the methods can be used as an initial condition by further processing stages in nodule segmentation algorithms to reduce the time required to reach convergence, or to reduce the search space by providing an estimate of the size of the region of interest of the nodule. The estimates provided by the methods were within approximately 5% of the size determined by the segmentation-based method as shown in Table II. Since the goal of these methods is only to get an estimate of the nodule size, the performance of both algorithms is acceptable as a pre-processing step.

Even though the LoG method performed well, there are several areas for improvement. The method is limited in its ability to accurately located and size elongated nodules; this can be addressed by using non-spherical kernels, but at the expense of additional computational complexity. Note that this limitation is not unique to the LoG method; the template-based method has a similar limitation. Further, the speed of the method can be improved by first performing a coarse estimate of size followed by a finer size estimate; in the current implementation, the method takes approximately 1 minute per nodule on a Xeon 3.0 ghz processor.

V. Conclusion

Estimating the location and size of a pulmonary nodule is useful as a pre-processing step for nodule segmentation algorithms; these parameters can be used to define a small volume of interest from the entire CT scan, reducing computation time and memory usage in later processing steps. In this work, we applied a Laplacian of Gaussian (LoG) filtering method to the task of estimating nodule size and location. To evaluate this method, the initial seed point provided to the algorithm was altered and the estimated nodule locations were compared. This new method was compared to a previously published template-based method

on the same set of nodules; the LoG filtering method showed improved stability to changes in the seed point compared to the template-based method, despite only requiring a seed point, whereas the template-based method required both a seed point and an estimate of the nodule radius. Although the LoG method works well for spherical nodules both isolated and attached to vascular structures or the pleural surface, nodules that are extremely lobulated or elongated still pose a challenge. Future work will improve upon these issues and incorporate this algorithm into a nodule segmentation algorithm.

References

[1] M. H. Nathan, V. P. Collins, and R. A. Adams, "Differentiation of benign and malignant pulmonary nodules by growth rate," *Radiology*, vol. 79, pp. 221–231, August 1962.

[2] D. F. Yankelevitz, A. P. Reeves, W. J. Kostis, B. Zhao, and C. I. Henschke, "Small pulmonary nodules: Volumetrically determined growth rates based on CT evaluation," *Radiology*, vol. 217, pp. 251–256, October 2000.

[3] H. T. Winer-Muram, S. G. Jennings, R. D. Tarver, A. M. Aisen, M. Tann, D. J. Conces, and C. A. Meyer, "Volumetric growth rate of stage I lung cancer prior to treatment: Serial CT scanning," *Radiology*, vol. 223, pp. 798–805, April 2002.

[4] S. G. Jennings, H. T. Winer-Muram, M. Tann, J. Ying, and I. Dowdeswell, "Distribution of stage i lung cancer growth rates determined with serial volumetric CT measurements," *Radiology*, vol. 241, pp. 554–563, Nov 2006.

[5] S. G. Jennings, H. T. Winer-Muram, R. D. Tarver, and M. O. Farber, "Lung tumor growth: Assessment with CT – comparison of diameter and cross-sectional area with volume measurements," *Radiology*, vol. 231, pp. 866–871, June 2004.

[6] K. Marten, F. Auer, S. Schmidt, G. Kohl, E. J. Runmeny, and C. Engelke, "Inadequacy of manual measurements compared to automated CT volumetry in assessment of treatment response of pulmonary metastases using RECIST criteria," *Eur Radiol*, vol. 16, pp. 781–790, 2006.

[7] B. Zhao, A. P. Reeves, D. F. Yankelevitz, and C. L. Henschke, "Three-dimensional multicriterion automatic segmentation of pulmonary nodules of helical computed tomography images," *Optical Engineering*, vol. 38, pp. 1340–1347, Aug. 1999.

[8] K. Okada, D. Comaniciu, and A. Krishnan, "Robust anisotropic gaussian fitting for volumetric characterization of pulmonary nodules in multislice CT," *IEEE Transactions on Medical Imaging*, vol. 24, pp. 409–423, March 2005.

[9] S. Diciotti, S. Lombardo, G. Coppini, L. Grassi, L. Petrolo, G. Picozzi, M. Falchini, and M. Mascalchi, "The characteristic scale as a consistent indicator of lung nodule size in CT imaging," in *CARS*, pp. S182–183, 2008.

[10] S. V. Fotin, A. P. Reeves, A. M. Biancardi, D. F. Yankelevitz, and C. I. Henschke, "A multiscale Laplacian of Gaussian filtering approach to automated pulmonary nodule detection from whole-lung low-dose CT scans," in *SPIE International Symposium on Medical Imaging*, vol. 7260, p. 72601Q, Feb 2009.

[11] A. Reeves, A. Chan, D. Yankelevitz, C. Henschke, B. Kressler, and W. Kostis, "On measuring the change in size of pulmonary nodules," *IEEE Transactions on Medical Imaging*, vol. 25, pp. 435–450, April 2006.

[12] W. J. Kostis, A. P. Reeves, D. F. Yankelevitz, and C. I. Henschke, "Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images," *IEEE Transactions on Medical Imaging*, vol. 22, pp. 1259–1274, October 2003.