5-2012

Modeling small objects under uncertainties : novel algorithms and applications.

Amal A. Farag
University of Louisville

MODELING SMALL OBJECTS UNDER UNCERTAINTIES:

NOVEL ALGORITHMS AND APPLICATIONS

By

Amal A. Farag
B.S., University of Louisville, 2008
M. Eng., University of Louisville, 2009

A Dissertation
Submitted to the Faculty of the
University of Louisville
J.B. Speed School of Engineering
As Partial Fulfillment of the Requirements
For the Professional Degree

DOCTOR OF PHILOSOPHY

Department of Electrical and Computer Engineering

May 2012

MODELING SMALL OBJECTS UNDER UNCERTAINTIES:

NOVEL ALGORITHMS AND APPLICATIONS

By

Amal A. Farag
M. Eng. University of Louisville, 2009

A Dissertation Approved on

April 23ʳᵈ, 2012

By the Following Reading and Examination Committee:

_____

James H. Graham, Ph.D., Dissertation Director

_____

John Naber, Ph.D.

_____

Roger Bradshaw, Ph.D.

_____

Robert Falk, M.D.

_____

Aly A. Farag, Ph.D.

ii

DEDICATION

This dissertation is dedicated to my wonderful parents Aly and Salwa, and my three younger brothers Ahmed, Mostafa and Ibrahim. They have been my support through everything. Thank you. I Love you all dearly.

ACKNOWLEDGMENTS

ABSTRACT

MODELING SMALL OBJECTS UNDER UNCERTAINTIES:

NOVEL ALGORTHMS AND APPLICATIONS

Amal A. Farag

April 23rd, 2012

Active Shape Models (ASM), Active Appearance Models (AAM) and Active Tensor
Models (ATM) are common approaches to model elastic (deformable) objects. These
models require an ensemble of shapes and textures, annotated by human experts, in order
identify the model order and parameters. A candidate object may be represented by a
weighted sum of basis generated by an optimization process. These methods have been
very effective for modeling deformable objects in biomedical imaging, biometrics,
computer vision and graphics. They have been tried mainly on objects with known
features that are amenable to manual (expert) annotation. They have not been examined
on objects with severe ambiguities to be uniquely characterized by experts.

This dissertation presents a unified approach for modeling, detecting, segmenting and
categorizing small objects under uncertainty, with focus on lung nodules that may appear
in low dose CT (LDCT) scans of the human chest. The AAM, ASM and the ATM
approaches are used for the first time on this application. A new formulation to object
detection by template matching, as an energy optimization, is introduced.

Nine similarity measures of matching have been quantitatively evaluated for detecting nodules less than 1 *cm* in diameter.

Statistical methods that combine intensity, shape and spatial interaction are examined for segmentation of small size objects. Extensions of the intensity model using the linear combination of Gaussians (LCG) approach are introduced, in order to estimate the number of modes in the LCG equation. The classical maximum a posteriori (MAP) segmentation approach has been adapted to handle segmentation of small size lung nodules that are randomly located in the lung tissue. A novel empirical approach has been devised to simultaneously detect and segment the lung nodules in LDCT scans. The level sets methods approach was also applied for lung nodule segmentation. A new formulation for the energy function controlling the level set propagation has been introduced taking into account the specific properties of the nodules.

Finally, a novel approach for classification of the segmented nodules into categories has been introduced. Geometric object descriptors such as the SIFT, ASIFT, SURF and LBP have been used for feature extraction and matching of small size lung nodules; the LBP has been found to be the most robust. Categorization implies classification of detected and segmented objects into classes or types. The object descriptors have been deployed in the detection step for false positive reduction, and in the categorization stage to assign a class and type for the nodules. The AAM/ASM/ATM models have been used for the categorization stage.   The front-end processes of lung nodule modeling, detection, segmentation and classification/categorization are model-based and data-driven. This dissertation is the first attempt in the literature at creating an entirely model-based approach for lung nodule analysis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Modeling entails a mathematical description to enable object recognition, for the purpose of identification and decision making. In lung screening using Computed Tomography (CT), the purpose is to detect and classify lung nodules and associate a specific pathology to them, if possible. In landfill surveillance, the purpose is to track illegal dumping and abuses of public sites such as forests, parks and highways. In deep space imaging, the purpose may be tracking space debris, planetary exploration and the status of communication satellites. In *Facial Recognition at a Distance* (FRAD), the purpose may include detecting and recognizing a particular individual from a faraway distance. As technology improves, the capability of sensors to obtain better images, i.e. better resolution, contrast, and tolerance to motion/noise that affect the image quality, enhances researchers' capabilities. In aerial photography and remote sensing, sensors involve more than one frequency of the spectrum, in order to provide better specificity of targets and enhancement of contrast between image contents, and better reaction to atmospheric conditions.

In biomedical imaging, the latest generation CT scanners provide faster image acquisition and better image quality. For example, it enables a full body scan in less than one breath. MRI scanners provide both structural and functional imaging at impressive specificity and sensitivity with no risk of ionizing radiations.

Ultrasound imaging is improving in terms of resolution and specificity and has no risk to the patients due to exposure. Optical imaging at a distance is possible using telescopes and various magnifying lenses. A target may be tracked and imaged at hundreds of meters with impressive image quality. Despite these impressive advances, one problem persists: how to detect and recognize small objects - that occupy only few pixels in the spatial support of an image? These scenarios are abundant and provide a real challenge to computational image analysis, which, in biomedical imaging, attempts to mimic the human expert.

A typical image analysis approach follows basic steps of image processing (or conditioning) that removes spurious details and motion artifacts. It may also provide extrapolation to resolve bias fields in MRI and partial volume effects in MRI and CT. The basic steps of image analysis are: object detection, to localize candidate targets or objects in the image; segmentation, to isolate the detected objects from the background; and classification of segmented objects into identifiable classes or categories. In these steps, attributes of shape, texture and a-priori information are sought in order to reach the desired conclusion. Small size objects provide a challenge in defining these attributes, and many image analysis algorithms fail due to size limitations, which limit precise description of the objects. Small size objects may also suffer from the uncertainties associated with the environment, such as occlusion, contrast and close similarity to background objects or constituents. These circumstances are indeed challenging for the human expert as well.

In the past several decades, image analysis algorithms have matured and reached a level of rigor that enables quantitative evaluation of performance and figures of merits

may be drawn on the confidence of the decisions made based on these algorithms. Unfortunately, small size targets which occupy a tiny portion of the spatial support still defy this progress and are still illusive to robust and automatic evaluation. Another difficulty with these scenarios is that algorithms developed for one problem does not hold to another; indeed, many image analysis tools are problem specific, which frustrates the point of view of mathematical modeling.

The common tradition is to try various approaches and adjust them to the specifics of the problem at hand and an integration of techniques may hold the best promise for a solution. This provides yet another challenge of how to compare with other researchers? The image analysis literature is aware of these difficulties; it is now common to have the so-called "[*specific problem*] challenge" in the workshops and scientific meetings, where the organizers of the "challenge" provide the imaging scenario and protocol and invite researchers to apply their algorithms and document their progress. Indeed, standard data sets from various imaging modalities are now available to test new image analysis algorithms.

Yet, there are still problems (too many unfortunately) that are not specific at the imaging protocol level, and differ at the definition levels. For example, in detecting nodules in the human lungs from low dose CT (LDCT) scans, there is a lack of precise definition of LDCT protocol, and human experts differ in the definition of the nodules, especially when they are less than few millimeters and are camouflaged with the tissue details. In image discoveries, it is often the case that "objects" have not been seen before; hence, defining the characteristics of the "unseen" is an unattainable task! This dissertation does not consider "discovery" since reasonable statistics can be obtained

about the objects from clinical studies, and phantom may be constructed based on plausible resemblance to the real world. In other words, uncertainty, in this case, does not mean complete lack or absence of information. The source of "uncertainties" in this work is limited to levels of human abilities and common imaging sensors; i.e., discovering "unknown" objects is not of concern, rather to attempt at mimicking human experts in reading the same images used in real life to make decisions is the objective.

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 <u>shape</u> is defined to be the information attributed to an object that is invariant to scale, origin and orientation [1][2]. A <u>texture</u> may be defined as the prevalence pattern of the interior of an object [3]. 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 [4].

A. Problem Motivation

Doctors use X-Ray, CT, MRI, US and other charts or screen prints or videos to diagnose lesions and highlight their spatial support with respect to the anatomical structures of the body. Doctors may use enhancing tools in order to enable them to distinguish the lesions from anatomical structures; for example, they may use backlighting, display enhancements, image enhancements tools, and they visualize forward and backward slices (in a CT or MRI scan) to decipher the pattern of the lesion and its extent. Uncertainties may develop due to resolution of the imaging protocol, size

4

of the lesion, motion and noise artifacts, scanner artifacts such MRI bias field, occlusion of lesions due to proximity or similarity to anatomical structures, and also due to the stress of over work on the part of the physician.

Computer assisted diagnosis (CAD) systems attempt to mimic the human expert using the same source of data; i.e., the visible information in an image. The uncertainties that may affect the performance of CAD systems are similar to those of the human expert. Human vs. machine vision may be judged fairly if objects have unique features. Three examples are provided below to highlight the common difficulties in human and machine vision – if the object cannot be properly described by humans, it is doubtful that the machine will be able to do so!

Figure 1 is a typical eye chart used to quantify/qualify human vision. Abilities of a person to recognize a letter 20 feet far, where most normal people can do, means "20/20" vision. If the person can read the big E at the top but none of the letters lower than that, the person's vision is considered 20/200; i.e., the person can read at 20 feet a letter that people with "normal" vision can read at 200 feet – hence, a 20/200, denotes vision (visual acuity) is very poor. Based on the status of the eye, people use lenses and other means to enhance their vision [5].

To measure the human vision, a person is asked to recognize something known (a letter from the alphabet or a letter from the Arabic numerals); hence, the object to be seen/recognized is well-described – that is has distinct features. Computer algorithms can be developed to automatically interpret the vision charts as well by translating the letters or numerals features into an algorithm. The distance issue can be translated as how much the object occupies in the field of view (FoV) of the eye/camera.

Arabic

English
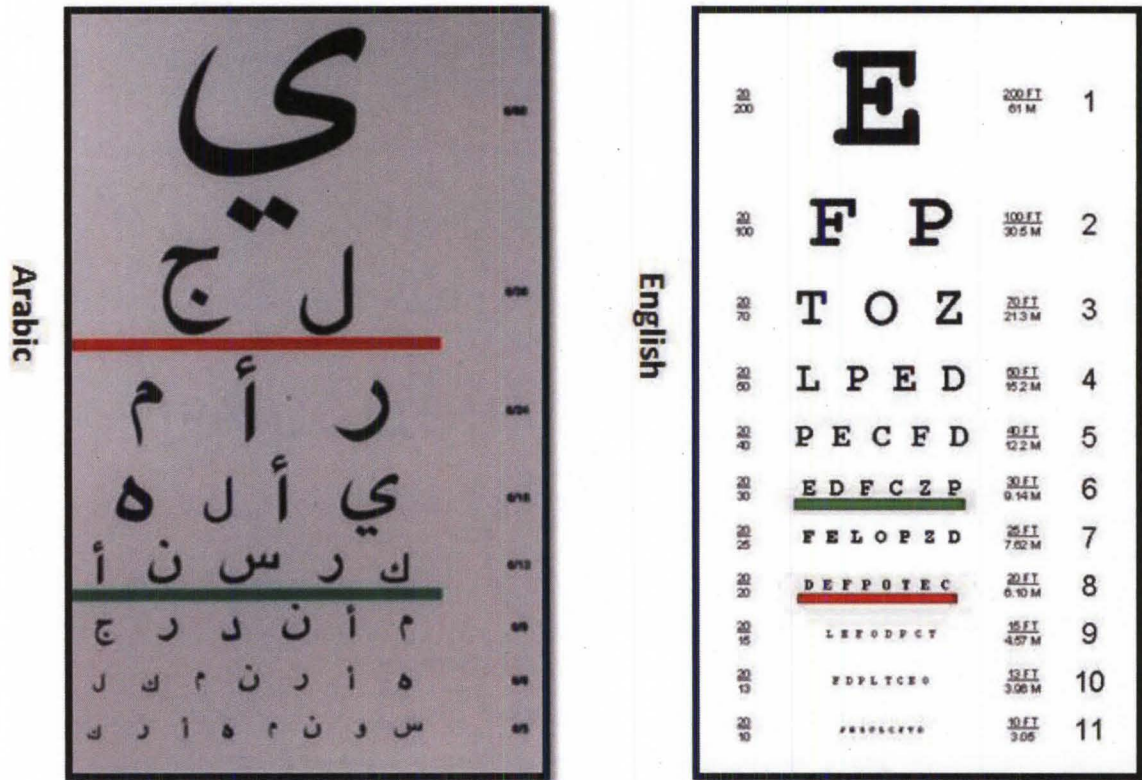
FIGURE 1: The Eye Chart demonstrating levels of vision based on abilities of letter recognition at a distance – Adopted from [5].

Figure 2 is a face recognition example. The comparison between humans and machines can be based on variable constraints. The above example of recognizing a letter at a distance can be simulated to measure the visual acuity or strength of eye vision, through recognizing a human face within a constant background as the area in which the face occupies in the field of view (FOV) changes. The face begins by occupying the entire FOV and then begins to change by variable percentages of that FOV. For the time being, consider that detection and recognition are *simultaneous* – it is one face in the FOV, and if it is detected then it is also recognized. Human and computer face detection are tested with respect to the following uncertainties: i) size only; ii) size + additive noise with variable signal to noise ratio (SNR); iii) size + blur; iv) size + focus; v) size +

occlusion. Of course ii-v could be combined as well. Computer face detection is performed using the Viola and Jones algorithm [6]. Figures 2-7 illustrate these scenarios.
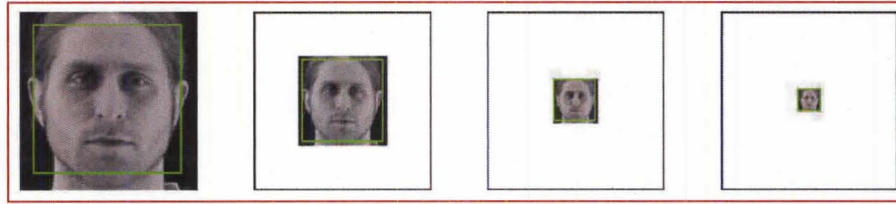
FIGURE 2 - Face detection under a clear imaging condition and change of size.
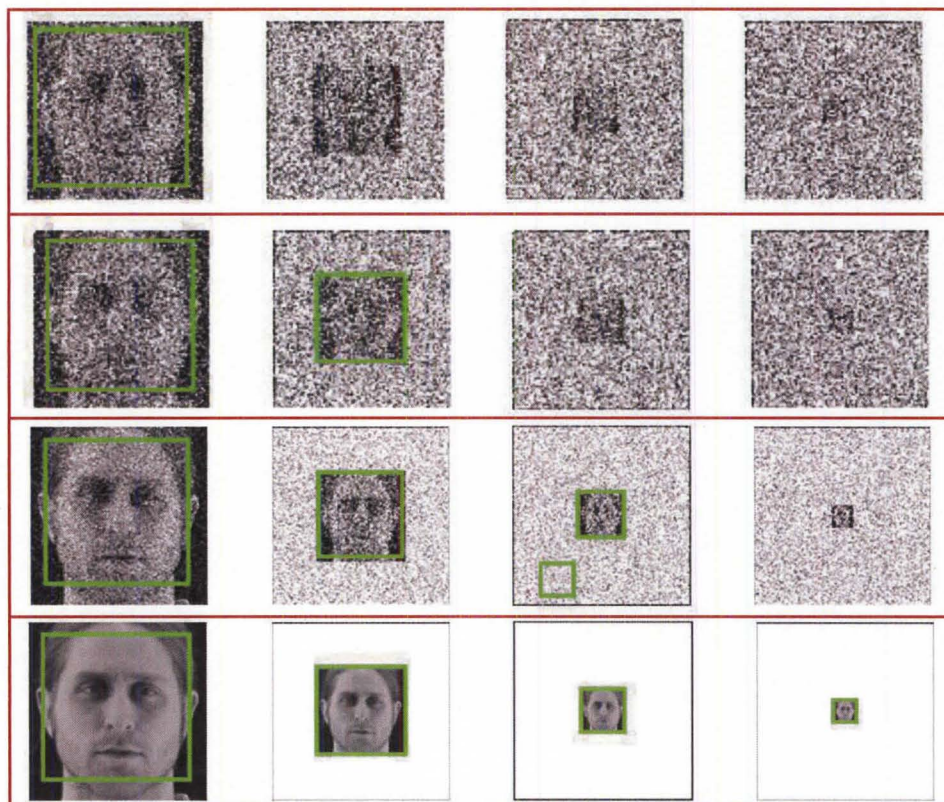
FIGURE 3 - Face detection with variable size and SNR (0, 1, 10, 100 dB - upper to lower rows); SNR $= 10 \, log \, (Contrast)$ [20]. The third row shows a false positive.

As shown in Figure 4, severe motion artifacts smear the features of the face which destroys the algorithm ability to perform. While human can still detect something in the middle of the image, one cannot say that it is a human face. Similarly, defocusing (Figure 5) and occlusions (Figure 6) affect the performance of the algorithm. The Viola-Jones

approach [6] is very popular method of face detection; many other approaches exist in the vast literature of face recognition (for a survey see [7]). A number of approaches in the image processing literature exist for simulating these effects. The occlusion was simulated based on psychology literature (e.g., [8]).
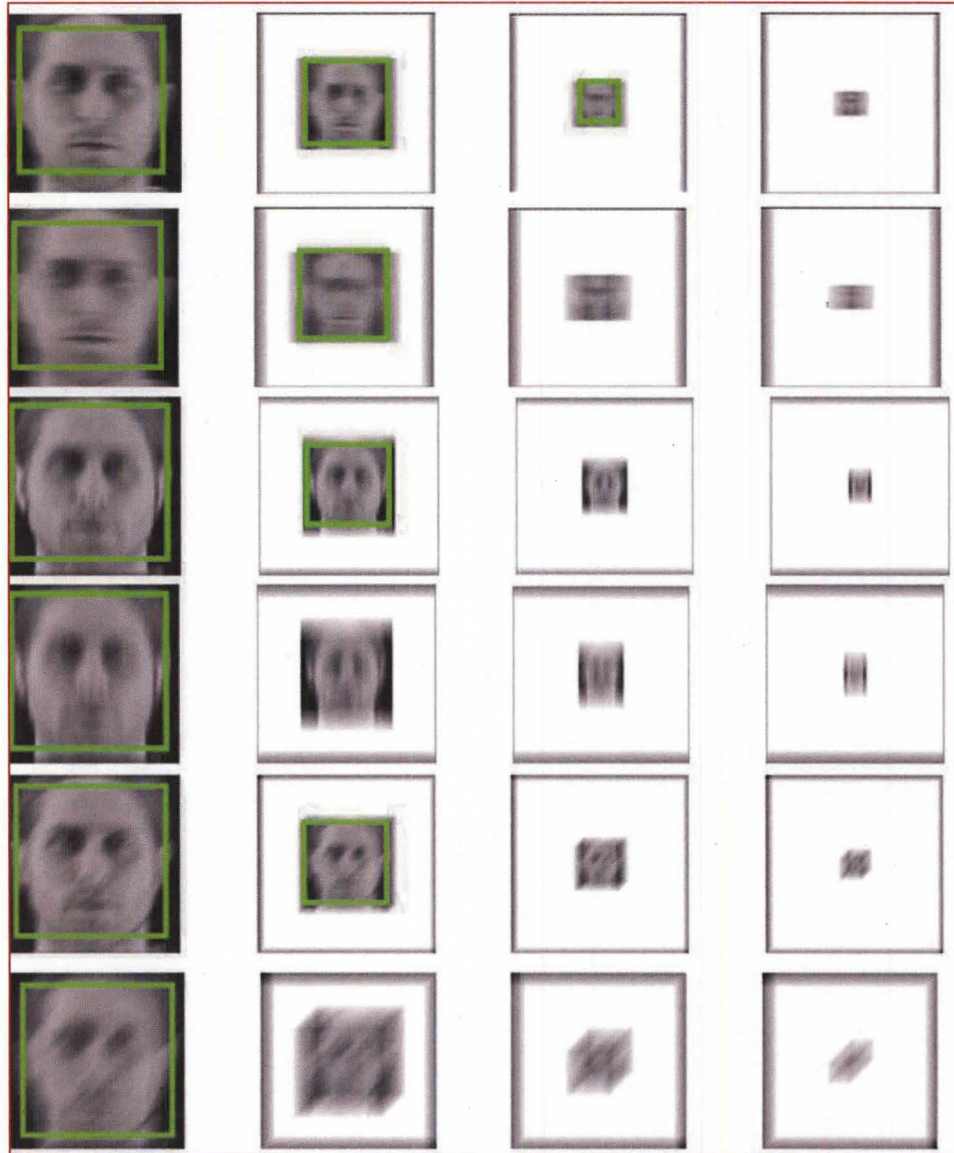
FIGURE 4 - Face detection with variable linear camera motion measured by number of pixels (20, 35 respectively). Top two rows represent horizontal direction motion; middle rows represent vertical direction motion, and bottom two rows represent 45° direction.

Note that motion artifacts are common in biomedical imaging due to breathing, heart contractions and uncontrolled patients movements. LDCT acquisition within one breath can minimize the impact of breathing on Lung nodules.
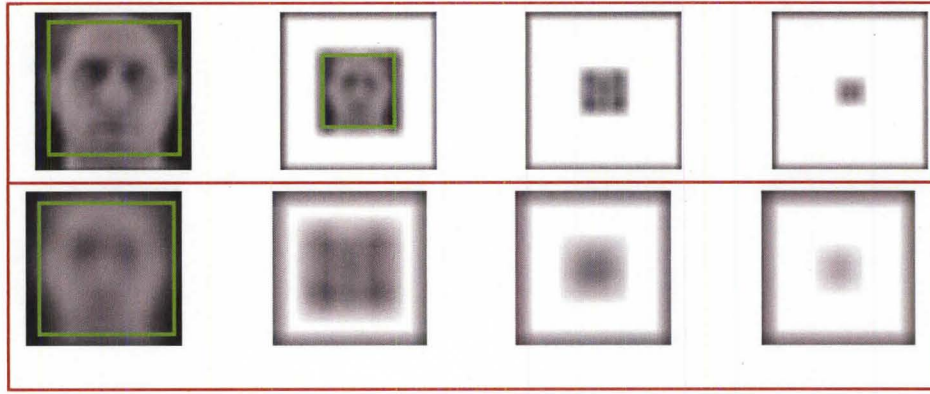
FIGURE 5 - Face detection with variable defocusing using a Gaussian filter with standard deviation of 30 pixels and window widths of 21 (upper) and 51 (lower).

In order to further motivate the research problem considered in this dissertation, the problem of chest tomography and the *visible* characteristics of lung CT scans, which the radiologists use in order to perform a diagnosis, are discussed below. The problem of lung nodules is chosen due to its importance and availably of clinical data that may enable front-end evaluation of a cohesive image analysis study. Also considered is the potential impact on healthcare cost, given that lung cancer is the leading cause of cancer-related deaths in the US and various parts of the world [9]. In this work a number of data sources are used: a) the ELCAP clinical database; b) the LIDC clinical database; c) Jewish Hospital, Louisville, KY, lung screening data; and d) Mansoura University, Mansoura, Egypt CT data from the El-Mogy Scan Center. Each of these data contains unique characteristics and challenges for the approach presented.

FIGURE 6 - Face detection with variable occlusion using a strip to hide a portion of the face. Removal of the forehead or chin (1st and 2nd rows) did not preclude detection/recognition unlike occluding the nose (3rd row) or an eye (bottom row) [20].

Described below are the characteristics of lung nodules as depicted in LDCT scans. The end goal, in this application, is to provide assistance for early detection of lung cancer; hence, the nodules are characteristically small.

Despite the wide range of nodule classifications among radiologists, the nodule classification of Kostis et al. [10] is found to be particularly useful in the algorithmic evaluation presented in this dissertation. Nodules in Kostis's work are grouped into four categories: i) well-circumscribed where the nodule is located centrally in the lung without being connected to vasculature; ii) vascularized where the nodule has significant connection(s) to the neighboring vessels while located centrally in the lung; iii) pleural

tail where the nodule is near the pleural surface, connected by a thin structure; and iv) juxta-pleural where a significant portion of the nodule is connected to the pleural surface.

Figure 7 shows examples of small size nodules (less than 10 mm in diameter) from the four categories. 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. Figure 8 shows sample of nodules and their histograms. These histograms are essentially bi-modal, for the nodule and background regions, and may be sharpened if the region of interest (ROI) is limited to be around the spatial support of the nodules.

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 9. Difference in manual annotation is common of small objects that have not 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 an order of magnitude more difficult than that of automatic face recognition.
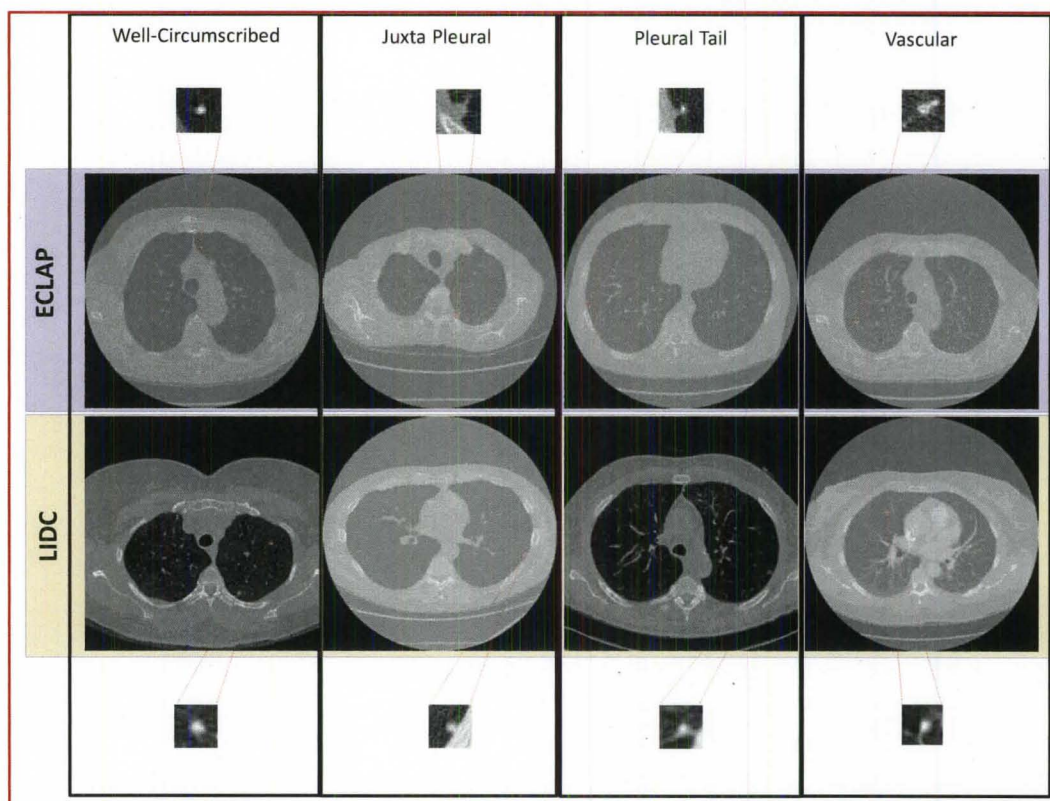
FIGURE 7 - Examples of lung nodules of size below 10 mm from two clinical studies. The upper and lower rows show zoomed pictures of the nodules.
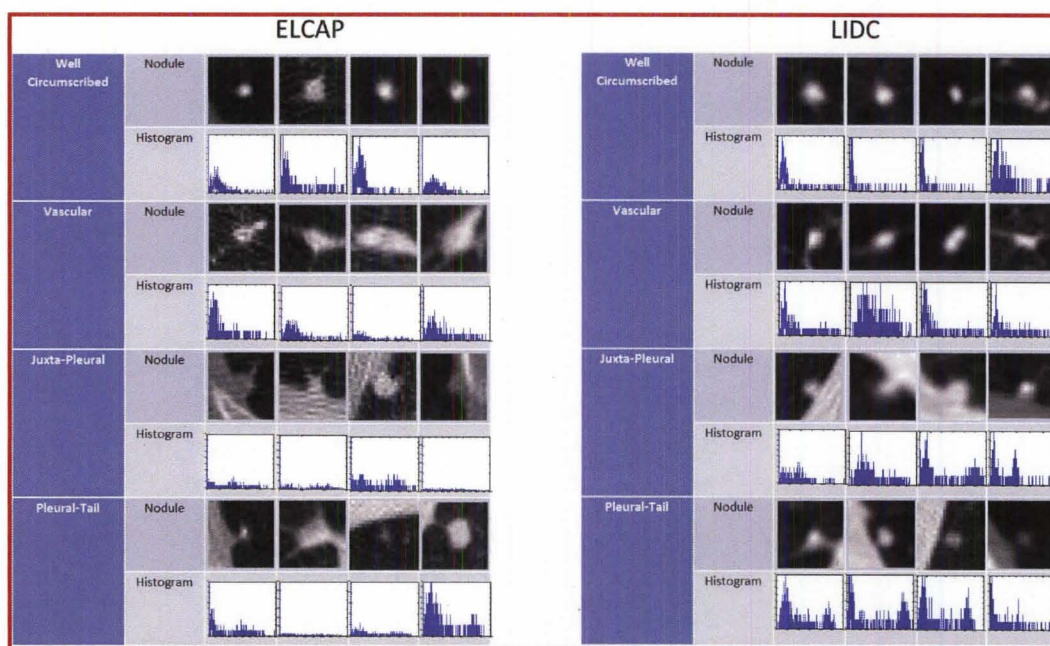
FIGURE 8 – Sample of nodules and their gray level (Hounsfield Units) histograms.
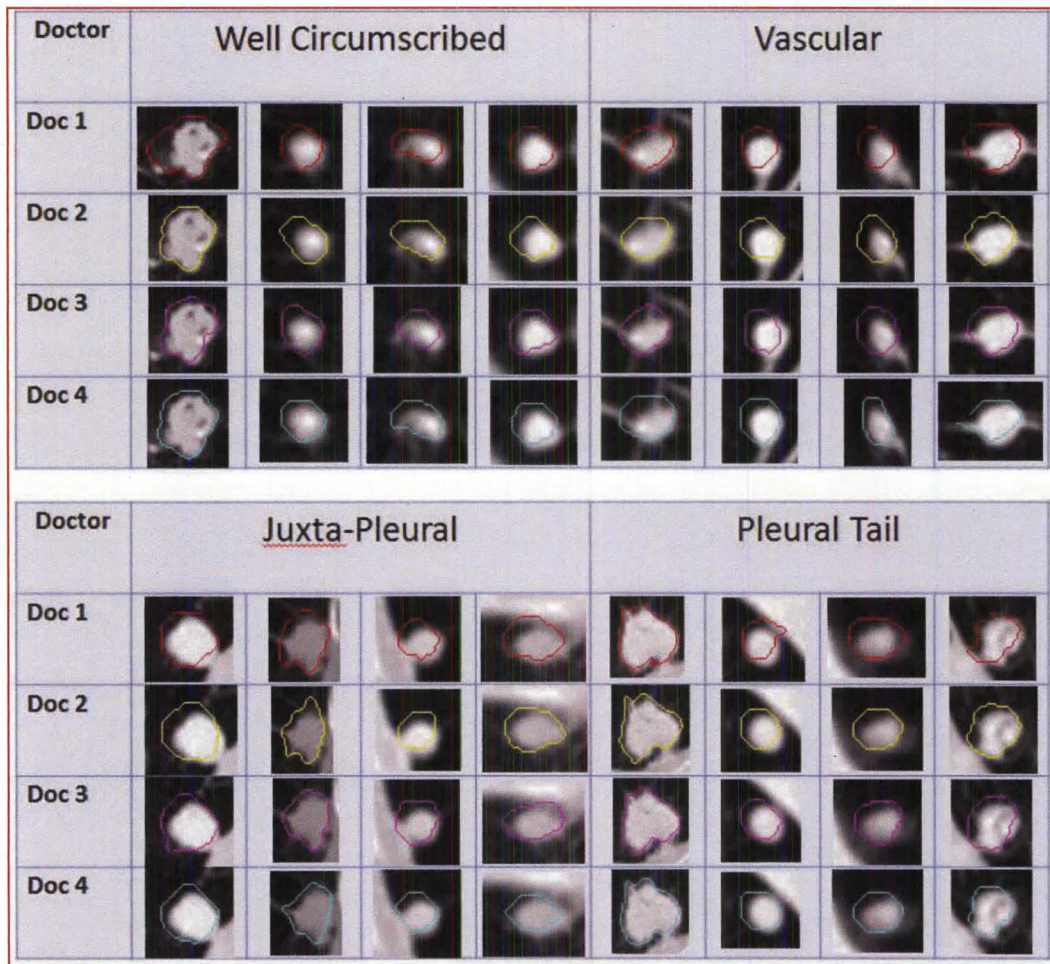
FIGURE 9 – Manual annotation of the main portion of the spatial support of lung nodules by four radiologists. Note the difference in size and shape of the annotations.

Farag, 2009 studied the behavior of the intensity vs. the radial distance of the nodule centroids [11]. The intensity vs. radial distance distribution for small nodules was shown to decay almost exponentially. An empirical measure of the region of support of the nodules was derived based on this distribution. This approach has been tested further on three additional clinical studies in this dissertation and has shown to hold true. The summation of the intensities of Hounsfield Units (HU) in concentric circles (or ellipses) beginning from the centroid of the nodule, decays in a nearly exponential manner with the distance from the centroid.

Figure 10 shows the radial distances for four nodule types from the ELCAP and the LIDC clinical studies [12][13]. This behavior provided a clue for empirically deciding the spatial support (ROI) of the nodules – which is used for auto cropping of the detected nodules. Of course a refinement step is needed to precisely define the exact ROI of the nodule – this is carried out in nodule segmentation.

Object segmentation is a traditional task in image analysis. Real world objects are hard to model precisely; hence the segmentation process is never an easy task. It is more difficult with the lung nodules due to the size constraints.

Figure 11 shows the average intensity (HU) histograms of the manually cropped nodules in the ELCAP and LIDC screening studies. The histograms are distinctly bimodal and a binary classifier (thresholding) may be used for separating the nodules and non-nodules regions. The decision boundary (threshold) 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 (e.g., [11]). Unfortunately, this approach does not work well due to the uncertainties associated with the physical nodules as previously described.

There is a vast literature on object modeling and considerably larger literature on the subsequent steps of modeling; e.g., synthesis, enhancements, detection, segmentation, recognition, and categorization. Farag, 2009 considered a five-step system for modeling of small lung nodules: i) Acquisition and Enhancement; ii) Parametric Modeling; iii) Detection; iv) Segmentation; and v) Categorization (Classification) [11]. By constructing a front-end system of image analysis (CAD system) for lung nodule screening, all of these steps must be considered.

14
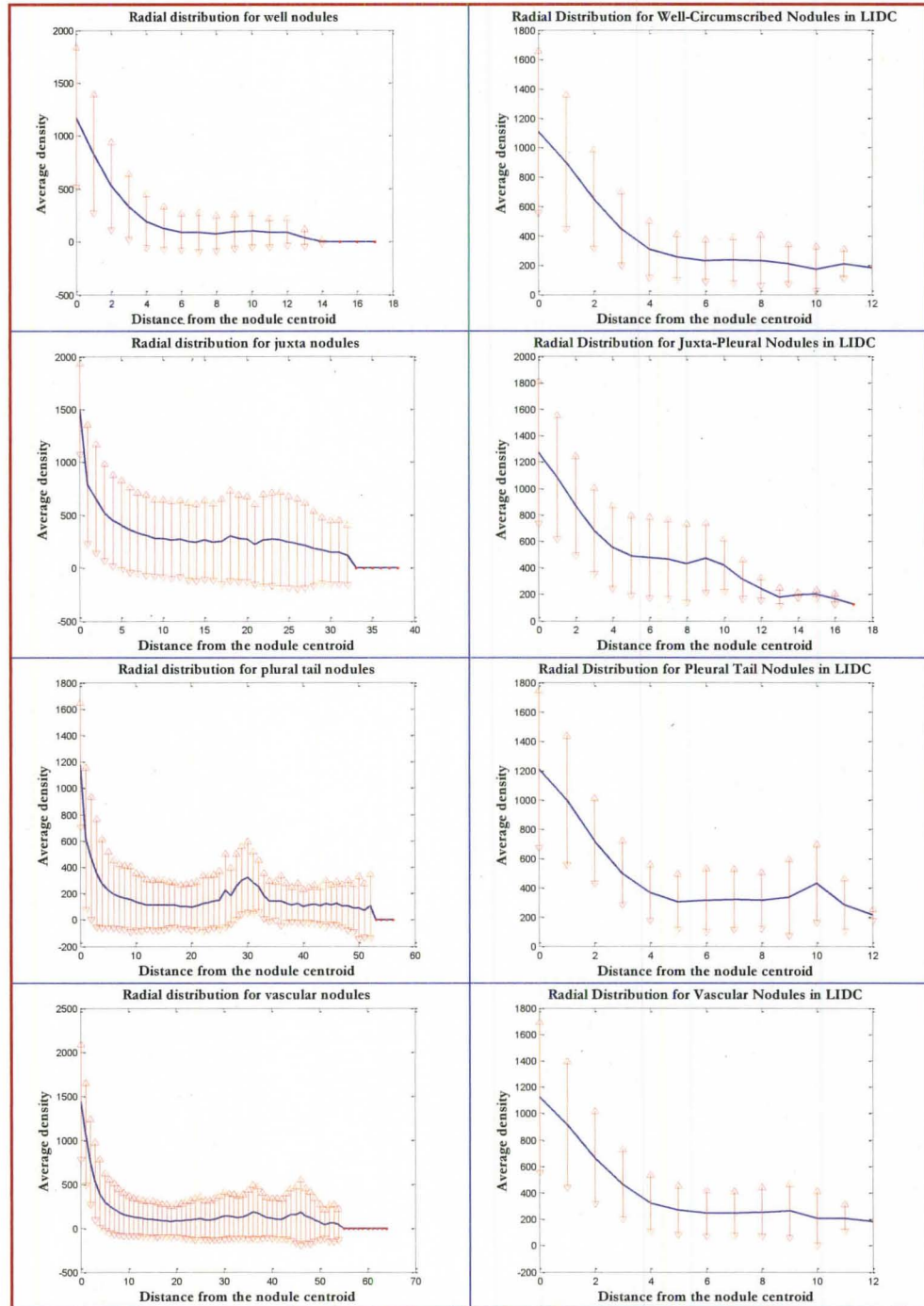
FIGURE 10 - Distribution of the nodule intensity (HU) for four nodule types manually cropped from the ELCAP (319 nodules) and LIDC (over 2000 nodules). For nodules less than 10 mm in diameter, an ROI of size 21 × 21 pixels may be used.
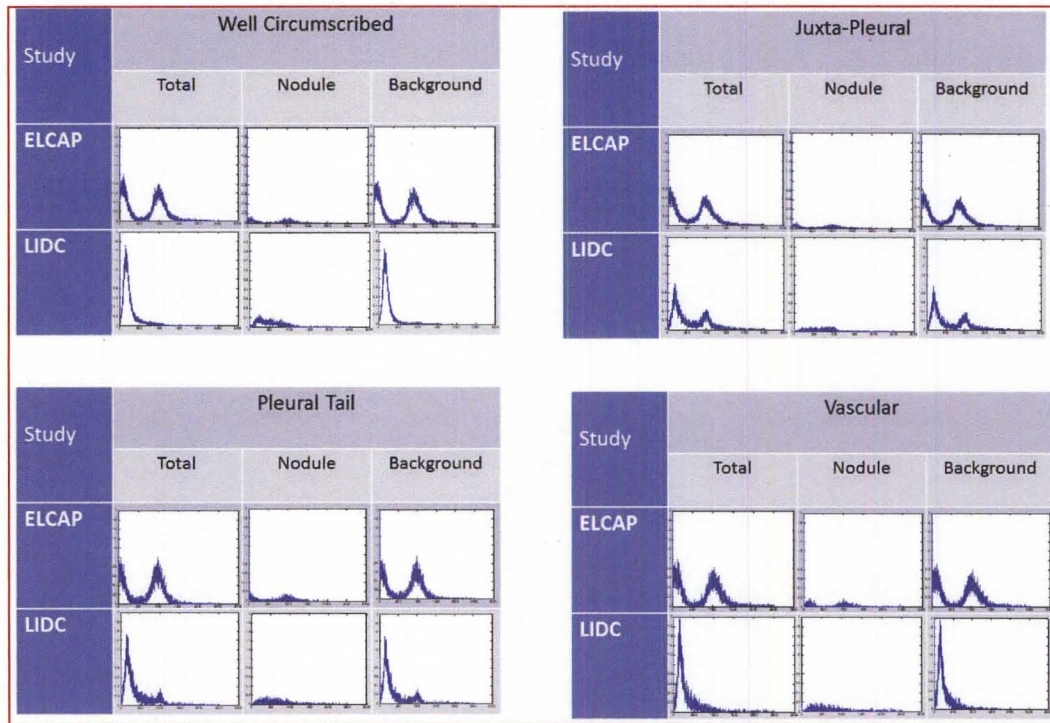
FIGURE 11 – The intensity (HU) histograms of the manually cropped nodules from the ELCAP and LIDC screening studies. These histograms are bio-modal showing the nodule and non-nodule regions in the ROI. These histograms are used as estimates of the probability density functions in the nodule segmentation process.

As the literature on lung nodule analysis is vast, the review will be limited to a sample of the references that bear adequate benefit to the problem of this dissertation. Once again, the focus is to examine the "data-driven" approach. Figure 12 encapsulates the overall focus of this research and the most pertinent tools and sample literature.

Activities in the past few years have led to the following discoveries: 1) Feature definitions on small size objects are hard to pin point, and correspondences, among populations, is very tough to obtain automatically; 2) Classical approaches for image segmentation based on statistical maximum a posteriori (MAP) estimation and the variational level sets approaches do not perform well on small size objects due to unspecific object characteristics; 3) Prior information is essential to guide the

segmentation and object detection algorithms - the more inclusive the a-priori knowledge, the better the performance of the automated algorithms; 4) An integration of attributes is essential for robust algorithmic performance; in particular shape, texture, and approximate size of desired objects are needed for proper definition of the energy functions outlining the MAP or the level sets approaches. These factors play a major motivational role of this dissertation.

Figure 12 summarizes the literature and the major mathematical and algorithmic tools deployed for small object modeling in this dissertation. It is by no means a comprehensive list, and many other tools are not included since they were not fully examined or because their performance was not particularly satisfactory.

B. Problem Statement

The problem that this dissertation is concerned with may be stated as follows: Given an ensemble of realizations of small objects that have, albeit ambiguous, shape and texture attributes, how to generate appropriate models for these objects that can be used for analysis and decision making.

The theoretical development in this dissertation 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., [14]-[18]). Similarly, tensor modeling (TM) allow for including various variations on the object to build a more comprehensive model (e.g., [19]). This dissertation builds on the attempt of Farag, 2009 [11] at using these models in

objects that are hard to annotate by experts. An overview of ASM, AAM and TM may be found in the CVIP Lab Technical Report [20].

Data Domain	Task	Approaches and Sample Literature
	Image Acquisition & Enhancements	Acquisition: CT imaging [12][13] - No Contrast - Contrast-Enhanced Imaging Image Processing [11] - Motion Artifacts removal - Noise removal by Anisotropic Filtering - Contrast enhancements
	Parametric Modeling	Image-Based [11][14][18][20] - Active Appearance Models (AAM) - Tensor Models Contour-Based [87] - Fourier Descriptors - Bezier Descriptors Parametric Models - Fourier Transform - Wavelet Transform
	Object Detection	Template Matching[27] [28][37] - Parametric Templates - Non-Parametric Templates - Exhaustive and Non-Exhaustive Search Matched Filtering - Rotation-invariant - Texture Features – LBP [82] Simultaneous Detection and Segmentation - Elastic registration
	Object Segmentation	Level Sets Method (LSM) [19][73][76] - Shape-based segmentation - Incorporation of shape and texture - Statistical methods of level sets Markov Random Field (MRF) [60][72] - Energy-based methods - Simulated annealing - Graph Cuts Intensity-Based Methods [41][42] [67] - Multimodal Histogram modeling - Mathematical morphology

Four Data Sources are Considered in this dissertation:	Object Categorization	Geometric Feature Extraction [85][88][89]
1. ELCAP 2. Jewish Hospital 3. Mansoura University 4. LIDC		- SIFT [32][89] - ASIFT [33] - Surface Point Signature (SPS) - SURF [84][88] - LBP [82][93] - Gabor Wavelets Object Matching - Rigid/Elastic registration [79][80] - Feature-based correspondence - Simultaneous Segmentation and Categorization - Evaluation of segmentations

FIGURE 12 – Summary of approaches and sample literature for the approaches in this dissertation pertaining to image analysis of small objects.

The practical domain of the dissertation is the lung nodule that shows visibly in LDCT scans. Figures 7-11 demonstrate the physical characteristics of the data to be considered in this dissertation. The nodule data decomposition may be considered based on strict data collection protocol, e.g., ELCAP in which no contrast was used in imaging, and the nodule variability will be mainly orientation (pose) per nodule type. Consideration of relaxed data collection protocols can also be used (e.g., LIDC), in which both intravenous (IV) contrast and no contrast imaging were used and was collected at multiple scanning sites using different hardware and settings.

C. Dissertation Contributions

The image analysis work devised in this dissertation focused on small nodules in LDCT for the purpose of early detection and classification of lung cancer from chest CT scans. The main contributions of the dissertation are the following:

1. Examined the lung nodule modeling approach using active appearance introduced in the master of engineering research of the author in 2009 [11], and tested it on four data sets. This approach has led to nodule models that possess the shape and

appearance of real nodules, which enables a) enhancement of detection performance in terms of sensitivity and specificity over parametric nodule models; b) application of a rotation-invariant matched filtering method for simultaneous detection and segmentation of the nodules; and c) application of registration-based approach for simultaneous detection, segmentation and categorization of nodules.

2. Proposed a new formulation to the detection approach, template matching. The approach is formulated as an energy optimization problem that allows for computation of a transformation that includes the weighting coefficients and rotation/scale parameters used in the generation of the models.

3. Constructed the probability density functions (PDF) for the HU of the lung nodules based on empirical evaluations of the region of supports of nodules below 10 mm in diameter. These PDFs have been crucial for successful model-based segmentation of the lung nodules.

4. Formulated and tested a variational level set approach for lung nodule segmentation using the shape and texture models of real nodules.

5. Formulated a statistical approach for lung nodule segmentation, following the traditional random field methods, and used the graph cuts method to optimize the resulting energy function.

6. Created an approach for lung nodule categorization based on geometric feature descriptors.

7. Developed a complete CAD system for lung nodule detection, segmentation and classification and evaluated it on several lung CT databases.

D. Dissertation Outline

This dissertation aims at creating a theory for modeling of small-size objects under uncertainties. The theory will examine the effects of object size constraints (and the uncertainties of imaging and the environment) on the estimation of parameters in three approaches for image modeling: statistical, variational and geometric modeling methods. These tools are useful for various classes of small objects under uncertainties, yet the application domain of the dissertation will be limited to the problem of lung nodule analysis from LDCT scans intended for early detection of lung cancer. The work reported here is divided as follows:

- Chapter II: Provides an overview of terminologies and definitions and the basic mathematical descriptions used in the work; The chapter also describes the concepts of shape, texture, active appearance models and the overall framework for statistical shape models;

- Chapter III: Considers lung nodule detection as a case study of small objects. The data-driven approach for lung nodule modeling using active shape models is detailed. A new template matching detection formulation is evaluated using these models on nine similarity measures.

- Chapter IV: Considers lung nodule segmentation as a case study of small objects. Three approaches for lung nodule segmentation are examined: a statistical method, where optimization is carried out by the graph cuts approach; a variational method where shape information is included in the energy minimization; and a novel method is devised using variational calculus (level sets) for lung nodule segmentation. Basic lung nodule segmentation is also

considered by a two-class classification, where the decision threshold is based upon two-normal classes with parameters estimated by the EM algorithm.

- Chapter V: Considers lung nodule classification as a case study of small objects. Various view-based object modeling approaches are detailed that extract texture and shape information to describe the nodules as feature vectors. The K-nearest neighbor approach for nodule classification of types is used for evaluation.

- Chapter VI: Provides a summary, conclusions and possible extensions

CHAPTER II

BASICS OF OBJECT MODELING

In this chapter two approaches are reviewed for statistical object modeling based on the shape and the texture (appearance) of the object in an image or volume. Active Shape Models (ASM) and Active Appearance Models (AAM) are among the most powerful tools for object analysis and are very popular in the computer vision and the biomedical imaging analysis literature. These models are based on ensemble or realizations of shape and texture of an object annotated by experts and analyzed to remove the redundancies in between the features of the annotations. The resulting co-registered ensemble is used to derive the parameters of the ASM and the AAM models. These models can be used separately or jointly, based on the application. This work reviews the basic foundation of the two approaches and provides illustrative examples to their effectiveness in object modeling. The two approaches depend on availability of distinct features in shape and texture. The major problems associated with using ASM and AAM for small object modeling are demonstrated.

A. Active Shape Models (ASM)

The field of statistical shape analysis deals with describing shapes of populations that are not deterministic. Similar to the concepts of random variables and random processes, shapes from a specific experiment are random.

A real-valued random variable is a mapping from the experiment's sample space to the real line such that the reverse mapping is an event. A random process is an indexed set of random variables. Probability distribution functions (PDF) are the most comprehensive form of describing random variables and random processes. The theory of random variables and random processes is an established one, and forms the foundation of signal analysis, data analysis and prediction [21]. Statistical shape analysis follows nearly the same style of building a cohesive description of shapes based on axioms, definitions and theorems [1-4].

A shape process may be established to generate shapes; similarly an appearance process may be established to generate appearances. Parameters of the shape and appearance processes will determine the details of the realizations, similar to random variables and random processes. Generative random processes are abundant in statistical theory, including Autoregressive (AR), Moving Average (MA), and combined Autoregressive Moving Average (ARMA) Models. Realizations can be generated from AR, MA and ARMA models [21]. Likewise, construction of generative shape and appearance models by linear representations of *modes* that can be constructed from an ensemble of shape and appearance processes may be possible.

Figure 13 illustrates a sample of small nodules, less than $1\,cm$ diameter, from the LIDC clinical study showing the variations that can be captured by shape and appearance models for the purpose of detection, segmentation and classification. These nodules may appear in various locations in the lung tissue, and may also be occluded by the anatomical structures in the lung. These nodules may be analyzed as one category, or may be subdivided into several categories based on particular shape characteristics.

This chapter aims at reviewing some of the basic results of statistical shape analysis. The concept of shape is defined and the tools for shape representation, modeling and analysis are studied– with focus on the tools that lend benefits for modeling small objects under uncertainty.
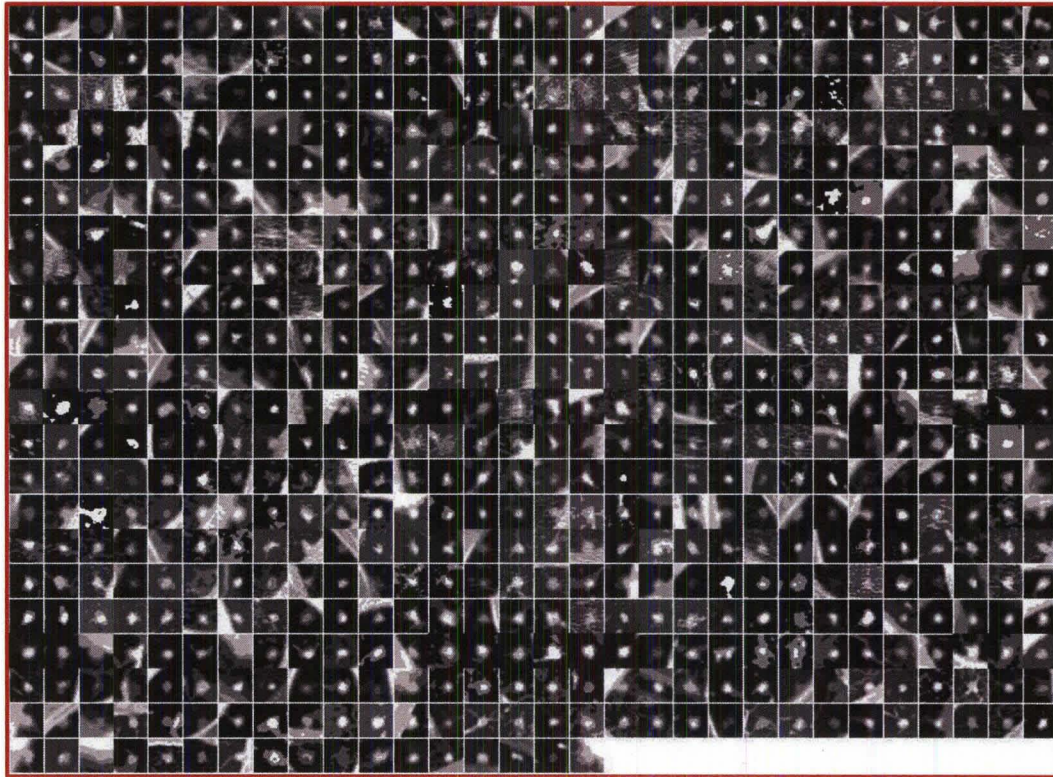
Figure 13 – An ensemble of small size lung nodules, less than 1 *cm* diameter, from the LIDC clinical study.

1. Basic Definitions

Adopting Kendall's definition [1][2], shape is formally defined as:

Definition 1: A Shape resembles all the geometrical information that remains when location, scale and rotational effects are filtered out from an object.

Shapes may be represented by various methods, mostly starting from certain landmarks that best describe the object. A landmark can be formally defined as:

Definition 2: A landmark is a point of correspondence on each object that matches between and within populations.

Real world objects may take various forms of details, and may be linear, planar or three dimensional. In [4], Dryden and Marida, define <u>anatomical landmarks</u> as points assigned by an expert that corresponds between organisms in some biologically meaningful way; <u>mathematical landmarks</u> as points located on an object according to some mathematical or geometrical property, i.e. high curvature or an extremum point; and <u>pseudo-landmarks</u> as constructed points on an object either on the outline or between landmarks. The computer vision literature is rich in approaches and techniques to extract features for object modeling, from various data representations. The focus here is limited to a few basic concepts from the vast literature of statistical shape analysis. The two papers of Cootes, 2000 [15], Stegmann and Gomez, 2002 [17] provides an adequate survey for the steps used in building an ASM model.

A shape is considered to be a set of n −vertices $x \in R^k$; for the two-dimensional case:

$$x = [x_1; x_2; \dots ; x_n; y_1; y_2; \dots ; y_n]^T \tag{1}$$

The shape ensemble (realizations of the *shape process* of a certain object) is to be adjusted (aligned) on 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; formally, defined as follows:

Definition 3: The *shape space* is the orbit shape of the non-coincident n point set configurations in R^k under the action of the *Euclidean* similarity transformations.

It can be noted that the translation and rotation processes associated with alignments of the shapes in the ensemble reduces the dimensions of the shape space from kn to $kn - k - 1 - k(k-1)/2$. Therefore, an ensemble of 100 vertices two-dimensional shape realizations would have a shape space of 197 instead of 200. This reduction would be significant with higher values of n and k.

To align the shapes in an ensemble, various procedures may be used. The *Procrustes* procedure, which is most common in *rigid* shape alignments, is discussed below. 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 (e.g., Figure 14) with one-to-one point (feature) correspondence is provided.
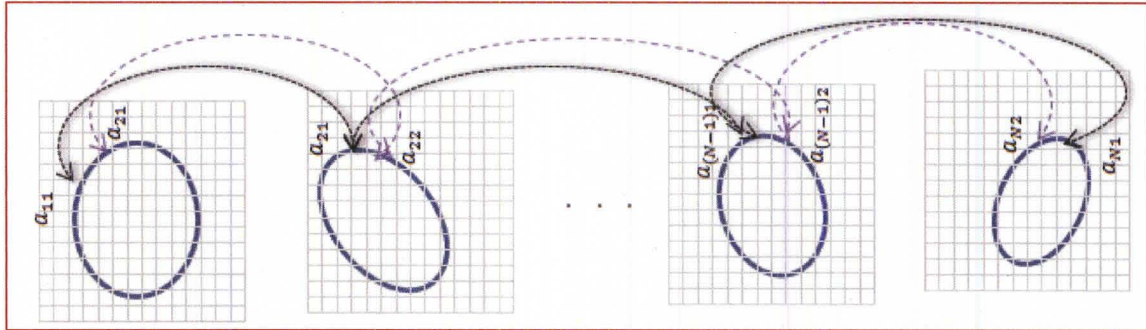
FIGURE 14 – A shape ensemble of size N with correspondence.

The alignment may be accomplished by the following steps [17]:

A General Shape Alignment procedure ...
1. Compute the centroid of each shape in the ensemble
2. Re-scale each shape to have equal size
3. Align the shapes, pair-wise, with respect to position at their centroids
4. Align the shapes, pair-wise, with respect to orientation by rotation.

The *Procrustes distance* between two shapes s_1 and s_2 is the sum squared distance (SSD)

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

The centroid $(\bar{x}, \bar{y})$ of the shape is its center of mass and is calculated by averaging the coordinates of the vertices; i.e., $(\bar{x}, \bar{y}) = \left(\frac{1}{n} \sum_{j=1}^{n} x_j, \frac{1}{n} \sum_{j=1}^{n} y_j \right)$.

Definition 4: The *shape-size metric* $S(\mathbf{x})$ is a positive real-valued function of the shape representation $\mathbf{x}$ such that: $S(a\mathbf{x}) = aS(\mathbf{x})$, where a is real number.

Common scale metrics are the Frobenius norm or the centroid size defined as:

$$\textit{Frobenius norm} = S(x) = \sqrt{\sum_{j=1}^{n}(x_j - \bar{x})^2 + (y_j - \bar{y})^2} \qquad (3)$$

$$\textit{Centriod size} = S(\mathbf{x}) = \sum_{j=1}^{n} \sqrt{(x_j - \bar{x})^2 + (y_j - \bar{y})^2} \qquad (4)$$

Shape alignment with respect to the centroid may be performed by the singular value decomposition (SVD) approach, using the following steps:

1) Arrange the size- and position-aligned shapes $\mathbf{s}_1$ and $\mathbf{s}_2$ as $n \times k$ matrices.
2) Calculate the SVD, $\mathbf{UDV}^T$ of the two shapes vertices $\mathbf{x}_1^T \mathbf{x}_2$; in order to maximize the correlation between the set of landmarks.
3) The rotation matrix needed to optimally superimpose $\mathbf{x}_1$ on $\mathbf{x}_2$ is $\mathbf{VU}^T = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

An iterative approach for alignment of a set of planar shapes may be performed by the following algorithm:

Shape Alignment Algorithm:
1. Chose an initial estimate of the mean shape – one may use the first shape in the ensemble
2. Align all the remaining shapes to the mean shape – The mean shape for an ensemble of size N may is the average or *Procrustes mean*, $\bar{\mathbf{x}} = \frac{1}{N} \sum_{j=1}^{N} \mathbf{x}_j$.
3. Re-calculate the estimate of the mean from the aligned shapes
4. If the mean changes return to step 2.

Experience has shown that the above algorithm converges in a few iterations.

Various methods may be deployed for shape alignment besides the above algorithm.

With rigid transformations the estimation of the translation, scale and rotation is straight forward.

2. Data Reduction by Principal Component Analysis (PCA)

Annotated data of an ensemble of shapes of a certain object carries redundancies due imprecise definitions of landmarks, and due to errors in the annotations. To perform proper shape alignment these redundancies may be reduced by a transformation step that flags the important features and filters out the highly correlated ones. PCA is a technique for data reduction that has been widely used due to its computational efficiency and theoretical appeal. In a nutshell, PCA involves linearly transforming the original data such that the important features may be weighted high with respect to the eigen vectors of the transformation. In PCA, the original shape vector $\mathbf{x}$ is linearly transformed by a mapping $\mathbf{M}$ such that $\mathbf{z} = \mathbf{M}\,\mathbf{x}$ has less correlated and highly separable features. The mapping $\mathbf{M}$ is derived for an ensemble of N shapes as follows:

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_i \text{ and } \Sigma_{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^{N} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \tag{5}$$

are the mean and covariance of $\mathbf{x}$. Therefore, the mean and covariance of $\mathbf{z}$ would be:

$$\bar{\mathbf{z}} = \frac{1}{N} \sum_{j=1}^{N} \mathbf{z}_j \text{ and } \Sigma_{\mathbf{z}} = \frac{1}{N} \sum_{i=1}^{N} (\mathbf{z}_i - \bar{\mathbf{z}})(\mathbf{z}_i - \bar{\mathbf{z}})^T = \mathbf{M}\Sigma_{\mathbf{x}}\mathbf{M}^T \tag{6}$$

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

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}\,\mathbf{b} \tag{7}$$

where $\mathbf{P} = (\mathbf{p}_1|\mathbf{p}_2|\dots|\mathbf{p}_m)$ matrix of m largest eigen vectors of $\Sigma_{\mathbf{x}}$ and $\mathbf{b}$ is $m \times 1$ vector given by

$$\mathbf{b} = \mathbf{P}^{\mathrm{T}}(\mathbf{x} - \bar{\mathbf{x}}) \tag{8}$$

By varying the elements of $\mathbf{b}$ one can vary the synthesized shape $\mathbf{x}$ in Eq. (7). The variance of the ith parameter $b_i \in \mathbf{b}$ can be shown across the training set to be equal to the eigen value λ_i [15]. The distribution of the PCA reduction is optimal if the training shapes have a Gaussian distribution - the *modes* of the $b_i \in \mathbf{b}$ may be limited to within seven times the standard deviation; i.e., $\pm 3\sqrt{\lambda_i}$, which will provide an adequate synthesized shape model. Of course, all of these modes may not be needed to generate a synthesized shape model; based on the complexity of the shape ensemble at hand, perhaps only a few modes will be enough. Visible inspection or SSD error measure may be used to select the adequate modes. The fact that the generated (synthesized) shape is mode dependent makes it deformable.

3. Example

The Procrustes shape alignment and the usefulness of the PCA method through an ensemble of shape contours from the LIDC study is demonstrated. In order to distinguish each nodule type a parametric model for each nodule type is used: the well-circumscribed nodule is an ellipse; the vascular nodule is a rectangle; the juxta-pleural is an outward open ellipse; and the pleural-tail is composition of a vascular plus a well-circumscribed nodule, as shown in Figure 15.
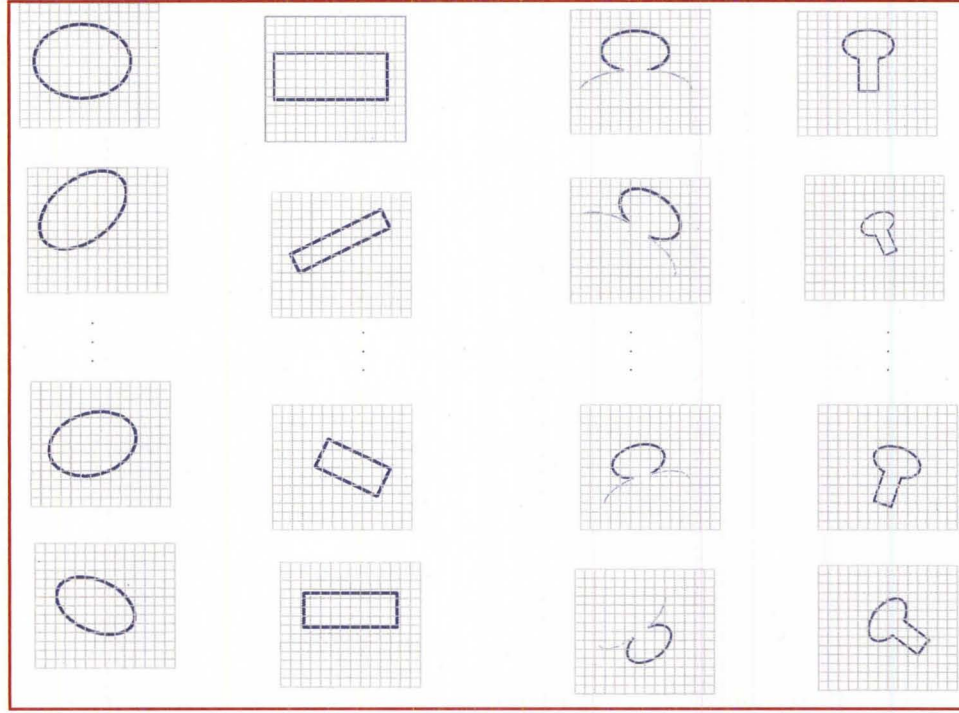
FIGURE 15 – Illustration of shape ensembles from four nodule types. 1st column represents the well-circumscribed nodule, 2nd column the vascular, 3rd column the juxta-pleural and 4th column represents the pleural-tail nodules. Features are defined from each nodule and used for annotation.

In the image, $\mathbf{x_t}$ can be related to $\mathbf{x}$ in the modeling space at a single pixel location (x, y) by the transformation:

$$T_{X_t,Y_t,s,\theta}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} X_t \\ Y_t \end{pmatrix} + \begin{pmatrix} s\cos\theta & -s\sin\theta \\ s\sin\theta & s\cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \qquad (11)$$

The parameters of $T_{X_t,Y_t,s,\theta}$ may be obtained by optimizing a similarity measure such as the SSD distance; i.e, the mapping between $\mathbf{x_t}$ and $\mathbf{x}$ would be such that:

$$SSD = |\mathbf{x} - \mathbf{x_t}|^2 = |\mathbf{x} - T_{X_t,Y_t,s,\theta}(\bar{\mathbf{x}} + \mathbf{P}\,\mathbf{b})|^2 \qquad (12)$$

is minimum. An iterative algorithm to perform the optimization is as follows [15]:

Algorithm: Matching model points to target points

1. Initialize the shape parameters, **b**, to zero (the mean shape).

2. Generate the model point positions using $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}\,\mathbf{b}$

3. Find the pose parameters (X_t, Y_t, s, θ) which best align the model points **x** to the current found points $\mathbf{x_t}$

4. Project $\mathbf{x_t}$ into the model co-ordinate frame by inverting the transformation T:

$$y = T^{-1}{}_{X_t, Y_t, s, \theta}(x_t) \tag{13}$$

5. Project **y** into the tangent plane to $\bar{\mathbf{x}}$ by scaling: $\mathbf{y'} = \mathbf{y}/(\mathbf{y}.\bar{\mathbf{x}}\ )$.

6. Update the model parameters to match to $\mathbf{y'}$

$$\mathbf{b} = \mathbf{P}^{\mathrm{T}}(\mathbf{y'} - \bar{\mathbf{x}}) \tag{14}$$

7. Return to step 2 if not converged.

Using the above procedure to fit a model shape to a target one in the image plane requires better initialization and an optimization procedure (e.g., the gradient descent). Initialization is best accomplished if one selects the important features that describe the shape. Detecting these features automatically is never an easy task. For small objects in particular, it is a very formidable one since the classic edge detectors and geometric feature descriptors do not perform well. A related approach is to model the full appearance of the object, including the internal texture. This is known as active appearance modeling (AAM). The quality of fit of such an appearance model can be assessed by measuring the difference between the target image and a synthetic image generated from the model. The AAM approach is considered next.

B. Active Appearance Models

Active appearance models (AAM) textures of images using an ensemble of test images. The test images are generated around labeled landmarks in the images; hence, AAM and ASM are tightly related. Following the approach of Cootes, Edwards and

32

Taylor, 1998 [14], the ASM results above are extended to highlight the basics of ASM. It should be pointed out that other developments of the AAM exist in the literature; in particular, the work of Mathews and Baker, 2004 [18].

To construct an ASM from image intensity, the test images in the ensemble are warped so that the feature points (control points) match the mean shape, using a triangulation algorithm. The region contained in the mean shape is then sampled and normalized to reduce the image variations effects. The notations in Cootes et al. [14], are continued to be used. Let the region from the ensemble members covered by the mean shape be $\mathbf{g}_{im}$, the normalized image regions $\mathbf{g}$ will be

$$\mathbf{g} = (\mathbf{g}_{im} - \beta\mathbf{1})/\alpha \tag{15}$$

The values α, β are chosen match members of the ensembles within the mean shape regions to the their average, $\bar{\mathbf{g}}$; i.e.,

$$\alpha = \mathbf{g}_{im}.\bar{\mathbf{g}} \text{ and } \beta = \mathbf{g}_{im}.\mathbf{1}/n \tag{16}$$

where n, is the number of the elements in the vectors (size of the region within the mean shape). To remove redundancies within the normalized members of the image ensembles the PCA approach is used; hence, obtaining a linear model for intensity

$$\mathbf{g} = \bar{\mathbf{g}} + \mathbf{P}_g\mathbf{b}_g \tag{17}$$

where $\bar{\mathbf{g}}$ is the mean normalized intensity vector, $\mathbf{P}_g$ (eigen matrix) is a set of orthogonal modes of variation and the $\mathbf{b}_g$ is a set of intensity parameters (eigen intensity). One may combine the ASM and AAM equations above. Let $\mathbf{b}_s$ (Eq. 8) and $\mathbf{b}_g$ represent the shape and intensity vectors. The two representations may be concatenated as follows:

$$\mathbf{b} = \begin{pmatrix} \mathbf{W}_s\mathbf{b}_s \\ \mathbf{b}_g \end{pmatrix} = \begin{pmatrix} \mathbf{W}_s\mathbf{P}_s{}^T(\mathbf{x} - \bar{\mathbf{x}}) \\ \mathbf{P}_g{}^T(\mathbf{g} - \bar{\mathbf{g}}) \end{pmatrix} \tag{18}$$

where W_s is a diagonal matrix of weights for each shape parameters, allowing for the difference in units between the shape and intensity models [14].

Further application of the PCA on the concatenated representation may reduce the redundancies. The results would be the following model:

$$b = Qc \tag{19}$$

where Q are the eigen vectors and c are the vector of appearance parameters controlling both the shape and intensity of the model [14]. The shape and intensity models can still be expressed, i.e.,

$$x = \bar{x} + P_s W_s Q_s c \quad , g = \bar{g} + P_g Q_g c \tag{20}$$

where

$$Q = \begin{pmatrix} Q_s \\ Q_g \end{pmatrix} \tag{21}$$

The above representation can be used to synthesize an image for a given appearance parameters c by generating the shape free intensity image vector g and warping it using the features or control points described by x.

Example:

AAM and ASM are widely applicable in face modeling research. The lung nodule example is used as an illustration of the AAM algorithm described above. Unlike the faces, the nodules have no distinct control points or features within the boundaries; hence, the output of the ASM will be easier to judge than the AAM since nodules may possess a distinct shape boundary.

Indeed, Farag, 2009 [11] used empirical models of appearance (nodule intensity) to fill out the interior of a shape model starting from its centroid. In the following example, both the geometric models of the nodule's shape (Figure 16) as well as annotations of

two concentric circles starting from the centroid are used. The ASM and the AAM approaches are applied. It is noted that the AAM and ASM capture the appearance and shapes of the nodules.

These models are used later in the dissertation for detection, segmentation and classification of nodules.
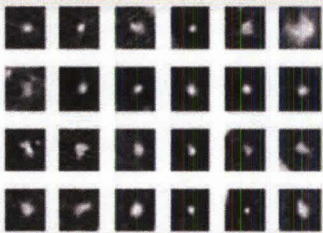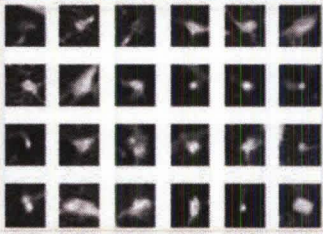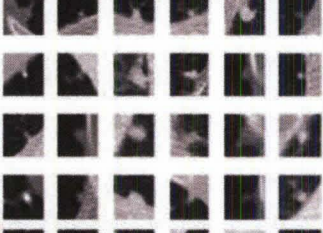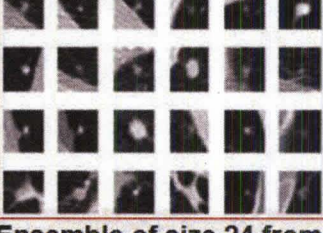
FIGURE 16: Data-driven nodule models by AAM and Level Sets. The templates carry the salient features of real nodules; resulting in higher sensitivity and specificity in detection.

C. Tensor Modeling Approach

Another approach for appearance and shape modeling is based on the use of tensor algebra. The same work introduced by Vasilescu and Terzopoulos, 2007 for face recognition [19] is followed. Tensors allow the incorporation of various geometric properties of the nodules and the effects of imaging conditions as well. In the face recognition, the tensor approach incorporates illumination, pose and expressions and multiple individuals (e.g., Figure 17(A)). The lung nodules may be arranged in a similar fashion as shown in Figure 17(B).
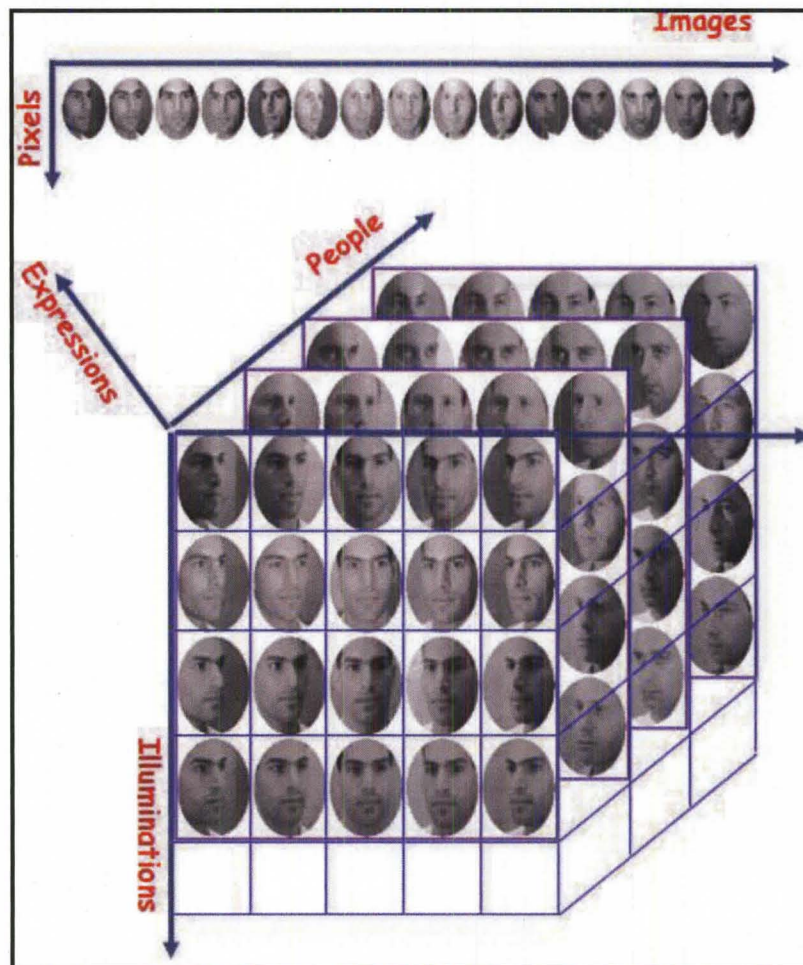
FIGURE 17(A) – Data arrangements for PCA (upper) and Tensor Analysis (lower) in face recognition (Adopted from Vasilescu and Terzopoulos [19]).
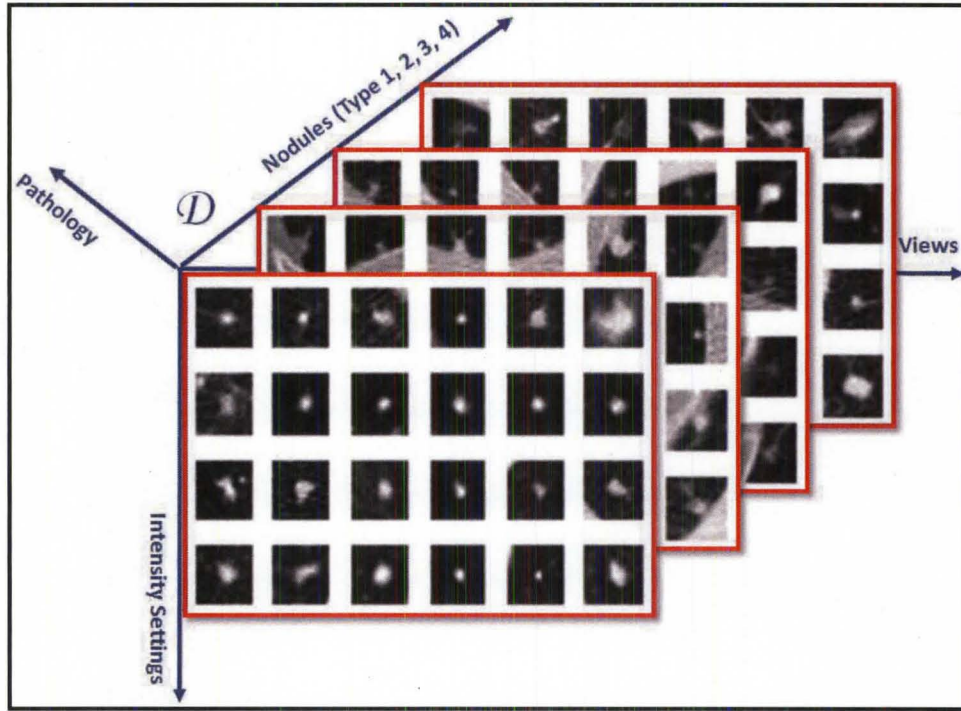
Figure 17(B) - A generic nodule data decomposition based on orientation (pose), and imaging parameters (intensity) for various nodule shape classifications.

FIGURE 17: Data arrangement for Active Tensor Modeling for face and nodule data.

Tensor analysis has been shown to provide significant improvement in face recognition compared to Principal Component Analysis (PCA) and Independent components Analysis (ICA) [16]. The nodule variability is not as much, but a tensor decomposition of nodules incorporating: a) nodule types; b) scanning protocol specifics, e.g., mA; c) vendor of imaging scanner – (one can expect slight variability in imaging using GE, Phillips and Siemens hardware); d) nodule size; e) use of contrast or lack of during imaging, etc. may be constructed. The LIDC enables tensor decomposition using some of these variables. In [20], experimentation with various arrangements of the nodules was conducted. Below, an example is shown representing one type nodule using PCA and Tenors, for illustration purposes. Of course, the objective is to be able to

synthesize a nodule by the least number of parameters, while maintaining the closest possible shape to the original image.

Figure 18 illustrates the performance of the AAM and ATM methods for sample nodules from the LIDC clinical study.

(a) AAM models of a well-circumscribed nodule, showing the average and 5 eigen nodules.

(b) ATM models of two well-circumscribed nodules, showing average and several tensors.
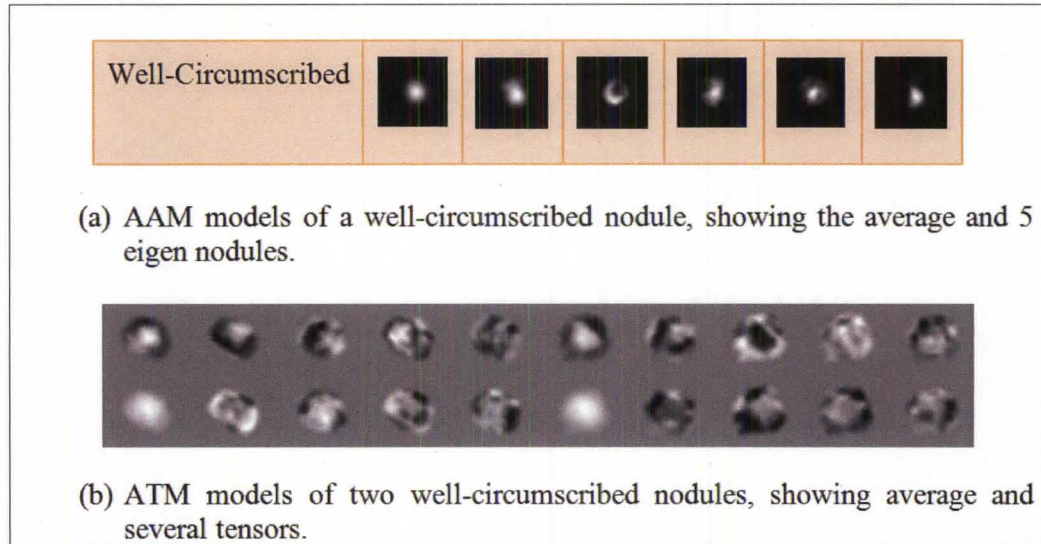
FIGURE 18 – AAM and ATM (small number of variations were used in the nodule acquisition) for sample nodules from the LIDC clinical study.

D. Summary

This chapter discussed a few of the mathematical concepts related to object modeling, which will be detailed in the following chapters. The concept of shape was defined. The basics of active appearance modeling approach were outlined. This chapter did not review the stochastic approaches of object modeling nor the variational methods. Chapter III is dedicated to the statistical approaches whereas Chapter V, nodule segmentation, considers the variational modeling approach.

CHAPTER III

DETECTION OF SMALL SIZE LUNG NODULES

Since the mid-eighties object detection has been an area of massive research, where an object is modeled by scale-space representations, affine image blobs, Active Appearance models (AAM), etc. [18, 22-25, 26]. The detection approaches in general are categorized as model-based or view-based methods depending on the representation of the models used [27, 28-30]. Detection is usually performed through matching a template or model with candidates in an image. The model-based template representations are used in numerous detection approaches including the widely used template matching approach. The object template is swept across the image in a raster fashion, then using a similarity measure, the intensity differences between the template and image region of the data underneath is calculated. If the result of the matching process between an unknown object (candidate target) in the image data and the template is sufficiently high, the unknown object is labeled as resembling the template. Other techniques use the view-based representations to generate the feature vector descriptions of an object and the image [31]. A matching technique is used to find the correspondence between the image feature description and the object feature vector description [32]. To perform this matching process, an affine transformation is used, which includes translation, rotation, scaling (and in some applications skew) effects, in order to project the object template information back into the image domain from the feature vector space.

The result is a transformed box around the object of interest found in the image [33].

In medical imaging applications template matching techniques are widely used, especially for lung nodule detection [27-28, 31, 34]. We focus a great deal of efforts in this chapter towards the problem of lung nodule detection as a most adequate example of small objects. A preliminary process of lung segmentation (e.g., [11]) is assumed to have been conducted, in order to isolate the lung tissues from CT imaging of the human chest. The detection of nodules is performed on the segmented lungs. The prime goal of lung nodule detection is of course early detection of lung cancer; hence, the emphasis on small-size lung nodules, which, clinically have been determined to be less than 1 cm in diameter.

Survival of lung cancer is strongly dependent on accurate and early diagnosis [35]. In the past two decades numerous screening protocols and studies have been conducted worldwide for the purpose of studying early indications of lung cancer. Low-dose computed tomography (LDCT) scans has lower radiation risk, hence became abundant in lung screening research. In the literature various approaches and methodologies can be found to model the lung nodules, exploiting geometric and texture attributes of the nodules (e.g., [34], [36-38]). In general, there are three key steps for precise candidate nodule detection: 1) a realistic model (template) of lung nodules; 2) an efficient detection methodology; and 3) ability to reduce or eliminate the false positives in the detected candidates.

In this chapter, a lung nodule modeling process that resemble the characteristic shape and texture information of real nodules will be investigated in detail, using Active Shape Modeling (ASM) and Active Appearance Modeling (AAM) approaches. These models

are then used for detection of candidate lung nodules via a formulated template matching technique using several correlation measures known in statistics. After candidate nodule detection, feature vector descriptors using various view-based approaches are constructed for false positive reduction to obtain the final candidate nodules. The first two steps will be explored in detail in this chapter.

A. Lung Nodule Modeling

Farag, 2009 and Farag et al., 2009 and 2010 [11][28][37] used the AAM and ASM methods for modeling the lung nodules, which produced superior results in terms of resembling actual nodule characteristics, and in the sensitivity and specificity of lung nodule detection than parametric methods. This section will examine the process of nodule modeling and simulation using an ensemble of nodules identified by radiologists.

1. Pulmonary Nodule Definitions

The definition of a lung nodule as a small lump of irregular or rounded shape is ambiguous from the perspectives of computer vision and machine learning. In the template design process of data-driven templates, the characteristic shape and texture information are essential in defining and distinguishing the various nodule types from one another. These characteristics must be utilized to extract the most information possible in order to obtain well-defined model representations that are robust to various deformations due to scale, noise, contrast, and local geometric distortion and acquisition artifacts. Samala et al. [38] defined nine feature descriptors that describe the nodule characteristics that were used in assessments by radiologists. These descriptors are: 1) subtlety; 2)

41

internal structure; 3) calcification; 4) sphericity; 5) margin; 6) lobulation; 7) speculation; 8) texture; and 9) malignancy.

The classification of Kostis et al. ([10]) is adopted in this dissertation, which groups the lung nodules into four classes: 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; well-circumscribed, where the nodule is located centrally in the lung without being connected to vasculature; 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.

2. Lung Nodule Model Simulation: Generation of Data-Driven Templates

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 (i.e. Chapter IV). A database of nodules was constructed using the Early Lung Cancer Action Program (ELCAP) [12] database of low-dose lung CT (LDCT) images, where radiologists provided location points corresponding to the nodule locations. The estimated probability density function of the Hounsfield (HU) vs. radial distance distribution of the nodules was examined in [11][37]. A nodule size representation of 10 pixels or about 5 mm from the centroid described the lung nodules in this database. Thus, automatic cropping of a bounding box of size 21×21 pixels was performed to extract the lung nodules from the surrounding tissue. These nodules where then classified into one of the four corresponding categories, described above, 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. Two different sub-databases of 96 nodules (24 nodules per type) are used in constructing the data-driven nodule templates using the AAM and ASM approaches (Figures 19 and 20). Note: the model databases are referred to as MB1 and MB2, respectively. Similar approach may be used with any other ensemble; the LIDC database has larger size nodule population [13].

3. Object Alignment for lung nodule modeling

In Chapter II the AAM approach was described in terms of the computer vision field. This approach commonly known in face recognition systems is manipulated and used to construct simultaneous shape and texture based data-driven lung nodule templates. The databases of nodules per type were annotated to highlight the basic geometric and structural features of the nodules. The Procrustes registration technique was employed to obtain co-registered nodules [14,18]
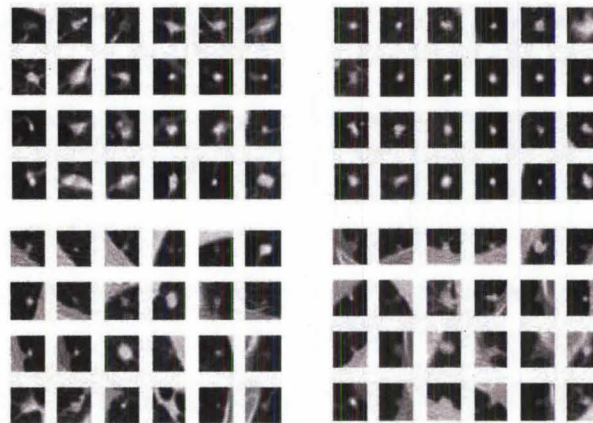
FIGURE 19 - An ensemble of 24 nodules from the ELCAP Clinical Study; Vascular (upper left); well-circumscribed (upper right); pleural-tail (lower left); and juxta-pleural (lower right).
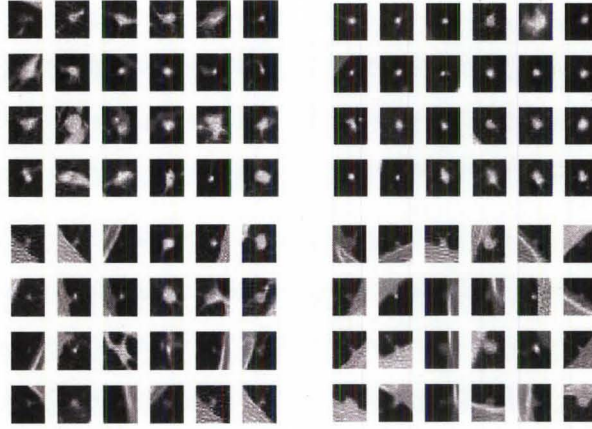
FIGURE 20 - A second ensemble of 24 nodules from the ELCAP Clinical Study; Vascular (upper left); well-circumscribed (upper right); pleural-tail (lower left); and juxta-pleural (lower right).

Let the annotated database of n nodule types be represented as $T = \{G_1, G_2, ..., G_n\}$ where each $G_n = \{s_1^n, s_2^n, ..., s_k^n\}$, where $s_k^n: R^2 \to R$ represents the sample nodule annotated image. In the analysis performed, $n = 4$ and $k = 24$. These images are aligned by employing the Procrustes approach [26], thus, producing n sets of aligned images where, $\tilde{T} = \{\tilde{G}_1, \tilde{G}_2, ..., \tilde{G}_n\}$. The average nodule image for each $\tilde{G}_n$ is computed as $\bar{s}^n = \left(\frac{1}{k}\right) \sum(\tilde{s}_k^n)$. For any set of nodule images indexed by i, its image deviations from its mean are, $\{\bar{G}_1^i = \tilde{s}_1^i - \bar{s}^i, \bar{G}_2^i = \tilde{s}_2^i - \bar{s}^i, ..., \bar{G}_k^i = \tilde{s}_k^i - \bar{s}^i\}$, where each image is of size $N_1 \times N_2$. Each image is converted into a column vector by taking row by row and setting them together as a one dimensional matrix, Ξ, forming the matrices $S_i = [\Xi_1^i \Xi_2^i \Xi_k^i]$ and $\Psi = \frac{1}{k} S_i S_i^T$.

An eigenvalue decomposition is employed to factor the matrix Ψ as $\Psi = U \sum U^T$, where U is an $N \times k$ matrix in which $N = N_1 \times N_2$. Each column vector of the matrix U indexed by j, is converted into a two dimensional image, $\hat{s}_j^i$. From these eigen images,

44

one can formulate an active appearance model of the nodule set images I, $AAM_i = \bar{s}^i + \sum_{j=1}^{m} w_j \hat{s}_j^i$ where $w_j \epsilon R$. The number m is the number of modes to consider. From this model nodule training images can be built as, $I_j^i = \bar{s}^i + \hat{s}_j^i$. The number of modes can be computed from the relation: $\sum_{z=1}^{m} \lambda_z \geq \frac{r}{100} \sum_{z=1}^{k} \lambda_z$, in this dissertation r = 90%, 95% and 97% are considered, and λ_z are the non zero eigen values. This will result in an average of $m = 2, 3$ and 5 eigen images, i.e. eigen nodules, per nodule type to be generated.

Figure 21 shows the average nodule templates and sample generated eigen nodule templates obtained from using MB1. The average shape and texture AAM models generated using MB1 and MB2 are shown in Figure 22. Variations in the developed average models can be seen, since the templates are depictive of the data information provided. The template(s) used for detection can be the average or the average plus a weighted sum of the eigen nodules, resulting in the usage of multiple models per type.

Nodule Type	Average Nodule	1st Eigen Nodule	2nd Eigen Nodule	3rd Eigen Nodule	4th Eigen Nodule	5th Eigen Nodule
Juxta						
Well						
Vascular						
Pleural- Tail						

FIGURE 21 - Generated average nodule templates per nodule type are depicted in column two, while sample eigen nodules per type are shown in columns three thru seven.
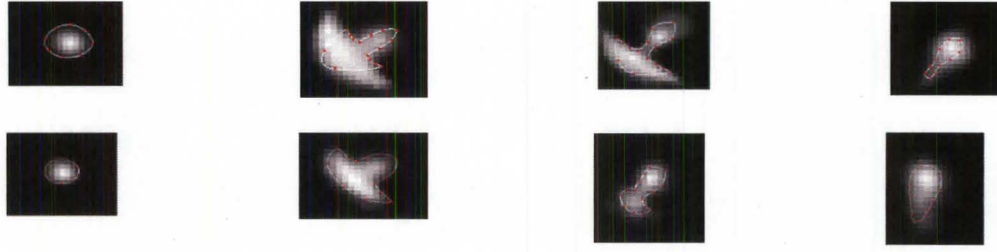
FIGURE 22 - Data-Driven Nodule Template using MB1 top row and MB2 bottom row-
From left to right: well-circumscribed, juxta-pleural, pleural tail and vascular nodule
types.

4. Variational level sets for lung nodule modeling

The ASM approach using variational Level sets [39] in this dissertation 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 (e.g., [40]) eliminates

sources of errors that can arise with manual annotation in which only placing the seed

point or points in the region of the nodule centroid is/are manually performed. Also, the

elasticity of this variational approach addresses the issue of shape variations that can arise

and handles these changes accordingly. A block diagram of the level sets process is

depicted in Figure 23.

Given two shapes represented by the vector functions Φ_1 and Φ_2, a transformation,

A, with scales, rotation and translation is to be calculated to transform the first object to

the second.
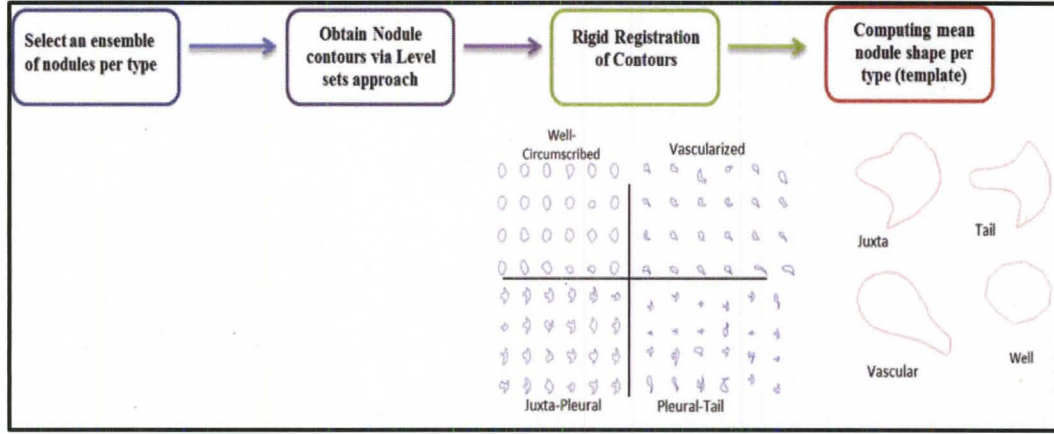
FIGURE 23 - Block diagram of the Level sets process used to generate the mean shapes.

The following dissimilarity, measures the difference between the vector and the other scaled vector function:

$$r = S R \, \Phi_1(X) - \Phi_2(A) \tag{22}$$

The following energy is formulated as a sum of squared differences

$$E = \int_\Omega \delta_\epsilon \, r^T r \, d\Omega \tag{23}$$

where δ_ϵ 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 this work, the contours of the lung nodules, are jointly registered with an evolving mean shape to find the corresponding global transformations $A_1, \ldots, A_n$. The dissimilarity measure is used as follows:

$$r_i = S_i \Phi_M - \Phi_i(A_i) \tag{24}$$

The energy function will be:

$$E(\Phi_M, \Phi_1, \ldots, \Phi_n) = \sum_{i=1}^{n} \int_\Omega \delta_\epsilon \, r_i^T r_i \, d\Omega \tag{25}$$

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

47

$$\Phi_p = \Phi_M + \sum_{i=1}^{n} w_i (\Phi_i - \Phi_M) \tag{26}$$

The shape parameters **w,** need also to minimize the energy function to process the registration. This results in the mean shape contours of the lung nodules. Since texture information is also needed for the template models, a synthetic texture modeling approach, typically used in the generation of parametric templates (e.g., [37]) , (i.e. models generated using geometric equations such as a circle), was implemented. The probability density distributions of the nodules intensities found in the ELCAP database was obtained for each nodule type. The intensity distributions were used to estimate the texture model to be filled into the generated ASM average shape contour templates, by the following equations (e.g., [37]):

$$q(r) = q_{max}e^{-(r/\rho)^2}, \quad 0 \leq r \leq R \tag{27}$$

$$\rho = R\big(\ln(q_{max}) - \ln(q_{min})\big)^{-1/2} \tag{28}$$

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 [11][37].

Figure 24 shows the contour mean shape templates generated and the final models with texture information produced using Equations 27 and 28. In the literature parametric templates, which are generally circular and semi-circular contour shapes with different radii and orientation are used as nodule models. The empirical forms of the intensity (Eq. 27 and 28) are used to fill these templates as performed in the Level sets approach. Figure 25 depicts samples of generated parametric templates.
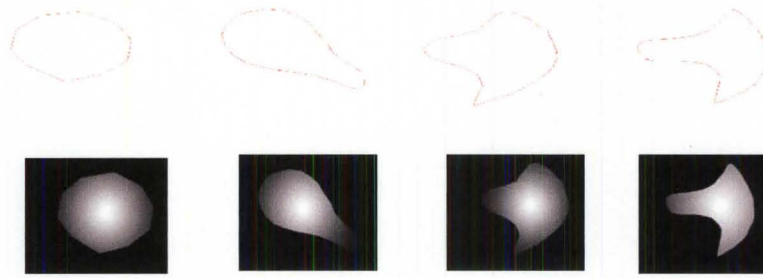
FIGURE 24 - Average shape of the nodules generated from the ASM approach, level sets (Top row). The nodule templates resulting from the intensity equations (lower row). From left to right: Well-circumscribed, vascular, juxta-pleural, and pleural-tail nodule types shape modeling process.
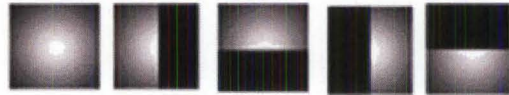
FIGURE 25 - An ensemble of generated circular and semi-circular templates with various orientations (e.g., [11][37]).

5. Lung Nodule Modeling Summary

In this section three approaches for generating lung nodule models or templates was examined. Of these methods two are non-parametric and data-driven while the other approach uses definitions of circular and semi-circular templates to generate parametric models that are filled with texture information. The ASM Level sets method of constructing the nodule models used the same Procrustes registration method implemented in the AAM approach [28] with the difference of the ensemble object being registered. In the level sets approach nodule contours were obtained and registered together to acquire templates containing shape information only, while the AAM approach had both shape and texture information in the resultant templates. The overall mean shapes of each nodule type from the two approaches showed extraordinary

49

resemblance as seen for MB1 in Figure 22 and Figure 24. In the ASM approach, once the mean shape contours were generated, gray level texture information was needed to fill the hollow shells, thus a technique used in [11][37] for filling parametric templates was conducted and the overall shape and texture nodule templates depicted in Figure 24 were produced. The overall bounding box size used for all three approaches was 21×21 pixels which was obtained by analyzing the Hounsfield (HU) vs. radial distance distribution.

B. Lung Nodule Detection: Template Matching

Template matching refers to the process of detecting an object with known prior information such as size, shape and orientation, where detection can be achieved by applying a filter, known as a template, having positive weights in a region that resembles the object to be detected and zero weights in other areas. If the result of the matching process between an unknown object and the template is sufficiently high, the unknown object is labeled as resembling the template, however due to image noise, spatial, amplitude quantization effects, and a priori uncertainty of the exact shape and structure of the object to be detected, an exact match rarely occurs. Hence, a similarity correlation is used, where the maximum or minimum value of unity that occurs if and only if the image function under the template exactly matches the template. Various similarity measures may be used, and various approaches may be taken to execute the matching process as well. Lee et al. 2001 [27], for example, used Genetic Algorithms for executing the matching process on parametric templates. Farag et al., 2009 and 2010 used the normalized cross-correlation (NCC) in evaluating the effectiveness of parametric and non-parametric templates for detection of lung nodules [11][28][37].

1. Similarity Measures for Template Matching

In this dissertation nine similarity measures are considered:

Normalized Cross Correlation (NCC)	$$NCC = \frac{\sum_{(i,j)\in W} I_t(i,j) * I_i(x+i, y+j)}{\sqrt[2]{\sum_{(i,j)\in W} I_t^2(i,j) * \sum_{(i,j)\in W} I_i^2(x+i, y+j)}}$$
Zero-mean Normalized Cross Correlation (ZNCC)	$$ZNCC = \frac{\sum_{(i,j)\in W}\left(I_t(i,j) - \overline{I_t}(i,j)\right) * \left(I_i(x+i, y+j) - \overline{I_i}(x+i, y+j)\right)}{\sqrt[2]{\sum_{(i,j)\in W}\left(I_t(i,j) - \overline{I_t}(i,j)\right)^2 * \sum_{(i,j)\in W}\left(I_i(x+i, y+j) - \overline{I_i}(x+i, y+j)\right)^2}}$$
Sum of Squared Distance (SSD)	$$SSD = \sum_{(i,j)\in W}\left(I_t(i,j) - I_i(x+i, y+j)\right)^2$$
Zero-mean Sum of Squared Distance (ZSSD)	$$ZSSD = \sum_{(i,j)\in W}\left(I_t(i,j) - \overline{I_t}(i,j) - I_i(x+i, y+j) + \overline{I_t}(x+i, y+j)\right)^2$$
Locally scaled Sum of Squared Differences (LSSD)	$$LSSD = \sum_{(i,j)\in W}\left(I_t(i,j) - \frac{\overline{I_t}(i,j)}{\overline{I_i}(x+i, y+j)} I_i(x+i, y+j)\right)^2$$
Sum of Absolute Differences (SAD)	$$SAD = \sum_{(i,j)\in W}\left|I_t(i,j) - I_i(x+i, y+j)\right|$$
Zero-Mean Sum of Absolute Differences (ZSAD)	$$ZSAD = \sum_{(i,j)\in W}\left|I_t(i,j) - \overline{I_t}(i,j) - I_i(x+i, y+j) + \overline{I_t}(x+i, y+j)\right|$$
Locally scaled Sum of Absolute Differences (LSAD)	$$LSAD = \sum_{(i,j)\in W}\left|I_t(i,j) - \frac{\overline{I_t}(i,j)}{\overline{I_i}(x+i, y+j)} I_i(x+i, y+j)\right|$$
Sum of Hamming Distances (SHD)	$$SHD = \sum_{(i,j)\in W} I_t(i,j) \; bitwise \; XOR \; I_i(x+i, y+j)$$

where $I_t(i, j)$ represents the template, $I_i(x, y)$ is the image slice or input image, W is the region of interest on which the similarity measure is evaluated, $\overline{I_t}(.)$ is the mean of the template image, and $\overline{I_i}(.)$ Is the mean of the input image slice. Below we derive the equations of the ZNCC similarity measure, the rest of the measures in the above table can be derived in a similar fashion.

2. Zero-mean Normalized Cross Correlation (ZNCC)

Assuming a template image $I_t: \Omega_t \subset R^2 \to R$ and an input image $I_i: \Omega_i \to R$ the ZNCC between the CT image slices and nodule template images is defined as:

$$\xi(X_i) = \int_{\Omega_t} (I_t(X_t) - \mu_t)(I_i(X_t + X_i) - \mu_i) d\Omega_t \Big/ \sqrt{\int_{\Omega_t} (I_t(X_t) - \mu_t)^2 (I_i(X_t + X_i) - \mu_i)^2 d\Omega_t} \quad (29)$$

Where $X_i \in \Omega_i$, $X_t \in \Omega_t$, the template image average is $\mu_t = \int_{\Omega_t} I_t(X_t) d\Omega_t \Big/ \int_{\Omega_t} d\Omega_t$, and the corresponding space intensity average in the input image is defined as:

$\mu_i = \int_{\Omega_i} I_i(X_t + X_i) d\Omega_t \Big/ \int_{\Omega_t} d\Omega_t$. Assuming the functional $\varphi: R^2 \to R$ such that for any two images I_1 and I_2, the function is written as:

$$\varphi = (I_1, I_2) = \int_{\Omega_1} (I_1(X_1) - \mu_1)(I_2(X_1 + X_2) - \mu_2) d\Omega_1$$

where the average parameters and image domains are defined in Section 2.

The similarity measure can be written as:

$$\xi(X_i) = \varphi\big(I_t(X_t), I_i(X_t + X_i)\big) \Big/ \sqrt{\varphi(I_t(X_t), I_t(X_t))\varphi(I_i(X_t + X_i), I_i(X_t + X_i))} \quad (30)$$

The goal is to maximize such a similarity measure over the whole domain of the input image Ω_i. Using the template image as is will have limited results on that measure especially when the problem encounters rotations and scale variations. However, one can incorporate a geometric transformation process on the template image such that the

parameters of that transformation are selected to maximize the similarity ξ_i. Assume a geometric transformation function $A(X_t) = SRX_t$ where S and R are the scales and rotation matrices. For each point, X_i, there exists a geometric transformation A_i that maximizes the ZNCC similarity measure. The similarity measure is re-written after incorporating the transformation as:

$$\xi(X_i) = \frac{\varphi(I_t(X_t), I_i(A_i + X_i))}{\sqrt{\varphi(I_t(X_t), I_t(X_t))\varphi(I_i(A_i + X_i), I_i(A_i + X_i))}} \tag{31}$$

Using, for example, the AAM eigen shapes as the models, I_t, the template function is expressed in terms of the weighted eigen shapes $I_t = \bar{I} + \sum_{j=1}^{m} w_j I_j$. The transformation parameters as well as N-weighting coefficients are required to maximize the similarity measure. The gradient descent approach is used in this optimization problem. The transformation parameters will then be computed as follows:

$$\frac{\partial p}{\partial t} = \frac{\varphi(I_t, I_i)\varphi(I_t, (\nabla I_i)^T \nabla pA) - \xi^2 \varphi(I_t, I_t)\varphi(I_t, (\nabla I_i)^T \nabla pA)}{\xi \varphi(I_t, I_i)\varphi(I_t, I_t)} \tag{32}$$

where p is the scale or rotation parameter of the transformation A_i, $I_t = I_t(X_t)$, and $I_i = I_i(A_i + X_i)$. Also, the weighting coefficients will change through the following PDE:

$$\frac{\partial w_n}{\partial t} = \frac{\varphi(I_t, I_i)\varphi(I_j, I_i) - \xi^2 \varphi(I_i, I_i)\varphi(I_t, I_j)}{\xi \varphi(I_t, I_i)\varphi(I_t, I_t)} \tag{33}$$

These parameters change from an initial position and iteratively reach a steady state solution where a maxima is achieved.

Summarizing the above approach: for every point in the input image domain, a ZNCC similarity measure is computed after finding the corresponding weighting coefficients and transformation parameters that maximizes such a criteria. Local maxima points are

determined by a thresholding criterion to find the best matches. The above formulation was derived for all similarity measures (ZSSD, NCC, etc.) in a straightforward manner.

C. Experimental Results

1. Nodule Modeling Evaluation

This first study is based primarily on the Early Lung Cancer Action Program (ELCAP) public database which was used in this dissertation 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 397 nodules are provided by 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. Of the 397 nodules 291 identified and categorized nodules are used in the detection process.

In the first set of experiments the sensitivity of the data-driven nodule templates constructed versus parametric templates are examined. 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 a value of 0.5 is approached. 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 which may be devised to select an optimal threshold. In this experiment the threshold of the NCC is set to 0.5.

Results using only the average (mean) template models generated from the data-driven approaches are examined in this first set of experiments. The overall sensitivity and specificity was computed using equations 34 and 35:

$$Sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negatives} \qquad (34)$$

$$Specificty = \frac{True\ Negatives}{True\ Negatives + False\ positives} \qquad (35)$$

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 shows the results with the templates centered with respect to the x-axis (i.e., zero orientation). The performance of the AAM approach is much more robust than the parametric templates and level sets method. The sensitivity of the well-circumscribed nodule types, which are nearly isotropic, shows improvements for both the Levels sets and AAM approach. Overall the Level sets approach was comparable to the parametric template method.

Experimentation on template orientation effect on the detection process was also examined. Table 2 shows the sensitivity and specificity of the non-parametric templates as compared with the parametric templates for orientation change (i.e. rotation transformation of the templates from $0°$ to $360°$ with step-size $90°$). Overall in the case of parametric templates sensitivity increased for all nodule types. The Level sets sensitivity also increased overall for each nodule type while specificity decreased with the addition

of orientation, the general results when compared to the parametric show the level sets approach more specific with orientation. Both sensitivity and specificity decreases were observed from the AAM approach when orientation was applied, yet, sensitivity and specificity rate were still much improved over the other two methods.

TABLE 1- Overall sensitivity and specificity analysis using ASM average templates, parametric models and AAM average templates, without accounting for 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 MB1and no orientation	72.16%	98.11%
AAM Approach using MB1 and no orientation	85.22%	97.81%
AAM Approach using MB2 and no orientation	83.51%	98.36%

TABLE 2 - Overall sensitivity and specificity of level sets, parametric, and AAM average templates averaging a number of orientations for the templates.

Algorithm	Sensitivity	Specificity
Parametric Approach with template radius 10 and orientation 0°- 360° with step-size 90° for semi-circular template	78.01%	96.41%
Level Sets Approach using nodule contours from dataset 1and orientation 0°- 360° with step-size 90°	76.98%	97.63%
AAM Approach using MB1 and orientation 0°- 360° with step-size 90°	86.94%	96.51%
AAM Approach MB2 and orientation 0°- 360° with step-size 90°	83.51%	97.40%

The AAM approach using the mean templates from MB1 provided better results than the parametric, level sets method and the AAM mean templates from MB2. Overall, the AAM algorithm using either set of mean templates generated from sub-database 1 or 2 yielded 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. Further studies were conducted using the AAM based approach on both sub-datasets used for template modeling (i.e. MB1 and MB2). The number of annotation points necessary for proper registration was found to be a function of how many were necessary to depict the main discriminatory shape information that withholds substantial texture information commonly found in each nodule for that particular type (i.e. the 24 nodules used per type). Also, in the modeling process 24 nodules per type were used, but what would be the effect of using fewer nodules in generating the mean templates and these templates be used in the detection process? This question was also examined using the AAM approach and MB1 and MB2 to generate average nodule representations with different sub-database sizes (i.e. 24 and less nodules per type).

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 from MB1 results in overall similar sensitivity and

specificity results; while in the case of using MB2, the results reduced in sensitivity as the number of nodules used for modeling decreased.

TABLE 3 - Overall sensitivity and specificity of the AAM approach using 8 and 16 nodules per type for nodule modeling instead of 24 nodules per type, respectively.

Algorithm	Sensitivity	Specificity
AAM Approach using 16 nodules from MB1 for modeling and no orientation	85.57%	97.84%
AAM Approach using 8 nodules from MB1 for modeling and no orientation	84.88%	97.99%
AAM Approach using 16 nodules from MB2 for modeling and no orientation	83.16%	98.42%
AAM Approach using 8 nodules from MB2 for modeling and no orientation	77.32%	98.57%

Overall the findings are dependent on how well the lung nodules are annotated and which nodules represented in the sub-database affected the generation of the mean templates per type; i.e. if the reference image used for alignment depicted a majority of the nodules in the larger database and all nodules were well annotated, then the overall mean templates generated yielded improved sensitivity and specificity results. If the nodules were not adequately annotated or the reference nodule did not represent the overall model database of lung nodules, then the mean templates generated did not always provide improvements in sensitivity and/or specificity. The results are expected to be further enhanced using larger ensemble sizes than the 24 per nodule types used in the experiments in this dissertation. Likewise, involvement of several radiologists to create the ensemble may also lead to further improvements. These results in this set of experiments are also based from only using the NCC similarity measure at a single

threshold; results can vary when using other similarity measures, thresholds and transformation parameters (i.e. rotation and scale).

2.Clinical Evaluation

This work is based on three lung CT database scans. The first is the ELCAP which as stated before, contains 50 sets of low-dose CT lung scans (tube current 30-40 mA) taken at a single breath-hold with slice thickness 1.25 mm and resolution 0.5x0.5mm. 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. Database 2 is the Lung Imaging Data Consortium (LIDC)[13], which contains 1018 helical thoracic CT scans from 1010 different patients. The tube current ranged from 40 to 627 mA with slice thickness range 0.6mm to 5.0 mm. The data consisted of CT lung images taken after intravenous (IV) dye contrast was administered to the patients, but most scans where imaged without any dye enhancing agents given to the patients. From the first database 291 nodules in the size range of 2-5 mm representing the four types of nodule categories specified were considered in the detection testing, while 316 nodules from database 2 in the range of 3-4 mm were considered for testing in this set of experiments. Database 3 contains 8 sets of LDCT lung scans with slice thickness 1.25 mm and 2.5 mm, with a small set of lung nodules provided by radiologists.

The average and eigen nodule templates generated using the AAM approach are used to test the effectiveness of the templates in respect to each other by implementing a decision fusion template matching approach where the four average nodule templates were one set of models. The 90% PCA criterion eigen nodules were a second set of

templates. The third set of models were the 95% PCA criterion eigen nodule representation, and finally the 97% PCA criterion eigen nodule templates. In the eigen nodule cases multiple images per type represented the nodules (i.e. if 90% PCA criterion resulted in two templates per type then eight total models were used in detection). The templates were used in a serial fashion and the final decision is the XOR of the binary outputs.

The Receiver Operating Curves (ROC) depicting true positive versus false positive rates were obtained using the different template models and nine similarity measures. True positives refers to the candidate nodule detected as correctly identified when the distance between the detected point and the closest ground truth point is smaller than the template radius, while all other incorrectly detected points are considered false positives.

Table 4 depicts the detection results for the all three databases. The first database is also correlated to the ROC curves shown in Figure 26 using the average templates (Figure 26a), 90% PCA criterion generated eigen nodule templates (Figure 26b), 95% PCA criterion generated eigen nodule templates (Figure 26c) and 97% PCA criterion generated eigen nodule templates (Figure 26d). The area under curve (AUC) was computed for each similarity measure.

Analyzing the results from the first database the detection process provided several conclusions. First, comparing the eigen nodule experiments to each other the overall conclusion can be made that more nodule templates does not provide a very large increase variation in the detection process except when using the SSD and ZSAD similarities. In the SSD case the 97% PCA criterion provided 0.10 increase in the detection AUC computation which is one of the largest increases seen. Thus, considering

60

additional eigen nodule template models per nodule type did not result in significant increases in the detection process and as such considering only 90% PCA criterion to produce the eigen nodule templates is sufficient.
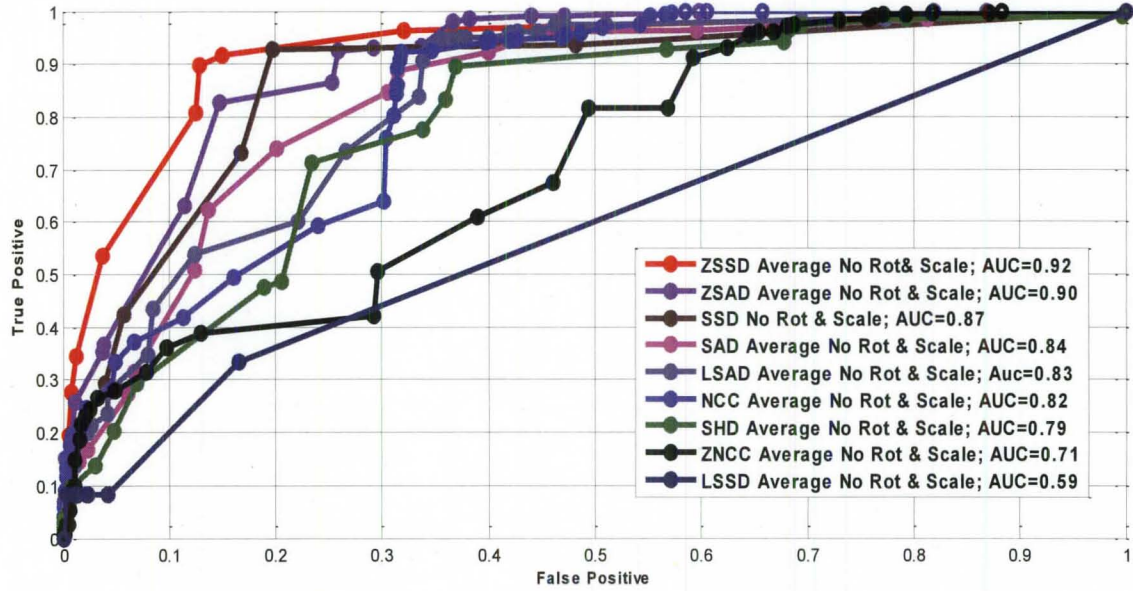
(a)
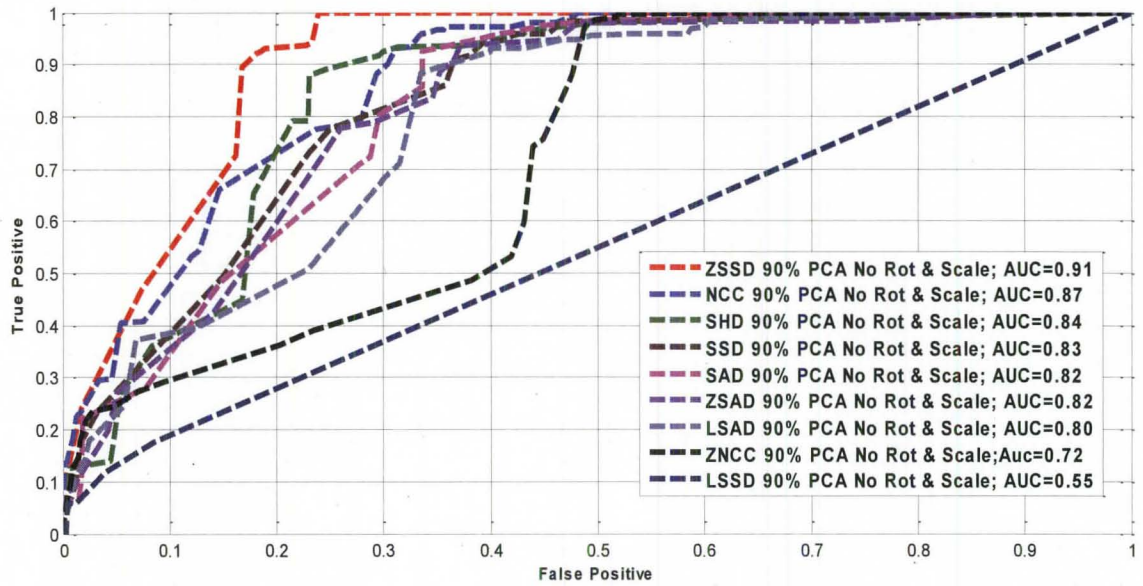
(b)
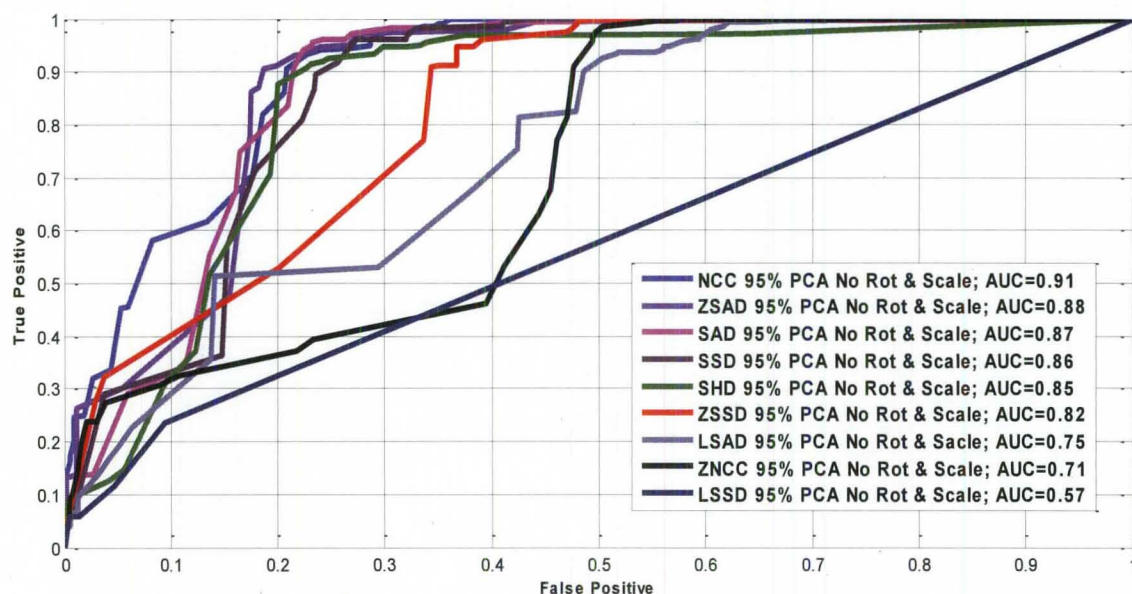
(c)
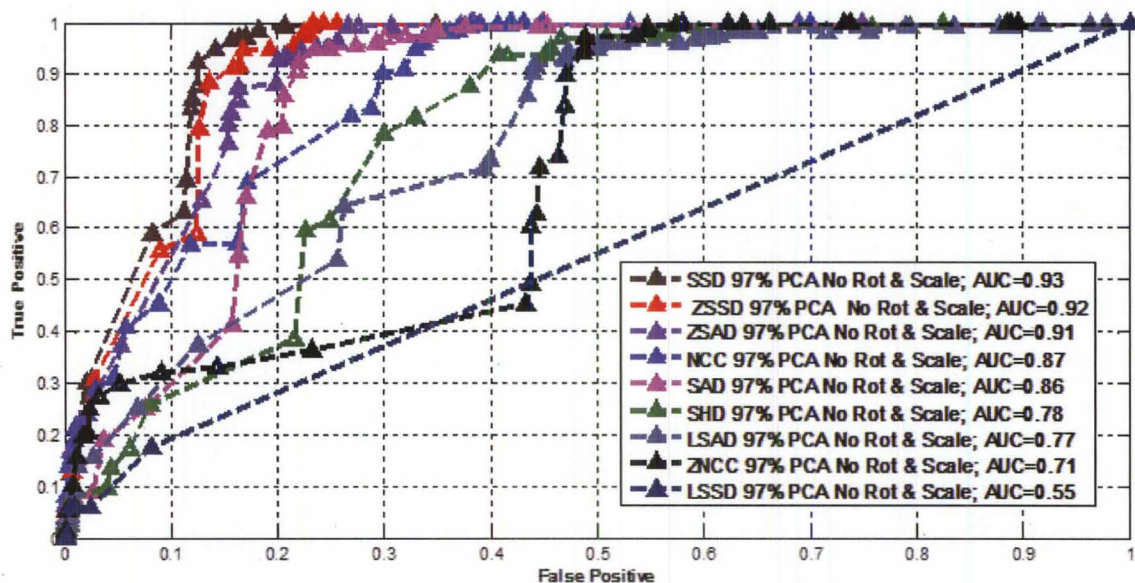
(d)

FIGURE 26 - ROC curves for template matching detection on Database 1 using nine similarity measures and a) the average nodule templates (4 templates one per type); b) eigen nodule templates generated using 90% PCA criterion; c) eigen nodule templates generated using 95% PCA criterion and d) eigen nodule templates generated using 97% PCA criterion.

Comparing the average and 90% eigen nodules for all similarity measures shows overall similar results; variations in AUC computation were less than 0.05. The top four similarity measures that provided the highest detection rates when testing on database 1 were the ZSSD, ZSAD, SSD and NCC. The NCC out-performed its zero-mean counterpart, ZNCC, in the four experiments conducted on database 1. Since the nodule models produced in this dissertation considered only training data from the first database and the results in Figure 26 test on the same low-dose CT datasets, it can be validated why the ZSSD and SSD provides better results. The ZSSD computation relies on the squared differences followed by the aggregation between the pixels within a square neighborhood between the template image and the image slice.

The results using database 2 showed the average nodule templates using the NCC and ZNCC 90% PCA nodule templates, were of the best detection rates. Overall the detection process for this database required more nodule model variations. As can be seen from the table, the 95% eigen nodule experiments provided better detection results using six out of the nine similarity measures.

Final detection results shown for database 3 are similar to that seen from database 1. When comparing the eigen nodule experiment cases, the 90% PCA criterion eigen nodule templates generated resulted in better detection rate over its 95% and 97% nodule template models counter-parts. The ZNCC and NCC similarity measures provided the best detection rates which were slightly higher in the average nodule template experiment then the 90% eigen nodule case.

Comparing the area under curve (AUC) results in Table 4 for all databases allows a very important conclusion to be made between the CT databases used in this dissertation:

databases 1 and 3 (DB1 and DB3) have CT data information that do not have very large distinctions in the scanning parameters among them. Thus, the nodule templates generated from DB1 where able to provide acceptable detection results on both studies. This is not the case with database 2 (DB2), where AUC showed values that do not exceed the mid-0.80 (mid 80%) range for all similarity measures and variations of nodule templates. In addition, since the image slices in DB2 were from both pre and post IV CT lung images, and the models where only generated from pre non-dye administered slices; this could further explain the discrepancies in the results as compared to DB1.

TABLE 4 - Computed AUC for the template matching detection process with nine similarity measures on databases 1, 2, and 3.

Measure	Database 1				Database 2				Database 3			
	Average	90%	95%	97%	Average	90%	95%	97%	Average	90%	95%	97%
NCC	0.82	0.87	0.91	0.87	0.81	0.76	0.74	0.79	0.91	0.90	0.88	0.84
ZNCC	0.71	0.72	0.71	0.71	0.78	0.80	0.75	0.76	0.93	0.92	0.90	0.91
SSD	0.87	0.83	0.86	0.93	0.63	0.73	0.75	0.69	0.81	0.88	0.94	0.87
ZSSD	0.92	0.91	0.82	0.92	0.72	0.76	0.78	0.74	0.79	0.95	0.79	0.85
LSSD	0.59	0.55	0.57	0.55	0.53	0.54	0.55	0.54	0.56	0.51	0.48	0.51
SAD	0.84	0.82	0.87	0.86	0.74	0.68	0.74	0.69	0.74	0.87	0.77	0.74
ZSAD	0.90	0.82	0.88	0.91	0.76	0.68	0.79	0.74	0.68	0.93	0.86	0.82
LSAD	0.83	0.80	0.75	0.77	0.78	0.77	0.71	0.84	0.86	0.83	0.82	0.78
SHD	0.79	0.84	0.85	0.78	0.73	0.68	0.76	0.69	0.72	0.83	0.79	0.78

The LSSD similarity provided the worst results in all experiments for each database and can thus be deemed not adequate for use in lung nodule detection. The NCC and

ZNCC correlation similarity measures showed their benefits in the 2nd and 3rd databases where the CT data did not have weight in generating the template models. In Table 4, the results are shown for the four sets of generated nodule template variations. Highlighted values represent the best computed AUC obtained when using the model described in the top row for the similarity measures in the first columns.

D. Conclusions

In this chapter the detection process was investigated. The design of data-driven templates for lung nodule detection were constructed using ASM and AAM models that provided realistic shape and texture representations of four types of lung nodules, found in CT images. The new templates were used for detection of candidate nodules via template matching, where sensitivity and specificity performance analysis was conducted using these templates and parametric templates generally known in the lung detection literature.

The effect of template shape on detection of different nodules types was studied. It is concluded from extensive experimentation in this dissertation that the data-driven AAM algorithm for template matching 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 all radii sizes between 1 and 20 pixels were tested the sensitivity was higher but the specificity in comparison to the data driven nodule templates was still lower.

The overall performance depended upon template shape and nodule type. The well-circumscribed nodule was the least sensitive nodule yet it emphasized the greatest improvement as shown in the above tables. The pleural tail in both the parametric and

data-driven templates yielded the greatest sensitivity. Overall, the Level sets approach is comparable to the parametric templates in terms of sensitivity and specificity for all nodule types using the ELCAP database.

In the second set of experiments the average nodules per type generated by the AAM approach and the eigen nodules using 90% PCA criterion, 95% and 97% PCA criterion were used for detection. The eigen nodule representations produced multiple models per nodule type. The effectiveness of the generated templates was tested for detection via a formulated weighted template matching approach using nine different similarity measures. Validation of the generated templates robustness for nodule detection was tested by implementing the detection process on three lung CT databases. The results showed that the LSSD similarity measure provided the worst detection results for all experiments conducted on each database. The NCC and ZNCC similarity measures when used in the experimentations on the 2nd and 3rd databases showed among the best detection performances.

CHAPTER IV

STATISTICAL AND VARIATIONAL OBJECT SEGMENTATION WITH APPLICATION TO NODULAR REGION SEGMENTATION

This chapter describes the segmentation process using statistical as well as variational approaches. This is one of the most challenging processes under the umbrella of computer vision. It is important for modeling and a basic step in object visualization. The chapter handles the intensity modeling process to describe region segmentation using a linear combination of Gaussians for probability density functions. Also, neighborhood pixels interaction will be take in consideration. Also, variational level set segmentation will be illustrated to handle the issue of adding prior shape information to the problem. Several experimental results will be demonstrated for segmenting lung nodules of different sizes ranging from small to large.

A. Introduction

Figure 27 provides illustrative examples to images with varying complexity. The first row is a star object inside a fixed background. The features of the object (i.e., corners of the star) are well-defined, and the gray level histogram (shown in the middle) contains only two peaks; hence, segmentation is straightforward, and can be accomplished by using a threshold in between these peaks (the binary image is shown in the right). The second row shows the star image with additive white noise.
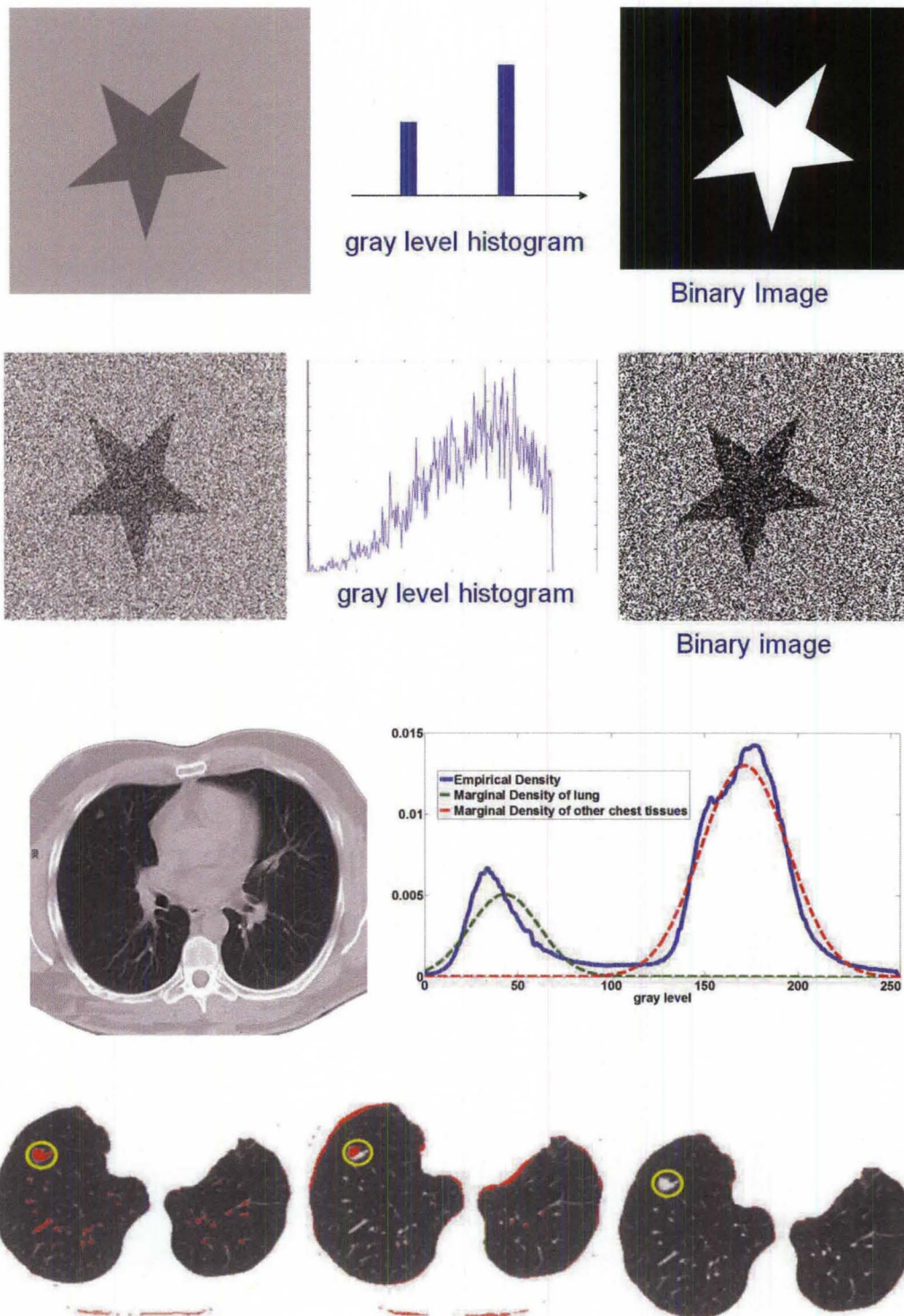
FIGURE 27 - Illustrative examples of multimodal images.(a) A star object with binary histogram; (b) The star object with added white Gaussian noise; (c) A slice from a low dose CT (LDCT) scan of the lung; (d) Lung nodule segmented by three different methods; yellow circles are inserted over the region of interests (ROI).

The uncertainties about the gray level distribution of the star shows in the disappearance of the sharp peaks in the histogram (middle) and, subsequently, a threshold-based segmentation won't be as accurate as the noise free case. Still, the corners of the star are apparent, and inclusion of a priori shape information about the star may be used to enhance the segmentation. The image in the third row is a CT slice of the human chest.

As the image contains chest (ribs), heart, fluids and lung tissues with different X-ray absorption in the CT scanning, the gray level histogram will be multimodal, yet the peaks are defined within modes corresponding to the major image constituents; lung tissue and others. The multimodal nature of the image histograms may suggest, at a first look, that a straight forward segmentation, such as thresholding or deterministic region growing, is possible (e.g., [41][42]). Unfortunately, this is not the case and a more elaborate scheme is required to segment the lung tissues from the rest of the chest (e.g., [43][73])

A severe complication comes into play when the ROI in the CT scan is the nodules within the lung tissue or attached to the pleural surface. Such nodules are random in shape, size and may be located anywhere in the lung tissue. Straight forward segmentation of these nodules will not be possible; indeed a series of steps are usually used for that purpose, and a different casting of the problem is used. In segmenting lung nodules, a detection step is used to narrow down candidate ROIs where these nodules may fall, then a false positive step is performed to reduce the number of candidate nodules, and finally, an elaborate operation (which in itself may involve several steps) is used to segment (produce an outline) of the candidate nodules (e.g., [74]). Bottom row of Figure 27 shows nodules detected and segmented in the CT slice of the third row.

For more than two decades, several research projects have been investigating the efficiency of lung cancer screening with low dose CT on asymptomatic people at high risk [44]. The main feature for the characterization of indeterminate small (< 1cm) screening detected lung nodules is the growth of size [45-46]. Nodule size is an important factor in volumetric analysis of lung nodules [47]. It has been shown clinically that size is linked to nodule malignancy, with non-calcified nodules larger than 2cm in diameter having a higher rate of malignancy than smaller nodules [48]. Size computation through a 3D measurement is usually performed by applying volumetric methods to a segmentation result. However, lung nodules segmentation in CT imaging is a complex and challenging process; one of the most important problems arises from the frequent attachments of the nodules to other anatomical objects.

The lungs are a complex organ which includes several structures, such as vessels, fissures, bronchi or pleura that can be located close to lung nodules. Also, the main "head" of the nodule is what radiologists consider when computing the size. In the case of detached nodules (i.e. well-circumscribed nodules) the whole segmented nodule is considered in size computations and growth analysis, while in detached nodules (i.e. juxta-pleural, vascularized and pleural-tail) the "head" is required to be extracted from the anatomical surrounds. For example, in the juxta-pleural case the main shape information attached to the pleural-wall is the "head" of the nodule and is used in size and growth follow-up computations which can aid in deciding malignancy.

Intensity-based segmentation [49-50] has been successfully applied to the nodule segmentation problem using local density maximum and thresholding algorithms. These classes of algorithms are primarily effective for solitary nodules (well-circumscribed),

however, fail in separating nodules from juxtaposed surrounding structures, such as the pleural wall (i.e. Juxta-Pleural and Pleural-Tail nodules) and vessels (Vascular), due to their similar intensities. More sophisticated approaches have been proposed to incorporate nodule-specific geometrical and morphological constraints to address this issue (e.g. [10][51-53]). However, juxta-pleural, or wall-attached, nodules still remain a challenge because they can violate geometrical assumptions and appear frequently.

In CT slices the rib areas tend to appear with high intensity values, which cause other difficulties when using only intensity-based segmentation approaches. These high-intensity regions near a given marker can bias the semi-automatic nodule center estimator resulting in inaccurate computations of size.

Robust segmentation of the juxta-pleural cases can be addressed in two approaches: a) global lung or rib segmentation (e.g., [54]) and b) local non-target removal or avoidance [50]. The first can be effective but also computationally complex and dependent on the accuracy of the whole-lung segmentation. The second is more efficient than the former but more difficult to achieve high performance due to the limited amount of information available for the non-target structures. Another set of approaches have been proposed in the literature but require excessive user interaction to obtain acceptable results (e.g. [44][55]). Some approaches assumed predefined lung walls before segmenting the juxta-pleural nodules but in cases where wall movement occurs, due to breath motion, the approach fails [56].

The rest of this chapter demonstrates statistical image modeling in addition to geometric image segmentation including shape priors. This work shows in detail the probability density function estimation as a linear combination of Gaussians and its use in

image segmentation. This approach works very well for homogeneous regions or when the image has little in-homogeneities. Beyond these circumstances, bad results are expected. Thus an approach that allows the use of shape priors is required. The level set segmentation helps include prior shape models to enhance the segmentation results especially with lung nodules which have many challenges as listed above. The intensity as well as prior shape segmentation using level sets will be demonstrated. Exhaustive experimental validation will show the efficiency of the proposed technique.

B. Image Modeling

The goal of image modeling is to describe the visual characteristics of the image. Stochastic approaches, particularly random field models, have been used with impressive success to model various types of images. These models have been incorporated in various image processing and image analysis tasks, including filtering, coding and segmentation, in the past three decades (e.g., [57-62]). Objects-of-interest in an image are characterized by geometric shapes and visual appearance, although it is very difficult to formally define these notions.

A priori information about the objects in the image is very crucial for robust performance of various tasks in image understanding. As a simple illustration, Figure 27(b) shows a star object in which an additive Gaussian noise would frustrate the segmentation process. A priori information about the objects' shape (i.e., star) will provide a clue for the segmentation process. Modeling the appearance (intensity information), spatial inter-dependence (spatial interaction), and a priori information is an active research area, and a number of issues related to incorporating these models into

image understanding have not been resolved due to the ill-posed nature of the specific problems, e.g., segmentation.

1. Modeling for Segmentation

The strategy followed in modeling images for the purpose of segmentation hinges on: a) proper estimation of the joint and marginal probability densities from the intensity information, b) proper modeling of the spatial inter-dependence of pixel intensities with respect to neighboring pixels, and c) formulating the segmentation process in terms of a sequential set of components that govern the objects characteristics in an image; e.g., intensity, texture, and shape, which may be implemented in an iterative, yet cooperative fashion. Of particular interest to the development in this work are approaches for modeling the intensity, spatial interaction and shape information, which enable a generic formulation of image segmentation as a minimization of an "energy" function. The segmentation approach in this chapter is unsupervised, and aims at obtaining a maximum a posteriori (MAP) estimate for the spatial support (labels) of the classes in the image through incorporation of region processes (low level) and class spatial support (high level), and a priori information about the shape of the objects.

The conventional approach for model-based image segmentation describes the input image and the desired map (labeled image) by a joint Markov-Gibbs random field (MGRF) model of independent image signals and interdependent region labels. Let $\mathbf{G} = \{0, ..., Q - 1\}$ and $\mathrm{L} = \{1, ..., \mathrm{K}\}$ denote sets of gray levels g and region labels l, respectively. Here, Q is the number of gray levels, and K is the number of image modes, i.e. peaks in the gray level frequency distribution. Assume that each dominant image mode corresponds to a particular class of objects to be found in the image. Let P be the

set of image pixels, then one can define the gray level image $I: P \rightarrow G$ and its desired

map $f: P \rightarrow L$. A two-level probability model of the original image and its desired map is

given by a joint distribution

$$P(I, f) = [P(f)P(I|f)] \tag{36}$$

where $P(I|f)$ is a conditional distribution of the original image given the map and $P(f)$

is an unconditional probability distribution of the map. The Bayesian *maximum-a-*

posteriori (MAP) estimate the map f, given the image I:

$$f^* = \arg\max_{f \in F} L(I, f) \tag{37}$$

where F is the set of all region maps with labels $l \in L$ on P, maximizes the log-likelihood

function:

$$L(I, f) = \log P(I|f) + \log P(f) \tag{38}$$

To find this log-likelihood function, one needs to estimate the conditional $P(I|f)$ and the

unconditional $P(f)$ image models, and identify their parameters. The following section

considers modeling the two components of Eq. 38 separately before embarking on the

techniques of segmentation. The first comonent is denoted as representing the texture or

appearnce of the region (sometimes this is referred to as the low level process in the

literature) and the second component the labeling (or the high level process).

2. **Density Estimation as a Linear Model**

This section examines parametric density estimation using a variable weighted sum of

Gaussian kernels, where the weights may take positive and negative values. Various

statistical properties of the estimator are studied as well as its extensions to

multidimensional probability density estimation. Identification of the estimator

parameters are computed by a modified EM algorithm and the number of kernels are estimated by information theoretic approach, using the Akiake Information Criterion (AIC). Empirical evaluation of the estimator is provided with respect to window-based estimators and the classical linear combinations of Gaussian estimator that uses only positive weights, showing its robustness (in terms of accuracy and speed) for various applications in image and signal analysis and machine learning. Numerical methods for estimating the probability density function (PDF) of a random variable X (random vector in general) are important in various signal and image analysis applications. Such estimates form the basis of optimal filtering, synthesis and segmentation of an image or a signal. Indeed, PDF estimation is foundational in Bayesian statistics and huge number of machine learning applications [63].

Given a random sample $D = \{X_1, X_2, X_3... X_N\}$ from a particular distribution with PDF $f_X(x)$ that is assumed to be continuous over a domain $\in [a, b] \subset R$. Hence, the modes of $f_X(x)$ (i.e., the minima and maxima) are in the closed interval [a, b]. The estimate $\hat{f}(x)$ of the PDF could be obtained by a number of methods, including the naïve estimator, the k-NN and variable (generalized) kernel estimator [63]; a kernel estimator has the following form: $\hat{f}(x) = \frac{1}{Nh} \sum_{i=1}^{N} K(\frac{x-x_i}{h})$ where $K(.)$ is a symmetric PDF and h is the "window" width. There are various kernels (windows) that have been studied in the literature including the uniform and Gaussian PDFs [63-64]. In particular, the differentiability, integrability and tail properties of the kernel-based estimator are well-studied in the statistical literature [63]. In addition, the quality of the estimator is examined in terms of various statistical properties such as the consistency and the effect

of samples size. An alternative to the kernel estimator is the so-called mixture model, where the underlying density is assumed to have the form:

$$\hat{f}(x) = \sum_{j=1}^{C} w_j \, \emptyset(x|\theta_j) \tag{39}$$

where $\emptyset(x|\theta_j)$ is the jth kernel function, w_j is the corresponding weight and C is the number of components of the mixture. Unlike kernel estimator where the number of components N is number of data points, in a mixture the number of components C can be pre-specified based on some a priori knowledge about the nature of the subpopulations in the data (dominant modes of the data). The Gaussian function is amongst the most used kernels in the above model; the resulting PDF estimate is known as linear combination of Gaussian (LCG).

The LCG possesses various computational and mathematical characteristics that makes it attractive, in particular its integrability, continuity, and the fact that a well-established suboptimal approach for estimation of the parameters $\theta_j, j \in [1, C]$ exists via the Expectation-Maximization (EM) algorithm (e.g., [65-66]). To provide a better fitting for the empirical density, Farag et al, 2006, [43], introduced a modification to the standard linear model above, in which the weights were allowed to take positive and negative values. The new LCG model (will be denoted by LCG$_2$, whereas the model in Eq. 39 will be denoted by LCG$_1$) has the following form:

$$\hat{f}(x) = \sum_{i=1}^{C_p} w_{p,i} \, \emptyset(x|\theta_i) - \sum_{j=1}^{C_n} w_{n,j} \, \emptyset(x|\theta_j) \tag{40}$$

Where, C_p and C_n are the number of positive and negative kernels used in the linear model, and $w_{p,i}$ and $w_{n,j}$ are the corresponding weights. In order for the LCG$_2$ estimate to be a real density model (i.e., positive and integrates to one), a condition on the weights is imposed such that:

76

$$\sum_{i=1}^{C_p} w_{p,i} - \sum_{j=1}^{C_n} w_{n,j} = 1. \tag{41}$$

It has been proven that the new model provides a better fitting for the empirical density

[43].

This section examines various statistical properties of the new LCG_2 model. In

particular the validity of the condition in Eq. 41 is studied by: evaluating the robustness

of the model with respect to mode resolutions in the empirical PDF, extensions to higher

dimensions are introduced, and the performance of the LCG_2 model with respect to other

traditional kernel-based PDF estimators are compared.

3. Generalized Linear Combinations of Gaussian PDF Estimator

As the models in Eq. 39 and Eq. 40 have been used for both marginal and joint PDFs,

below a clear distinction is shown for the two scenarios.

a. *Marginal PDF Estimation*

Let the sample $D = \{X_1^k, X_2^k, \cdots, X_N^k\}$ be drawn from a *single* distribution; random

variable $\mathbf{X}^k$ with PDF $f_{X^k}(x)$ or $f_k(x)$. The LCG_2 model in Eq. 40 may be rewritten as:

$$\hat{f}_k(x) = \sum_{i=1}^{C_{p,k}} w_{p,i} \, \emptyset(x|\theta_i) - \sum_{j=1}^{C_{n,k}} w_{n,j} \, \emptyset(x|\theta_j). \tag{42}$$

b. *Composite Density Modeling (set of marginals)*

Suppose there is a composite density function of m random variable. A common

scenario, the histogram of realizations represents weighted marginal densities of a

number of classes. The goal would be to separate the marginal densities, each may be

modeled by LCG_1 or LCG_2, in order to obtain segmentation threshold, for example.

Using LCG_2, this composite density model will be written as:

$$\hat{f}(x) = \sum_{i=1}^{C_p} w_{p,i}\,\emptyset(x|\theta_i) - \sum_{j=1}^{C_n} w_{n,j}\,\emptyset(x|\theta_j) \tag{43}$$

In this case, one can relate Eq. 40 and Eq. 43 by defining $C_p = \sum_{k=1}^{m} c_{p,k}$ and $C_n = \sum_{k=1}^{m} c_{n,k}$.

Integrability: As both sides in Eq. 40 must integrate to 1; hence, $\sum_{i=1}^{C_{p,k}} w_{p,i} - \sum_{j=1}^{C_{n,k}} w_{n,j} = 1$ is immediate since $\emptyset(x|\theta_i)$ is a PDF; however, the integration should be conducted over the entire real line. In practice, the ensemble D is finite-length, i.e., $f_k(x)$ has a finite domain; $x_i^k \in [a,b] \subset R$; hence, the integrability of $\hat{f}_k(x)$ will be assured over the domain [a, b], which provides further credence for the condition imposed on the set of weights.

Convergence: A weak form of convergence (in distribution) may be obtained through evaluating the distribution function of the LCG$_2$ which minimizes the *Levy distance* from the empirical distribution function [43]. Indeed, a stronger form of convergence may be obtained and can be shown to provide least square error from the empirical density at the optimum set of weights. One way to look at this is the fact that the model is a regression to the empirical density $f_X(x)$.

Minimum variance: Either form of the LCG satisfies the consistency requirements as the number of samples in D becomes large, and there exist an integer number of kernels and a set of weights that provides minimum variance estimation. The effect of the number of samples will be discussed in the next section.

Model Identification: For either the marginal or the joint density models, to estimate the parameters of the LCG$_2$ model, the modified EM algorithm [43] is used to deal with the

positive and negative components. However, most shortcomings in that approach are the lack of an optimal approach to select the number of dominant modes C (number of random variables), and the inability to resolve modes of the empirical density that are close to each other. Also, the performance of the algorithm depends on the initialization; a typical problem with the EM algorithm. To overcome these problems, an information theoretic approach based on the Akiake Information Criterion (AIC) was used in order to estimate these values (C and the required initial parameters) [67], which enabled automatic estimation of the number of kernels in the LCG$_2$ model (Eq. 40) by optimization of a likelihood function incorporating the *Akaike Information Criterion* (AIC). The approach may be summarized as follows: Suppose one has a joint density function of C random variable; i.e., $f_{X_1}f_{X_2}f_{X_3} \cdots f_{X_C}$, of random sample $D = \{X_1, X_2, \cdots, X_N\}$. The AIC criterion may be written as:

$$\text{AIC} \propto \sum_{i=1}^{N} \sum_{j=1}^{C} \Delta_j \log \emptyset(x|\theta_j) - \mathfrak{B}C \tag{44}$$

where $\mathfrak{B}$ is model component penalty and Δ_j is the posterior probability of the mode j given x. The component penalty $\mathfrak{B}$ of the model can be selected greater than the *increase* in the likelihood when a Gaussian distribution is *iteratively* added to the model. This criterion can be used for higher dimensions as well.

4. Gibbs-Markov Modelling of the Region/Labeling Process

In order to estimate the region map unconditional probability distribution $P(\boldsymbol{f})$, the region map $\boldsymbol{f} = \{f_1, \dots, f_P\}$, is presented as realizations of random variables, and the probability measure representing the joint distribution of all pixel labels on an image grid is presented as a Markov-Gibbs Random Field (MGRF) with respect to a neighborhood system **N**. Fitting an Markov Random Field (MRF) model to an image requires

estimating its parameters from a sample of the image. The literature is rich with works that propose different Markov-Gibbs Random Field (MGRF) models, which are suitable for a specific system behavior. Let Θ be a vector of potential parameters (e.g. in a second order neighborhood system) for a homogeneous anisotropic Potts model with pairwise cliques[1]:$\Theta = [\gamma_1, \gamma_2, \gamma_3, \gamma_4]$. The Gibbs probability distribution is represented as a function of Θ as follows:

$$P(f) = \frac{1}{Z}\exp\left(-\Sigma_{\{p,q\}\in N} V(f, \Theta)\right), \tag{45}$$

where Z is a normalizing factor, V is the potential function, and f is a realization of the MGRF. Thus, the MLE of Θ is defined by

$$\Theta^* = \arg\max_\Theta \frac{-1}{|P|}\left(\Sigma_{\{p,q\}\in N} V(f, \Theta) + \log(Z(\Theta))\right), \tag{46}$$

Equation (46) cannot be solved by the differentiation of the log-likelihood because the second term $\log(Z(\Theta))$ is intractable. Thus, numerical techniques are usually used to find a solution for this problem. Coding Method (CM); method-3 and Least Square Error (LSQR) method-6 are of the most popular MRF parameters' estimators. However, CM's performance varies widely for different data and CM's estimations sometimes need adjustment. Also, to estimate the model parameters using LSQR, one needs to solve an overdetermined system of linear equations. This is not practical in the case of realizations with many colors, where the overdetermined system of equations has up to $\#colors$ to the power of 8 equations. Empirical approaches similar to those Picard et al., 1992 [69] and the Gibbs energy histogram (e.g., El-Baz and Farag, 2006 [70]) may be possible for

---

[1]A clique is defined as a set of sites (e.g. image pixels) such that all pairs of sites are mutual neighbors in accord with a given neighborhood system N [68].

incorporation into the MAP segmentation framework; here the estimation may be conducted off line.

5. Maximum A posteriori (MAP) Segmentation

Once the image models are estimated, the goal is to find the desired map f by maximizing the likelihood in Eq. 38. Unfortunately, this problem has no analytical solution. However, using Eq. 39 and 40, and considering pairwise Gibbs potential, it can easily be proven that maximizing the likelihood in Eq. 38 is equivalent to minimizing the following energy function:

$$E(f) = \sum_{\{p,q\} \in N} V(f_p, f_q) - \sum_{p \in P} \log(P(I_p|f_p)) \qquad (47)$$

The energy function may be optimized by various approaches, including Simulated Annealing, Dymanic Programming and Graph-cuts [71]. Ali et al., [72][75] and Farag et al. [73-74] evaluated some of these approaches for biomodal image segmentation.

C. Variational Image Segmentation

1. Intensity Region Modeling by level sets

The level set function as a signed distance map is able to capture complicated topological deformations. A level set function $\emptyset: \Omega \subset R^2 \rightarrow R$ can be defined as the minimum Euclidean distance between the point $X \in \Omega$ and the shape boundary points. A curve can be initialized inside an object, and then evolves to cover the region guided by image information. The evolving curve within the level set formulation is a propagating front embedded as the zero level of a 3D scalar function $\emptyset(X, t)$. It should be pointed out that there is a slight change of notation in this section from the previous; X represents a

location in space here, whereas in the previous section it represents image intensity values.

To formulate the intensity segmentation problem, it is necessary to involve the contour representation. Given an image $I: \Omega \subset R^2 \rightarrow R$, the segmentation process aims to partition the image into two regions: object (inside the contour denoted by o) and background (outside the contour denoted by b). An error term can be computed by counting the number of correctly classified pixels and then measuring the difference with respect to the total number of pixels. This can be done by summing up the probabilities of the internal pixels to be *object* and the external pixels probabilities to be classified as *background*. This is measured by the term:

$$Error = 1 - \pi_o \int_{\Omega_o} p_o(I(X)) \, d\Omega - \pi_b \int_{\Omega_b} p_b(I(X)) \, d\Omega \qquad (48)$$

where p_o and p_b are the probabilities of the object and background according to the intensity values (Gaussian distributions are used to model these regions). Prior probabilities of regions (π_o and π_b) are involved in the formulation as well. Minimizing this error term is equivalent to minimizing the energy functional:

$$E(\emptyset) = -\pi_o \int_{\Omega_o} p_o H_\epsilon(\emptyset) \, d\Omega - \pi_b \int_{\Omega_b} p_b H_\epsilon(-\emptyset) \, d\Omega \qquad (49)$$

where H is the Heaviside step function and $\epsilon \in R^+$ represents the narrow band region width. An extra term is added to the energy function to represent the contour arc-length (L) which also needs to be minimal to guarantee a smooth evolution. The new energy will be:

$$E(\emptyset) = -\pi_o \int_{\Omega_o} p_o H_\epsilon(\emptyset) \, d\Omega - \pi_b \int_{\Omega_b} p_b H_\epsilon(-\emptyset) \, d\Omega + \lambda L \qquad (50)$$

where $\lambda \in R^+$. The level set function evolves to minimize such a functional using the Euler-Lagrange formulation with the gradient descent optimization:

$$\frac{\partial \emptyset}{\partial t} = \delta_\epsilon(\emptyset)(\pi_o p_o - \pi_b p_b) + \lambda \kappa \tag{51}$$

where δ is the derivative of the Heaviside function and κ is the curvature. Thus, the evolution depends on the local geometric properties (local curvature) of the front and the external parameters related to the input data I. The function $\emptyset$ deforms iteratively according to the above equation, while solving $\emptyset(X, t) = 0$ gives the position of the 2D front iteratively. Let $\emptyset_g$ denote the intensity segmented region function representation. The Gaussian distribution and prior probabilistic parameters are computed according to the method in [76].

2. Shape Alignment

This process aims to compute a transformation $\mathbf{A}$ that moves a source shape (α) to its target (β). The in-homogeneous scaling matching criteria from [76-77] is adopted, where the source and target shapes are represented by the signed distance functions $\emptyset_\alpha$ and $\emptyset_\beta$, respectively. The transformation function is assumed to have scaling components:

$S = diag(s_x, s_y)$, rotation angle, θ (associated with a rotation matrix $\mathbf{R}$) and translations: $T = [T_x, T_y]^T$. A dissimilarity measure to overcome the scale variance issue is formulated by assuming that the signed distance function can be expressed in terms of its projections in the coordinate directions as: $\mathbf{d}_\alpha = [d_x, d_y]^T$ at any point in the domain of the shape α. Applying a global transformation $\mathbf{A}$ on $\emptyset_\alpha$ results in a change of the distance projections to $\mathbf{d}'_\alpha = \mathbf{RSd}_\alpha$ which allows the magnitude to be defined as: $\emptyset'_\alpha = \|\mathbf{Sd}_\alpha\|$ which implies that $\emptyset'_\alpha \leq \max(s_x, s_y)\emptyset$. Thus, a dissimilarity measure to compute the

difference between the transformed shape and its target representation can be directly formulated as:

$$r(\mathbf{X}) = \|\mathbf{RSd}_\alpha(\mathbf{X})\| - \emptyset_\beta(\mathbf{A}) \tag{52}$$

by summing-up the squared difference between the two representations, an energy function can be formulated as:

$$E_1 = \int_\Omega \delta'_\epsilon(\emptyset_\alpha, \emptyset_\beta) r^2 \, d\Omega \tag{53}$$

where δ'_ϵ reduces the complexity of the problem and ϵ is the width parameter of the band around the shape contour. The given measure r, from the shown derivations, satisfy the relation $r \leq s\emptyset_\alpha(\mathbf{X}) - \emptyset_\beta(\mathbf{A})$, where $s = \max(s_x, s_y)$. Thus, an energy function can be obtained where $E \leq E_1$;

$$E = \int_\Omega \delta'_\epsilon(\emptyset_\alpha, \emptyset_\beta) \left(s\emptyset_\alpha(\mathbf{X}) - \emptyset_\beta(\mathbf{A}) \right)^2 d\Omega \tag{54}$$

The above functional better describes the registration since it incorporates a scaled version of the source shape representation. In this dissertation, the gradient descent optimization is used to solve the problem, which requires the involved functions to be differentiable. A smeared version of, $s(s_x, s_y) = \max(s_x, s_y)$, is used at the line $(s_x = s_y)$ since, the function is not differentiable there, which is based on its original definition:

$$s(s_x, s_y) = \max(s_x, s_y) = s_x H_\epsilon(s_x - s_y) + s_y \left(1 - H_\epsilon(s_x - s_y) \right) \tag{55}$$

which will return s_x if $s_x - s_y \geq 0$, otherwise s_y. The smeared Heaviside step function H is used to obtain a smooth transition around the line $s_x = s_y$ allowing the function to be differentiable everywhere. The function derivatives will be calculated as

$$\frac{\partial s}{\partial s_x} = H_\epsilon(s_x - s_y) + (s_x - s_y)\delta_\epsilon(s_x - s_y) \tag{56}$$

$$\frac{\partial s}{\partial s_y} = H_\epsilon(s_y - s_x) + (s_y - s_x)\delta_\epsilon(s_y - s_x) \qquad (57)$$

The parameters $\{s_x, s_y, \theta, T_x, T_y\}$ are required to minimize the energy functional E.

3. Level Set Segmentation Algorithm with Shape Prior

The above steps have resulted in an algorithm whose input are LDCT scans and output are segmented lung nodules. The algorithm can be summarized as follows:

> *Lung Nodule Segmentation:*
> 1. Segment the Lungs from their surroundings – Lung tissue segmentation
> 2. Train the lung nodule modeling step on a portion of the data at hand – Lung Nodule Modeling
> 3. Apply the lung nodule detection approach discussed in Chapter IV to compute the positions of the candidate nodules and hence crop them for classification. Cropping here means setting a box around the nodule center and extracts its neighbor area from the surroundings; i.e., a region of interest, ROI, is cropped around the detected nodules – Nodule detection and ROI determination
> 4. Based on the input image size, construct the initial prior shape circle and its shape model representation $\emptyset_p$.
> 5. Solve Eq. 50 to compute the intensity segmentation region representation $\emptyset_g$. Solution is iterative until the function converges – reaches a certain state. Note the function keeps the sign distance property by following the approach in [76].
> 6. Initialize the transformation parameters to $s_x = 1$, $s_y = 1$ and $\theta = 0$. At this moment the nodule center location is manually selected which initializes the translation parameters t_x and t_y.
> 7. Solve the gradient descent approach to minimize the energy in Eq. 53. Parameters converge to their steady state values and hence the final boundaries of the ellipse are computed.
> 8. Threshold the region inside the ellipse to accurately mark the nodule pixels. The resulting region may under-go a median filter smoothing step to remove noisy pixels.

D. Experimental Results

1. Datasets

This work is validated using four different databases. The first is the ELCAP [12] public database, DB1, described in Chapter III. This database has nodules of diameter ranging from $2mm$ to $5mm$. The second database (DB2) contains 108 nodules from LDCT scans of slice thickness 2.5mm and a pixel-spacing of $0.72461mm \times 0.72461mm$ (diameter from $2.9mm$ to $6mm$). The third database (DB3) has 28 nodules, 1.25mm and 2.5mm slice thickness, and nodules diameter ranging from $7mm$ to $20mm$. The fourth dataset is the LIDC (DB4) which contains nodules ranging in sizes. The slices are both low-dose and high-dose CT images [13].

2. Comparative Study

The discussed segmentation approaches in this chapter were examined on all four databases described. A small sample of results from DB4 with various nodule sizes, positions and segmentation challenges are depicted in Figure 28. The figure aids in identifying the benefits of using intensity with shape-based approaches versus intensity-based only segmentation methods.

Figure 28 demonstrates that the EM, variational intensity-based level sets, and graph cuts approaches kept the pleural surface for connected nodules. Since the focus of the nodule segmentation in this chapter is for size computation and, when applicable, growth rate measurements for malignancy, the pleural surface is considered undesired information. Figure 28 (b) and (d) are examples of nodule cases with visible noise artifacts and non-nodule information (i.e. anatomical structures). These illustrate further

some of the difficulties discussed before about small objects, and how such challenges affect the ability for proper segmentation.

Overall the variational shape-based level set method provided the best segmentation results for obtaining the "head" of the nodule region. A post segmentation step was conducted on the elliptical region extracted, (Note to re-iterate: the shape model defined was an ellipse since the "head", as previously described, can be encompassed in a circular region of interest.) to obtain a well-defined shape of the main nodule "head".

(a)

(b)

	Original	EM	Variational	Shape Variational	Shape Variational + EM	Graph Cuts (Mask)
Radiologist 1						
Radiologist 2						
Radiologist 3						
Radiologist 4						

(c)

	Original	EM	Variational	Shape Variational	Shape Variational + EM	Graph Cuts (Mask)
Radiologist 1						

(d)

FIGURE 28 - An LIDC Nodule: a) centrally located in the lung tissue; b) connected to the pleural surface; c) connected to the pleural surface and d) lung nodule located in the peripheral lung region. The first column is the cropped regions of interests. Column two depicts the EM segmentation followed by morphological operations. Column three illustrates the variational intensity level set segmentation, while column four shows the results of the variational shape-based level set segmentation. The final column represents the intensity-based graph cuts segmentation results.

The results concluded that the intensity-based approaches can be used as an initial or post segmentation process to the variational shaped-based level sets. Also, approaches where a shape model can be embedded into the formulation of the segmentation method are necessary for such cases as nodule segmentation.

3. Performance of the Variational Shape-based Segmentation

The developed approach uses a region of interest (ROI) image that contains the lung nodule as input. Image intensity segmentation using level sets (as described above) is used to extract the non-lung regions from the lung tissue regions and represents the slices by a level set function ($\emptyset_g$) (see Figure 29). The first four rows are from DB1 while the rest belong to DB2. The shape model is combined with the sign distance functions $\emptyset_g$ to segment the nodule region in a variational registration step. The shape prior model is a signed distance function of a circle denoted by $\emptyset_p$. The shape prior model $\emptyset_p$ is registered with the intensity model $\emptyset_g$ to mark the boundaries of the lung nodule using the approach shown above. The model is initialized and then the alignment parameters are estimated using the gradient descent optimization.

Different scales, rotation, and translation parameters are computed in each case to obtain an ellipse exactly around the nodule head (see Figure 30). Changing of the shape model can be noticed until the steady state around the nodule boundaries is reached. Also, the axis of the ellipse rotates and varies in size to include the most boundary information of the nodule. The approach is robust for various nodule sizes from larger nodules ($< 1cm$) to nodules that occupy smaller spatial support regions (i.e. $> 1cm$).

Sample results from the third and fourth databases are shown in Figures 31 and 32 respectively. These databases have a variety of nodule sizes and types. In all cases, the hyper ellipse is capable of covering the prominent nodule shape. The elliptic region can be segmented later by a simple thresholding technique to result in an accurate measure of the nodule size by removing the lung tissues which have low intensity values. From Table 5, the number of dubbed failure cases is less than 6%. Result cases are judged by

checking if the model fits the boundaries, if nodule is found not to fit the boundary it is identified as a failure case.

The ultimate goal of nodule segmentation is to specify the malignancy by following up the nodule size in different scans taken at different time intervals. The proposed nodule segmentation approach was used to segment 3D nodules in different patient scans, as a follow-up (see table 6).
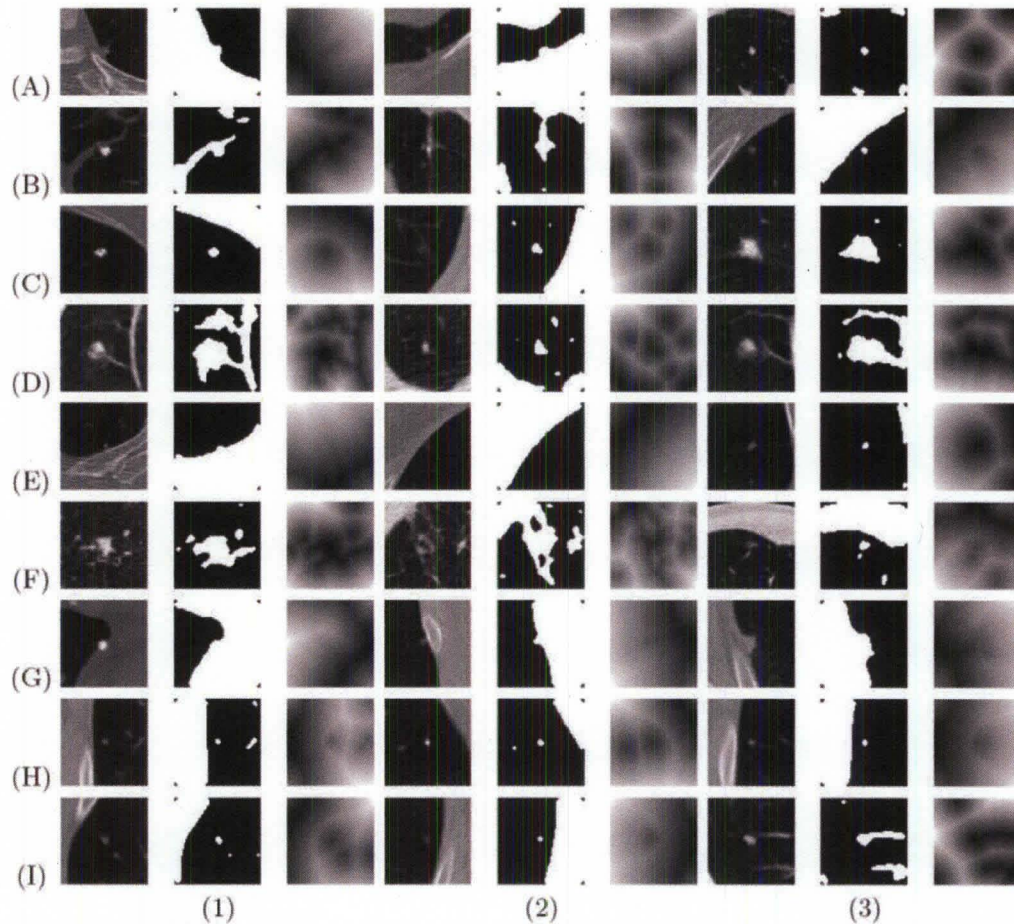
FIGURE 29 - Different ROI images of lung nodules and their corresponding binary segmentation results: First fourth, seventh columns represent intensity images. Second, fifth, eighth columns demonstrate the binary segmentation results. The remaining columns illustrate the sign distance function $\emptyset_g$ of the resulting regions. Juxta pleural nodules are shown in examples A.1 and A.2. Well circumscribed nodule examples are given in: A.3, B.3 and C.1. Vasculature nodule examples can be seen in B.1, C.3, D.1, and D.3, while a pleural tail nodule is shown at B.2. Other examples are given in rows E to I.

The table further emphasizes the correlation between malignancy and size-volume in the corresponding case. Volume changes for chronological scans of a given patient are illustrated, which corresponds to the malignancy estimate results provided by an expert radiologist. Nodule growth/decay rates can be observed from scan to scan over time
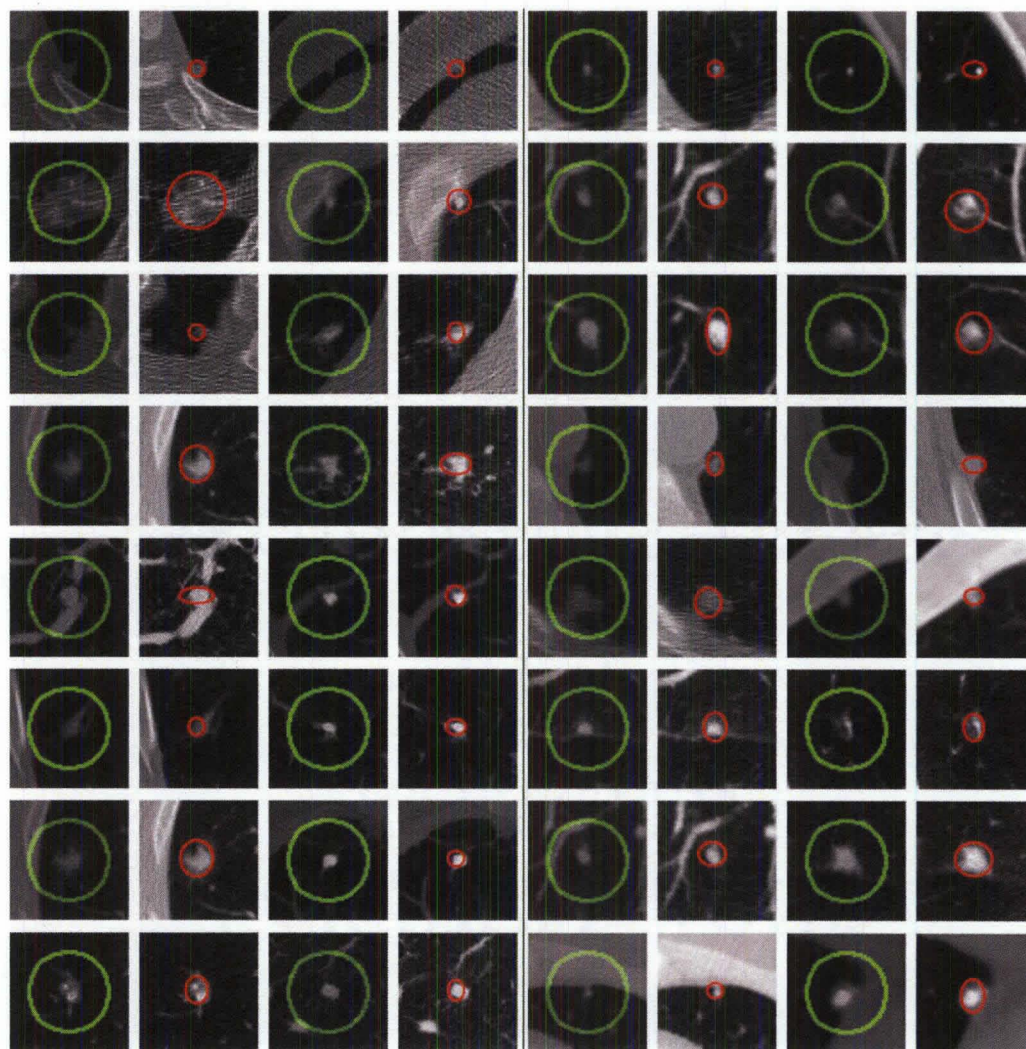
FIGURE 30 - Nodule segmentation results from DB1 (left block-first four columns) and DB2 (right block-last four columns). Initialization is given in green while final nodule boundaries are shown in red.
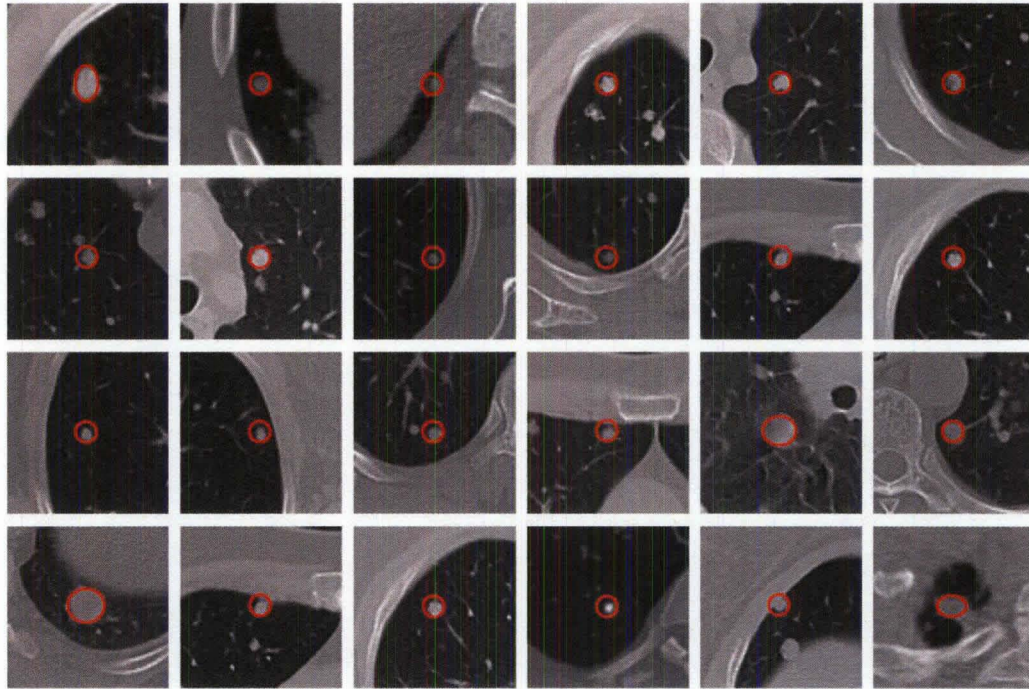
FIGURE 31 - Sample nodule segmentation results are illustrated from DB3. Nodules range in size as well as vary in positions with respect to the pleural wall and vasculature.

TABLE 5 - Overall segmentation results found for the four datasets. Nodules are separated into the four nodule categories. The total numbers of nodules from the database considered in the nodule class are the numbers to the left of the brackets, while the failure cases are in the brackets. First two rows provide information on the four databases.

Nodules	DB1	DB2	DB3	DB4	Summary
	291	108	28	315	742
Nodule Diameter(range in mm)	2-5	2.5-6	7-20	2-4	2-20
Juxta-Pleural	115{4}	27{2}	6{0}	59{5}	94.66%
Well-Circumscribed	91{3}	50{3}	21{1}	174{4}	96.73%
Pleural-Tail	45{1}	10{1}	0{0}	34{4}	93.26%
Total	291{10}	108{8}	28{1}	316{18}	94.12%
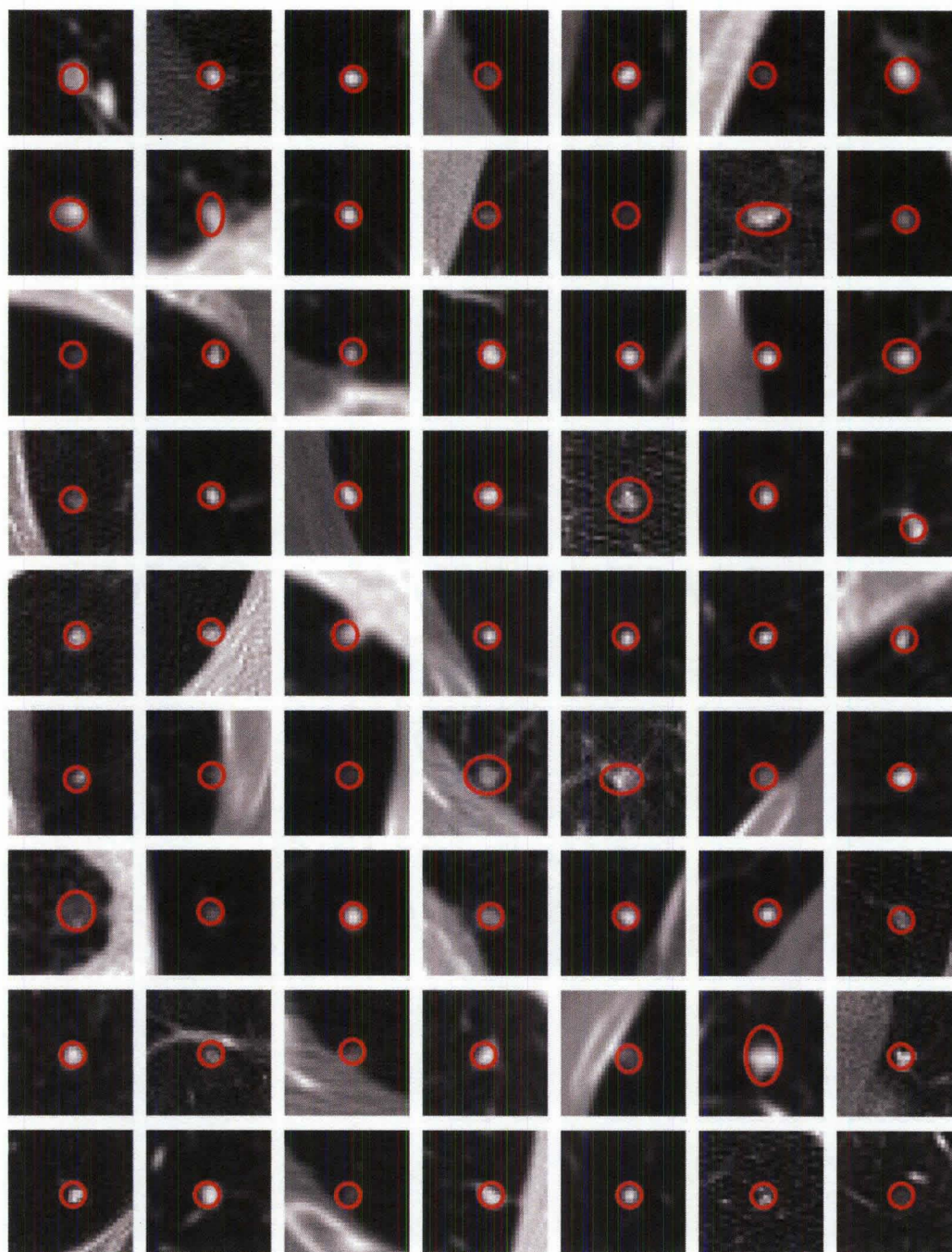
FIGURE 32 - Sample nodule segmentation results are illustrated from DB4. Nodules have sizes varying from 2-4 mm as well as varying positions with respect to the pleural wall and vasculature.

TABLE 6 - Follow-up evaluation for 5 patients on several scans is shown column-wise. Nodule size values are shown in mm^3.

Patient ID #	1	1	2	4	6	6	6	7
Nodule ID	3	5	2	2	1	2	3	1
Scan #1	824.76	3092.80	1052.30	2771.70	828.78	579.56	724.45	1639.50
Scan #2	1021.80	5608.90	1164.90	2346.60	1051.30	700.86	1051.30	965.52
Scan #3					1305.9	473.14	520.46	

E. Summary

This chapter examined statitical and vartaional approaches for segmentation of lung nodules in LDCT scans. The cropped regions around the nodules have bimodal histograms (see Figure 11 in Chapter 1 of this dissertation). Intensity-based segmentation is plausible and various methods may be used to select a segmentation threshold. Unfortunately, due to the uncertainties associated with the nodules in LDCT, more elaborate segmentation methods are required. A further extension to the approach of Farag et al., 2006 [43] was introduced, in the intensity models based on the linear combinations of Gaussians (LCG) model, and in the paramter identification of the Gibbs-Markov model used for modeling the spatial interaction.

A generalized linear model for probability density estimation using linear combinations of Gaussians (LCG) with general weights was used. Approaches to estimate the number of modes in the histogram (components of the LCG model) and the corresponding parameters of the Gaussian kernels were presented. An analytical approach to estimate spatial interaction potentials in the Markov-Gibbs random field (MGRF) model is presented and evaluated. A maximum a posteriori (MAP) approach

was presented which uses the LCG and the MGRF models, resulting in an energy function that can be minimized iteratively using the Graph Cuts approach.

A new variational shape-based level set formulation for the lung nodules segmentation problem was described. A shape prior model in the form of a circle was embedded into the image domain by means of registration, allowing the use of inhomogeneous scales. The usage of such scales in registering the shape and intensity models allowed the template to deform as an adaptive ellipse that tended to better fit around the head of the nodule boundaries. The registration process incorporates, in the implicit spaces, both a prior shape model and image intensity information.

Among the possible extensions are the following: a) How to model the statsitical shape models and include the shape priors in the segmentation process? Since, the statistical shape-prior segmentation in the literature is known for its speed; b) How to generalize the transformation paramters that embed the shape model into the image domain, thus, avoiding the post EM step; c) Study incorporatation of the shape priors into the energy function, of general topological cliques in the MGRF models, and evaluation of the segmentation algorithm with respect to variational shape-based techniques such as level sets; and d) Perform exhuastive follow-up study for diagnosing nodule malignacy .

CHAPTER V

CLASSIFICATION OF SMALL SIZE LUNG NODULES

In the computer vision and biomedical imaging literature the terms categorization, classification, identification, and recognition share a lot of commonality of methods and purpose. In the face recognition setup, the general approach is to detect faces, extract signature information and then perform recognition. In the lung nodule example, one may also denote the classification step as recognition. However, classification may indeed entail two aspects: assigning segmented objects into types (classes), or assigning them into a definitive group (e.g., pathology in the lung nodule case).

Chapter II studied the deformable modeling methods based on ASM, AAM and ATM. Chapter III used these models for small object detection. Chapter IV used the models also for segmenting the detected object. This chapter aims at evaluating the effectiveness of the modeling approaches devised in the dissertation for classification, and completing the loop in computer-aided diagnosis/decision (CAD) system. The chapter further examines approaches for classification of lung nodules using modern object descriptors that have been introduced in the computer vision literature during the past decade.

As objects of interest have been described in terms of their shape and appearance, this work shall focus on descriptors that adhere to shape and appearance contexts.

As discussed in Chapter II, texture, color, and intensity of an object will connote its appearance. Boundaries, specific unique features that are invariant to size, orientation and translation will connote its shape. This examination study shall be confined to feature descriptors of appearance and shape that are most suitable for classification of small object; again, small lung nodules (less than 10 mm in diameter) are the focus of this case study.

Given precise object models, categorization may be performed by registration. That is the segmented image (or volume in 3D) of the candidate object may be registered to the available models, and the categorization will be assigned to the closet model. Various registration methods may be used. If appearance is captured in the models and images, the mutual information approach may be used (e.g., [78]). On the other hand, if the boundaries (shape) information is the main form for the models and candidates, then various matching methods, which are distance based, may be used; e.g., the iterated closest point (ICP) algorithm [79]. Various enhancements and modifications have been introduced to the MI and ICP algorithms during the past decade and half, which improved speed and accuracy of the registration process. Likewise, elastic registration of shapes and volumes has been enhanced a great deal during the past decade. This chapter will not cover this literature. Recent surveys of modern registration methods may be found in [80][81]. If the object models and candidates are represented in terms of signatures (vector of features), then categorization may be performed using a distance classifier, such the kNN (e.g., [63]).

On the other hand, if classification is categorization of objects into classes + groups, then a two-stage process may be conducted, the first may use appearance/shape

registration or distance classification, and the second may also repeat the same step but on a narrower population.

To clarify this matter, two examples in face recognition and nodule analysis are given. Suppose there is a face recognition system tilted towards identifying young Asian teens. The system may use a typical face recognition system to identify young teen faces first. Then a classification stage, based on morphometric measurements, may be used to classify teens into Asian and non-Asian. In the lung nodule example, the first stage can detect and segment the lung nodules, then classification can provide: a) benign or malignant; or b) nodule category (well-circumscribed, vascular, juxta-pleural, and pleural-tail) + benign or malignant. In all cases, the most important aspects of categorization are two: features + matching criterion.

The remainder of this chapter examines the issues of feature descriptors and lung nodule classification into nodule category. Planned future work will address the subsequent step to classify into benign or malignant. Figure 33 illustrates the entire framework of nodule analysis formed of three steps: modeling, detection, segmentation and categorization.

A. Object Feature Descriptors

In the past decade, several object descriptors have been introduced in the computer vision literature, including the local binary pattern (LBP) [82], the scale-invariant feature transform (SIFT) [32], its color extension CSIFT [83], and the speeded-up robust features (SURF) [84]. A comprehensive evaluation of the geometric feature descriptors may be found elsewhere, in particular Mikolajczyk and Schmid, 2005 [85]. Various enhancements to the SIFT descriptor were introduced (e.g., the ASIFT [33]). The

Daugman code method (e.g., [86]) - used to discriminate between human irises – may be useful to quantify the intensity distribution of the lung nodules. If the objects are outlined in a closed contour, the classical Fourier Descriptor (FD) may lend benefit for shape classification [87]. This work applies the FD method to lung nodule classification.
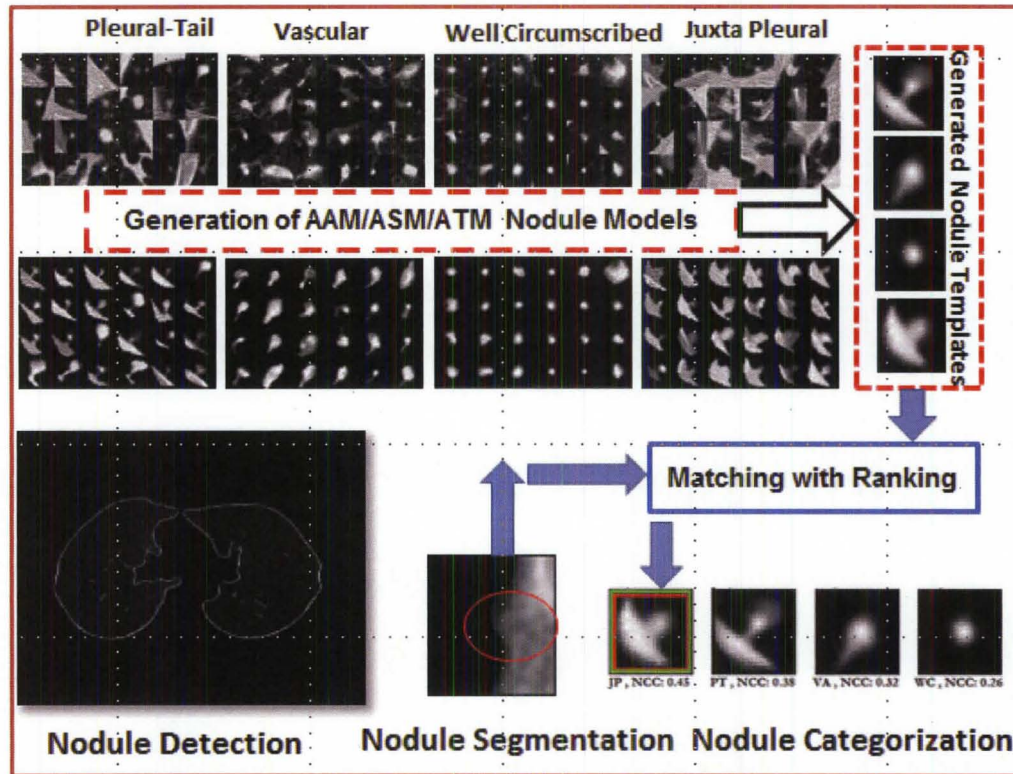
FIGURE 33 – Summary of the lung nodule analysis system. The off-line models using ASM/AAM/ATM are used in detection, segmentation and categorization.

In all, these descriptors have not been used before in the literature for the study of lung nodules. Highlighted below are the major characteristics of the descriptors and evaluation of their performance on the lung nodule class of objects as an example of small object classification under uncertainties.

1. Multi-Resolution Local Binary Pattern (LBP)

The Local Binary Pattern is an operator invariant to monotonic changes in grayscale and can resist illumination variations as long as the absolute gray-level value differences are not badly affected (e.g., [82]). The original operator labeled the pixels of an image by thresholding the 3×3 neighborhood of each pixel with the center value and considered the result as a binary number. At a given pixel position (x_c, y_c), the decimal form of the resulting 8-bit word is

$$LBP(x_c, y_c) = \sum_{i=0}^{7} s(I_i - I_c) 2^i \tag{58}$$

where, I_c corresponds to the center pixel (x_c, y_c), I_i to gray level values of the *eight* surrounding pixels and function $s(o)$ is a unit-step function.

The LBP operator was extended to a circular neighborhood of different radius size to overcome the limitation of the small original 3×3 neighborhood size failing to capture large-scale structures [82]. Each instance is denoted as (P, R), where P refers to the equally spaced pixels on a circle of radius R. The parameter P controls the quantization of the angular space and R determines the spatial resolution of the operator.

An LBP pattern is considered uniform if it contains at most two bitwise transitions from 0 to 1 and vice-versa, when the binary string is circular. The reason for using uniform patterns is that they contain most of the texture information and mainly represent texture primitives. The operator is derived on a circularly symmetric neighbor set of P members on a circle of radius R denoting the operator as LBP_{PR}^{u2}.

Figure 34 illustrates examples of circularly symmetric neighbor sets for various (P, R). The LBP operator was further enhanced by combining it with a rotation invariant measure $VAR_{P,R}$, which characterizes the contrast of local image texture. The

combination of the LBP_{PR}^{u2} operator and the variance measure produces a powerful operator that is rotation and gray-scale invariant.
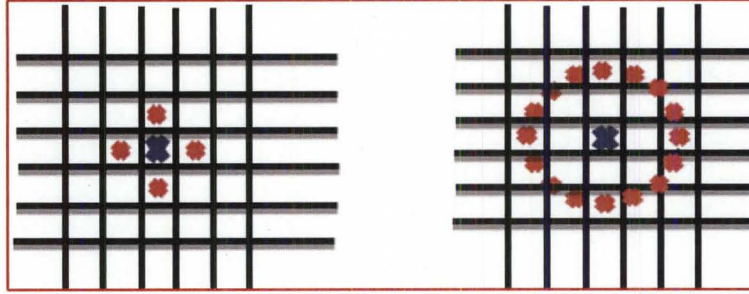
FIGURE 34 - Circularly symmetric neighbor sets for different values of (P, R); left (a) $P = 4, R = 1.0$; right $P = 16, R = 2.0$.

In the multi-resolution analysis the responses of multiple operators realized with different (P, R) are combined together and an aggregate dissimilarity is defined as the sum of individual log-likelihoods computed from the responses of individual operators [82]. The notation LBP_{PR}^{u2} used in this chapter refers to the extended LBP operator in a (P, R) neighborhood, with only uniform patterns considered.

The LBP is used to generate a feature vector which describes the nodule region of interest in a LCDT slice. The LBP is applied to one of three scenarios on: i) the original nodule images; ii) the gradient of the nodule image or, iii) an addition of the original and gradient nodule images. The gradient image was computed by first obtaining each individual image in the x- and y-spaces by filtering the corresponding directional-space original image with the corresponding parameter vector identified in the author's work [88][94]; the overall gradient nodule image is:

$$\nabla_{nodule} = \sqrt{\nabla_x{}^2 + \nabla_y{}^2} \qquad (59)$$

101

A similarity measure is then used to classify these nodules to one of the four classes: juxta, well-circumscribed, pleural tail and vascularized. Principle component analysis (PCA) and linear discriminant analysis (LDA) are used to project the extracted LBP descriptors to a low-dimensional subspace where noise is filtered out.

Figure 35 illustrates the formation of the LBP descriptors on lung nodules.
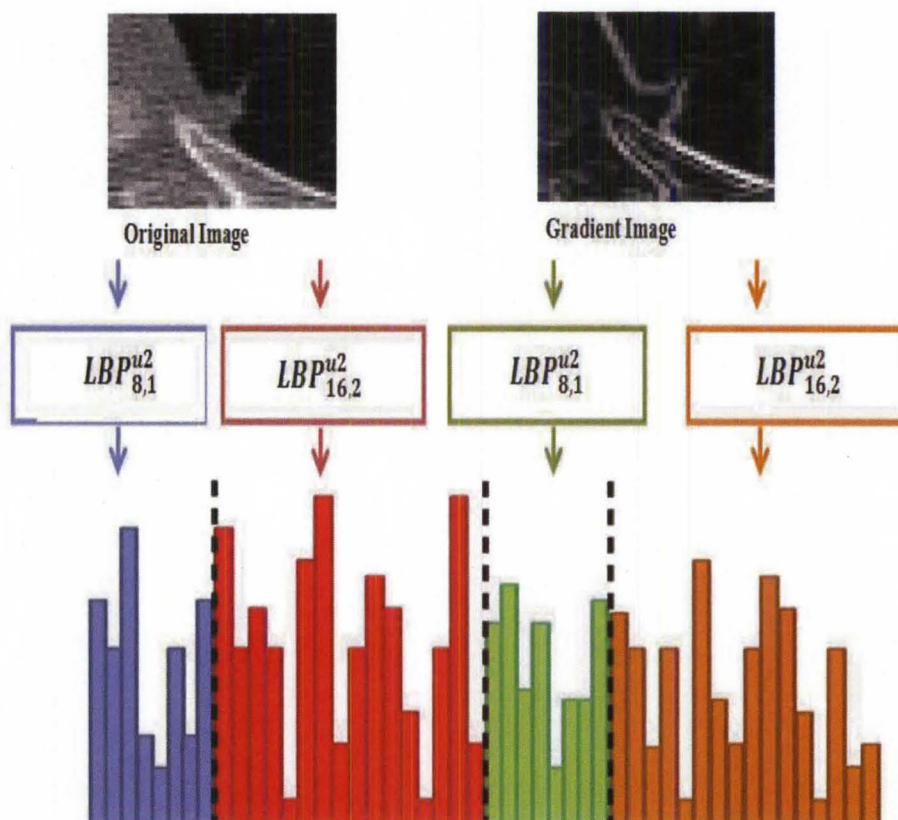
FIGURE 35 - Block Diagram of generating the LBP for a juxta-pleural nodule. The equation for the above picture is: $\mathbf{LBP}_{8,1}^{u2} + \mathbf{LBP}_{16,2}^{u2} + \mathbf{LBP}_{8,1}^{u2} + \mathbf{LBP}_{16,2}^{u2}$, where the first two terms represent the original image and the last two terms represent the gradient image.

2. The Signed Distance Transform

The distance transform is a shape-based feature descriptor that represents each pixel of the binary edge map image with a distance to the nearest obstacle pixel i.e. binary

pixel. The extracted Signed Distance transform images were projected to a lower-dimensional subspace using PCA and LDA.

The LBP of the signed distance image results were also obtained, thus, resulting in a combinational shape and texture feature descriptor representation of the nodules and non-nodules. The relevance of combining shape and texture feature vectors is described in the recognition stage. Figure 36 illustrates the approach which combines the LBP and Signed Distance Transform.
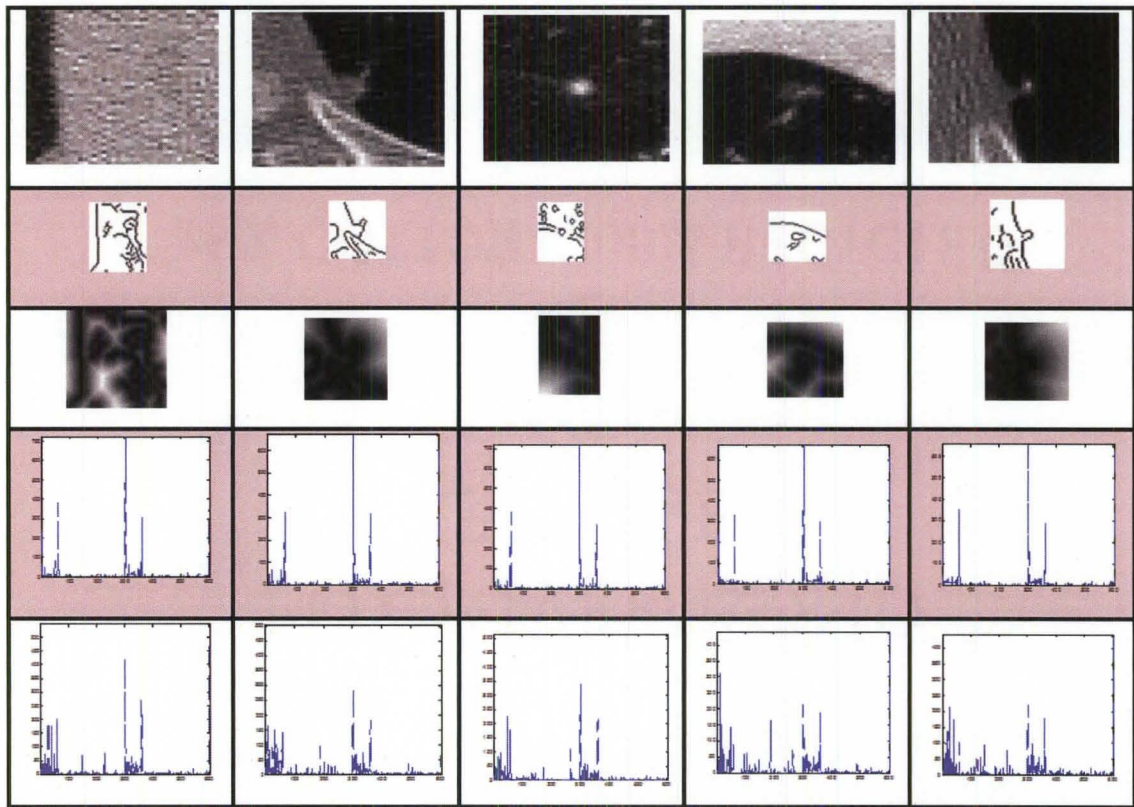
FIGURE 36: First row shows typical non-nodule (first column) and nodule textures (juxta-pleural, well-circumscribed, vascularized and pleural tail, respectively). Second row shows edge maps (using the Canny Operator). Third row is the signed distance. Fourth row is the LBP of the nodules. Final results depict the LBP + Signed distance features.

3. The Scale-Invariant Feature Transform (SIFT)

The SIFT is a combinational detector and descriptor approach introduced by Lowe [32] that allows extraction of distinctive scale and rotation invariant features from images. The SIFT is a combination of a scale invariant region detector known as the difference of Gaussian (DoG) detector and a proper descriptor referred to as SIFT-key. The approach consists of four major steps of computation to generate the set of image features: Scale Space extrema detection, Key-point Localization, Orientation assignment and Key-point descriptor. In the first stage of computation all scales and image locations are searched over using a DoG function to identify potential interest points that are invariant to orientation and scale. Once the potential interest points are found at each candidate location a detailed model is fitted to determine scale and location.

The keypoints selected are based on the stability measures. To each keypoint location one or more orientations are assigned based on the local image gradient directions. All future operations are performed on image data that has been transformed relative to the assigned scale, location and orientation for each feature. At the selected scale in the region around each keypoint the local image gradients are measured and transformed into a representation that allows for significant levels of change illumination and local shape distortion.

The scale-space of an image defined as a function, $L(x, y, \sigma)$, was shown by Koenderink (1984) [90] and Lindeberg (1994) [91] as follows: The only possible scale-space kernel, under reasonable assumptions, is the Gaussian function, thus the scale-space of an image $L(x, y, \sigma)$ is produced from convolving a variable-scale Gaussian, $G(x, y, \sigma)$, with an input image, $I(x, y)$:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \tag{60}$$

Lowe proposed using scale-space extrema in the difference-of-Gaussian function, to accurately detect stable keypoint locations in scale-space, convolved with the image, $D(x, y, \sigma)$, which from the difference of two nearby scales separated by a constant multiplicative factor k can be computed:

$$D(x, y, \sigma) = \big(G(x, y, k\sigma) - G(x, y, \sigma)\big) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{61}$$

In order to detect the local maxima and minima of $D(x, y, \sigma)$, each sample point is compared to its eight neighbors in the current image and nine neighbors in the scale above and below. The keypoint is selected if it larger or smaller than all of these neighbors. Once the keypoint candidate is obtained a detailed fit to the nearby data for location, ratio of principal curvatures and scale is performed to reject points with low contrast or poorly localized along an edge. Consistent orientation assignment to each keypoint based on local image properties allows the keypoint descriptor to be represented relative to this orientation and thus achieve invariance to image rotation. The scale of the keypoint is used to select the Gaussian smoothed image, L, with the closest scale. Each image sample, $L(x, y)$, at this scale, the gradient magnitude, $m(x, y)$ and orientation $\theta(x, y)$ is pre-computed using pixel differences:

$$m(x, y) = \sqrt{(L(x + 1, y) - L(x - 1, y))^2 + (L(x, y + 1) - L(x, y - 1))^2} \tag{62}$$

$$\theta(x, y) = \tan^{-1}(L(x, y + 1) - L(x, y - 1)/(L(x + 1, y) - L(x - 1, y))) \tag{63}$$

An orientation histogram of 36 bins covering the 360° range of orientations is formed from the gradient orientation of sample points within a region around the keypoint. Additional samples added to the histogram is weighted by its gradient magnitude and by

a Gaussian-weighted circular window with a σ that is 1.5 times that of the scale of the keypoint.

All the weighted gradients for the descriptor are normalized to the main orientation of the circular region around the keypoint which is divided into 4x4 non-overlapping patches. The histogram gradient orientations within the patches are computed and then histogram smoothing is performed to avoid sudden orientation changes and bin size reduction to eight bins to limit the descriptor's size results into a $4 \times 4 \times 8 = 128$ dimensional feature vector for each key-point. The feature vector is finally normalized to unit length and thresholded to reduce the effects of linear and non-linear illumination changes.

In nodule analysis framework, it is assumed that nodules have been already detected which correspond to interest/key points in Lowe's algorithm; hence, this step can be bypassed. In order to obtain a nodule SIFT descriptor which is invariant to orientation, a consistent orientation should be assigned to the detected nodule which is represented by its centroid, x_0. This orientation is based on the gradient of the nodule's local image patch. Considering a small window surrounding x_0, the gradient magnitude and orientation can be computed using finite differences. Local image patch orientation is then weighted by the corresponding magnitude and Gaussian window. Eventually the orientation is selected to be the peak of the weighted orientation histogram.

Building a nodule SIFT descriptor is similar to orientation assignment, for example a 16x16 image window surrounding the nodule centroid point x_0 is divided into sixteen 4x4 sub-windows, then an 8-bin weighted orientation histogram is computed for each sub-window, hence, $16x8 = 128$ descriptors for each nodule is obtained. Thus, each

detected nodule can now be defined at location (x_0, y_0), specific scale σ, explicit orientation θ and descriptor vector, $x_o = \{x_0, y_0, \sigma, \theta, d\}$.

Thus the SIFT operator $\mathcal{S}: I(x) \rightarrow X$ can be viewed as mapping a CT slice $I(x)$ to the nodule space with n-nodules, $X = \{x_i\}_{i=1}^n$ detected from $I(x)$, where $x_i = \{x_0^i, y_0^i, \sigma_i, \theta_i, d_i\}$.
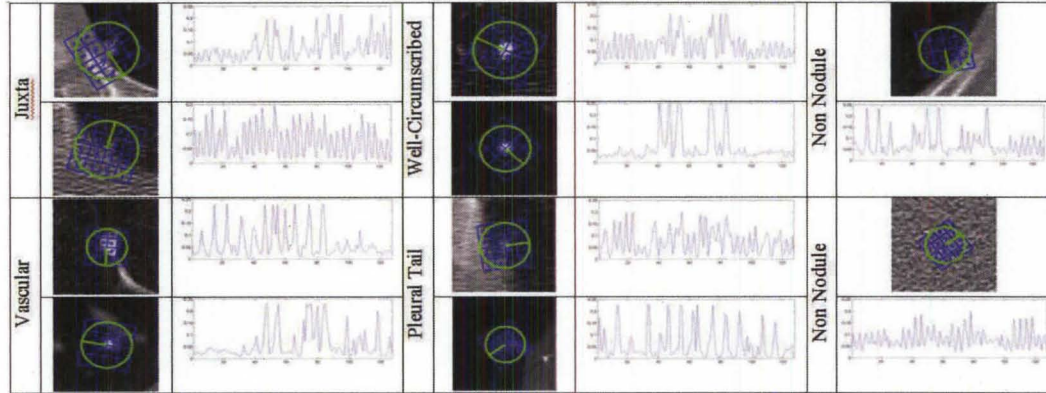
FIGURE 37 - Visualization of the SIFT Recognition process for a sample Non-nodules and Nodules: Juxta, Well-Circumscribed, Pleural-Tail and Vascularized.

Principle component analysis (PCA) and linear discriminant analysis (LDA) are used to project the extracted SIFT descriptors to a low-dimensional subspace where noise is filtered out. Figure 37 shows the SIFT descriptors of lung nodules as well as non-nodule objects.

4. The speeded-up robust features (SURF)

This approach has numerous adaptations over the SIFT algorithm and the Harris-Laplace feature detector [84]. This algorithm was implemented to improve execution time without compromising feature detection efficiency. The detector is based on the Hessian matrix, (e.g., Tao and Dickinson, 2000 [92]), and relies on integral images to

reduce computation time. The descriptor is a distribution of Haar-wavelet responses within the neighborhood of interest. The Hessian matrix is defined as:

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \tag{64}$$

where $x = (x, y)$ represents a point in the image I, and σ is a scalar value. $L_{xx}(x, \sigma)$ is the convolution of the second order Gaussian derivative with the image in point x. The SURF descriptor consists of several steps; a square region is constructed around the interest point and oriented either in a rotation invariant method, where the Haar-wavelet response in the $x-$ and $y-$ directions are computed and weighted with a Gaussian centered at the interest point, or a non-rotation invariant method. The wavelet responses in both directions are then summed-up over each sub-region. The total number of descriptors for each point is 64.

The nodules are the input images, in this work, and the region of interest is mainly the texture information concentrated around an area where texture information is not sparsely found since the spatial support of lung nodules are relatively small in size. The "blob-response threshold" was reduced to 240; the number of octaves desired is only 1; and the step-response is 1 instead of 2.

Principle component analysis (PCA) and linear discriminate analysis (LDA) are used to project the extracted SURF descriptors to a low-dimensional subspace where noise is filtered out.

5. The Daugman Code for Iris Recognition

John Daugman in 1994 [86] developed an algorithm for use in iris recognition which is also the basis of all current iris recognition systems. An image of the iris is first acquired, the iris region is then segmented, a fixed number feature points in the iris regions are selected to be encoded in an iris template, and then feature matching is used to identify or reject the presented subject. The overall Daugman Iris Recognition framework is depicted in Figure 36. In the image acquisition process iris images can be obtained under visible light; however, due to dark colored eyes absorbing the light and revealing less texture information infrared light is used in capturing iris images.
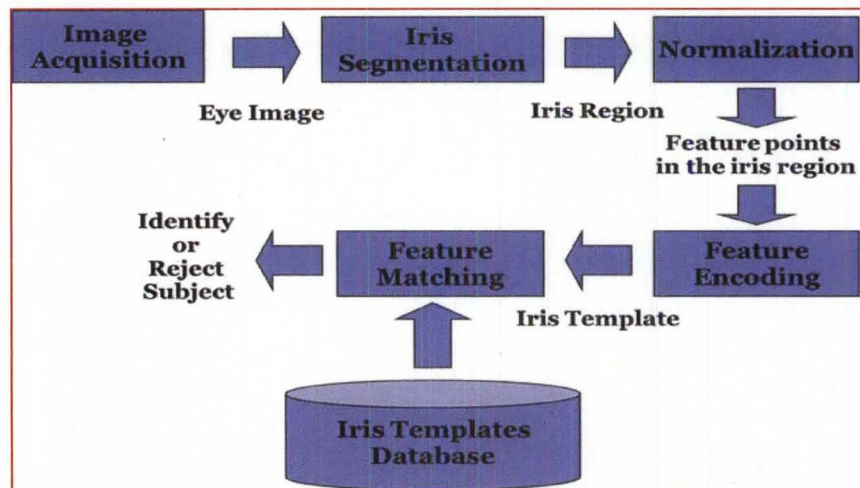
FIGURE 38 - The Daugman model for automated iris recognition.

The first stage of iris recognition is to isolate the actual iris region in a digital eye image. As illustrated in Figure 37, the iris region can be approximated by two circles, one for the iris boundary and another, interior to the first, for the pupil boundary. The eyelids and eyelashes normally occlude the upper and lower parts of the iris region, while specular reflections can also occur within the iris region, corrupting the iris pattern.

In [86], an *integro-differential* operator, defined below, was used to find both the pupillary boundary and the outer iris limbic boundary:

$$max_{(r,x_o,y_o)} \left| G_\sigma(r) * \frac{\partial}{\partial r} \oint_{r,x_o,y_o} \frac{I(x,y)}{2\pi r} ds \right| \tag{65}$$

where $I(x,y)$ is an image such as that seen in Figure 38, $G_\sigma(r)$ is a smoothing function. The operator searches over the image domain (x,y) for the maximum with respect to increasing radius r in the blurred partial derivative of the normalized contour integral of $I(x,y)$ along a circular arc ds, with parameters r as the radius and (x_o, y_o) corresponding to the circle center point.
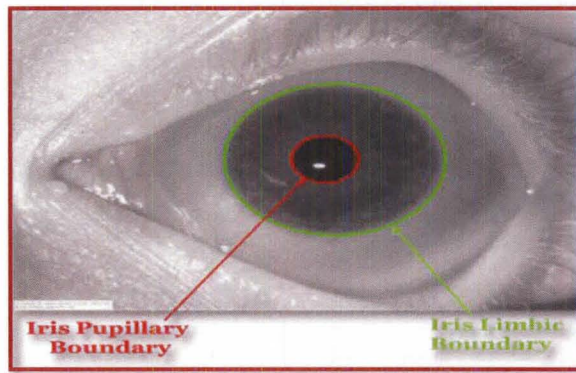
FIGURE 39 - The geometry of the iris region lends itself to approximating the iris region as two circles; the interior represents the iris pupillary boundary while the outer represents the iris limbic boundary. The iris region is the area encapsulated by the two contours.

This operator functions as a coarse-to-fine iterative boundary search method. Once the iris and pupil boundaries are located the operator is manipulated to fit spline parameters to describe each optimal eyelid boundary.

Quadrature 2D Gabor wavelets are used to demodulate each isolated iris pattern its phase information. This process encodes the binary texture information by identifying in

which quadrant of the complex plane each resultant phase lies, when a given area of the iris is projected onto complex-valued 2D Gabor wavelets:

$$h_{\{Re,Im\}} = = sgn_{\{Re,Im\}} \int_\rho \int_\emptyset I(\rho,\emptyset) e^{-i\omega(\theta_0-\emptyset)} e^{-(r_0-\rho)^2/\alpha^2} e^{-(\theta_0-\emptyset)^2/\beta^2} \rho d\rho d\emptyset \qquad (66)$$

where $h_{\{Re,Im\}}$ can be regarded as a complex-valued bit as either 1 or 0 depending on the sign of the 2D integral representing the real and imaginary parts; $I(\rho,\emptyset)$ is the raw nodule image in a dimensionless polar coordinate system that is translation and size invariant; α and β are the multi-scale 2D wavelet size parameters; ω is the wavelet frequency, spanning three octaves in inverse proportion to β; and $(r_0-\rho)$ represents the polar coordinated of each region of the iris for which the phasor coordinates $h_{\{Re,Im\}}$ are computed. The fractional Hamming Distance (HD) is computed, and is used as the dissimilarity measure between any two irises, where 0 would represents a perfect match.

$$HD = (\| (\text{nodule A} \otimes \text{nodule B}) \cap \text{mask A} \cap \text{mask B} \|)/(\| \text{mask A} \cap \text{mask B} \|) \qquad (67)$$

where the XOR, $\otimes$, operator detects disagreement between any corresponding pair of bits while the AND, $\cap$, operator ensures that the compared bits are both deemed to have been uncorrupted by the lung nodule surroundings.

Given two descriptors $d_i = [d_{i1}, d_{i2}, ..., d_{iN}]^T$ and $d_j = [d_{j1}, d_{j2}, ..., d_{jN}]^T$, ED can be defined as follows;

$$ED_{ij} = \sqrt{\sum_{k=1}^n (d_{ik} - d_{jk})^2} \qquad (68)$$

Figure 39 depicts the process for a sample of juxta-pleural and well-circumscribed nodules. The different binarized and un-bianarized lung nodule codes obtained from the Daugman method are shown.

In [89], several descriptors, including the Daugman code approach, were examined for lung nodule classification.
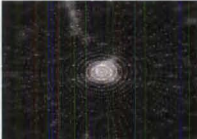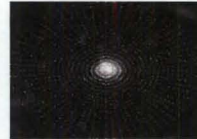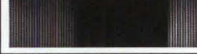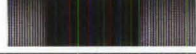
	Juxta		Well-Circumscribed	
Radial Distribution				
Rubber Sheet				
Binarized Code				
Unbinarized Code				

FIGURE 40 - Visualization of Daugman Recognition process for a sample Juxta and Well-Circumscribed Nodules.

6. The Fourier Descriptors

The Fourier Descriptors is a classic shape analysis approach (e.g., [87]). It is invariant under scale and rotation and may be used to describe shapes. Fourier descriptors (FD) are presentations of vertices of a polygon in terms of the basis functions of the discrete Fourier transform.

Let N represents the number of points of a given shape. The initial points can be represented as $\mathbf{z}$. The Fourier descriptors ($\{\mathbf{c(k)}: k = -M/2 + 1, ..., M/2\}$) are the coefficients of the Fourier transform of input points $\{(\mathbf{z(i)}: i = 0, ..., N-1)\}$. The variables of the function can be as below.

- o N: The number of points of a given shape, (input).
- o M: The number of the Fourier descriptors you want to use (user selection).
- o L: The number of points needed to be reconstructed.
- o $\tilde{z}_\ell$: The reconstructed points where $\ell = \{0, ..., L-1\}$.
- o z_i: The given (original) points where $i = \{0, ..., N-1\}$.
- o c_k: The Fourier descriptors where k= $\{-M/2 + 1, ..., M/2\}$.

For a given closed contour, the points represented as $z(i) = x(i) + j.y(i)$, where $i = 0, ..., N - 1$. These points may be selected in a clockwise or counterclockwise direction as shown in Figure 40. From the start and to the end point, coordinate pairs (x_0, y_0), $(x_1, y_1), ..., (x_{N-1}, y_{N-1})$ are constructed from which the descriptor is evaluated.

The Fourier descriptor can be defined as the typical Discrete Fourier Transform (DFT) summation:

$$c_k = \frac{1}{N} \sum_{i=0}^{N-1} z_i \, exp(-j2\pi \frac{ik}{N}) \tag{69}$$

where $k = \{-M/2 + 1, ..., M/2\}$. The reconstructed points is the inverse Discrete Fourier Transform (IDFT) are calculated as follows:

$$\tilde{z}_\ell = \sum_{k=-\frac{M}{2}+1}^{\frac{M}{2}} c_k \, exp(j2\pi \frac{\ell k}{N}) \tag{70}$$

where $\ell = \{0, ..., L - 1\}$.

The number of the Fourier descriptor is not dependent on the number of the given and reconstructed points. It is related with how much information one wants to reconstruct from the given points.

In the Fourier descriptor representation, the c_k coefficients around c_0 (DC term) describe the low-frequency information. The high frequency information is carried in the coefficients around $c_{-M/2}$ and $c_{M/2}$ terms. The Fourier descriptors at lower frequency represent the general shape, whereas the Fourier descriptors at higher frequency describe the details and sharp corners.
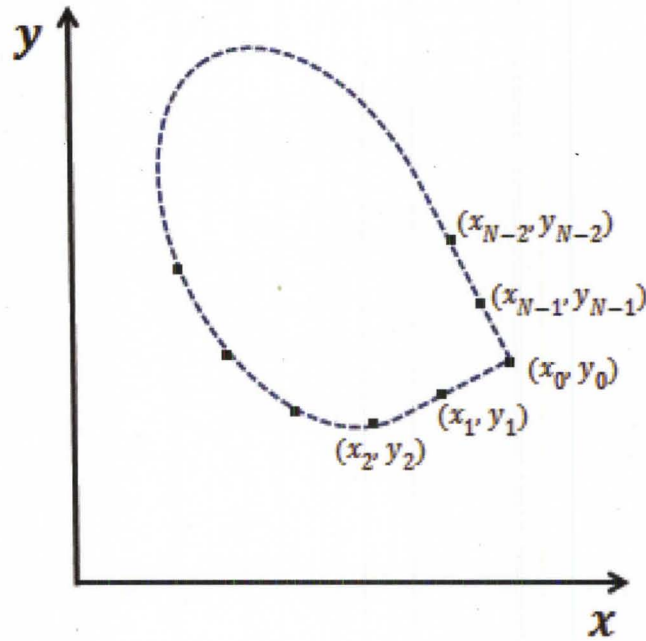
FIGURE 41 – The Fourier Descriptor is the DFT of sampled points on a closed contour.

The effectiveness of FD approach on lung nodule models is illustrated below. Figure 41 shows four lung nodule types from LDCT scans. Each nodule is manually annotated and evaluated the FD. Using 30% of the FD coefficients, the nodules contours can be reconstructed quite well. The reconstructed contours degrade in quality using fewer FD coefficients. The Fourier descriptors are effective for representing smooth non intersecting contours. They are invariant under scale and translation, and may provide a degree of invariance under rotation; thus incorporating them in a segmentation or registration frame work may lend some benefits.
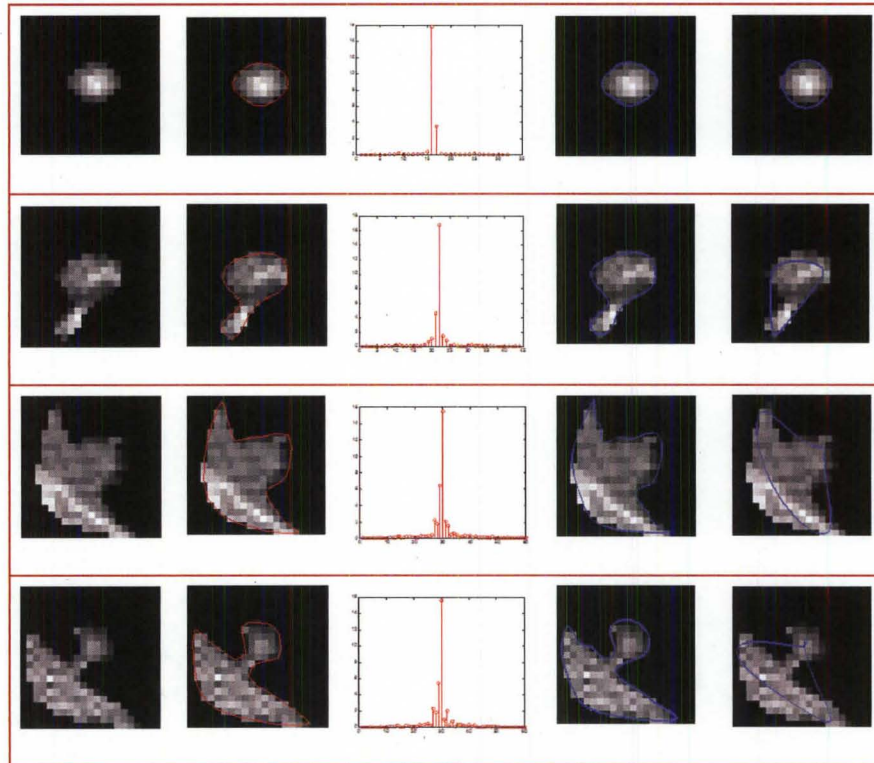
FIGURE 42 – Illustration of the effectiveness of the Fourier Descriptors for representation of four lung nodule types. The 1st column shows the original nodules; manually outlined contours on 2nd column, FD on 3rd column, reconstruction using 30% of the FD coefficients are shown in 4th column and using 10% on the 5th column.

In the nodule example, shown in Figure 41, the FD is concentrated at the DC value, yet there are still distinctions among the various nodules.

7. Feature distance Measures

The feature distance measurement is a numerical description of how far apart the feature vectors are from one another. Numerous methods found in the literature can be used; below are described three different distance measurements.

i)   *The Euclidean distance*

The Euclidian distance (ED) between feature point vectors p and t in the Euclidean n-space

$$ED = \sqrt{\sum_{i=1}^{n}(p_i - t_i)^2} \tag{71}$$

Note that the Euclidean distance is rotation invariant but not scale-invariant.

ii)   The Mahalanobis Distance

The Mahalanobis Distance is a scale-invariant distance measure based on correlations between variables by which variations can be identified for analysis. A multivariate vector $X = [x_1, x_2, \ldots, x_N]^T$ from a group of values with mean $\mu = [\mu_1, \mu_2, \ldots, \mu_N]^T$ and covariance matrix, S, is defined as:

$$D_M(X) = \sqrt{(X - \mu)^T S^{-1}(X - \mu)} \tag{72}$$

iii)   The Chebyshev Distance

This distance is a metric defined on a vector space where the distance between two vectors is the greatest of their differences along any coordinate dimension. The distance between two vector points p and t with standard coordinates p_i and t_i is defined as:

$$D_{Chebyshev}(p, t) = max(|p_i - t_i|) \tag{73}$$

Evaluation of these distance measures for shape analysis exist elsewhere (e.g., [63]).

B. Lung Nodule Classification

The above descriptors form the basis for the classification process to be examined in the following section.

116

1. General Approach

The general approach for nodule classification may be summarized by the following algorithm.

1. Construct a statistically sufficient database of pathological nodules;

2. Co-Register members of the nodule database to create the templates used for nodule detection, as described before

3. Generate the feature vectors using the geometric descriptors (e.g., SIFT, ASFIT, SURF, LBP and Gabor Wavelet) for all members of the nodule database and store offline. Machine learning algorithms may be used such as PCA, RANSAC and Adaboost for optimal selections of the feature vector in terms of discrimination as well as execution time;

4. Perform the nodule detection using template matching;

5. Crop ROIs of sizes NxN over detected nodules (e.g., N = 21) – these will be used for categorization;

6. Segment the nodule regions with the ROIs using the variational approach described in Chapter IV, enhanced with a priori information about shape and intensity, using the nodule database;

7. Repeat step # 3 on the candidate nodules after segmentation;

8. Calculate the distance between the feature vectors of candidate nodules and those in the pathological database, and assign the nodule category based on minimum distance.

The above algorithm may be carried out by various ways, depending the features available. Below feature-based and registartion-based nodule classifcation implemented in the author's recent work [93-94] are discussed.

2. Feature-based classification

The classification stage using various feature descriptors was extensively analyzed in the author's work [88-89, 93-94]. The most significant classification results were obtained when the shape based signed distance transform was combined to the texture based LBP approach. The results in Tables 7-9 illustrate the classification results of the signed distance transform versus the multi-resolution local binary pattern (LBP). A third feature descriptor using the combination of the methods is also shown.

Higher true-positive rates can be seen from the LDA projection in tables 7 and 8 when more training is conducted using either the LBP or distance descriptors separately. When comparing the PCA results less training data resulted in better true-positive classification of nodules. In the non-nodule distance transform experimentations more training data was needed to obtain in some instances perfect results. This is understandable since the non-nodules do not have specific shape characteristics that can be defined or manipulated as in the nodules case.

TABLE 7 - Classification Results for various nodules using Raw LBP, LDA LBP and PCA LBP with variable training percentages.

Nodule Type	Raw LBP				LDA LBP				PCA LBP			
	100%	75%	50%	25%	100%	75%	50%	25%	100%	75%	50%	25%
Juxta Pleural	52	50	47	38	100	86	65	50	64	64	59	67
Well-Circumscribed	40	41	40	26	65	80	63	36	64	60	66	82
Vascular	22	29	32	10	32	76	56	32	20	22	37	56
Pleural Tail	22	20	17	11	100	76	52	39	33	17	33	46
Non Nodule	78	77	74	68	100	88	60	44	86	87	83	96

TABLE 8 - Classification Results for various nodules using Raw Distance Transform, LDA LBP and PCA Distance Transform with variable training percentages.

Nodule Type	Raw Distance Transform				LDA Distance Transform				PCA Distance Transform			
	100 %	75 %	50 %	25 %	100 %	75 %	50 %	25%	100 %	75 %	50 %	25 %
Juxta Pleural	38	39	35	34	100	88	61	45	62	54	60	68
Well-Circumscribed	33	33	36	34	74	83	63	45	46	59	48	55
Vascular	12	12	15	15	29	76	54	29	37	22	61	63
Pleural Tail	17	1 7	17	15	100	85	54	33	17	24	35	52
Non Nodule	63	68	68	49	100	87	65	49	83	89	85	79

Overall, the PCA combinational shape and feature description of nodules resulted in a drastic true-positive rate increase in classification. All of the results depicted in Tables 7 and 8 allow the conclusion to be made that non-nodules do in-fact contain descriptor variations that allow them to be correctly classified. Also, combination of shape and texture feature information allows for better object representation to be obtained, thus improved results in classification.

TABLE 9- Classification Results obtained from Raw Combinational Feature Transform and PCA Combinational Feature Transform with variable training percentages.

Nodule Type	Raw Combinational Feature Descriptor				PCA on Combinational Feature Descriptor			
	100%	75%	50%	25%	100%	75%	50%	25%
Juxta Pleural	40	41	39	37	78	76	76	79
Well-Circumscribed	40	37	36	34	73	68	71	68
Vascular	24	20	22	12	51	54	44	76
Pleural Tail	22	26	22	20	33	35	41	54
Non Nodule	63	57	58	49	100	99	100	98

Table 9 depicts impressive results when the LBP was obtained from the distance transform images. A 20% true-positive rate increase was found, in the PCA 25% training combinational vascular nodule case when comparing it to the PCA LBP results obtained

when only the texture information was used for classification, and a 13% increase over the distance transform results alone. Variations of percentage increases were seen for each nodule category.

3. Registartion-Based Classification

The idea of the registartion-based classification is to compare the segmented nodules with nodule models, using a registration algorithm. Since the ASM, AAM and ATM generated impressive nodule models (Chapter II) which resembled both the shape and appreance of real nodules, it is plausible to use the normalized nodule models as templates to compare with candidate nodules for classifcation. From the face recognition analogy, a probe (test face) is compared to a gallary using either direct matching (by registration) or through the use of features.

The following terminologies are relavant to the categorization process:

i) <u>Target</u> set $\mathcal{T}$: a set of textured regions containing the nodule models generated by ASM/AAM/ATM approaches for all nodule types.

ii) <u>Gallery</u> set $\mathcal{G}$: a subset of $\mathcal{T}$ containing template(s) to be matched in a certain matching setup.

iii) <u>Query</u> set $\mathcal{Q}$: a set of textured regions of unknown nodule type, where nodule type identification is performed by matching all elements in the query set to the target set.

iv) <u>Probe</u> set $\mathcal{P}_{\mathcal{G}}$:a subset of $\mathcal{Q}$, where each element has a match in the gallery set.

v) <u>Imposter</u> set $\mathcal{P}_{\mathcal{N}}$: a subset of $\mathcal{Q}$, which contains elements that don't have a match in the gallery set.

As an example, again, using the face recognition terminology, a region centered at a well circumscribed nodule is considered an imposter to a gallery containing only juxta pleural nodules. Also a non-nodule region is always considered as an imposter. Comparing the feature vector for all nodule models in the gallery set with the feature vector for all regions in the probe set results in a similarity matrix S, where the ijth element is the similarity between the ith element of the gallery and the jth element of the probe. The following metrics can be defined according to a similarity score: Normalized cross-corerlation (NCC), the mutual information (MI) or the output of descriptors such as SFIT, LBP, etc.

Identification Rate/Probability: It is calculated as the proportion of testing nodules correctly matched to its own type, i.e. probe p_j is identified correctly in the top N gallery nodule types, where $N = \text{rank}(p_j)$, such that:

$$\text{rank}(p_j) = |\{g_k : s_{kj} \geq s_{ij}, \text{id}(g_i) = \text{id}(p_j)\}| \quad \forall g_k, g_i \in G \qquad (72)$$

For each probe p in the probe set $\mathcal{P}_G$, the similarity measures are sorted against the gallery G, and obtain the rank of the match. Identification performance is then stated as the fraction of probes whose gallery match is at rank r or lower. Thus the probability of identification at specific rank is defined as:

$$P_I(r) = \frac{|\{p_j : \text{rank}(p_j) \leq r\}|}{|\mathcal{P}_G|} \quad \forall p_j \in \mathcal{P}_G \qquad (73)$$

These quantities have been calculated for all the nodules in the ELCAP data.

Table 10 shows the results for the four nodule categories.

In measuring the ranking, the cropped nodules are used in two fashion; without segmentation (i.e., no extraction of the nodule part in the cropped region) and with segmentation. The segmentation of nodules were conducted by various homegrown

methods (including use of shape and intensity priors in an energy model optimized by graph cuts; also experimented with were basic segmentation using adaptive thresholding of the cropped regions by median filtering and anisotropic diffusion filtering, etc.).

TABLE 10 – Results of the nodule categorization using registration/matching nodule candidates to nodule models.

Nodule Model	Nodule and background				Nodule region segmented			
	Rank 1	Rank 2	Rank 3	Rank 4	Rank 1	Rank 2	Rank 3	Rank 4
Juxtal Pleural	0.4606	0.9217	0.9826	1.0	0.4261	0.9217	0.9826	1.0
Well-Circumscribed	0.764	0.7978	0.8427	1.0	0.8876	0.9663	0.9775	1.0
Vascularized	0.4146	1.0	1.0	1.0	0.5122	1.0	1.0	1.0
Pleural Tail	0.3261	0.7609	0.8261	1.0	0.3913	0.5217	0.5435	1.0

Figure 42 is the ROC for 291 nodules specified in the ELCAP dataset. Both the well-circumscribed and the viscularized nodules provide the best performance. This is because both nodule types possess the best texture and shape information that enhances the correlation between the nodules and the models.

In general, the results of the ranking (i.e., matching models with nodules) improved by segmentation of the nodule portion in the cropped region. Model-based approaches such level sets and combinations of Gibbs-Markov models enhance the segmentation at severe computation cost. Nodule segmentation is a work in progress issue. A code or signature for the models and the nodules will provide better matching than using the classic image registration methods on regions with small spatial support. The conclusion, however is that the cropped regions have always been correctly categorized within second ranks by a simple computational approach such the normalized cross-correlation. This indeed is very encouraging for moving into using context based image processing

and the ability to invoke advanced machine learning approaches to perform the matching process.
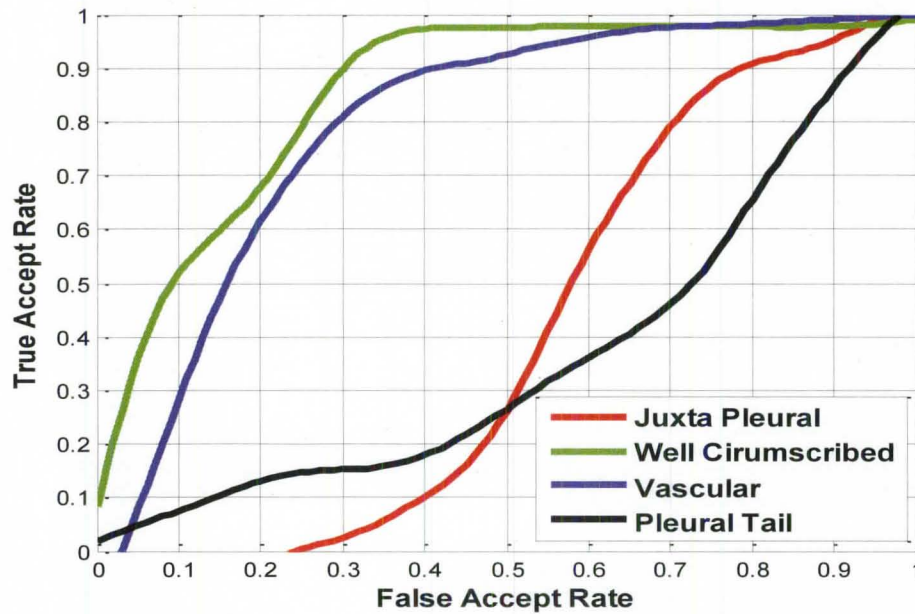
FIGURE 43 - ROC of automatic categorization on the ELCAP data. Well-circumscribed and vascular nodule types possess the best ranking for automatic categorization.

The extensive analysis using the approaches described in this chapter has allowed several conclusions to be made:

1) Texture and shape feature information separately are not sufficient for lung nodule categorization, since the combination of the approaches yielded great improvements.

2) In all of the approaches used, the non-nodule features generated and projected by PCA or LDA provided excellent classification results; thus, non-nodules contain descriptor variations that allow them to be correctly classified and not confused with nodules.

3) Intensity-based registration methods did not provide accurate categorization of small objects; a more appropriate similarity measures may be needed for these types of objects.

4) Signatures of nodules – based on multiple approaches – may be generated and used for categorization; similar to face recognition methods. However, more extensive annotated databases of nodules are needed.

C. Summary

In this chapter, a system for nodule candidate detection and classification was created to show the robustness and accuracy of the produced models. Detection using a template matching method with normalized cross-correlation similarity measure without false positive reduction was implemented to show the robustness of the data-driven templates formulated from the AAM and ASM approaches over the known parametric template generation. Detection using the data driven template matching approach, after false positive reduction via SIFT and LBP feature extraction, was also implemented, further enhancing the detection process.

Classification of the nodules and non-nodules were examined using a k-NN leave-one-out algorithm with the Euclidean distance as the similarity measure, in order to test whether or not significant distinctions between the nodule classes exist. An overall 12% true-positive rate increase was found in the PCA combinational classification results over using the PCA LBP or the PCA distance transform separately.

CHAPTER VI

CONCLUSIONS AND FUTURE DIRECTIONS

A. Problem Overview

This dissertation considered the problem of modeling of small objects under uncertainties of the imaging process. Modeling entails a mathematical description of the features and characteristics of the object in a way to allow proper representation of the key characteristic shape and/or texture information of the object. Deterministic and stochastic approaches of modeling are abundant; no approach is adequate to model all types of models. The developed model can then be utilized for accurate object analysis, detection, segmentation, classification, etc.

From the human visual system (HVS) perspective, if the object is not visible, then it would not be possible to describe/model, and therefore no inference/recognition would be feasible. Small objects, which occupy a small portion of the field of view (FOV) of the eye - after proper fixation, magnification and enhancements – resemble a challenge to the HVS. This challenge increases in difficulties and consequences with added uncertainties such as motion, contrast, occlusion and noise.

Training adds a clue for modeling and recognition in these circumstances. For example, if the imaging process is known, then specific and consistent characteristics/features about small objects may be used for modeling and recognition.

125

In the biomedical imaging field, the image formation has been improving steadily, in order to provide structure and function about the human body, even at the cellular level. At the end, the human visual system makes a decision based on detection of objects and associating a classification/diagnosis to them.

This dissertation attempted at mimicking the HVS as it views and recognizes small objects. The difficulty, from the computational perspectives, is to translate the human expertise associated with target characteristics which are used for detection and recognition. The approaches taken are to first, quantify the information about the objects using the experts' annotation, and then develop the geometric models for detection, segmentation and classification. Hence, the availability of an inclusive ensemble is crucial for the success of the approach presented here.

This dissertation developed a unified theory, and set of algorithms, for image-based modeling, detection, segmentation and categorization/classification of small objects under uncertainty. Modern approaches based on modeling shapes and appearance were examined. Two approaches were studied in detail, using active shape models (ASM) and active appearance models (AAM). A third approach using active tensor models (ATM) was briefly examined. The dissertation also studied the concepts of shape alignment for objects without distinct characteristics.

The detection, segmentation and classification processes for small objects using the generated models were examined, and new formulations to approaches were developed and tested. As such a front-end process of modeling, detection, segmentation and classification/categorization that is model-based and data-driven in nature was established.

B. Summary of contributions

From the theoretical point of view, this dissertation developed a general approach for statistical inference about small objects that is entirely model-based.

- Detection of small objects using template matching with various similarity metrics were thoroughly examined in Chapter III. The ROC of various similarity measures were examined on two clinical databases (ELCAP and LIDC), and on two other studies of small size.

- Segmentation of small objects using statistical and variational methods was studied in Chapter IV. A novel energy formulation was introduced for the Gibbs-Markov segmentation approach. Also, a novel energy formulation for the variational segmentation approach which incorporates the shape information was described.

- Classification of small nodules using registration and feature matching was performed for the first time. Assigning the detected objects to categories was examined. A subsequent step to categorization based on type/pathologies is being pursued.

From the algorithmic point of view, the dissertation considered the front-end deployment of model-based approaches for small object recognition.

- The ASM, AAM and ATM approaches have been programmed and tested on an ensemble of objects annotated by experts.

- The geometric descriptors were used in object detection for false positive reductions, and were also used for classification into categories.

- The algorithms were validated using phantom and synthetic information, and also compared with experts results whenever available.

From <u>applications point of view,</u> the dissertation focused on small lung nodules, of sizes less than 10 mm. Early detection of nodules as they develop in the lung tissue will assist in early diagnosis of lung cancer, thus assist in jump starting the treatment early, which enhances the chances for a cure. Lung cancer is the leading cause of deaths among all cancers. The end goal of the work presented in this dissertation is to create "biopsy from images" for lung cancer using LDCT scans. The work has pursued creating a fully automated CAD system that can detect, segment and classify the lung nodules. Pathologic databases for small size nodules are not available; an immediate focus of the CVIP Lab future work is to construct such a database. Indeed, the author's software system, which will be optimized for use on various computerized media, will enable the creation of such a database.

C. Possible Extensions

This dissertation has discovered many of the difficulties associated with modeling small objects. <u>Among the theoretical extensions to be considered are the following:</u>

1. Combine the variational and statistical methods for segmentation in a decision fusion framework.

2. Quantify the effects of target pose, intensity and categories on the tensor model.

3. Quantify the performance of ASM/AAM vs. ATM for object detection with respect to pose, illumination/intensity, size and occlusion.

4. Discover the use of simultaneous detection, segmentation and classification (SDSC) based on rotation-invariant matched filtering, followed by elastic registration.

Among the algorithmic extensions to be considered are the following:

5. Optimize the annotation software to facilitate its use by experts; e.g., in the lung nodule project

6. Create a web-based small object categorization approach to enhance the participation of various users

7. Deploy modern programming languages and paradigms of machine learning for analysis of small objects.

Among the extensions to be considered for the Lung Nodule CAD system are the following:

8. Create an annotated clinical nodule database for LDCT scans that address variations in scanning hardware, demographic locations and wide range of settings in the LDCT imaging protocol

9. Create a lung nodule consortium which encourages uploading de-identified LDCT scans with validated diagnosis

10. Upload the software system of the lung CAD onto smart phones and various IT media which encourages and facilities dissemination of expertise, and re-training the software on the local imaging conditions.

The dissertation has shown the difficulties of working with small objects; yet, it also showed that modern computer vision methodologies hold a real promise in solving an extremely difficult problem from the point of view of human experts as well as the

machine. The problem addressed in this dissertation has also shown the value of multidisciplinary research and how disparate disciplines can be combined to solve a real world problem.

REFERENCES

1. D. G. Kendall. The diffusion of shape. Advances in Applied Probability, 9:428– 430, 1977.

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

3. B. Julesz, E.N., Gilbert, L.A. Shepp, and H.L. Frisch, "Inability of humans to discriminate between visual textures that agree in second-order statistics--revisited. Perception," 2, 391-405, 1973.

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

5. Eye chart - http://en.wikipedia.org/wiki/File:Snellen_chart.svg

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

7. Zhang and Z. Zhang, "A Survey of Recent Advances in Face Detection," Microsoft Research, Technical Report, MSR-TR-2010-66, 2010.

8. P. Sinha, B. Balas, Y. Ostrovsky, R. Russell, " Face Recognition by Humans: 19 Results All Computer Vision Researchers Should Know About," Proceedings of the IEEE, Vol. 94, No. 11, pp. 1948-1962, November 2006.

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

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 Farag, Lung Nodule Modeling and Detection for Computerized Image Analysis of Low Dose CT Imaging of the Chest. Master of Engineering, University of Louisville, May 2009.

12. ELCAP public lung image database. www.via.cornell.edu/databases/lungdb.html

13. G. Armato, G. McLennan, M. F. McNitt-Gray, C. R. Meyer, D. Yankelevitz, D. R. Aberle, C. I. Henschke, E. A. Hoffman, E. A. Kazerooni, H. MacMahon, A. P. Reeves, B. Y. Croft, and L. P. Clarke, "Lung image database consortium: developing a resource for the medical imaging research community," Radiology, vol. 232, no. 3, pp. 739–748, 2004. (LIDC)

14. T. Cootes, G. Edwards and C. Taylor, "Active Appearance Models," Proc. European Conference on Computer Vision, ECCV'98, pp. 484-498, 1998.

15. T. Cootes, "An introduction to active shape models," – Chapter 7 in Model-Based Methods in Analysis of Biomedical Images, Ed. R. Baldock and J. Graham, Oxford University Press, 2000, pp223-248.

16. 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.

17. M. B. Stegmann and D. D. Gomez, "A brief introduction to statistical shape analysis," 2002; {mbs,ddgg}@imm.dtu.dk - http://www.imm.dtu.dk/~mbs/

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

19. M.A.O. Vasilescu, D. Terzopoulos, "Multilinear Projection for Appearance-Based Recognition in the Tensor Framework," Proc. Eleventh IEEE International Conf. on Computer Vision (ICCV'07), Rio de Janeiro, Brazil, October, 2007, 1–8.

20. Amal Farag, Ham Rara, Cambron Carter, Ahmed El-Barkouky, and Aly Farag, Experiments with Active Shape, Appearance and Tensor Modeling, for Faces and Objects - TR-April 2012, CVIP Lab, University of Louisville, April 2012.

21. A. Papoulis, Probability, Random Variable, and Stochastic Processes, 3rd Edition, McGraw-Hill, New York, 1991.

22. Rothganger, F., Lazebnik, S., Schmid, C., and Ponce, J. 2003. 3D object modeling and recognition using affine-invariant patches and multi-view spatial constraints. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Madison, Wisconsin, USA, pp. 272–277.

23. Crowley, J. L. and A. C. Parker: 1984, 'A representation of shape based on peaks and ridges in the difference of low-pass transform'. IEEE Transactions on Pattern Analysis and Machine Intelligence 6, 156–170.

24. Lindeberg, T.: 1998, "Feature Detection with Automatic Scale Selection." International Journal of Computer Vision 30(2), 77–116.

25. Mikolajczyk, K. and C. Schmid: 2002, "An affine invariant interest point detector." In: European Conference on Computer Vision, Vol. I. pp. 128–142.

26. Edwards G.J. Taylor C.J. Cootes, T.F. Active appearance models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(6), 2001.

27. 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 pp. 595—604, 2001.

28. Amal Farag, James Graham, Aly Farag, Salwa Elshazly and Robert Falk. Parametric and Non-Parametric Nodule Models: Design and Evaluation, Proc. of Third International Workshop on Pulmonary Image Processing in conjunction with MICCAI-'10, Beijing, September 2010.

29. Xujiong Ye, Xinyu Lin, Jamshid Dehmeshki, Greg Slabaugh, and Gareth Beddoe, "Shape Based Computer-Aided Detection of Lung Nodules in Thoracic CT Images", IEEE Transactions on Biomedical Engineering, 56(7), pp: 1810-20, 2009.

30. McCulloch, C.C., Kaucic, R.A., Mendonça, P.R.S., Walter, D.J., Avila, R.S.: Model-based detection of lung nodules in computed tomography exams. Academic Radiology 11(3) (2004) 258–266

31. 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.

32. D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, 60, 2, pp. 91-110, 2004.

33. J.M. Morel and G. Yu., "ASIFT: A New Framework for Fully Affine Invariant Image Comparison," SIAM Journal on Imaging Sciences, pp. 438–469, 2009.

34. K. Murphy, B. van Ginneken, A.M.R. Schilham, B.J. de Hoop, H.A. Gietema, M. Prokop, "A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification", Medical Image Analysis, Volume 13, Issue 5, October 2009, Pages 757–770.

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

36. Namin, S.T., Moghaddam, H.A., Jafari, R., Esmaeil-Zadeh M., and Gity, M., "Automated detection and classification of pulmonary nodules in 3D thoracic CT images", Systems Man and Cybernetics (SMC), 2010 IEEE International Conference

37. 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.

38. Samala, R., et al., "A Novel Approach to Nodule Feature Optimization on Thin Section Thoracic CT." Acad. Radiology. Vol. 15, pp.1181–1197.2009

39. N. Paragios, R. Deriche, "Active Contours and Level Sets for the Detection and Tracking of Moving Objects", PAMI, No. 3, March 2000, pp. 266-280.

40. 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.

41. P. K. Sahoo, A. A. Farag, and Y. P. Yeap, "Threshold Selection based on Histogram Modeling," 1992 IEEE International Conference on Systems, Man, and Cybernetics, Chicago, IL, pp. 351¬356, October 1992.

42. P. C. Chen and T. Pavlidis, "Segmentation by texture using a cooccurrence matrix and a split-and-merge algorithm,"Comput. Graphics Image Processing, vol. 10, pp. 172-182, 1979.

43. A. Farag, A. El-Baz, G. L. Gimel'farb, "Precise segmentation of multimodal images," IEEE Transactions on Image Processing Vol. 15, no. 4, April 2006, pp. 952-968.

44. Stefano Diciotti, Simone Lombardo, Massimo Falchini, Giulia Picozzi, Mario Mascalchi. "Automated segmentation refinement of small lung nodules in CT scans by local shape analysis." IEEE Transactions on Biomedical Engineering, 58(12):3418–28, 2011.

45. D.M. Libby, J.P. Smith, N.K. Altorki, M.W. Pasmantier, D. Yankelevitz, and C.I. Henschke, "Managing the small pulmonary nodule discovered by CT." Chest, 125:1522-1529, 2004.

46. K. Okada, D. Comaniciue, and A. Krishnan, "Robust anisotropic Gaussian fitting for volumetric characterization of pulmonary nodules in multislice CT." IEEE Transactions on Medical Imaging, 24(3):409–423, 2005.

47. A.P. Reeves, A.B. Chan, D.F. Yankelevitz, C.I. Henschke, B. Kressler, and W.J. Kostis, "On measuring the change in size of pulmonary nodules." IEEE Transactions on Medical Imaging, 25(4):435–450, 2006.

48. MacMahon H, Austin JH, Gamsu G, et al. "Guidelines for management of small pulmonary nodules detected on CT scans: a statement from the Fleischner Society," Radiology 2005; 237:395-400.

49. Zhao B, Yankelevitz D, Reeves A, Henschke C., "Two-dimensional multi-criterion segmentation of pulmonary nodules on helical CT images," Med Phys., 1999 Jun;26(6):889–95.

50. Jane P. Ko, Henry Rusinek, Erika L. Jacobs, James S. Babb, Margrit Betke, Georgeann McGuinness, and David P. Naidich, "Small Pulmonary Nodules: Volume Measurement at Chest CT-Phantom Study," Radiology, pp: 864–870, 2003.

51. J.M. Kuhnigk, V. Dicken, L. Bornemann, A. Bakai, D. Wormanns, S. Krass, and H.O. Peitgen, "Morphological segmentation and partial volume analysis for volumetry of solid pulmonary lesions in thoracic CT scans."IEEE Transactions on Medical Imaging, 25(4):417434, 2006.

52. T. Kubota, A.K. Jerebko, M. Dewan, M. Salganicoff, and A. Krishnan, "Segmentation of pulmonary nodules of various densities with morphological approaches and convexity models." Medical Image Analysis, 15(1):133–154, 2011.

53. Dijia Wu, Le Lu, Jinbo Bi, Yoshihisa Shinagawa, Kim Boyer, Arun Krishnan, and Marcos Salganicoff1, "Stratified Learning of Local Anatomical Context for Lung Nodules in CT Images." Computer Vision and Pattern Recognition (CVPR), San Francisco, California, USA, June, 2010.

54. 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," Journal of Radio Graphics, Vol. 19 pp.1303–1311 (1999).

55. Stefano Diciotti, Giulia Picozzi, Massimo Falchini, Mario Mascalchi, Natale Villari, and Guido Valli, "3-D Segmentation Algorithm of Small Lung Nodules in Spiral CT Images," IEEE Transactions on Information Technology in Biomedicine, 12(1):7–19, 2008.

56. Artit C. Jirapatnakul, Yury D. Mulman, Anthony P. Reeves, David F. Yankelevitz, and Claudia I. Henschke, "Segmentation of Juxtapleural Pulmonary Nodules Using a Robust Surface Estimate." International Journal of Biomedical Imaging, Volume 2011, Article ID 632195, 2011.

57. J. E. Besag, "Spatial interaction and the statistical analysis of lattice systems," J. Roy. Stat. Soc. B, vol. 36, pp. 192–236, 1974.

58. J. E. Besag, "On the statistical analysis of dirty pictures," J. Roy. Stat. Soc. B, vol. 48, pp. 259–302, 1986.

59. S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," IEEE TPAMI, vol. 6, pp. 721–741, 1984.

60. H. Derin and H. Elliott, "Modeling and segmentation of noisy and textured images using gibbs random fields," IEEE TPAMI, vol. 9, no. 1, pp. 39–55, 1987.

61. S. Zhu, Y. Wu and D. Mumford, "Filters, random fields and maximum entropy (FRAME): Towards a unified theory of texture modeling," Int. Journal of Computer Vision, IJCV, Vol. 27(2), pp. 107-126, 1998.

62. Y. Zhang , M. Brady, S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," IEEE Trans. On Med Imaging, Vol. 20(1): pp. 45-57, Jan, 2001.

63. R. O. Duda, P. E. Hart, and D. G. Stork, Pattern Classification New York: Wiley, 2001.

64. B. W. Silverman, Density Estimation for Statistics and Data Analysis, Chapman and Hall, London, 1986.

65. P. Dempster, N. M. Laird, and D. B. Rubin, Maximum likelihood from incomplete data via the EM algorithm, J. Roy. Stat. Soc. B, vol. 39, pp. 1-38, 1977.

66. R. Redner and H. Walker, Mixture densities, maximum likelihood and the EM algorithm (review), SIAM Rev., vol. 26, pp. 195-237, 1984.

67. A. Ali and A. A. Farag, "Density Estimation using a new AIC type Criterion and the EM algorithm for a Linear Combination of Gaussians," ICIP-08, pp. 3024-3027, 2008.

68. S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," IEEE TPAMI, vol. 6, pp. 721–741, 1984.

69. R. W. Picard and I. M. Elfadel, "On the Structure of Aura and Co-occurrence Matrices for the Gibbs Texture Model," Journal of Mathematical Imaging & Vision, No. 2, pp. 5-25, 1992.

70. A. El-Baz, A. Farag, and G. Gimel'farb, and A. E. Abdel-Hakim, "Robust Image Registration Based on Markov-Gibbs Appearance Model," *Proc. of IEEE International Conference on Pattern Recognition (ICPR'06)*, Hong Kong, August 20–24, 2006, pp. 1204 - 1207.

71. Y. Y. Boykov and M. P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images," in Proc. of ICCV, vol. 1, 2001, pp. 105–112.

72. A. Ali, Amal Farag, and A. Farag, "Multimodal Imaging- Modeling and Segmentation with Biomedical Applications," Journal of IET Computer Vision, submitted October 2010.

73. Amal Farag, James Graham, and Aly Farag, "Robust Segmentation of Lung Tissue In Chest CT Scanning," Proc. of 2010 IEEE International Conference on Image Processing (ICIP), pp. 2249-2252, 2010.

74. Amal Farag, James Graham and Aly Farag, "Statistical Modeling of the Lung Nodules in Low Dose Computed Tomography Scans of the Chest," Proc. of 2010 IEEE International Conference on Image Processing (ICIP), pp. 4281-4284, 2010.

75. Asem Ali, Amal Farag and Aly Farag, "Labelling Color Images by Modelling the Colors Density Using A Linear Combination of Gaussians and EM Algorithm" 2009 IEEE International Conference on Image Processing (ICIP), Nov. 7 – Nov. 10, 2009, Cairo, Egypt.

76. Hossam Abd EL Munim and Aly A. Farag, "Curve/Surface Representation and Evolution using Vector Level Sets with Application to the Shape-based Segmentation Problem," IEEE Transactions on Pattern Analysis and Machine Intelligence – Vol. 29, No. 6, pp. 945-958, June 2007.

77. Hossam Abdelmunim and Aly Farag, "A New CAD System for the Evaluation of Renal Rejection Using DCE-MRI," Proc. of IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA'08), Anchorage, Alaska, June 27-28, 2008.

78. W. M. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal Volume Registration by Maximization of Mutual Information," Medical Image Analysis, vol. 1, no. 1, pp. 35-51, 1996.

79. Paul J. Besl, Neil D. McKay: A Method for Registration of 3-D Shapes. IEEE Trans. Pattern Anal. Mach. Intell. 14(2): 239-256 (1992).

80. J. Maintz and M. Viergever, "A survey of medical image registration", Medical Image Analysis, Vol. 2, No. 1, pp. 1-36, 1998.

81. J. Pluim, J. Maintz and M. Viergever, "Mutual-information-based registration of medical images: a survey," IEEE Trans. On Medical Imaging, Vol. 22 (8), pp. 986-1004, 2003.

82. T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns" in IEEE Transactions on Pattern Analysis and Machine Intelligence, 24, 2002, pp. 971-987

83. A. E. Abdel-Hakim and A. A. Farag, "CSIFT: A SIFT Descriptor with Color Invariant Characteristics," Proc. of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, 17-22 June 2006, pp. 1978-1983.

84. H. Bay, A. Ess, T. Tuytelaars et al., Speeded-Up Robust Features (SURF), Computer Vision and Image Understanding, vol. 110, no. 3, pp. 346-359, 2008.

85. K. Mikolajczyk, C. Schmid, "A performance evaluation of local descriptors," IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(10):1615-1630, 2005.

86. Daugman, J. How iris recognition works. Proceedings of 2002 International Conference on Image Processing, Vol. 1, 2002.

87. C.T. Zahn and R.Z. Roskies, "Fourier descriptors for plane close curves", IEEE Trans. Computers, Vol C-21, March 1972, pp. 269-281.

88. Amal Farag, Asem Ali, Shireen Elhabian, James Graham, Aly Farag, and Robert Falk, "Feature-Based Lung Nodule Classification," Proc. of International Symposium on Visual Computing (ISVC), pp. 79-88, 2010.

89. Amal Farag, Shireen Elhabian, James Graham, Aly Farag, and Robert Falk, "Toward Precise Pulmonary Nodule Descriptors for Nodule Type Classification," Proc. of the 13th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), pp. 626-633, 2010.

90. J. J. Koenderink, "The structure of images. Biological Cybernetics," Vol. 50:363-396, 1984.

91. T. Lindeberg, "Detecting salient blob-like image structures and their scales with a scale-space primal sketch: a method for focus-of-attention. International Journal of Computer Vision, 11(3):283-318.

92. B. Tao and B. Dickinson, "Texture recognition and image retrieval using gradient indexing," Journal of Visual Communication and Image Representation, Vol. 11(3), pp. 327–342, 2000.

93. Amal Farag, Asem Ali, James Graham, Aly Farag, Salwa Elshazly and Robert Falk. "Feature-Based Nodule Classification and Detection in Low Dose CT scans of the human chest." ISBI 2011

94. Amal A. Farag, Aly A. Farag, Hossam Abdelmunim, Asem M. Ali, James Graham, Salwa Elshazly, Ahmed Farag, Sabry Al Mogy ,Mohamed Al Mogy, Sahar Al Jafary, Hani Mahdi, Robert Falk and Rebecca Milam. "Feature Descriptors for Nodule Type Classification". CARS 2011.

CURRICULUM VITAE

Amal A. Farag

Objective: Interested in research and may pursue an academic career. Currently, graduate research assistant at the Computer Vision and Image Processing Laboratory, University of Louisville, USA, working on biomedical image analysis with focus on lung cancer.

Education

Degree	Year	Place	Area
Bachelor of Science	May 2008	University of Louisville	Electrical and Computer Engineering
Master of Engineering	May 2009	University of Louisville	Electrical and Computer Engineering
Doctoral Student	May 2009	University of Louisville	Electrical and Computer Engineering

Master of Engineering Thesis:

Lung Nodule Modeling and Detection for Computerized Image Analysis of Low Dose CT Imaging of the Chest. Master of Engineering. Advisors: Dr. James Graham and Dr. Aly A. Farag, May 2009.

Employments

Date	Place
2007 - Current	Research Assistant, Computer Vision and Image Processing Laboratory, University of Louisville, USA
2006	Three Semester Engineering Co-OP, Computer Vision and Image Processing Laboratory, University of Louisville, USA

Honors

| 2009 | NASA Graduate Student Fellowship |
| 2009 | ICIP 2009 Travel Grant, Cairo, Egypt November 2009 |

2009	Delegate, among world-wide competition, to the 21st International School for Computer Science Researchers on Molecular and Medical Image Analysis and BioInformatics; Lipari ,Italy; July 11th-18th 2009
2009	Master of Engineering with Highest Honors, University of Louisville; Graduate GPA 4.0/4.0
2009	2nd place, overall achievements in the Engineering Day Competition, University of Louisville, March 2009
2008	1st place winner of Robotics Design Competition, University of Louisville

Professional Service

2010	ECE Department Graduate School Student Representative
2009	Volunteer at ICIP 2009, Cairo, Egypt November 2009
2008-Present	Student Member, IEEE

Course Work: ABET accredited Electrical and Computer Engineering curricula with research focus on signals and systems, computer vision, biomedical imaging and computing. Software tools include MatLab, C/C++ and optimization techniques.

Patents

1. Title: Data-Driven Models for Detection and Classification of Pulmonary Nodules. Application No. 61/257,579 (Draft to be resubmitted to UofL OTT)

Publications

Journals

1. Asem Ali, **Amal Farag** and Aly Farag, "Multimodal Imaging- Modeling and Segmentation with Biomedical Applications," IET Computer Vision. To Appear

2. Hossam Abdelmunim, **Amal Farag** and Aly Farag, "Shape Representation and Registration in Vector Implicit Spaces: Adopting a Closed Form Solution in the Optimization Process," IEEE Transactions on Pattern Analysis and Machine Intelligence, Revised Manuscript Submitted April 2012.

3. **Amal Farag**, James Graham, Hossam Abdelmunim and Aly Farag, "On Lung Nodule Detection Under Uncertainty", IEEE Transactions on Image Processing (In preparation)

4. __ "Variational Approach for Lung Nodule Segmentation," IEEE Transactions on Image Processing (In preparation)

5. __ "Nodule Classification and Categorization from LDCT Scans," IEEE Transactions on Image Processing (In preparation)

6. __ "A Data-Driven CAD System for Lung Nodule Detection and Classification from LDCT Scans," IEEE Transactions on Medical Imaging (In preparation)

Conferences

7. **Amal Farag,** Hossam Abdelmunim, James Graham, Aly A. Farag, Cambron Carter, Salwa Elshazly, Mohamed El-Mogy, Sabry El-Mogy and Robert Falk. An AAM-Based Detection Approach of Lung Nodules from LDCT Scans," International Symposium on Biomedical Imaging (ISBI-12) – Barcelona, Spain, May 1-3, 2012 (Oral Presentation).

8. **Amal Farag,** Hossam Abdelmunim, James Graham, Aly A. Farag, Salwa Elshazly, Asem M. Ali, Ahmed Farag, Sabry Al Mogy,Mohamed Al Mogy , Robert Falk ,Sahar Al Jafary, Hani Mahdi, and Rebecca Milam, "Variational Approaches For Segmentation Of Lung Nodules," IEEE International Conference on Image Processing (ICIP-2011), pp. 2157 - 2160, Sept 2011.

9. **Amal Farag,** Aly A Farag, Hossam Abdelmunim, Asem Ali, James Graham, Salwa Elshazly, Ahmed Farag, Sabry Al Mogy, Mohamed Al Mogy , Sahar Al Jafary, Hani Mahdi, Robert Falk and Rebecca Milam, "Feature Descriptors For Nodule Type Classification," CARS-2011 (Oral Presentation).

10. **Amal Farag**, Asem Ali, James Graham, Aly Farag, Salwa Elshazly and Robert Falk, "Evaluation of geometric feature descriptors for detection and classification of lung nodules in low dose ct scans of the chest," International Symposium on Biomedical Imaging (ISBI-11), pp. 169 - 172, Chicago, March 2011.

11. **Amal Farag**, Asem Ali, Shireen Elhabian, and Aly A Farag, "Probability density estimation by linear combinations of Gaussian kernels-generalizations and algorithmic evaluation, IEEE International Conference on Multimedia Technology (ICMT), pp. 6491-6494, 2011.

12. **Amal Farag,** Asem Ali, Shireen Elhabian, James Graham, Aly Farag and Robert Falk, Feature-Based Lung Nodule Classification. International Symposium on Visual Computing (ISVC- 10), Las Vegas, November 2010, pp. 79-88.

13. **Amal Farag**, Shireen Elhabian, James Graham, Aly Farag and Robert Falk, Toward Precise Pulmonary Nodule Descriptors for Nodule Type Classification, 13th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI-10), Beijing, pp. 626-633, September 2010.

14. **Amal Farag**, James Graham, Aly Farag, Salwa Elshazly and Robert Falk, Parametric and Non-Parametric Nodule Models: Design and Evaluation, Proc. of Third International Workshop on Pulmonary Image Processing in conjunction with MICCAI-'10, Beijing, pp. 151-162, September 2010.

15. **Amal Farag**, James Graham and Aly Farag, Robust Segmentation of Lung Tissue In Chest CT Scanning, 2010 IEEE International Conference on Image Processing (ICIP-10), Hong Kong, pp. 2249-2252, September 2010.

16. **Amal Farag,** James Graham and Aly Farag, Statistical Modeling of the Lung Nodules in Low Dose Computed Tomography Scans of the Chest , 2010 IEEE International Conference on Image Processing (ICIP-10), Hong Kong, pp. 4281-4284 September 2010 (Oral Presentation).

17. **Amal Farag,** James Graham, Salwa Elshazly and Aly Farag, Data-Driven Lung Nodule Models for Robust Nodule Detection in Chest CT, 20th International Conference on Pattern Recognition (ICPR-10), Istanbul, pp. 2288-2291, August 2010.

18. **Amal Farag,** Shireen Elhabian, James Graham, Aly Farag, Salwa Elshazly, Robert Falk, Hani Mahdi, Hossam Abdelmunim and Sahar Al-Ghaafary, Modeling of the Lung Nodules for Detection in LDCT Scans, 32nd IEEE Engineering in Medicine and Biology Society (EMBC-10), August 2010 (Oral Presentation).

19. **Amal A. Farag**, Shireen Y. Elhabian, Salwa A. Elshazly and Aly A. Farag, Quantification of Nodule Detection in Chest CT: A Clinical Investigation Based on the ELCAP Study, Proc. of Second International Workshop on Pulmonary Image Processing in conjunction with MICCAI-09, London, September 2009, pp. 149-160.

20. **Amal Farag**, James Graham, Aly Farag and Robert Falk, Lung Nodule Modeling – A Data-Driven Approach, International Symposium on Visual Computing (ISVC- 09), Las Vegas, December 2009, pp. 347-356 (Oral Presentation).

21. Asem Ali, **Amal Farag** and Aly Farag , Labelling Color Images by Modelling the Colors Density Using A Linear Combination of Gaussians and EM Algorithm, 2009 IEEE International Conference on Image Processing (ICIP-09), Cairo, November 2009.

22. **Amal A Farag**, Shireen Elhabian, Abdelrehim Ahmed and Aly A Farag, "Noise Analysis of SFS Algorithm Formulated Under Various Imaging Conditions," 4th International Symposium on Visual Computing (ISVC-08), Las Vegas, Nevada, USA, pp. 803-812, December 1-3, 2008.

23. Shireen Y. Elhabian, **Amal A. Farag**, Salwa Elshazly and Aly A Farag, Sensitivity of Template Matching for Pulmonary Nodule Detection: A Case Study, 4th Biomedical Engineering Conference, CIBEC 2008, Cairo 18-20 Dec. 2008, pp. 110 -114.

PROFESSIONAL REFERENCES

1. Dr. James Graham, Professor and Chairman, Department of Electrical and Computer Engineering, University of Louisville, Kentucky, USA; E-mail: jhgrah01@lousiville.edu (Dissertation Advisor).

2. Dr. John Naber, Professor and Associate Chairman, Department of Electrical and Computer Engineering, University of Louisville, Kentucky, USA; E-mail: jfnabe01@lousiville.edu

3. Dr. Robert Falk, Director of Medical Imaging, Jewish Hospital, Louisville, Kentucky, USA; E-mail: rfalk@3drinc.com

4. Dr. Mohamed N. Ahmed, Senior Technical Lead - IBM, Washington, DC; E-mail: ahmoham@us.ibm.com

5. Dr. Aly A. Farag, Professor of Electrical and Computer Engineering, University of Louisville, Kentucky, USA; E-mail: aly.farag@louisville.edu

6. Mr. Mike Miller, Senior Research Scientist, Computer Vision and Image Processing Laboratory, University of Louisville, Kentucky, USA; E-mail: mike@cvip.uofl.edu