SEMI-AUTOMATED MEASUREMENT OF PULMONARY NODULE GROWTH WITHOUT EXPLICIT SEGMENTATION

A. C. Jirapatnakul[1], *A. P. Reeves*[1], *A. M. Biancardi*[1], *D. F. Yankelevitz*[2], *C. I. Henschke*[2]

[1]School of Electrical and Computer Engineering, Cornell University, Ithaca, NY
[2]Weill Cornell Medical College, New York, NY

ABSTRACT

Many nodule measurement methods rely on accurate segmentation of the nodule and may fail with complex nodule morphologies; often slight variations in segmentation result in large volume differences. A method, growth analysis from density (GAD), is presented that measures nodule growth without explicit segmentation through the application of a Gaussian weighting function to a region around the nodule, avoiding the drawbacks of segmentation-based methods. The resulting mean density is used as a surrogate for volume when computing growth. A zero-change nodule dataset was used to establish the variability of the method, followed by testing on datasets of stable, malignant, and complex nodules. There was no significant difference in percent volume change between the methods (p=0.55), and while GAD showed similar measurement variability and discriminative performance as a segmentation-based method (GAS), it was able to successfully measure growth on complex nodules where GAS failed.

Index Terms— X-ray tomography, pulmonary nodule growth, lung cancer, density change

1. INTRODUCTION

Improvements in CT scanner technology have enabled the earlier detection of smaller nodules than previously possible. Small nodules are often difficult to biopsy, and thus, other indicators must be used to diagnose such nodules. The growth rate of a pulmonary nodule on CT scans has been shown to be a good indicator of malignancy [1]. Many methods of measuring nodule volume segment the nodule and count the number of voxels in the segmentation to compute a volume estimate. Although these methods have steps that address attached structures, some nodules with substantial and complicated attachments can be difficult for these algorithms to segment, resulting in a substantial error in their volume estimate. Further, measuring tumor growth based on volume alone may not accurately characterize the true growth of the nodule – non-solid and part-solid nodules are prime examples where measuring the density change in addition to the volume change would be desirable.

Quantifying the uncertainty in volume estimation is critical for the proper interpretation of the results of an automated measuring tool. Studies have attempted to address this through the use of scans obtained in "coffee break" studies where a patient is scanned several times, leaving the scanning table between scans. Since these scans are separated by only a few minutes, no change should have occurred, and the nodule volume measurements should be identical. Gietema et al showed a 95% limits of agreement for the difference

in measured volumes of (-21.2%, 23.8%) [2] in their study using Siemens LungCare. Wormanns et al assessed the precision of a pre-release version of Siemens LungCare on 151 nodules from ten subjects [3], and the limits of agreement for repeatability were -20.4% to 21.9%.

We propose an improved nodule growth measurement method that takes into account the radiodensity of the nodule and does not require explicit segmentation. One method has attempted to measure nodule volume without explicit segmentation [4], relying instead on the size of a Gaussian kernel determined using a multi-scale approach; another method relied on the CT density histogram to compute nodule growth rates [5]. In contrast, the method described in this paper, growth analysis from density (GAD), measures the mean density of a region containing the nodule after applying a weighting function to reduce the influence of structures far from the nodule center. GAD was compared to a segmentation-based method, growth analysis from segmentation (GAS), on the basis of measurement variability on a zero-change dataset and diagnostic performance on stable and malignant nodules.

2. MATERIALS AND METHODS

The GAD method is comprised of several stages. Initial preprocessing of the CT scans is performed to select a region of interest around the nodule and generate an isotropic image. Next, the center and size of the nodule are estimated from a manually specified seed point using an iterative optimization approach. In the third step, the regions surrounding the nodule on each CT scan are registered. In the final step, a weighting function is applied to the region and the mean density computed; this mean density is used as a surrogate for volume in the computations for growth rate. The first three steps rely on previously published work [6]; the novelty of this method is in the final step and evaluation method. The method is evaluated for variability and discrimination performance using several datasets.

2.1. Preproccessing

The input to the algorithm is a pair of CT scans, I_1 and I_2, containing the nodule, and seed points located within the nodule on both scans. Based on this, for each scan, a region of fixed size is extracted from the original CT scan around each seed point. These regions are re-sampled into isotropic space, and these images, I_{R_1} and I_{R_2}, along with the coordinates of the seed points in the resampled space are provided to the next stage of the algorithm. Resampling the image into isotropic space enables subvoxel precision for locating and sizing the nodule and makes the process of registration easier, with the drawback of increasing the image size and computation time. An additional step of juxtapleural detection is performed to aid in the next step.

(a) I_1 (b) I_2
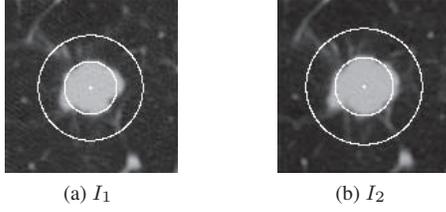
Fig. 1: Estimate of nodule location and size for a) scan at T1 and b) scan at T2. Only the central slice of the region is shown. The location of the nodule is at the center of the circle; the inner circle indicates the estimated size of the nodule.

(a) I_{RR_1} (b) I_{R_2} (c) $I_{R_2} - I_{RR_1}$

Fig. 2: Result of region registration; central slice of a) nodule region on first scan registered to b) second scan and c) difference image between the second scan and registered image. Gray indicates no difference.

2.2. Estimation of nodule center and size

Given the resampled regions and seed points, estimates of the nodule center, C_1 and C_2, and size, S_1 and S_2, on both regions are determined using an algorithm described by Reeves et al [6], but in contrast to their method, the GAD method performs this step on the re-sampled isotropic image. In their algorithm, two template functions are defined: a localization template, which is the negative Laplacian of Gaussian, and a sizing template, which is a Gaussian. An iterative process is used to determine the nodule location and size with the highest response. An example of the result of the algorithm on the first and second scans of a nodule is shown in Figure 1. A guess of the radius may be passed to the algorithm; this guess is used as an initial condition and the same process described above is used to determine the nodule size. The computed nodule center and size from this method are only estimates used to determine the appropriate location and kernel size for later steps in the algorithm. The region from the second scan is reduced in size based on the estimated nodule size; the region of the first scan is not altered in this step.

2.3. Nodule region registration

The I_{R_1} image from Section 2.1 is registered to the second scan, I_{R_2}, using a three-dimensional rigid body registration algorithm [6]. Although the reference image could be swapped without significantly changing the result, I_{R_2} is chosen as the reference here since further steps compute parameters from the nodule on I_{R_2}. In brief, the algorithm uses a Gaussian-weighted mean-square-difference matching metric and conducts a search over all parameters to minimize the metric using Powell's method. The initial translation parameter is derived from the computed difference in nodule centers, $C_2 - C_1$, and the standard deviation of the Gaussian weighting function is derived from the estimated size of the nodule on the second scan (S_2). In Figure 2, the nodule on the first scan has been registered to the nodule on the second scan. The registered image I_{RR_1} is shown on the left with the difference between the registered image and the nodule on the second scan on the right ($I_{R_2} - I_{RR_1}$). Note that this step does not alter the image from the second CT scan.

2.4. Growth determination

The inputs to this stage are nodule locations C_1 and C_2, the size on the second image S_2, the registered image I_{RR_1}, and resampled image I_{R_2}. To measure nodule growth, instead of computing the volume by segmenting the nodule, a novel method of computing the mean density is used. This is driven by the observation that the region around a nodule consists of lung parenchyma, some part-solid components and partial voxels, the solid component of the nodule, and high-intensity attached structures, such as vessels and the pleural wall. Although the mean density could be computed directly from the image, high-intensity structures in the lung not part of the nodule but included in the region will contribute equally to the mean density of the nodule, which is undesirable for nodule growth measurement. The attached structures are often located at the periphery of the nodule; to de-emphasize these structures, the densities of the entire region are weighted by a Gaussian. The equation for a Gaussian used here is:
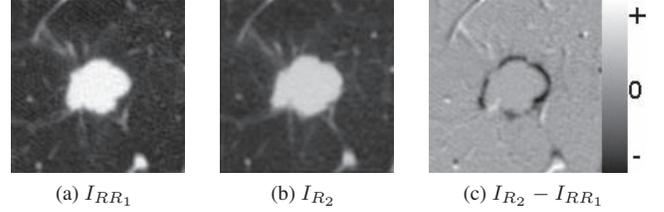
$$G(x, y, z, C, \sigma) =$$
$$\frac{1}{(2\pi)^{3/2}\sigma^3} \exp\left(\frac{-(x - C_x)^2 - (y - C_y)^2 - (z - C_z)^2}{2\sigma^2}\right) \quad (1)$$

where C is the coordinate of the center of the nodule, and σ is the standard deviation of the kernel.

The standard deviation of the Gaussian used to weight the densities in the region has an effect on the performance of the method – if the size is too small, the method will not be very sensitive to changes, while if the size is too large, other structures have a greater influence on the result. For this method, the Gaussian was truncated at 3σ and σ was set to 66% of the estimated nodule size. The Gaussian was centered at C_2 on both I_{R_2} and I_{RR_1}. Thus, the mean density was computed according to the following equation on both images:

$$M = \frac{\sum_x \sum_y \sum_z I(x, y, z)G(x, y, z, C_2, S_2)}{\sum_x \sum_y \sum_z G(x, y, z, C_2, S_2)} \quad (2)$$

By doing this, the mean density was computed on as similar regions as possible. The difference in the cube of the mean density between the scans was used as a measure of nodule change. Growth rates were computed in the same way as for the volumetric method, with the cube of the mean density substituted for the volume.

2.5. Data

Four datasets were used in the evaluation of the method. Unless otherwise specified, cases were selected that had a nodule with solid consistency, as determined by a radiologist, a successful segmentation by the GAS method, and a diameter less than 20 mm on the first scan, according to the GAS method. Two scans through the entire nodule were available for each case to allow for growth measurement. A summary of the parameters for all datasets is presented in Table 1; only information not listed in the table is described below.

The first dataset consisted of 20 cases, each with a single nodule, with scans several minutes apart which were imaged during the preliminary stages of a biopsy; the expected volume change of the

Table 1: Dataset parameters

Dataset	# Nodules	Mean, SD of Size (mm)	Slice Thickness (mm)	kVp	Current (mA)	Median, Range of Interval (days)
Zero-change	20	12.69, 3.63	1.25 - 5.0	120, 140	40 - 250	0, 0 - 0
Stable	38	6.91, 3.16	1.0, 1.25	120, 140	80 - 300	393.5, 91 - 1918
Malignant	19	6.68, 3.03	1.0 - 5.0	120, 140	40 - 300	165, 90 - 756
Complex	4	17.61, 3.81	1.0 - 5.0	120, 140	40 - 300	50, 29 - 98

nodules between these scans was zero. Ten cases had two scans of the same slice thickness (1.25 mm), while ten cases had two scans of different slice thicknesses – seven of these cases had one scan with a slice thickness of 1.25 mm and the other 2.5 mm, and in three cases, one scan was 1.25 mm while the other was 5.0 mm. Scans were obtained using either a GE LightSpeed QX/i or LightSpeed Ultra scanner.

The second dataset was comprised of 38 cases with a stable nodule, as confirmed by either biopsy or 2 years of no clinical change, selected from the Weill Cornell Medical College (WCMC) database. All of the nodules but one had scans of the same slice thickness. Scans were acquired with either a GE LightSpeed Ultra, LightSpeed Pro 16, LightSpeed VCT, HighSpeed CT/i, or Genesis scanner.

The third dataset had 19 cases with a malignant nodule selected from the WCMC database. Cases were selected to have an interval of at least 90 days and a volume change of at least 30%, as measured by the GAS method. These criteria were selected to ensure that the volume change was greater than the measurement uncertainty measured by previous studies. Malignancy was confirmed by biopsy or resection. A GE LightSpeed Ultra, LightSpeed QX/i, HighSpeed CT/i, or Genesis scanner was used.

In the fourth dataset, malignant nodules with a complex appearance were selected from the WCMC database where the GAS method was unsuccessful in segmenting the nodule. These nodules were solid or part-solid in appearance. Since the nodules were malignant, the measured growth rate should correspond to malignancy.

2.6. Experiment

The first task was to determine if GAD was measuring the same quantity as GAS. The percentage change of the estimator of GAD $(M_2^3 - M_1^3)/M_1^3$ was compared to the percentage change in volume of GAS to determine if the methods were correlated. The methods were then evaluated on the basis of measurement repeatability and discrimination performance for malignant nodules. For each nodule in the zero-change dataset, both the GAS and GAD methods computed the percentage change in their respective estimators. The 95% limits of agreement (LoA) for the percentage change was used as a metric of the method variability; smaller LoA imply less variability.

Measurement variability is not the sole consideration of the effectiveness of a method; it must be able to accurately measure nodule growth to enable the diagnosis of malignant nodules. To accomplish this, the growth indices (GI) are measured for a set of stable nodules using the following equation:

$$GI = 100 \left[(E_2/E_1)^{30.4375/\Delta t} - 1 \right] \qquad (3)$$

where E_1 and E_2 are either the volumes V_1 and V_2 computed by GAS on the first and second scans or M_1^3 and M_2^3 measured by GAD on the first and second scans, and Δt is the interval in days between scans. The GI represents the percentage change per month assuming the growth follows an exponential model. A 95% LoA

Table 2: Limits of agreement (LoA) on the percentage volume change (PVC) for the subset of zero-change dataset with same slice thickness scans for both images (10 nodules).

Method	Mean PVC (%)	SD PVC (%)	LoA on PVC
GAD	1.98	13.2	-24.0, 27.9
GAS	-3.3	11.2	-25.2, 18.6

Table 3: LoA of growth index (GI) for stable nodules.

Method	Mean GI	SD GI	LoA of GI
GAD	-0.07	2.18	-4.35, 4.21
GAS	0.07	2.45	-4.74, 4.87

was established for the GI for each method on the stable nodules, and nodules with growth greater than the upper limit of the LoA were considered malignant. The GI for nodules in the malignant dataset were compared with this threshold value; all of the malignant nodules should have GI values consistent with significant growth.

3. RESULTS

Linear regression was performed between both methods, using the percent volume change (PVC) of the GAS method and the percent change in the cube of the estimator of the GAD method over the all datasets except for the set of complex nodules. The GAD method was highly correlated with the GAS method ($r^2 = 0.97$). The Wilcoxon signed-rank test also showed no significant difference in the PVC values of the two methods ($p = 0.55$).

The measurement variability of both methods on the zero-change dataset was compared. The mean PVC was -3.5% with a standard deviation (SD) of 18.5%, resulting in a 95% limits of agreement (LoA) of -39.9% to 32.8% for the GAD method. In contrast, the GAS method had a mean PVC of -12.2% with a SD of 21.6%, resulting in a LoA of -54.6% to 30.3%. In two cases, the nodule location and size estimation step failed; for these cases, the algorithm was also given an estimated radius in addition to the seed point. The PVC difference between both methods was not statistically significant ($p = 0.26$). The 95% LoA was also computed on just the subset of ten nodules with scans of the same slice thickness in Table 2.

On the dataset of stable nodules, the GAD method measured growth rates that were consistent with stability, defined as a growth index < 5.3 or a doubling time (DT) > 400 days, for all cases. The GAS method measured growth rates consistent with stability in all but one case which had a GI of 6.4 (DT=335.6 days). The differences in GI were not statistically significant ($p = 0.75$). The 95% LoA of the GI for both methods are shown in Table 3. In five cases, an approximate radius had to be specified to the nodule locating and

Table 4: Summary of growth index (GI) values for both methods on the datasets of stable and malignant nodules

Dataset	Method	GI Range	Median GI
Stable	GAD	-7.3 - 3.7	0.37
	GAS	-6.2 - 6.4	0.10
Malignant	GAD	4.0 - 36.9	15.2
	GAS	3.0 - 45.5	16.7
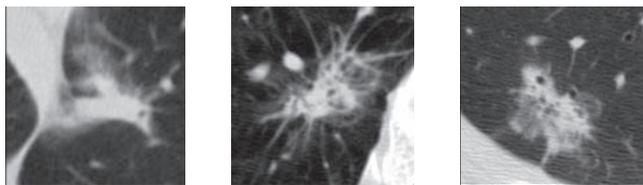
Fig. 3: Central slice of three complex nodules.

sizing function.

For the set of malignant nodules, the upper limit of the 95% LoA of the GI from the stable cases was used to establish a threshold for malignancy. The GAD method identified 94.7% (18/19) cases as malignant, while the GAS method identified 89.5% (17/19) of the cases as malignant. The differences in growth index were not statistically significant (p = 0.77). The range and mean of GI values are shown in Table 4 for both methods on the datasets of malignant and stable nodules. Three cases required the specification of a radius estimate in addition to the seed point.

On the dataset of complex, malignant nodules not able to be segmented by the GAS method, the GAD method measured GI values of 6.8, 13.9, 25.1, and 31.2 which indicated malignancy for all of the nodules. Three of the nodules are shown in Figure 3.

4. DISCUSSION

The preliminary results of the GAD method are promising. The PVC values computed by the GAD method are well correlated with the GAS method ($r^2 = 0.97$), and the signed-rank test indicated that there was no significant difference between the two methods. The new method had a smaller LoA on the entire zero-change dataset compared to GAS, though a slightly larger LoA on the subset of zero-change cases on the same slice thickness. These results are comparable to previously published work in the area which report a LoA of approximately -20% to 20%.

The performance of the method on the zero-change dataset is not sufficient to ensure the proper behavior of the method, as the goal of these methods is to accurately measure pulmonary nodule change. On the stable nodule dataset, the GAD method had a smaller range of GI values and a smaller LoA compared to the GAS method. In addition, the growth rates of the GAD method were all under the threshold of malignancy (GI < 5.3). These results are similar to the results on the zero-change dataset. We can use these LoA to assess what GI represent significant growth.

Using these LoA on the dataset of malignant nodules, the GAD method identified one additional case as malignant (18/19) compared to the GAS method (17/19). The same case was misidentified by both methods, though the GAD GI was higher (4.0) for the case than the GAS GI (3.0). All four cases of the set of complex malignant nodules had a GI measured by GAD consistent with malignancy.

These results suggest that GAD is just as sensitive at detecting malignant nodules but has the added advantage of working on complex nodules that are problematic for other methods.

Although the GAD method works well on this dataset, there are a few areas for future improvement. One problem experienced on this dataset was a failure of the nodule location and sizing algorithm in 13% of cases (10/77). In all of these cases, providing an estimate of the radius enabled the correct estimation of the size and location. In most of these cases, the radius was underestimated; future work will aim to make the method more robust. Another concern is the effect of high-intensity structures on the estimated mean density. While these structures contribute little to the mean density if they are located far from the nodule center and are sparse in number, if they are located near the nodule center, as may be the case in some juxtavascular nodules, or if they have many attached vessels, the mean density may be affected, though it is not clear if this is a desirable effect – a change in the number of vascular attachments is also relevant to the characterization of nodule growth.

5. CONCLUSION

A method for measuring nodule growth without explicit segmentation (GAD) was described and extensively compared to a segmentation-based method (GAS) on the basis of measurement variability on a zero-change dataset, ability to discriminate between stable and malignant nodules, and performance on a set of complex malignant nodules. GAD exhibited similar measurement variability as the GAS method and slightly better discriminative ability than the GAS method. GAD has the advantage of being more robust than GAS, successfully measuring a GI consistent with malignancy for complex nodules not successfully segmented by GAS.

6. REFERENCES

[1] 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, no. 1, pp. 251–256, October 2000.

[2] H. A. Gietema, C. M. Schaefer-Prokop, W. P. T. M. Mali, G. Groenewegen, and M. Prokop, "Pulmonary nodules: Interscan variability of semiautomated volume measurements with multisection CT influence of inspiration level, nodule size, and segmentation performance," *Radiology*, vol. 245, no. 3, pp. 888–894, 2007.

[3] D. Wormanns et al., "Volumetric measurements of pulmonary nodules at multi-row detector CT: In vivo reproducibility," *European Radiology*, vol. 14, pp. 86–92, 2004.

[4] K. Okada, D. Comaniciu, and A. Krishnan, "Robust anisotropic gaussian fitting for volumetric characterization of pulmonary nodules in multislice CT," *IEEE Trans. on Medical Imaging*, vol. 24, no. 3, pp. 409–423, March 2005.

[5] Y. Kawata et al., "A computerized approach for estimating pulmonary nodule growth rates in three-dimensional thoracic CT images based on CT density histogram," in *Medical Imaging 2005: Image Processing*, J. Michael Fitzpatrick and Joseph M. Reinhardt, Eds. 2005, vol. 5747, pp. 872–882, SPIE.

[6] A. Reeves, A. Chan, D. Yankelevitz, C. Henschke, B. Kressler, and W. Kostis, "On measuring the change in size of pulmonary nodules," *IEEE Trans. on Medical Imaging*, vol. 25, no. 4, pp. 435–450, April 2006.