DEFORMABLE MODELS FOR RANDOM SMALL-SIZE OBJECTS: CASE OF LUNG NODULES IN CT TOMOGRAPHY

Amal A. Farag[1], James H. Graham[2] and Aly A. Farag[2]

[1]Imaging Biomarkers and Computer-Aided Diagnosis Laboratory, National Institutes of Health, Bethesda, MD
[2]Computer Vision and Image Processing Laboratory, University of Louisville, Louisville KY
E-mail: Amal.aly1@gmail.com, james.graham@louisville.edu, aly.farag@louisville.edu

ABSTRACT

Deformable models are common in image modeling and analysis. Random objects provide major challenges as shapes and appearances are hard to quantify; hence, formulation of deformable models are much harder to construct and validate. In this paper, we examine the effect of randomness on building the shape and appearance models for small-size lung nodules ($\leq 1cm$) *which appear in computed tomography (CT) of the human chest. We devise an approach for annotation, which lends a standard mechanism for building traditional active appearance (AAM), active shape (ASM) and active tensor models (ATM). We illustrate the effectiveness of AAM for nodule detection.*

1. INTRODUCTION

Automatic approaches for image analysis require precise quantification of object attributes such as shape and texture. These concepts have precise definitions, but their descriptors vary so much from one application to another. A shape is defined to be the information attributed to an object that is invariant to scale, origin and orientation [1]. A texture may be defined as the prevalence pattern of the interior of an object [2]. Geometric descriptors identify "features" that are "unique" about an object. Shape, texture and geometric descriptors are major concepts in this work; they will be defined and used in the context of modeling small size objects under uncertainties [3].

1.1 Lung Nodules in Low Dose CT

Figure 1 shows examples of small size nodules ($\leq 1cm$) from four categories [4][5]. The upper and lower rows show zoomed images of these nodules. Notice the ambiguities associated with shape definition, location in the lung tissues, and lack of crisp discriminatory features. Modeling aims at representing the objects with mathematical formulation that captures their characteristics such as shape, texture and other salient features. The histogram of the object's image provides some information about its texture – the modes of the histogram indicate the complexity of the texture of the object. Another difficulty of small objects lies with inabilities of exact boundary definition. For example, radiologists may differ in outlining the lung nodules spatial support as shown in Figure 2. Difference in manual annotation is common of small objects that have no well-defined description. This adds another dimension of difficulty for automatic approaches, as they are supposed to provide outputs that mimic human experts. In other words, human experts differ among themselves, how would they judge a computer output? Validation of automatic approaches for lung nodule detection, segmentation and classification - using only the visible information in an image - is much more difficult than that of automatic face recognition, for example.

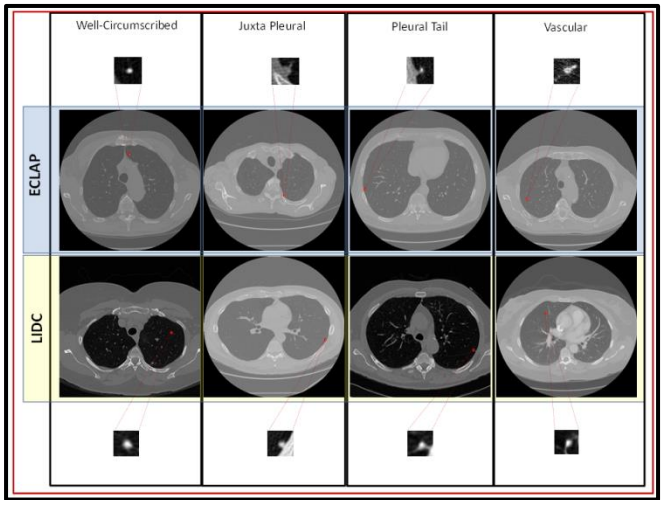

Figure 1 - *Examples of lung nodules of size below 10 mm from two clinical studies. The upper and lower rows show zoomed pictures of the nodules.*

Figure 3 shows histograms of the average intensity measured in Hounsfield Units (HU) of manually cropped nodules from the ELCAP and LIDC screening studies. The histograms are distinctly bimodal; a binary classifier may be used for separating the nodules and non-nodules regions. The decision boundary may be selected by various techniques, including fitting one-dimensional Gaussian density for the nodule and non-nodule regions and using the expectation-maximization approach (EM) to estimate the parameters. Unfortunately, this approach has been shown

not to perform well due to the uncertainties associated with the physical nodules as previously described (e.g., [6]). Most approaches of lung nodule modeling do not consider shape or texture information; some approaches utilize features that rely heavily on the dataset at hand [13], thus, comparison of the same approach using different data is difficult. In [6], for example, a five-step system for modeling of small lung nodules was introduced: i) CT Acquisition and Enhancement; ii) Parametric Modeling; iii) Nodule Detection; iv) Nodule Segmentation; and v) Categorization (Classification). By constructing a *front-end* system of image analysis (CAD system) for lung nodule screening, all of these steps must be considered. The focus is to examine the "data-driven" modeling approach.

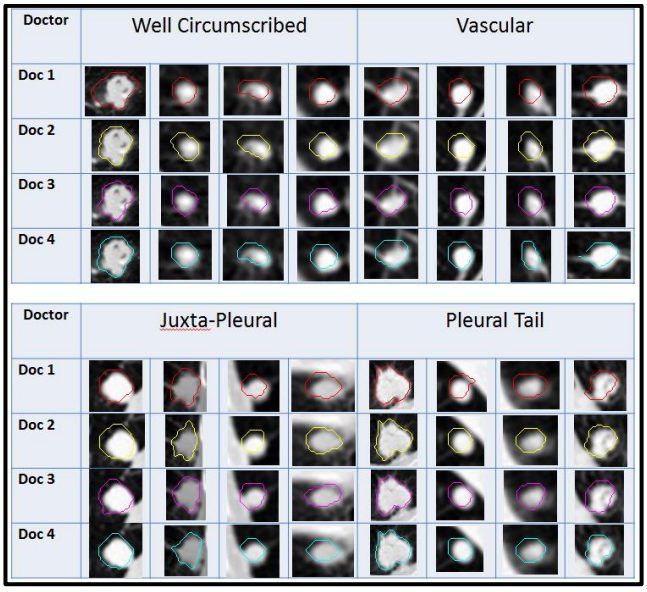

Figure 2 – *Manual annotation of the main portion of the spatial support of lung nodules by four radiologists.*

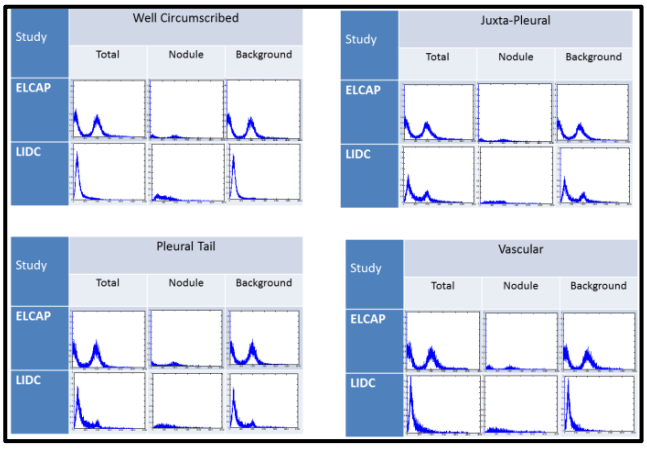

Figure 3 – The intensity (HU) histograms of the manually cropped nodules from the ELCAP and LIDC screening studies.

The theoretical development in this paper falls under the modern approaches of shape and appearance modeling. These models assume the availability of an ensemble of objects annotated by experts – the ensemble includes variations in the imaging conditions and objects attributes to enable building a meaningful statistical database. Active shape models (ASM) and active appearance models (AAM) have been powerful tools of statistical analysis of objects (e.g., [8][9]). The main contribution of the paper is to validate the *data-driven* approach for lung nodule modeling and analysis, developed by the authors (e.g., [6] [7]), on larger size nodule databases, and to examine the issue of auto-nodule cropping and annotation, which is essential for successful building of active appearance models (AAM).

2. LUNG NODULE MODELING

Real world objects may take various forms of details, and may be linear, planar or three-dimensional. In [3], Dryden and Marida, define anatomical landmarks as points assigned by an expert that corresponds between organisms in some biologically meaningful way; mathematical landmarks as points located on an object according to some mathematical or geometrical property, i.e. high curvature or an extremum point; and pseudo-landmarks as constructed points on an object either on the outline or between landmarks. Figure 4 is a sample of small-size nodules smaller than $1\,cm$ in diameter from the LIDC [5] clinical study, showing the variations that can be captured by shape and appearance models. From a computer vision prospective, AAM and ASM modeling have been used with great successes in objects having distinct landmarks (e.g., Cootes et al. [8] and Mathews and Baker [9]). A shape is considered to be a set of n −vertices $\boldsymbol{x} \in R^k$; for the two-dimensional case:

$$\boldsymbol{x} = [x_1; x_2; \dots ; x_n; y_1; y_2; \dots ; y_n]^T \qquad (1)$$

The shape ensemble (realizations of the *shape process* of a certain object) is to be adjusted (aligned) on the same reference to enable filtering of scale, orientation and translation among the ensemble, per the shape definition. This alignment generates the so-called *shape space*, which is the set of all possible shapes of the object in question. To align the shapes in an ensemble, various procedures may be used. The *Procrustes* procedure is common for *rigid* shape alignments. The alignment process removes the redundancies of *scale*, *translation* and *rotation* using a similarity measure that provides the minimum *Procrustes distance*. Suppose an ensemble of shapes is available with *one-to-one* point (feature) correspondence is provided. The *Procrustes distance* between two shapes $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ is the sum of squared distance (SSD)

$$P_d^2 = \sum_{j=1}^{n} \left(x_{j_1} - x_{j_2}\right)^2 + \left(y_{j_1} - y_{j_2}\right)^2 \qquad (2)$$

Annotated data of an ensemble of shapes of a certain object carries redundancies due to imprecise definitions of landmarks and due to errors in the annotations. *Principal Component Analysis (PCA)* may be used for reducing these redundancies. In PCA, the original shape vector $\mathbf{x}$ is linearly transformed by a mapping such that $\mathbf{z} = \mathbf{M}\,\mathbf{x}$ has

less correlated and highly separable features. The mapping **M** is derived for an ensemble of N shapes as follows**:**

$$\bar{\mathrm{x}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i \;;\; \Sigma_{\mathbf{x}} = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i - \bar{\mathrm{x}})(\mathbf{x}_i - \bar{\mathrm{x}})^T \qquad (3)$$

are the mean and covariance of **x.**Therefore, the mean and covariance of **z** would be**:**

$$\bar{\mathrm{z}} = \frac{1}{N}\sum_{j=1}^{N}\mathbf{z}_j \qquad (4a)$$

$$\Sigma_{\mathbf{z}} = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathrm{z}})(\mathbf{z}_i - \bar{\mathrm{z}})^T = \mathbf{M}\Sigma_{\mathbf{x}}\mathbf{M}^T \qquad (4b)$$

If the linear transformation **M** is chosen to be orthogonal; i.e., $\mathbf{M^{-1}} = \mathbf{M^T}$, and selecting it as the eigenvectors of the symmetric matrix $\Sigma_{\mathbf{x}}$, this would make $\Sigma_{\mathbf{z}}$ to be a diagonal matrix of the eigenvalues of $\Sigma_{\mathbf{x}}$. The eigenvectors corresponding to the small eigenvalues can be eliminated, which provides the desired reduction. Therefore, **x** may be expressed as:

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}\,\mathbf{b} \qquad (5)$$

where $\mathbf{P} = (\boldsymbol{p_1}|\boldsymbol{p_2}|\ldots|\boldsymbol{p_m})$ matrix of m largest eigen vectors of $\Sigma_{\mathbf{x}}$ and $\mathbf{b} = \mathbf{P^T}(\mathbf{x} - \bar{\mathbf{x}})$ is an $m \times 1$ vector. Equation (5) is the statistical shape model, which is derived using PCA. By varying the elements of **b** one can vary the synthesized shape **x** in Eq. (5). The variance of the i^{th} parameter $b_i \in \mathbf{b}$ can be shown across the training set to be equal to the eigenvalue λ_i [8].

2.3 Annotation of Random Objects

In order to construct the active appearance or active tensor models, we need an annotated ensemble of objects. In case of random objects, the annotation process becomes extremely difficult; it takes yet another level of difficulty with small-size. Yet, the major goal of this work is to address such objects, specifically, small size lung nodules, which are used for early detection screening of possible lung cancer. We used the fuzzy description of lung nodules from Kostis et al. [10] to devise a feature definition approach for four categories of nodules; *well-circumscribed*, *vascularized, juxta-pleural* and *pleural-tail* nodules. Figure 4 illustrates the landmarks that correspond to the clinical definition of these four nodule categories.

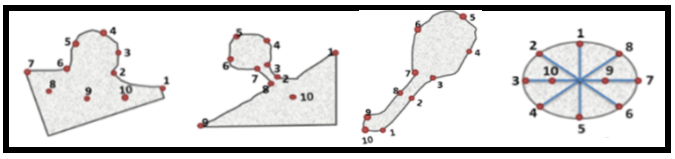

Figure 4 – *Definition of Control points (landmarks) for nodules. Right-to-left: juxta-pleural, pleural tail, vascular, and well-circumscribed nodule models.*

Using these definitions, we created a manual approach to annotate the nodules. First, we take the experts' annotation, zoom it and manually register it to a template defining the nodule type/category, and then we select the control points on the actual nodule using the help of the template. This annotation enabled creation of active appearance models, which mimics largely the physical characteristics of lung nodules that cannot be modeled otherwise. Figure 5 shows examples for the nodule models generated by ensembles from the ELCAP and LIDC clinical lung screening studies**.** The average nodules (shown in Figure 5) capture the main features of real nodules. Incorporation of other basis has been studied in Farag et al., 2012 [7].

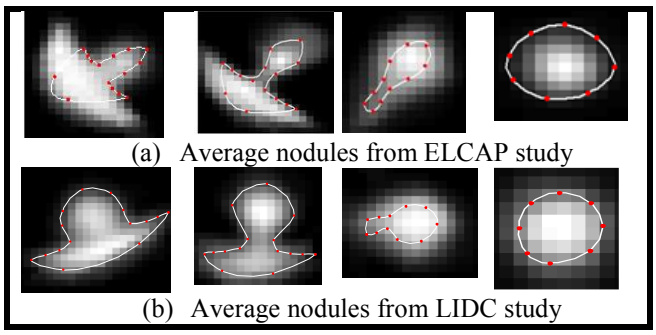

Figure 5 – *AAM Models for lung nodules from clinical CT scans. Right-to-left: juxta-pleural, pleural tail, vascular, and well-circumscribed nodule models.*

3. MODEL-BASED NODULE DETECTION

The above modeling approach has provided tremendous promise in three subsequent steps of lung nodule analysis: detection, segmentation, and categorization. Due to space limitations, we only consider lung nodule detection using the AAM nodule models. Further, we use only a basic detection approach that is based on template matching with normalized cross-correlation (NCC) as similarity measure. Other measures have been examined in our related work (e.g., [7]). We report the detection performance by constructing the ROC of both the ELCAP and LIDC clinical studies for the first time. We chose to limit the ensemble size for modeling to be 24 *per nodule type* for the two studies, to provide a comparison with our earlier work [6].
The ROCs are built to show the overall *sensitivity* and $1 - specificity$ of the detection process. The textures of the parametric nodules were generated by the analytical formulation in our earlier work (e.g., [6]).

3.1 Clinical Evaluation on ELCAP database

The ELCAP database [4] contains 397 nodules, 291 identified and categorized nodules are used in the detection process. Results using only the average (mean) template models generated from the AAM approach is examined against parametric nodule models, (i.e. circular and semi-circular) of radius 10, templates in this first set of experiments. Figure 6 shows the ROC; illustrating a superb performance of the AMM-based approach.

3.2 Clinical Evaluation on LIDC database

The Lung Imaging Data Consortium (LIDC) [5] contains 1018 helical thoracic CT scans from 1010 different patients. We used ensembles of 24 nodules *per nodule type* to design the nodule models (templates) and the rest to test the detection performance. Figure 7 shows the ROC of $1 -$

specificity vs. *sensitivity*. The results once again showed the AAM-models provided better detection results over the parametric models.

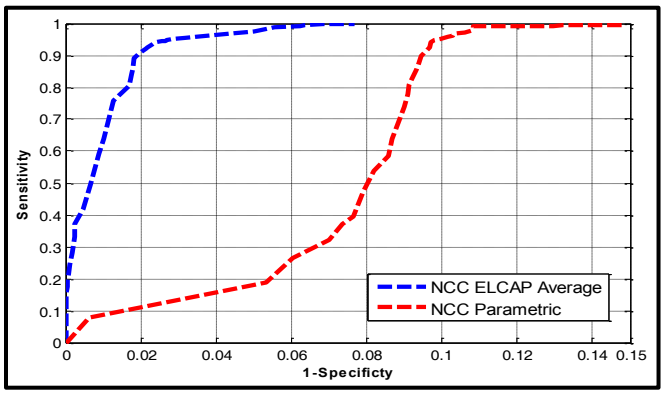

Figure 6 - *ROC curves for template matching detection on the ELCAP database using ELCAP models versus the circular and semi-circular models (i.e.parametric).*

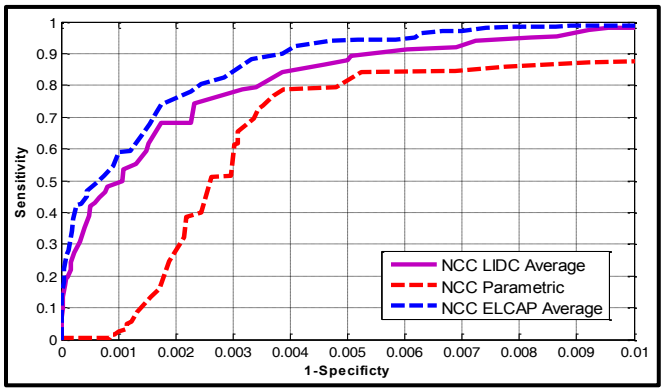

Figure 7 - *ROC curves for template matching detection on the LIDC database using LIDC versus parametric and ELCAP models.*

4. DISCUSSION

In generating the above ROCs, we used the mean in the AAM models as the nodule template (note: in Farag et al., 2012 [7] we used other *eigen-nodules* besides the mean). From Figures 6 and 7, the templates from the ELCAP ensemble showed better performance than those from the LIDC ensemble. This could be a result of the wide range of variations in texture information found in the LIDC database, which affect the appearance of the resulting nodule model (template). We used 24 nodules, per nodule type, in both ELCAP and LIDC.

We note that in the ELCAP database, the data acquisition protocol was the same throughout; very low resolution. That was reflected in the AAM model, showing a texture that is relatively more homogenous than that in the LIDC case, which uses data from various imaging centers and various imaging scanners, with somewhat variable range of Hounsfield Units (HU).

In general, if we include more nodules in the design, we expect a better appearance modeling; the LIDC database allows such choice.

5. CONCLUSIONS

The paper dealt with modeling of small-size lung nodules using two clinical studies, the ELCAP and LIDC. We discussed the process of nodule annotation and the steps to create AAM nodule models. These models resemble the real nodules, thus using them as templates for nodule detection is more logical than the non-realistic parametric models. These types of models add two additional distinctions over the parametric approaches; it can automate the processes of nodule segmentation and categorization. Our preliminary results show advantage on both issues – this is a natural extension of the work reported here (e.g., [11]).

We plan to carry out the tensor modeling approach, and study the effect of nodule ensemble size. From the algorithmic point of view, an *adaboost* strategy for carrying out the detection may lend speed advantage over the typical cross-correlation implementation used in this paper. In addition, with proper lung nodule models, an approach similar to Viola-Jones Face Detector may be possible [12].

6. REFERENCES

[1] D.G. Kendall, "Shape Manifolds, Procrustean Metrics, and Complex Projective Spaces". Bulletin of the London Mathematical Society 16 (2): 81–121, 1984.

[2] B. Julesz, et al. "Inability of humans to discriminate between visual textures that agree in second-order statistics--revisited. Perception," 2, 391-405, 1973.

[3] I. L. Dryden and K. V. Mardia, Statistical Shape Analysis, John Wiley & Sons, 1998.

[4] The United States Department of Health and Human Services (NIH); www.nih.gov

[5] G. Armato, et al., "Lung image database consortium: developing a resource for the medical imaging research community," Radiology, vol. 232, no. 3, pp. 739–748, 2004.

[6] Amal A. Farag, James Graham, Aly A. Farag, Salwa Elshazly and Robert Falk, "Parametric and Non-Parametric Nodule Models: Design and Evaluation," Workshop on Pulmonary Image Processing in conjunction with MICCAI-'10, Beijing, September, pp. 151-162, 2010.

[7] Amal Farag, H. Abdelmunim, J. Graham, Aly Farag, C. Carter, Salwa Elshazly, S. El-Mogy, M.S. El-Mogy, and R. Falk, "An AAM-Based Approach for Lung Nodule Detection from low dose CT (LDCT) Scans," ISBI'12 – Barcelona, Spain, pp. 1040 – 1043, 2012.

[8] T. F. Cootes and C. J. Taylor, "Anatomical Statistical Models and their role in feature extraction," The British Journal of Radiology, Vol. 77, pp. S133-S139, 2004.

[9] I. Mathews and S. Baker, Active Appearance Models Revisited, International Journal of Computer Vision, Vol. 60, No. 2, November, 2004, pp. 135 - 164.

[10] W. J. Kostis, et al., "Small pulmonary nodules: reproducibility of three-dimensional volumetric measurement and estimation of time to follow-up," Radiology, Vol. 231, pp. 446-52, 2004.

[11] Amal A. Farag, "A Variational Approach for Small-Size Lung Nodule Segmentation," ISBI'13, San Francisco, April 2013.

[12] P. Viola and M. Jones, "Robust Real-Time Face Detection," International Journal of Computer Vision 57(2), 137–154, 2004.

[13] Namin, S.T., et.al. "Automated detection and Classification of pulmonary nodules in 3D thoracic CT images." IEEE SMC, 2010.