International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCICCT' 14 Conference Proceedings

Automated 3-D Segmentation of Lungs with Lung Cancer in CT Data using a Watershed Algorithm

A.Niranjana
Assistant Professor
Department of Computer Science and Engineering,
Easwari Engineering College, Anna University,
Ramapuram, Chennai-89, India
nithu6289@gmail.com

***Abstract*— Segmentation of lungs with lung cancer regions is a complicated problem. By presenting a new fully automated approach for segmentation of lungs with such high-density pathologies. This method consists of three main dispensation stepladders. A novel robust active shape model (RASM) matching technique is utilized to roughly segment the outline of the lungs. The method allows the detection of a large number of element points in the illustration. Robust point matching is used to search for the communication between feature and model points while the model is being deformed along the modes of variation of the active shape model. The initial position of the RASM is found by means of a rib cage detection technique. Second, an optimal surface finding approach is utilized to promote the original segmentation consequence to the lung. Left and right lungs are segmented independently. Third, watershed algorithm is utilized to segment the diseased region. Experiments on the same 30 data sets showed that our methods delivered statistically significant better segmentation results, compared to the two commercially existing lung segmentation approaches. In adding up, the approaches are generally applicable and suitable for large shape models.**

Keywords—Segmentation, Lung segmentation, Optimal Surface finding, Rib Detection, Robust Active Shape Model.

I. INTRODUCTION

Lung cancer is one of the solemn cancer which causes more death globally. Statistics says 28 percent of overall cancer deaths are due to lung cancer. It is the number one cause of death from cancer every year and the second most diagnosed after breast and prostate cancers for women and men respectively. Lung cancer is typically originated in older persons because it develop over a long period of time. Lung cancer is usually visible as small round lesion called 'nodules' through Medical images like CT. However the most important crisis in identifying the nodules with CT image is that the distinctiveness of nodules are similar to the distinctiveness of blood vessels and bronchi in terms of size, shape and density. Consequently it is obligatory to enhance and segment the nodules using some unique image processing techniques. Early detection of lung cancer using the CT image has the largest chance of saving the patient's life. It is important to augment and perceive the nodules in CT images in order to identify the lung cancer at early period. Different methods and algorithms are developed to effectively detect the nodules.

Segmentation is one of the most imperative problems in image processing. It consists of constructing a symbolic representation of the image. The image is described as homogeneous areas according to one or several apriority attributes and also used to find the various segmentation algorithms. LUNG cancer represents a foremost health problem. Worldwide, lung cancer is dependable for 1.3 million deaths annually, according to the WHO. Tomographic imaging modalities like multi-detector X-ray computed tomography (CT) play an important role in diagnosis, treatment and research of lung cancer. State-of-the-art CT imaging technology enables physicians to create high-resolution volumetric scans recounting lung anatomy and pathology. Higher resolution benefits diagnostic capabilities, but on the other hand, the increased amount of image data to be analyzed represents a burden for physicians.

CT scan is a computed tomography (CT) scan uses X-rays to make detailed pictures of structures inside of the body. The CT scanner sends X-rays through the body area. Each rotation of the scanner provides a picture of a thin slice of the organ. All of the pictures are saved as a group on a computer. They also can be printed in some cases; a dye called contrast material may be used. It may be place in a vein in an arm or it may be placed into other parts of your body to see those areas better. For some types of CT scans you drink the dye. The dye makes structures and organs easier to see on the CT pictures. A CT scan can be used to cram all parts of the body such as the chest, belly, pelvis, an arm and leg. It can take pictures of body organs, such as the liver, pancreas, intestines, kidneys, bladder, adrenal glands, lungs and heart. It also can tidy blood vessels, bones and the spinal cord.

Robust point matching is used to search for the correspondence between feature and model points while the model is being deformed along the modes of variation of the active shape model. Both synthetic and in vivo results have shown that the proposed algorithm is immune to feature noise and can tolerate over 50% of outliers.

II. RELATED WORKS

To address this problem, an automated lung image analysis methods are required. Segmentation of a lung with cancer using a conventional approach. (a) Axial CT image shows normal right and cancerous left lung tissue.(b) Corresponding segmentation result generated with a conventional lung segmentation technique. Segmentation errors are indicated by arrows. Many approaches to automated quantification of lung disease require the segmentation of lung parenchyma in an initial processing step. In the case of normal lungs imaged with CT, a large density difference between air-filled lung parenchyma and surrounding tissues can be observed.

A number of algorithms can be found in the literatures that rely on this observation for the segmentation of lungs. In the case of lungs with lung cancer [1] or other high density pathologies (e.g., pneumonia), lung segmentation becomes a nontrivial task, and frequently, conventional algorithms fail to deliver suitable segmentation results [2]. Thus, to enable computer-aided cancer treatment planning (e.g., surgery or radiation treatment) and to facilitate the quantitative assessment of lung cancer masses (e.g., evaluation of treatment response), robust lung segmentation methods are needed. In this work, we present a new approach for the fully automated segmentation of lungs with lung cancer regions which addresses the limitations of existing methods like robustness or processing speed.

Approach is based on a robust model matching method for 3-D active shape models (ASM). It builds on preliminary work, which required a manual initialization of the ASM. To address this limitation, we propose a model initialization method which is based on a novel rib detection approach that is suitable for normal or contrast enhanced CT scans. The performance of our fully automated lung segmentation system is assessed on 30 lung CT scans with 40 abnormal (lung cancer) and 20 normal (no signs of lung disease) left/right lungs.

In addition, we provide a performance comparison with two commercially available methods on the same image data. Both methods are utilized routinely in the context of lung radiation treatment planning. The first method is based on a region growing algorithm and the second method utilizes a deformable template-based segmentation approach. In terms of computing time, the model-based 3-D segmentation of lungs is particularly challenging, because of the size of lungs and the amount of image data to be processed [3].

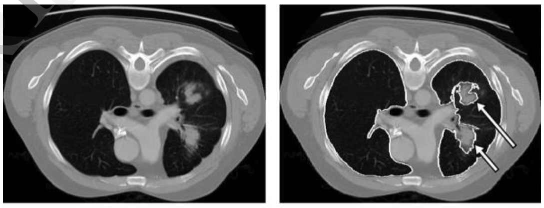
Additionally watershed algorithm is used to segment the lung region. The watershed transform is the method of choice for image segmentation in the field of mathematical morphology. The watershed transform can be classified as a region-based segmentation approach. The intuitive idea underlying this method comes from geography: it is that of a landscape or topographic relief which is oozed by water, watersheds being the divide lines of the domains of attraction of rain falling over the region.

An alternative approach is to imagine the landscape being immersed in a lake, with holes pierced in local minima. The Watershed Transform will fill up with water starting at these local minima, and, at points where water coming from different basins would meet, dams are built. When the water level has reached the highest peak in the landscape, the process is stopped. As a result, the landscape is partitioned into regions or basins separated by dams, called watershed lines or simply watersheds [4].

Computed Tomography as a Screening Method for Lung Cancer: The successes in screening for cancers such as cancer of the cervix, breast and the prostate, along with the widespread introduction of MDCT scanning into large and small medical centres, has lead to the re-discussion of lung cancer screening [5]. A number of projects investigating the feasibility and best methodology of lung cancer screening are being conducted both in the United States of America and in Europe.

The Early Lung Cancer Action Project (ELCAP) began in 1992 to assess the usefulness of annual MDCT screening for lung cancer. In 2002 the National Lung Screening Trial (NLST) was launched to investigate if lung cancer screening by either x-ray or MDCT actually saves lives [6]. A similar study was initiated in the Netherlands i[7]. These lung cancer screening studies have shown MDCT can detect early stage lung cancer [8,9]. While the complete mortality data for these studies is not yet available, an evaluation in 2004 by the U.S. Preventive Services Task Force of all available lung cancer screening data found they ‘could not determine the balance between the benefits and harms of screening for lung cancer’ [10].

One of the primary concerns with lung cancer screening using MDCT is what to do with an identified new nodule and in many cases the recommendation is to watch the nodule for signs of growth over 3 to 24 months. This time period is required, as a large number of early detected nodules (97%) are false positives and will not require treatment. Tracking growth is one way of separating non-cancerous nodules, which are unlikely to change over this time period, from cancerous ones. Thus, it is important to develop the most effective possible ways of staging these early detected nodules as well as reliable methods for tracking nodule growth.



(a) (b)

Fig. 1. Segmentation of a lung with cancer using a conventional approach

III. PROPOSED METHODOLOGIES

System architecture is the theoretical or conceptual model that defines the structure and performance of the system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structure of the system. It provides a way in which products can be consists of three major components. Procured, systems can be developed an architectural indication of the overall system. There have been efforts to formalize languages to describe system architecture, collectively these are called architecture description languages (ADLs). The system architecture for the system is given in Figure1.The system architecture

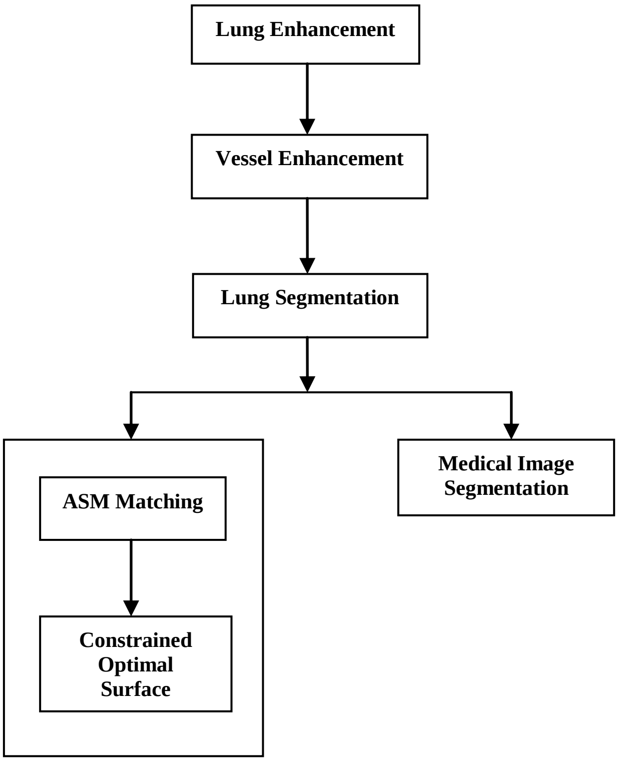


Fig. 2. System Architecture

A. Lung and vessel enhancement

Vessel enhancement method initially pre-processing a CT image is done to remove the noises present in it. Then the vessels are enhanced using the morphological operators followed by lung segmentation. The fissure regions are identified, enhanced and verified. Lung nodules are segmented using adaptive threshold. Through the project they have developed a segmentation algorithm for identifying the fissures and nodules from CT images. Using the statistical analysis such as mean difference between the automatic and manual segmentation, the results are verified.

B. Image segmentation

In computer vision, image segmentation is the process of partitioning a digital image into multiple segments. The aspiration of segmentation is to abridge and change the representation of an image into something that is more consequential and easier to explore. Image segmentation is typically used to locate objects and boundaries in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image (see edge detection). Each of the pixels in a region is similar with respect to some characteristic or computed property, such as colour, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristics. When applied to a stack of images, typical in medical imaging, the resulting contours after image segmentation can be used to create 3D reconstructions with the help of interpolation algorithms like marching cubes. Segmentation is an important step in medical image analysis and classification for radiological evaluation or computer aided diagnosis.

C. Medical Image Segmentation

Image segmentation refers to the process of partitioning an image into distinct regions by grouping together neighbourhood pixels based on the some predefined similarity criterion. The similarity criterion can be determined using specific properties or features of pixels representing objects in the image. Segmentation is a pixel classification method that allows the pattern of regions of similarity in the image. Segmentation has been remained as an important tool in medical image processing and it has been useful in many applications. The applications include detection of the coronary border in angiograms, multiple sclerosis lesion quantification, surgery simulations, surgical planning, measuring tumor volume and its response to therapy, functional mapping, automated classification of blood cells, studying brain development, detection of micro calcification on mammograms, image registration, atlas matching, heart image extraction from cardiac cine angiograms, detection of tumours etc. In medical imaging, segmentation is important for feature extraction, image measurements, and image display.

In some applications it may be useful to classify image pixels into anatomical regions, such as bones, muscles, and blood vessels, while in others into pathological regions, such as cancer, tissue deformities and multiple sclerosis lesions. In some studies the goal is to divide the entire image into sub regions such as the white matter, gray matter, and cerebrospinal fluid spaces of the brain, while in others on specific structure has to be extracted, for example breast cancer from Magnetic Resonance images. The goal of the lung segmentation required for the computer aided diagnosis from CT scan images is to essentially separate the voxels corresponding to the lung cavity in the axial CT scan slices from the surrounding lung anatomy.

D. ASM matching

For ASM (Active Shape Model) based segmentation, utilize a novel robust ASM matching approach that extends the standard ASM matching scheme described below.

1) Standard ASM Matching: The PDM can be used for lung segmentation by matching the model to the target structure. This could be achieved by utilizing a standard ASM matching framework. The matching procedure consists of four steps. a) An instance of the shape model (e.g., mean shape) is generated and placed in proximity to the target structure. b) To match the model to the target, shape points are updated by searching from the current landmark location along a profile normal to the model surface with length. c) Once all shape points are updated, pose parameters are adjusted to map the updated shape points in the target image coordinate frame to the mean shape in the model coordinate frame. d) Steps a–c are repeated until the model converges.

2) Robust ASM Matching: We are interested in the segmentation of pathological lungs that contain large areas of lung cancer (high density). Thus, it is very likely that some update points are found during the model matching procedure

that do not represent lung surface (outliers) and belong to an area of transition between normal and diseased lung tissue. Consequently, the standard matching approach will fail, because it is a least squares optimization procedure that is not suitable to handle outliers. Therefore, a robust shape model matching approach is required.

E. Constrained Optimal Surface

Depending on the training data utilized for model building, the model might not be able to describe smaller local shape variations. To capture this information, we generate the final lung segmentation by applying a global optimal surface finding method, which allows finding a smooth surface related to a shape prior. The algorithm transforms the segmentation problem into a graph optimization problem, which is solved by means of a maximum- flow algorithm. An edge-weighted directed graph is built, and weights derived from the volume are assigned to the graph edges to reflect local image properties. For this purpose, the final ASM mesh is utilized. Since the ASM vertices are sparse, the mesh is restructured by adding triangles, before the graph is built. For graph generation, columns along the surface normal of each vertex are generated. The length of the profile is utilized to constrain the segmentation to the proximity of the initial ASM segmentation.

IV. METHODOLOGIES (ALGORITHM)

In the proposed work, the algorithm used is Watershed Algorithm. Watershed Algorithm is the purpose of the segmentation of the lung region in the CT image is to achieve a better orientation in the image. For separating the regions of interest or objects of interest from other parts of the image a region growing approach is used to distinguish between the specific nodules of lungs and other suspicious region. Region growing is a technique for extracting a region of the image that is connected based on some predefined criteria. This criterion can be based on intensity information and/or edges in the image. In its simplest form, region growing requires a seed point that is manually selected by an operator, and extracts all pixels connected to the initial seed with the same intensity value. The watershed transform is a popular segmentation method coming from the field of mathematical morphology.

I. Algorithmic Steps for Watershed Segmentation:

1. Accept feature map as input and build an image boundary around the borders of the image.
2. Threshold small fluctuations on the low end of the input feature map.
3. Locate and label all single pixel regional minima (pixels those are lower in value than all their 4-neighbors). Pixels are labeled with unique integer value, beginning with 1.
4. Identify all flat regions, give them a distinct label and locate the lowest pixel surrounding the boundary of the region.
5. Trace all pixels that are not members of flat region.
6. Trace all remaining unlabeled flat regions to their respective regional minima (starting with lowest pixel that has already been labeled. By this step all the pixels in the image are labeled with a distinct region label).
7. Mark the flat regions that are regional maxima and mark boundary pixels between region.
8. Depth of each region i.e. Watershed is computed as a difference between the largest and smallest valued pixel in that watershed.
9. Watersheds are threshold according to their depths.
10. Sequentially labeling all individual regions bounded by the marked edges in the binary image to create a segmented image.

Figure.1 illustrates the original image and steps of the segmentation process applied on Lung CT scan training data set.

```
1. Make a 3-D binary image containing two overlapping
   spheres.
2. centre1 = -10;
3. centre2 = -centre1;
4. dist = sqrt(3*(2*centre1)^2);
5. radius = dist/2 * 1.4;
6. lims = [floor(centre1-1.2*radius)
   ceil(centre2+1.2*radius)];
7. [x,y,z] = meshgrid(lims(1):lims(2));
8. bw1 = sqrt((x-centre1).^2 + (y-centre1).^2 + ...
9.      (z-centre1).^2) <= radius;
10.bw2 = sqrt((x-centre2).^2 + (y-centre2).^2 + ...
11.     (z-centre2).^2) <= radius;
12.bw = bw1 | bw2;
13.figure, isosurface(x,y,z,bw,0.5), axis equal,
   title('BW')
14.xlabel x, ylabel y, zlabel z
15.xlim(lims), ylim(lims), zlim(lims)
16. view(3), camlight, lighting gouraud

17.  Compute the distance transform.
18.E = bwdist(~bw);
19.figure, isosurface(x,y,z,E,radius/2), axis equal
20.title('Isosurface of distance transform')
21.xlabel x, ylabel y, zlabel z
22.xlim(lims), ylim(lims), zlim(lims)
   view(3), camlight, lighting gouraud

23.  Complement the distance transform, force non-object
   pixels to be -Inf, and then compute the watershed
   transform.
24.E = -E;
25.E(~bw) = -Inf;
26.L = watershed(E);
27.figure
28.isosurface(x,y,z,L==2,0.5)
29.isosurface(x,y,z,L==3,0.5)
30.axis equal
31.title('Segmented objects')
32.xlabel x, ylabel y, zlabel z
33.xlim(lims), ylim(lims), zlim(lims)
   view(3), camlight, lighting gouraud
```

Let $u(x,y)$ with $(x,y) \in R$, be a scalar function describing an image I. The morphological gradient of I is defined in Beucher by,

$$\delta Eu = (c \oplus E) - (c \ominus E)$$

Where $(c \oplus E)$ and $(c \ominus E)$ are respectively the elementary dilation and erosion of c by the structuring element E.The morphological Laplacian is given by,

$$\Delta Ec = (c \oplus E) - 2c + (c \ominus E)$$

Note that this morphological Laplacian allows us to distinguish influence zones of minima and uprema: regions with $\Delta Ec < 0$ are considered as influence zones of suprema, while regions with $\Delta Ec > 0$ are influence zones of minima. Then $\Delta Ec = 0$ allows us to interpret edge locations, and will represent an essential property for the construction of morphological filters. The zone of a minimum or a maximum. Basic idea is to apply either dilation or erosion to the image I, depending on whether the pixel is located within the influence.

V. PERFORMANCE EVALUATION

The performance evaluation can be done with watershed algorithm. The watershed lines always correspond to the most significant edges between the markers. So this technique is not affected by lower-contrast edges, due to noise, that could produce local minima, and, thus, erroneous results, in energy minimization methods. Even if there are no strong edges between the markers, the watershed transform always detects a contour in the area. This contour will be located on the pixels with higher contrast. An important task was to identify what features must be taken into consideration of a Dicom image for successfully detecting the lung cancer.

It is of great significance to find casual association, as well as, correlation between and among the cancer indicating features extracted from image for developing highly successful diagnostic system. Therefore, nature of data collected from various sources (various patients and subjects images). Therefore, the nature of the data was analyzed using regression technique. The main objective for doing analysis of this dataset was to explain variation of one variable (called dependent variable) defined in medical fraternity as state of cancer having either malignant or benign state represented as 0 and 1, respectively in our thesis, based on variations of one or more other variables called independent variables, which includes:

1. Image number
2. Subtletly (difficult to detect the nodule)
3. Number of lesions
4. Sphericity (roundness of a 3-D object)
5. Malignancy
6. Solidity (texture)
7. Lobulation (boundary marked or not)

Where, y is the dependent variable and x1, State (benign or malignant).

There are basically two approaches to regression:

- A hit-and-trial approach.
- A pre-conceived approach.

The general regression model (linear) is of type:

$$Y = a + b1x1 + b2x2 + \ldots\ldots + bnxn$$

x2,……x3 are the independent variables expected to be related to y and explain or predict y. b1, b2…………bn are the coefficients of the respective independent variables.

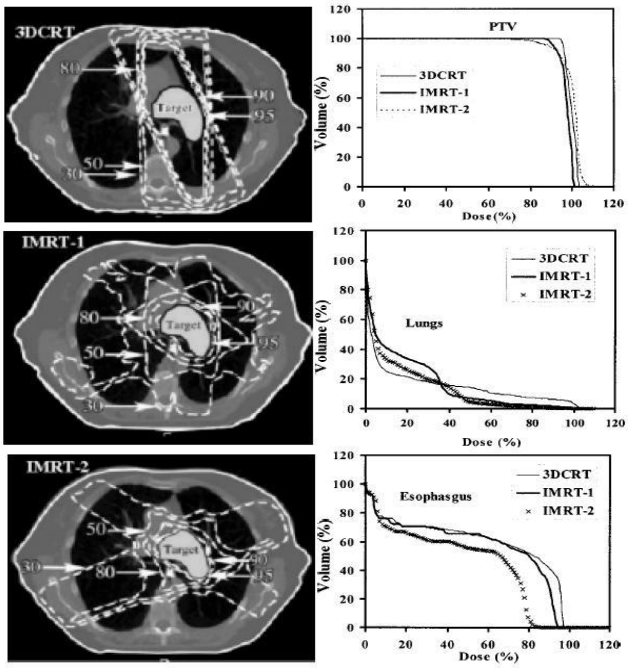


Fig. 3. Segmentation of a lung with cancer using throughput

Where,

CRT - Cathode Ray Tube

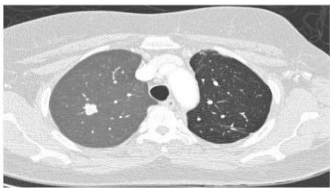
IMRT - Intensity Modulated Radiation Therapy

Wherever Times is specified, Times Roman or Times New Roman may be used. If neither is available on your word processor, please use the font closest in appearance to Times. Avoid using bit-mapped fonts if possible. True-Type 1 fonts are preferred.

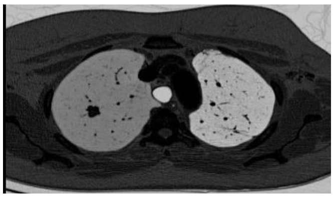
VI. CONCLUSION AND FUTURE WORK

A core component of our method is a novel robust ASM matching method. The approach not only allows coping with disturbances but it is also well suitable for large shape models and parallel implementation, allowing low computation times. Our robust ASM framework is also applicable to other segmentation problems as well as imaging modalities, requiring mainly an adaption of the matching cost function. Medical image segmentation definitely has a large potential in the medical domain. Watershed Segmentation method can be used on a large variety of images and in a wide area of applications. By applying segmentation tools on several pulmonary CT images of lung.

- First, the algorithm cost is interesting.
- Second, the experimental results show that the over segmentation problem, which usually appears with the watershed technique, can be attenuated, and the segmentation results can be performed using the topological gradient approach.
- Another advantage of this method is that it splits the segmentation process into two separate steps: first we detect the main edges of the image processed, and then we compute the watershed of the gradient detected. This methodology has many advantages, particularly in real life application.

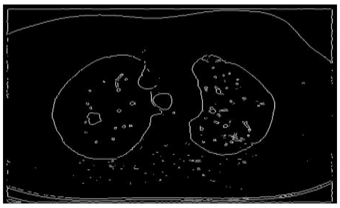


a)

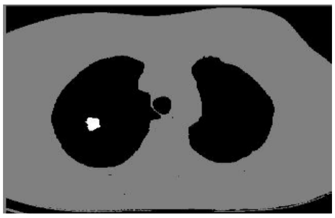


b)

c)



d)



e)

Fig. 4. Final results of feature extraction displayed on Lung CT scan images: (a) Original images, (b) image after ROI, (c) Results of image superimposition, (d) effects of watershed segmented, (e) extracted cancerous nodule from the image

ACKNOWLEDGMENT

I hereby greatly acknowledge thanks to the Head of the Department, Computer Science and Engineering, Easwari Engineering College, Chennai, to give me support and patient enough to make me understand the complexities in this research and relentlessly supported me till now.

REFERENCES

[1] S. Beucher and F. Meyer, “The morphological approach to segmentation: The watershed transform,” in Mathematical Morphology in Image Processing, E. R. Dougherty, Ed. New York: Marcel Dekker, 1993, vol. 12, pp. 433–481.

[2] J. Weickert, “Fast segmentation methods based on partial differential equations and the watershed transform,” in Proc. DAGM Symp., 1998,pp. 93–100.

[3] J. B. T. M. Roerdink and A. Meijster, “The watershed transform: Definitions, algorithms and parallelization strategies,” Fundamental Informaticae, vol. 41, pp. 187– 228, 2000.

[4] S. G. Armato and W. F. Sensakovic, “Automated lung segmentation for thoracic CT,” Acad. Radiol., vol. 11, no. 9, pp. 1011– 1021, 2004.

[5] J. K. Leader, B. Zheng, R. M. Rogers, F. C. Sciurba, A. Perez, B. E. Chapman, S. Patel, C. R. Fuhrman, and D. Gur, “Automated lung segmentation in X-ray computed tomography,” Acad. Radiol., vol. 10, no. 11, pp. 1224–1236, 2003.

[6] T. R. Church, "Chest radiography as the comparison for spiral CT in the National Lung Screening Trial," Acad Radiol, vol. 10, pp. 713-5, 2003.

[7] D. M. Xu, H. Gietema, H. de Koning, R. Vernhout, K. Nackaerts, M. Prokop, C. Weenink, J. W. Lammers, H. Groen, M. Oudkerk, and R. van Klaveren, "Nodule management protocol of the NELSON randomised lung cancer screening trial," Lung Cancer, vol. 54, pp. 177-84, 2006.

[8] B. Karacali, A. P. Vamvakidou, and A. Tozeren, "Automated recognition of cell phenotypes in histology images based on membrane- and nuclei-targeting biomarkers," BMC Med Imaging, vol. 7, pp. 7, 2007.

[9] S. J. Swensen, "Screening for cancer with computed tomography," Bmj, vol. 326, pp. 894-5, 2003.

[10] L. L. Humphrey, S. Teutsch, and M. Johnson, "Lung cancer screening with sputum cytologic examination, chest radiography, and computed tomography: an update for the U.S. Preventive Services Task Force," Ann Intern Med, vol. 140, pp. 740-53, 2004.