

## Parametric and Non-Parametric Nodule Models: Design and Evaluation

Amal A. Farag, James Graham, Aly A. Farag, Salwa Elshazly and Robert Falk\*  
Department of Electrical and Computer Engineering, University of Louisville  
(\* ) Medical Imaging Division, Jewish Hospital, Louisville, KY, USA

**Abstract.** Lung nodule modeling quality defines the success of lung nodule detection. This paper presents a novel method for generating lung nodules using variational level sets to obtain the shape properties of real nodules to form an average model template per nodule type. The texture information used for filling the nodules is based on a devised approach that uses the probability density of the radial distance of each nodule to obtain the maximum and minimum Hounsfield density (HU). There are two main categories that lung nodule models fall within; parametric and non-parametric. The performance of the new nodule templates will be evaluated during the detection step and compared with the use of parametric templates and another non-parametric Active Appearance model to explain the advantages and/or disadvantages of using parametric vs. non-parametric models as well as which variation of non-parametric template design, i.e., shape based or shape-texture based yields better results in the overall detection process.

**Keywords:** Nodule modeling, Sensitivity and Specificity of CAD systems, Data-driven nodule models, Level Sets

### 1 Introduction

This paper focuses on modeling of the lung nodules which appear in low dose computer tomography (LDCT) of the human chest. In the past two decades numerous screening studies in Europe, Japan and the US have been conducted for studying the enhancements of early detection of lung cancer using CT vs. X-ray and for studying the correlation of early detection and possible enhancement in lung cancer related mortality. Suffices to say, lung cancer is a major problem worldwide [1]. The survival of lung cancer is strongly dependent on diagnosis [2]. Research studies to reach an optimal detection rate for early detection of lung cancer, is the hope for improved survival rate [3]-[5].

Machine learning and computer vision methodologies have been used for image analysis of low dose CT (LDCT) of the chest (e.g., [2] [5]-[8]). A computer-aided (CAD) system for interpretation of nodules is formed of four major steps: scan filtering to remove acquisition artifacts; segmentation to isolate the lung tissue from the rest of the chest region; nodule detection to isolate candidate nodules; and nodule classification which categorizes detected nodules into possible pathologies. The literature is rich in approaches to segment the lung from the rest of the chest tissues, but the majority of the nodule modeling methods are based on parametric descriptions

of the nodules (e.g., in 2D circular or semicircular models are used, while in 3D volumes spherical or hemispherical models are used [5]). This paper will focus on the third component of the CAD system used for early screening of lung cancer; more specifically how to properly model lung nodules by using the actual data information to create the models/templates that will improve detection rate. The approach of nodule detection hinges mainly on proper modeling of nodule templates and much less on the computational approach to carry out the detection. Extensive surveys on automatic lung nodule segmentation and detection may be found in [8][9]. The long history of work in the communications literature (e.g., matched filtering) and computer vision literature (e.g., active shape and active appearance models) has demonstrated the value of proper modeling of the objects to be detected. This is especially the case in biomedical applications as the anomalies are often camouflaged or occluded by anatomical structures and the limitations of the scanning device or protocols. To the best of our knowledge none of the studies on lung screening conducted worldwide has resulted in identifiable databases of nodules listing their types and pathologies. Therefore, the need is persistent for reliable nodule models based on the actual scans; this is one of the goals of the authors of this paper.

This paper will investigate two main topics: First, whether parametric template models are more effective than non-parametric models in terms of sensitivity and specificity. We use the term data-driven for the non-parametric templates since the raw data information is used. Second, we study the effect of shape and texture on the non-parametric models. Shape only models were obtained by co-registering the contours of an ensemble of nodules obtained by a level set approach. The shape and texture models were obtained by active appearance models (AAM) (e.g., [17]). Our focus in particular, is on four nodule types (e.g., Kostis et al. [6]) that possesses discriminatory features of shape and to some extent texture.

The closest related work to this paper are: 1) Kostis et al. [6] provides a description of four major types of lung nodules based on identifiable landmarks which will enable automatic annotation in our methods. 2) Lee et al. [5] established an empirical relationship, or behavior, for the intensity (or Hounsfield Units) of the nodules as a function of the radial distance from the centroid of the nodule; this is beneficial for texture/intensity/gray level estimation of the inside of a nodule shape. 3) Farag et al. [7][10][11] established a parametric form for the relationship between the radial distance and the Hounsfield units in Lee's work; this is very useful for estimating the intensity of a nodule model given the statistics of an ensemble and the size (radius of the bounding box containing the nodule). Other works include: the use of 3D Markov random models [13] to compute the optimal states over the cells from the relationship of the neighboring cells and Bayesian voxel labeling [14] which labels the image according to three categories; anatomy, pathology and miscellaneous. The probability that the voxel belongs to each class is computed and a decision is made. The approaches described here assume parametric shape representations of the nodules, whether circles, spheres, etc. or fused ellipsoids to represent the nodule and its outer surroundings. Parametric modeling of the nodules (e.g. circles and/or semicircles) will not capture the shape variations in the lung nodules. Likewise, assuming a uniform (or binary) HU for the nodule doesn't resemble reality. Shape and texture models have shown great promise in a number of computer vision and biomedical imaging analysis applications (e.g. [15][16]). To the

best of our knowledge Farag et al. in 2009 [17] is the first reported attempt at employing shape and appearance modeling into the problem of automatic detection and segmentation of lung nodules in LDCT scanning. This paper is organized as follows: section 2 describes the level sets shape based approach for obtaining a mean template representation for each nodule type and how these templates are filled with texture. The parametric and AAM methods are not described in great details in this paper. Section 3 discusses performance evaluation; and section 4 concludes the paper.

## 2 Nodule Modeling

This section will examine the process of nodule modeling and simulation using an ensemble of nodules identified by radiologists. The level sets method for generating nodule models is the main focus of this section, while the parametric [7][11] nodule modeling method and data-driven using AAM [17] to model the nodules in terms of both shape and intensity method will briefly be described.

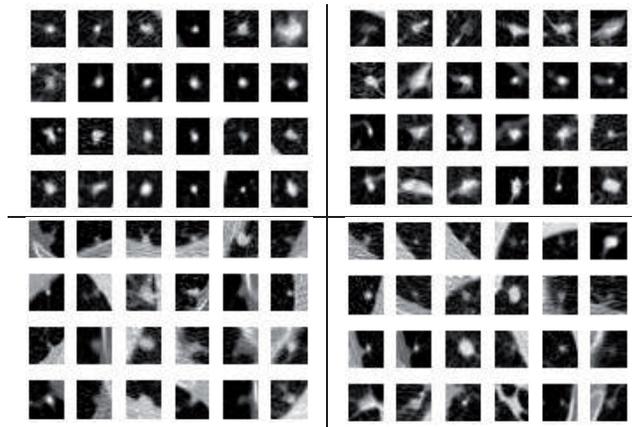
### 2.1 Pulmonary Nodule Definitions

In radiology, a pulmonary nodule is a mass in the lung usually spherical in shape; however it can be distorted by surrounding anatomical structures such as the pleural surface. This paper uses the classification of Kostis et al. [6], which groups nodules into four categories (Fig. 1): *well-circumscribed* where the nodule is located centrally in the lung without being connected to vasculature; *vascularized* where the nodule has significant connection(s) to the neighboring vessels while located centrally in the lung; *juxta-pleural* where a significant portion of the nodule is connected to the pleural surface; and *pleural tail* where the nodule is near the pleural surface, connected by a thin structure; in all of these types there is no limitations on size or distribution in the lung tissue. These definitions are used in our approach. These nodule types are characterized mainly by shape, and location with respect to the anatomy of the chest; the appearance of an individual nodule may not hold too much discrimination. The ELCAP [12] database provides the LDCT scans where the center location of the nodule was specified by radiologists. We constructed a database of nodules using a semi-automatic method of cropping and categorizing each nodule into one of the four types.

### 2.2 Lung Nodule Model Simulation: Level Sets Approach

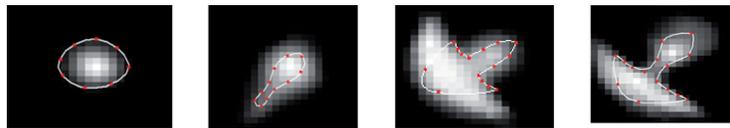
The goal of the modeling process is twofold: generate a template for each nodule type to be used in the detection process, and establish a procedure for simulation of nodules to be used for nodule classification. Our focus in this paper is on generating an overall mean template per type that best represents the characteristics of the real nodules. A database of nodules was constructed using the Early Lung Cancer Action Program (ELCAP) lung scans datasets [12]. As described in Farag, et al. 2009 [7][11], the probability density function of the Hounsfield (HU) vs. radial distance

distribution of the ELCAP dataset had an exponentially decaying form concentrated in distances of about 10 pixels from the centroid of the nodules (i.e., about 5mm); thus cropping boxes of size 21x21 pixels were used to obtain the nodule ensemble, given their location in the LDCT scans. These nodules were then classified into one of the four corresponding categories described in sec. 2.1, constructing a nodule database that contains variations in intensity distribution, shape/structural information and directional variability which the cropped regions, within the determined bounding-box, maintain. A sub-database of 96 nodules (24 nodules per type) is used in both the level sets approach and the AAM method to generate a mean nodule template for each type that depicts shape or shape and texture information of the nodules (Fig.1).



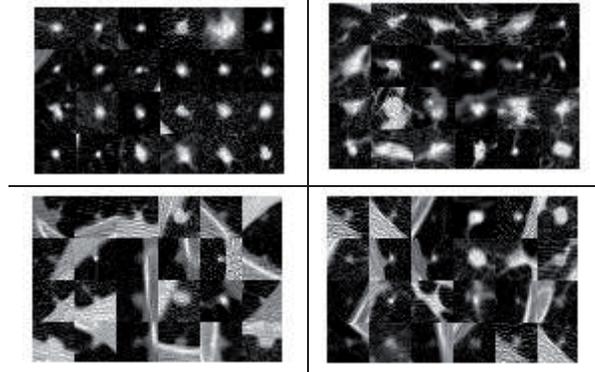
**Fig. 1.** An ensemble of 24 nodules from the well-circumscribed (upper left), vascular (upper right), juxta-pleural (lower left) and pleural-tail (lower right) nodule types.

The Procrustes registration-based AAM approach [15] for lung nodule modeling [17], required manual annotation by trained experts of the 96 nodules to employ Procrustes registration to obtain co-registered nodules. A combinational shape and texture AAM algorithm was used in [17] to generate a mean template that contained both shape and texture information of the lung nodules shown in Fig. 2.

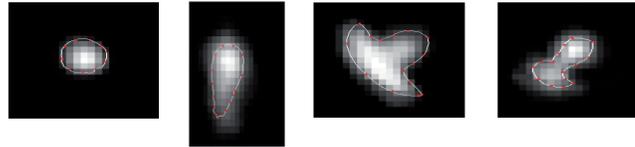


**Fig. 2.** Nodule models generated using the Procrustes based AAM method. From left to right: Well-circumscribed, vascular, juxta-pleural, and pleural-tail nodule types.

A second sub-database of 96 nodules, different from that in Fig. 1, was used to generate another set of four mean nodules using the AAM approach, the sub-database and mean nodules generated are depicted in figures 3 and 4, respectively.



**Fig. 3.** Second ensemble of 24 nodules from the well-circumscribed (upper left), vascular (upper right), juxta-pleural (lower left) and pleural-tail (lower right) nodule types.



**Fig. 4.** Nodule models generated using the Procrustes based AAM method. From left to right: Well-circumscribed, vascular, juxta-pleural, and pleural-tail nodule types.

The usage of variational Level sets (e.g., [18]) in this paper replaces the step of manual annotation performed in the AAM method by a semi-automatic approach that generates the contours of the lung nodules depicting the shape information. These contours are then co-registered using the Procrustes method; the contour boundary points are obtained and used for registration. The Level sets approach eliminates sources of errors that can arise with manual annotation since only placing the seed point or points in the region of the nodule centroid is manually performed. Also, the elasticity of this variational approach addresses the issue of shape variations that can arise and handles these changes accordingly. In this paper, we used the approach of Abdelmunim and Farag, 2007 ([18]) for rigid and elastic shape representation via Level sets. The mean shape templates generated from the contours is shown in Fig. 5. Given two shapes represented by the vector functions  $\Phi_1$  and  $\Phi_2$ , a transformation  $A$  with scales, rotation and translation is to be calculated to transform the first object to the second. The following dissimilarity measures the difference between the vector and the other scaled one:

$$r = SR\Phi_1(X) - \Phi_2(A) \quad (1)$$

The following energy is formulated as a sum of squared differences

$$E = \int_{\Omega} \delta_{\epsilon} r^T r d\Omega \quad (2)$$

where the delta is an indicator function with value 1 inside the shape and zero otherwise. The two shapes are aligned by minimizing the energy function using various approaches including the gradient descent method. The training shapes in our case the contours of the lung nodules are jointly registered with an evolving mean

shape to find the corresponding global transformations  $A_1, \dots, A_n$ . The dissimilarity measure is used as follows:

$$r_i = S_i \Phi_M - \Phi_i(A_i) \tag{3}$$

The energy function will be:

$$E(\Phi_M, \Phi_1, \dots, \Phi_n) = \sum_{i=1}^n \int_{\Omega} \delta_{\epsilon} r_i^T r_i d\Omega \tag{4}$$

The shape model is a function of the registered training shapes:

$$\Phi_p = \Phi_M + \sum_{i=1}^n w_i (\Phi_i - \Phi_M) \tag{5}$$

The shape parameters  $w$ , need also to minimize the energy function to process the registration.

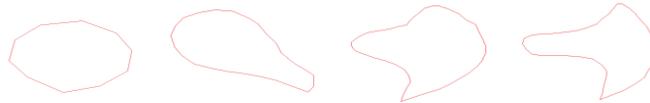
Once the mean shapes were generated estimation of texture information was required, thus using a synthetic approach described in [5][7] and [10] for parametric templates was performed. The probability density of the intensity for the nodules is the same as that in [7] since the ELCAP database is used in both.

The intensity of a nodule model was found to be estimated by the following equations [7][10]:

$$q(r) = q_{max} e^{-(r/\rho)^2}, \quad 0 \leq r \leq R \tag{6}$$

$$\rho = R (\ln(q_{max}) - \ln(q_{min}))^{-1/2} \tag{7}$$

where  $R$  is the radius of the circle interior to the bounding box containing the nodule model (mean shape). The parameters  $q_{min}$  and  $q_{max}$  are the lower and upper bounds of the intensity (Hounsfield Units) in the probability density function of each nodule type; which is estimated from the ensemble of nodules (e.g., [7][11]). Fig. 6 shows the templates generated based on the mean shape and the empirical forms of the intensity (Eq. 6 and 7).



**Fig. 5.** Average shape of the nodules in Fig. 1. From left to right: Well-circumscribed, vascular, juxta-pleural and pleural-tail.



**Fig. 6.** The nodule templates resulting from the intensity equations Eq. 6 and 7 and the mean shapes in Fig.4. From left to right: Well-circumscribed, vascular, juxta-pleural, and pleural-tail nodule types shape modeling process.

A set of parametric templates (e.g. circular and semi-circular) were generated using equations for circular and semi-circular shapes for various radii of empirical nodules and filled with texture information using equations 6 and 7. Template size

was the same 21x21 pixels bounding box area, similar to the AAM and level sets approaches. The same probability density of the nodule intensity observed for the nodule database for  $q_{\min}$  and  $q_{\max}$  were used once more and the templates were filled using the same technique implemented for the level sets approach. Circular or isotropic templates are defined in terms of radius and gray level distribution as a circular symmetric Gaussian function while the semi-circular template have an additional orientation parameter. Fig. 7 shows as set of isotropic and non-isotropic templates generated by this approach.



**Fig.7.** An ensemble of generated circular and semi-circular templates with various orientations (adopted from [7]).

### 2.3 Lung Nodule Modeling Summary

In this paper we examined three approaches for generating lung nodule models or templates. Of these methods two are non-parametric and data-driven, while the other approach was parametric as shown in Fig. 7. Parametric nodule models and the non-parametric nodule models, based on shape only, may be used in template matching in a binary form (after we segment the original lung images to generate a binary image) or in gray scale form applied straight to the original lung images (after we remove the non-lung tissues in the segmentation step preceding the detection). In this paper, as indicated above, the intensity information in the shape models (using level sets) as well as the parametric models were obtained using Eq. 6 and 7.

We should point out that the Procrustes approach was used in obtaining the mean shape (generated from level sets) and the mean shape and texture (generated from the AAM approach). Other methods may be used to carry out these co-registrations and may lend enhanced efficiencies with respect to orientations, shapes, etc. in the ensemble. The overall mean shapes of each nodule type from the two approaches showed extraordinary resemblance as seen in Fig 2 and Fig. 5.

In the next section we evaluate the performance using the nodule templates from the methods described above for the detection of candidate nodules.

## 3 Results

The Early Lung Cancer Action Program (ELCAP) public database [12] was used in this paper for nodule modeling, classification and detection. The database contains 50 sets of low-dose CT lung scans taken at a single breath-hold with slice thickness 1.25 mm. The locations of the 397 nodules are provided by the radiologists, where 39.12% are juxta-pleural nodules, 13.95% are vascularized nodules, 31.29% are well-circumscribed nodules and 15.65% are pleural-tail nodules. In the detection stage all slices containing nodules are used unlike the modeling stage where only 96 nodules of

the total 397 were used for generating the templates. Due to the mean template per nodule type is found, in the cases of the data-driven models, we use the slices from the modeling stage in the detection process as well since the templates are not biased towards any particular one nodule in Fig. 1. There are numerous methods for performing lung nodule detection in the literature; our goal for this paper is to depict the effectiveness of our nodule modeling, thus a generic implementation of template matching with the normalized cross-correlation (NCC) as the similarity measure is executed. The nodule model (template) is swept across the scan (2D slices or the 3D volume) in a raster fashion and a similarity measure is calculated between the intensity (or HU) of the template and the region of the CT data underneath, each template produces a binary image that represents candidate nodule locations, the four images are xored together to obtain one image that depict overall candidate nodule location. We use the widely known form of the NCC in the literature for the normalized cross-correlation of a template,  $t(x,y)$  with a sub-image  $f(x,y)$ :

$$NCC = \frac{1}{n-1} \sum_{x,y} \frac{(f(x,y) - \bar{f})(t(x,y) - \bar{t})}{\sigma_f \sigma_t}, \quad (8)$$

where  $n$  is the number of pixels in template  $t(x,y)$  and sub-image  $f(x,y)$  which are normalized by subtracting their means and dividing by their standard deviations. The probability density functions (pdf) of nodule and non-nodule pixels are computed using the normalized cross correlation coefficients resulting from templates with varying orientations.

The NCC behavior with the data-driven nodule models takes the same general shape as with the parametric nodules except the distribution function decays faster as we approach a value of 0.5. Setting a suitable threshold for the NCC is important as lower thresholds will increase detection rate but increases the false positives, and vice versa. Various methods can be used for an optimal threshold including modeling the normalized histogram as two classes, nodules and non-nodules and a Bayesian approach may be devised to select an optimal threshold. In this paper we set the threshold to a NCC of 0.5 for the sake of comparison between the three methods.

For nodule recognition (i.e., deciding the pathology of the nodule), features from the detected nodules need to be compared with pathological counterparts. This issue will not be considered in this paper as its significance really depends on the availability of pathological nodule database which is under construction by this research group. Also, the validation of pathology requires three human experts at least (two independent reader and a third to decide on opposing decisions).

The overall sensitivity and specificity was computed using equations 9 and 10:

$$Sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (9)$$

$$Specificity = \frac{True\ Negatives}{True\ Negatives + False\ positives} \quad (10)$$

True positive rate refers to the number of actual nodules that are detected as nodules while false negatives are the number of nodules that were not detected as nodules. Thus sensitivity depicts how well the detection was able to recognize nodules from other lung features using the desired designed templates (parametric and data-driven). True negatives are the number of nodules that are truly not nodules while false positives are the number of non-nodules that were detected as nodules.

Specificity rate is more subjective in its computation since it depends on how the true negatives and false positive rates are computed during detection. The specificity represents the negative rate that is correctly identified.

Table 1 depicts the overall sensitivity and specificity results of using the templates generated by the parametric method, the Level-sets algorithm and AAM method using the sub-database in Fig. 1 and Fig. 3, centered with respect to the x-axis (i.e., zero orientation). While Table 2 depicts the overall sensitivity and specificity results when the templates are rotated from  $0^\circ$  to  $360^\circ$  with step-size  $90^\circ$ . From Tables 1 and 2 several conclusions can be drawn, first the data-driven method using the Level-sets to generate template that depict overall shape only which are then filled, as described in section 2.2, yields comparable results to that of the parametric templates in terms of sensitivity and slightly higher specificity results. Second, rotation of the templates improved sensitivity in only the parametric and level-set based template approaches and overall specificity slightly decreased for all approaches. Third, template generated by the AAM approach provides better results than the parametric, level sets method and the AAM mean templates depicted in Fig 4. Overall, the AAM algorithm using either set of mean templates generated from sub-database 1 or 3 yield better results in terms of both sensitivity and specificity. Thus, data-driven models are more robust and an enhanced method of lung nodule modeling over the use of parametric templates, since the actual data is used in modeling and generating mean templates to represent each nodule type. Also, shape and texture based approaches give a more accurate and precise representation to the true nodule that provides improved detection results. So, shape information alone does not suffice, both shape and texture information is required.

**Table 1:** Overall sensitivity and specificity of level sets, parametric, AAM using dataset 1 and AAM using dataset 2 without accounting for template orientation.

Algorithm	Sensitivity	Specificity
Parametric Approach with template radius 10 and single orientation for semi-circular template	72.16%	97.12%
Level Sets Approach using nodule contours from dataset 1 and no orientation	72.16%	98.11%
AAM Approach using dataset 1 and no orientation	85.22%	97.81%
AAM Approach using dataset 2 and no orientation	83.51%	98.36%

**Table 2:** Overall sensitivity and specificity of level sets, parametric, AAM using dataset 1 and AAM using dataset 2 templates averaging a number of orientations for the templates.

Algorithm	Sensitivity	Specificity
Parametric Approach with template radius 10 and orientation $0^\circ$ - $360^\circ$ with step-size $90^\circ$ for semi-circular template	78.01%	96.41%
Level Sets Approach using nodule contours from dataset 1 and orientation $0^\circ$ - $360^\circ$ with step-size $90^\circ$	76.98%	97.63%
AAM Approach using dataset 1 and orientation $0^\circ$ - $360^\circ$ with step-size $90^\circ$	86.94%	96.51%
AAM Approach using dataset 2 and orientation $0^\circ$ - $360^\circ$ with step-size $90^\circ$	83.51%	97.40%

Further studies were conducted using the AAM based approach on both sub-datasets used for template modeling. The number of annotation points necessary for proper registration was found to be a function of how many was necessary to depict the main discriminatory shape information and withholds substantial texture information commonly found in each nodule for that particular type (i.e. the 24 nodules used per type). Also, 24 nodules per type were used in the modeling of each nodule type but if fewer nodules were used what will be the effect on the generated mean nodule templates formulated and used in the detection process was examined.

Table 3 depicts the results obtained when the first 16 and 8 nodules from each of the sub-databases are used. From the table it is seen that overall sensitivity and specificity using half or one-third of the nodules in figure 1 results in overall similar sensitivity and specificity results while in the case of using the sub-database in figure 3 results in reduced sensitivity as the number of nodules used for modeling decrease. Overall we found that depending on how well the lung nodules are annotated and which nodules are represented in the sub-database effected the generation of the mean templates per type; i.e. if nodules used in the modeling depicted a majority of the nodules in the larger database and annotated well then the overall mean templates generated yielded improved sensitivity and specificity results, if the nodules were not adequately annotated the mean templates generated not always gave improvements in sensitivity and/or specificity likewise if the nodules in the modeling database did not depict a vast majority of those in the original database then detection rates reflected that.

The results are expected to be further enhanced using larger ensemble sizes than the 24 per nodule types which we used in our experiments. Likewise, involvement of several radiologists to create the ensemble may also lead to further improvements.

**Table 3:** Overall sensitivity and specificity of AAM approach using 8 and 16 nodules from datasets 1 and 2 for mean template modeling, respectively.

Algorithm	Sensitivity	Specificity
AAM Approach using 16 nodules from dataset 1 for modeling and no orientation	85.57%	97.84%
AAM Approach using 8 nodules from dataset 1 for modeling and no orientation	84.88%	97.99%
AAM Approach using 16 nodules from dataset 2 for modeling and no orientation	83.16%	98.42%
AAM Approach using 8 nodules from dataset 2 for modeling and no orientation	77.32%	98.57%

## 4 Conclusions and Extensions

In this paper, a data-driven approach using level sets was devised to model and simulate typical lung nodules. The modeling procedure of parametric and non-parametric template models was examined and used for nodule detection. The effect of template shape and texture on detection of different nodules types was studied.

From our extensive experimentation we can conclude that the data-driven AAM algorithm for lung nodule modeling yielded an overall higher sensitivity and specificity rate, yet, the Level sets approach showed instances of improvement for specificity and/or sensitivity over the usage of parametric templates. In the parametric case where we tested on all radii sizes between 1 and 20 pixels the sensitivity was higher but the specificity in comparison to the data driven nodule templates was still lower.

This paper has shown that approaches where both shape and texture information is simultaneously computed for modeling is more robust and an accurate approach than relying only on shape information for precise nodule descriptions. Current efforts are directed towards constructing and testing the data-driven modeling approach on a large clinical database and extending this work into the 3D space. The nodule databases will be made available to the research community in order to measure the enhancements made in the detection as well as recognition/classifications based on a common standard. Also, evaluating these templates to classify the nodules into designated pathologies (e.g. benign and malignant) is a key and important step that will be examined. Other algorithms of constructing shape and texture based models will be explored.

Therefore, we have established a systematic approach to model and simulate the lung nodules in LDCT scans which is applicable to any data protocol, and any nodule definition. This contribution is very crucial and may be the building block for all work on CAD systems applied to lung nodules; indeed, it may be also used for all similar approaches that generate templates to be detected in data of various types and formats. The main power of this approach is the fact that it is *data-driven*; hence, various attributes of the data may be incorporated in the template design.

**Acknowledgements:** Dr. Asem Ali and Ham Rara, members of the CVIP Lab, provided various insights and assistance in this research. The first author has been supported by a fellowship from the United States National Space and Aeronautics Agency, NASA. This research has been supported by the Kentucky Lung Cancer Program.

## References

1. United States National Institute of Health [www.nih.gov](http://www.nih.gov)
2. B. Zaho, G. Gamsu, M.S. Ginsberg, L. Jiang, L.H. Schwartz, "Automatic Detection of small lung nodules on CT utilizing a local density maximum algorithm," *Journal of Applied Clinical Medical Physics* 4 (2003).
3. Armato, S. G. 3rd, Giger, M. L., Moran C. J., Blackburn, J. T., Doi, K., MacMahon H.: Computerized detection of pulmonary nodules on CT scans. *Radio Graphics* 19 pp.1303--1311 (1999).
4. S. Hu, E. A. Hoffman and J. M. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images," *IEEE Transactions on Medical Imaging*, Vol. 20, pp. 490–498, 2001.

5. Y. Lee, T. Hara, H. Fujita, S. Itoh and T. Ishigaki, "Automated Detection of Pulmonary Nodules in Helical CT Images Based on an Improved Template-Matching Technique," *IEEE Transactions on Medical Imaging*, Vol. 20, 2001.
6. 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, 2003.
7. Amal A. Farag, S.Y. Elhabian, S.A. Elshazly and A.A. Farag, "Quantification of Nodule Detection in Chest CT: A Clinical Investigation Based on the ELCAP Study," *Proc. of International Workshop on Pulmonary Image Processing in conjunction with MICCAI-09*, Sept. 2009, pp. 149-160.
8. I. Sluimer, A. Schilham, M. Prokop, and B. van Ginneken, "Computer Analysis of Computed Tomography Scans of the Lung: A Survey," *IEEE Transactions on Medical Imaging*, vol. 25, No. 4, pp. 385–405, April, 2006.
9. B. Ginneken, B. Romeny and M. Viergever, "Computer-Aided Diagnosis in Chest Radiography: A Survey," *IEEE Transactions on Medical Imaging*, Vol. 20, 2001.
10. A. Farag, A. El-Baz, G. L. Gimel'farb, R. Falk, M. Abou El-Ghar, T. Eldiasty, S. Elshazly, "Appearance Models for Robust Segmentation of Pulmonary Nodules in 3D LDCT Chest Images," *Proc. of International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'06)*, Copenhagen, Denmark, October 1-6, 2006, pp. 662-670
11. Amal Farag, Lung Nodule Modeling and Detection for Computerized Image Analysis of LDCT Imaging of the Chest, Master of Engineering Thesis, University of Louisville, April 2009.
12. ELCAP public lung image database - [www.via.cornell.edu/databases/lungdb.html](http://www.via.cornell.edu/databases/lungdb.html)
13. H. Takizawa, S. Yamamoto, T. Matsumoto, Y. Tateno, T. Iinuma and M. Matsumoto, Recognition of lung nodules from X-ray CT images using 3d mrf models. In: *Proceeding of International Congress of Computer Assisted Radiology and Surgery (2001)*, pp. 605–614
14. Mendonça, P.R.S., Bhotika, R., Zhao, F., Miller, J.V.: "Lung nodule detection via Bayesian voxel labeling," *Int. Conf. on Information Proc. and Medical Imaging. (2007)* 134–145
15. Edwards G.J. Taylor C.J. Cootes, T.F. Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6), 2001.
16. I. Matthews, and S. Baker, "Active Appearance Models Revisited," *Int. J. of Computer Vision*, pp. 135-164, 2004.
17. Amal Farag, James Graham, Aly Farag and Robert Falk, "Lung Nodule Modeling – A Data-Driven Approach," *5th International Symposium on Visual Computing (ISVC'09)*, Nov. 30 – Dec. 2, 2009, Las Vegas, Nevada, USA.
18. H. Abdelmunim and A. A. Farag, "Shape Representation and Registration using Vector Distance Functions," *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07)*, Minneapolis, MN, June 18-23, 2007.