Published in final edited form as:
IEEE Trans Med Imaging. 2012 February ; 31(2): . doi:10.1109/TMI.2011.2171357.

Automated 3-D Segmentation of Lungs With Lung Cancer in CT Data Using a Novel Robust Active Shape Model Approach

Shanhui Sun,
Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, IA 52242 USA shanhuisun@uiowa.edu

Christian Bauer, and
Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, IA 52242 USA christian-bauer@uiowa.edu

Reinhard Beichel* **[Member, IEEE]**
Department of Internal Medicine and the Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, IA 52242 USA

Abstract

Segmentation of lungs with (large) lung cancer regions is a nontrivial problem. We present a new fully automated approach for segmentation of lungs with such high-density pathologies. Our method consists of two main processing steps. First, a novel robust active shape model (RASM) matching method is utilized to roughly segment the outline of the lungs. The initial position of the RASM is found by means of a rib cage detection method. Second, an optimal surface finding approach is utilized to further adapt the initial segmentation result to the lung. Left and right lungs are segmented individually. An evaluation on 30 data sets with 40 abnormal (lung cancer) and 20 normal left/right lungs resulted in an average Dice coefficient of 0.975 ± 0.006 and a mean absolute surface distance error of 0.84 ± 0.23 mm, respectively. Experiments on the same 30 data sets showed that our methods delivered statistically significant better segmentation results, compared to two commercially available lung segmentation approaches. In addition, our RASM approach is generally applicable and suitable for large shape models.

Keywords

Lung segmentation; optimal surface finding; rib detection; robust active shape model

I. Introduction

LUNG cancer represents a major health problem. World-wide, lung cancer is responsible for 1.3 million deaths annually, according to the WHO.[1] Tomographic imaging modalities like multidetector 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 describing lung anatomy and pathology. Higher resolution benefits diagnostic capabilities, but on the other hand, the

*(reinhard-beichel@uiowa.edu). .

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

[1]http://www.who.int/mediacentre/factsheets/fs297/en/;

increased amount of image data to be analyzed represents a burden for physicians. To address this problem, automated lung image analysis methods are required.

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 literature (e.g., [1]–[5]) that rely on this observation for the segmentation of lungs. We will denote such methods as conventional lung segmentation approaches. In the case of lungs with lung cancer [Fig. 1(a)] or other high density pathologies (e.g., pneumonia), lung segmentation becomes a nontrivial task, and frequently, conventional algorithms fail to deliver suitable segmentation results [Fig. 1(b)]. 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.

Only a few papers have been published that deal with segmentation of diseased lungs. None of the existing methods directly targets the segmentation of lungs with large cancer regions at arbitrary locations. For example, a Bézier surface-based method was proposed in [6] to deal with lesions adjacent to the chest wall and mediastinum. First a shape model of lung side walls is fitted to the target image by using an affine transformation. Second, an active contour model was utilized to refine the initial segmentation. Because the lung apex and base are not included in the model, the resulting segmentation was combined with a conventional segmentation method. Thus, lesions adjacent to the lung apex or diaphragm can result in segmentation errors. Pu *et al.* [7] proposed an automated lung segmentation approach based on a 2-D adaptive border marching algorithm to deal with juxtapleural lung nodules. Larger areas of under-segmentation were reported in hilar and pulmonary consolidation regions. In recent work, Pu *et al.* describe a shape “Break-and-Repair” strategy which was utilized to segment lungs with juxtapleural lung nodules [8]. A method for the robust segmentation of lung parenchyma based on the curvature of ribs was presented in [9]. The method is based on an adaptive thresholding scheme and utilizes a comparison of the curvature of the lung boundary to the curvature of the ribs to select thresholds. Because lung pathologies like cancer can have density values similar to other tissues surrounding the lung, the method will likely show errors in such cases. Recently, Wang *et al.* proposed a method for the segmentation of lungs with interstitial disease [10]. First, an initial segmentation was generated by utilizing a threshold-based conventional lung segmentation method. Second, interstitial lung tissue regions were identified based on texture features. The resulting segmentations were then combined to form the final segmentation result. Sluimer *et al.* proposed a segmentation by registration approach for the segmentation of pathological lungs [11]. While delivering promising results, not all pathological cases could be handled successfully [11]. In addition, the authors also identified the need to reduce computing time from 3 h to a clinically more acceptable processing time [11]. To solve this problem, a hybrid lung segmentation method was presented in a recent publication of the same group [12]. The basic idea is to first use a conventional lung segmentation method, assess the correctness of the segmentation based on volume and shape features, and utilize the segmentation by registration approach similar to [11] only if the conventional method failed.

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. Our 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 [13]. 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. Our approach addresses this issue—the robust matching algorithm is specifically designed to take advantage of general-purpose computation on graphics processing units, which reduces the execution time considerably.

II. Methods

An overview of our segmentation approach is shown in Fig. 2. First, ribs are detected and utilized to initialize (place) the ASM [14] in the lung CT scan. Second, a robust ASM (RASM) matching algorithm is applied to generate a coarse segmentation of the (diseased) lungs. Third, the segmentation result is adapted by means of a constraint optimal surface finding approach. In the following sections, we describe our approach in detail.

A. Lung Model Generation

Our lung segmentation approach requires learning shapes to build a lung model. In this context it is desirable to have a learning set which is representative for the targeted population. For our experiments, a set of $n = 41$ different total lung capacity (TLC) lung CT scans without contrast enhancement, which showed no signs of lung disease or other pathology, were available. Clearly, the utilized learning set size is limited. However, additional learning samples can be easily added, if needed. The details of the model generation process are described in the following paragraph.

Learning data sets were segmented using the commercial lung image analysis software Pulmonary Workstation 2.0 (PW2) (VIDA Diagnostics, Inc., Coralville, IA). The selected segmentation algorithm generates smooth surfaces which “cut” across main bronchi and pulmonary arteries/veins in the area near the mediastinum. In addition, all segmentation results were manually inspected and corrected, if needed. To produce left and right lung models, the below outlined process was applied to segmented left and right lungs, respectively. From the segmentations, triangle meshes were generated by utilizing a marching cube algorithm [15]. A set of corresponding points (landmarks) $\{s_1, s_2, \ldots. s_m\}$ with $m = 2562$ were automatically identified on all meshes by means of a minimum description length (MDL) approach [16] based on shape index and curvedness [17]. The result of this processing step is a set of n meshes with m corresponding vertices. In this context, we found that the selected number of landmarks represents a good trade-off between computing time and surface point density.

All n landmark sets were aligned in a common coordinate frame by using Procrustes analysis [19], resulting in a mean shape vector $\bar{\mathbf{x}}$. For each learning shape, a shape vector $\mathbf{X}_i$ with $i = 1, 2, \ldots, n$ was generated by concatenating the coordinates: $\mathbf{X}_i = [x_{i,1}, y_{i,1}, z_{i,1}, x_{i,2}, y_{i,2}, z_{i,2}, \ldots, x_{i,m}, y_{i,m}, z_{i,m}]^T$. A principal component analysis (PCA) was applied to the covariance matrix $S = 1/(n-1)\sum_{i=1}^{n}\left(\mathbf{x}_i - \bar{\mathbf{x}}\right)\left(\mathbf{x}_i - \bar{\mathbf{x}}\right)^T$ to generate a point distribution model (PDM) [14]. An instance of a left or right lung shape can be generated from the corresponding PDM by the linear model

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}\mathbf{b} \quad (1)$$

where **P** denotes the shape eigenvector matrix derived from *S*, and **b** represents the shape coefficients. The statistic model (1) allows describing lung shapes in terms of a mean shape and variation about the mean.

B. Automated Model Initialization

It is well known that active shape models (ASMs) have typically a limited capture range. Consequently, they need to be initialized in proximity to the target structure. Therefore, initial shape **(b)** and pose (size, rotation, and location) parameters of the ASM need to be determined. For initialization, we set the lung model to its mean shape (**b = 0**) and utilize an automated method which detects rib centerlines in the CT volume to determine isotropic scale and location (pose) parameters for a given CT data set (Fig. 3). Note that instead of segmenting ribs, we utilize a shape-based rib centerline detection approach, which is more robust against rib density variations due to age, etc. A detailed description is given in the next paragraphs.

1) Centerline-Based Representation of Rib-Like Structures—In a first step, tubular structures that are comparable in size, density, and scale to rib structures are detected and a centerline-based representation is generated as follows.

a. The density values of the volume data set are truncated to a gray-value range of interest between 0 and 500 HU [Fig. 3(a)]. Subsequently, Frangi's "vesselness measure" [18] at a scale of $\sigma = 5$ mm is computed to identify tubular structures of appropriate size [Fig. 3(b)]. Note that at this scale, the cross-section of ribs appears as a bright "blob," thus the darker (less dense) bone marrow is no longer visible. For the other parameters of Frangi's approach, values of $\alpha = 0.5$, $\beta = 0.5$, and $c = 5$ are used. Note that the cross-section of ribs varies significantly, but the choice of parameters allows us to overcome this problem.

To speed up the computation of Frangi's vesselness measure, the data set is first downsampled by a factor of 4 in each dimension. In addition, a volume of interest is generated by means of region growing to exclude the volume (e.g., air) outside of the human body from calculation of the vesselness measure. For this purpose, the border of the CT image is set as seed region and a threshold of −500 is utilized.

b. From the vesselness response image [Fig. 3(b)], a centerline description is extracted for each tubular structure by utilizing a height ridge traversal with hysteresis thresholding, which we have published previously in [20] [Fig. 3(c)]. In a postprocessing step, centerlines with less than ten voxels are discarded, because they typically are caused by image noise or imaging artifacts. In addition, centerlines are cut at furcations of multiple centerlines to avoid problems in areas where other structures like bones (e.g., shoulder blade) or contrast enhanced vessels are in close proximity to ribs, and thus, appear connected at the scale of $\sigma = 5$ mm.

The centerlines found [Fig. 3(c)] are then utilized in the subsequent analysis step to detect centerlines representing ribs.

2) Detection of Rib Centerlines—Our approach is based on the observation that ribs show similar centerline patterns among each other, compared to other structures like (contrast enhanced) vessels, for example. Thus, we utilize a clustering-based approach outlined in Fig. 3 to find rib centerlines. The algorithm consists of two stages. First, potential candidates are selected [Fig. 3(d)]. Second, a more fine-grain pattern analysis is performed

[Fig. 3(e)]. This two step approach has been chosen to save computing time, because the second step is computationally more expensive. The two steps are described in detail below.

In both stages, a mean shift [21] clustering approach is utilized to detect repeating patterns of ribs. For analysis of feature vectors $\mathbf{q} = \{q_1, q_2, \ldots q_d$, an exponential kernel with profile $k(a) = \exp(-a/2)$ and kernel size h_j is used. h_j corresponds to the feature dimension j and is utilized to set the scale of the mode detection. Using the above defined kernel, the mean shift algorithm [21] is utilized to find the mode points for all centerlines. After this analysis, a quantification step is performed to group close-by mode points. If two mode points are closer than the smallest kernel size they are combined to a new mode point represented by the average of both modes. This process is repeated until convergence.

The features used for clustering are based on geometric properties of the centerlines. For this purpose, a PCA analysis is applied to all points of each centerline, and the resulting eigenvalues $|e_1| \leq |e_2| \leq |e_3|$ and corresponding eigenvectors $\mathbf{w}_1$, $\mathbf{w}_2$, and $\mathbf{w}_3$ are further analyzed.

As outlined above, the actual rib detection is performed in two stages.

- **a.** The goal of this stage is to reject centerline objects that do not show a typical spatial extent of ribs (long curved structures). For all detected centerlines, a feature vector $\mathbf{q} = (e_1, e_2, e_3)^T$ is generated. For analysis, the mean-shift kernel sizes are set as follows: $h_1 = 5.0e-1$, $h_2 = 5.0e-4$, and $h_3 = 3.0e-5$. The above described mean shift clustering typically results in one large and a number of smaller clusters. The feature points of the large cluster represent irregular and small-size centerlines corresponding to structures like vessels, vertebrae, etc. After removing this cluster, the remaining feature points relate to ribs and large-size structures like spine, aorta, or parts of the shoulder blade Fig. 4(b), which are further analyzed in the next stage.
- **b.** For rib centerlines, the eigenvectors $\mathbf{w}_2$ and $\mathbf{w}_3$ span a slanted plane in which the rib centerline is located [Fig. 4(a) and (b)]. In addition, the eigenvector $\mathbf{w}_1$ is approximately oriented along the z-axis. Because the eigenvectors of similar rib centerlines can point in opposite directions [Fig. 4(a) and (b)], we utilize a tensor-based representation ($\mathbf{T}_1$, $\mathbf{T}_2$, and $\mathbf{T}_3$) of eigenvectors with $\mathbf{T}_i = [t_{i_{k,l}}]_{k=1,2,3;\ l=1,2,3}$, where $\mathbf{T}_i$ corresponds to $\mathbf{w}_i$. Note that each tensor matrix $\mathbf{T}i$ is symmetric. Consequently, only the upper triangular part of the matrix is utilized for further analysis: $\mathbf{t}i = \{t_{i1,1}, t_{i1,2}, t_{i1,3}, t_{i2,2}, t_{i2,3}, t_{i3,3}$. To describe the orientation of centerlines, the term $\tau = |\mathbf{w}_1 \cdot (0, 0, 1)^T|$ is calculated. A feature vector is generated for each centerline by concatenation: $\mathbf{q} = \{\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3, \tau\}$. For cluster analysis, the kernel sizes are set as follows: $h_{\mathbf{t}_1} = 0.82$, $h_{\mathbf{t}_2} = 0.9$, $h_{\mathbf{t}_3} = 1.0$, and $h_\tau = 0.07$. After applying mean shift clustering, the majority of ribs are located in the largest cluster, because of the similarity of their feature vectors, and all other clusters are discarded.

The kernel sizes for both stages were determined on the learning data set. We observed that occasionally false positives (e.g., included clavicles or costal cartilage) and false negatives cases (missing ribs) occur, but we found that such minor errors still allow deriving model pose parameter which are suitable for a rough initialization of the shape model. As demonstrated in Fig. 3, our approach is also suitable to deal with contrast enhanced lung CT scans.

3) Model Pose Parameter Estimation

Based on the detected rib centerlines, isotropic scale and center location for the ASM are derived. For this purpose, a bounding box for the ribcage is calculated. First, the smallest

and largest *x*- and *y*-coordinate as well as the median *z*-coordinate for each centerline in the detection result are calculated. Second, the bounding box $\mathbf{B} = (x_{min}, x_{max}, y_{min}, y_{max}, z_{min}, z_{max})$ is determined. To robustly estimate x_{min}, x_{max}, y_{min}, and y_{max}, the third largest (max) or third smallest (min) value of all centerline extremes in *x*- and *y*-direction is selected, respectively. This allows us to deal with occasionally occurring false positive centerline parts (see discussion in previous paragraph). z_{min} and z_{max} are directly calculated from the median *z*-coordinates of all centerlines. Finally, the scale and position are estimated for the left and right lung separately. For this purpose the rib bounding box is split in the middle, perpendicular to the *x*-axis. The initial position for the left and right lung model is found by calculating the center of the left and right bounding box, respectively. An isotropic scale factor is calculated for each lung by averaging the two *x*- and *y*-size ratios between left/right mean shape and corresponding bounding box. If needed, the algorithm can be extended to estimate rotation parameters, but based on the utilized CT scan protocols this was not necessary.

C. ASM Matching

For ASM-based segmentation, we utilize a novel robust ASM matching approach (Section II-C2) that extends the standard ASM matching scheme described below.

1) Standard ASM Matching

The PDM (Section II-A) 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, similar as described in [14]. 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 (Section II-B).

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 l_{ASM}. To find update points **y**, we use the following cost function:

$$c_i = \begin{cases} \text{ignore point}, & \text{if } \mathbf{n}_i \cdot \mathbf{g}_{dir_i} < 0 \\ g_{\mathrm{mag}_i}, & \text{otherwise.} \end{cases} \quad (2)$$

c_i represents the cost of the *i*th column element, and the associated sampled gradient magnitude, gradient direction, and surface normal vector are denoted as g_{mag_i}, $\mathbf{g}_{dir_i}$, and $\mathbf{n}i$, respectively. The gradient calculation is based on Gaussian derivatives with a standard deviation of σ_{ASM}, and the calculation of $\mathbf{g}_{\mathrm{dir}}$ and g_{mag} is done for each voxel of the volume before the model matching is started. These precalculated gradient values are then used to interpolate gradient vectors during model matching. If a gradient value outside the volume is required during model matching, the position closest to the boundary is utilized. In the case that no new update point can be found, the old position is used.

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. For this purpose, a Procrustes alignment step is used to estimate transformation matrix **T**, which consists of scaling, rotation, and translation parameters, by minimizing $\left(\mathbf{T}[\mathbf{y}] - \bar{\mathbf{x}}\right)^T \left(\mathbf{T}[\mathbf{y}] - \bar{\mathbf{x}}\right)$. Model parameters **b** are updated using

$$\mathbf{b}=\mathbf{P}^T\left(\mathbf{T}\left[\mathbf{y}\right]-\bar{\mathbf{x}}\right) \quad (3)$$

and a new instance of the model is calculated utilizing (1) and transformed to the image space by $\mathbf{T}^{-1}$.

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.

The basic idea behind robust ASM matching is to only use inlier components of **y** to update model parameters. In this context, Rogers *et al.* investigated M-estimators and random sampling-based robust parameter estimation techniques for 2-D ASM matching [22]. It is well known that the effectiveness of M-estimators strongly depends on the selection of the weighting function and its parameters. Usually, this selection is not trivial, and the optimal selection might change from case to case. Random sampling techniques try to find a subset of inliers by evaluating a number of randomly sampled subsets of update points. Such approaches work well, if the required subset of inliers is quite small. In the case of large ASM models, this strategy can lead to suboptimal results, because a small set of inliers might not be representative enough to describe a complex lung shape (many landmark points), and thus can negatively impact the matching result. For our application, it is desirable to use as many inliers as possible for the model update. Lekadir *et al.* proposed a robust 3-D active shape model matching method based on local shape dissimilarity defined by point triplet ratios [23]. In the case of large ASMs, a large number of possible triplets exist and selecting an optimal subset is not trivial.

Our approach utilizes a robust PCA coefficient estimation scheme that builds on the work of Storer *et al.* [24]. Storer's method was designed for robust image reconstruction and targets a predefined number of inliers. In this paper, we propose a novel voting scheme that does not require to specify a targeted number of inliers. Our method consists of two processing steps. First, normal shape patterns of landmark subsets are learned. Second, these patterns are then utilized during ASM matching to identify and reject outliers.

Offline Learning—The overview of the robust shape pattern learning process is shown in the Fig. 5. Corresponding landmark points of all learning shapes are partitioned randomly into k shape subsets of approximately equal size. This process is repeated l-times, resulting in a set of subsets: $\Omega = \{w_{i,j} | i \in 1,2,\ldots,l; j \in 1,2,\ldots,k\}$. Note that corresponding landmark points of all n learning data sets are always assigned to the same subset. Consequently, each subset $w_{i,j}$ consists of n subset samples. For each subset $w_{i,j}$, a mean shape $\bar{\mathbf{x}}_{\omega_{i,j}}$ is calculated by using Procrustes analysis, and all shapes of the subset are aligned. The subset shapes are then converted to shape vectors by concatenating their x-, y-, and z-components. By means of PCA, the corresponding eigenvector matrices $\mathbf{P}_{w_{i,j}}$ and eigen-value vectors $\lambda_{w_{i,j}}$ are calculated. $\bar{\mathbf{x}}_{\omega_{i,j}}$, $\mathbf{P}_{w_{i,j}}$, and $\lambda_{w_{i,j}}$ are stored and utilized for robust ASM matching.

Robust Matching—In order to match the model to the target image, outlier components of the update position **y** in each iteration need to be identified. This is accomplished by analyzing the subset combinations of **y** and utilizing a voting scheme. Let $\mathbf{y}_{w_{i,j}}$ represent the components of **y** that are corresponding to the landmark points that constitute subset $w_{i,j}$. Let $m_{w_{i,j}}$ represent the number of update points in each subset. A subset reconstruction error is defined

$$e_{\omega_{i,j}} = \left\| \mathbf{T}_{\omega_{i,j}} \left[\mathbf{y}_{\omega_{i,j}} \right] - \left[\bar{\mathbf{x}}_{\omega_{i,j}} + \mathbf{P}_{\omega_{i,j}} \tilde{\mathbf{b}}_{\omega_{i,j}} \right] \right\| \quad (4)$$

where $\mathbf{T}_{w_{i,j}}$ is a transformation matrix that aligns $\mathbf{y}_{w_{i,j}}$ to the corresponding mean $\bar{\mathbf{x}}_{\omega_{i,j}}$. The vector $\tilde{\mathbf{b}}_{\omega_{i,j}}$ is derived from

$$\mathbf{b}_{\omega_{i,j}} = \mathbf{P}^{T}_{\omega_{i,j}} \left[\mathbf{T}_{\omega_{i,j}} \left[\mathbf{y}_{\omega_{i,j}} \right] - \bar{\mathbf{x}}_{\omega_{i,j}} \right] \quad (5)$$

by constraining $\mathbf{b}_{w_{i,j}}(v)$ in $\left[-\xi \sqrt{\lambda_{\omega_{i,j}}(v)}, \xi \sqrt{\lambda_{\omega_{i,j}}(v)} \right]$, $v \in \{1, 2, \ldots, m_{w_{i,j}}\}$. A large reconstruction error $e_{w_{i,j}}$ is an indication that subset$w_{i,j}$ is very likely contaminated by one or more outliers. The *l*-times repeated subdivision increases the possibility of outlier free point combinations. To identify the outliers, the reconstruction error $e_{w_{i,j}}$ is interpreted as a vote, which is casted for all update points that are included in the subset $\mathbf{y}_{w_{i,j}}$. This voting process is carried out for all subsets $w_{i,j} \in \Omega$. Outliers frequently get large vote values. The casted votes are collected in a matrix $\mathbf{V}_{\text{err}}$ of size $m \times l$, in which rows correspond to shape points in ***y***. After all votes are casted, $\mathbf{V}_{\text{err}}$ is analyzed to detect outliers. First, to increase robustness, a rank order statistics filter is applied to each row; the values are sorted, and the *g*-lowest value is selected to represent the filter result. This filtering step reduces $\mathbf{V}_{\text{err}}$ to a vector $\mathbf{v}_{\text{err}} = [v_1, v_2, \ldots, v_m]^T$ and helps to reject accidentally occurring point constellations that contain outliers, which are similar to constellations of inlier points. A typical histogram of vector component values is shown in Fig. 6. Second, a threshold δ is derived from $\mathbf{v}_{\text{err}}$ by analyzing the distribution of vector components v_i: $\delta = \mu + \beta\sigma$ with $\mu = \text{median}_{i \in 1, 2, \ldots, m}$ and $\sigma = \sqrt{1/m \sum_{i=1}^{m} (v_i - \mu)^2}$, where β represents a constant. Third, the threshold is applied to $\mathbf{v}_{\text{err}}$ to yield a selection vector: $\mathbf{p}_{\text{sel}} = [p_1, p_2, \ldots, p_m]^T$ with

$$p_i = \begin{cases} 1 & : \quad v_i < \delta \\ 0 & : \quad v_i \geq \delta \end{cases} \quad (6)$$

to discriminate between inliers ($p_i = 1$) and outliers ($p_i = 0$). Once the inliers are identified, the transformation matrix **T** is obtained by aligning selected update points $\mathbf{y}\{p_i = 1\}$ to the selected mean shape points $\bar{\mathbf{x}}_{\{p_i=1\}}$ and then shape parameters **b** are calculated using

$$\mathbf{b} = \mathbf{P}^{T}_{\{p_i=1\}} \left(\mathbf{T} \left[\mathbf{y}_{\{p_i=1\}} \right] - \bar{\mathbf{x}}_{\{p_i=1\}} \right). \quad (7)$$

An advantage of our outlier detection algorithm is that it is well suited for parallel processing, because the analysis of the *k* shape subsets which is repeated *l*-times is independent from each other, and thus, can be done in parallel. For example, we utilize a GPGPU-based implementation to speed up the model-based segmentation algorithm.

To achieve a robust behavior of the ASM during matching, all outliers present must be rejected. In addition, we would like to utilize as many inliers as possible to achieve a good

match between image and model. This has several implications for the selection of parameters for the RASM matching algorithm. For example, when selecting the number of shape subsets *k*, a trade-off must be made. On the one hand, the value for *k* must be small enough such that the shape points within each shape set can form distinctive point patterns. On the other hand, a larger number of shape subsets is preferable, because it becomes more likely that outlier-free shape subsets can be found, which reduces the number of required subset evaluation iterations (parameter *l*). The parameter β should be selected conservatively, to make sure that all outliers are rejected, even if this implies that a small number of inliers is not utilized for matching the ASM.

For our application, we used the following parameters: $l_{\mathrm{ASM}} = \pm 40$mm, $k = 200$, $l = 60$, $\xi = 2$, $g = 10$, and $\beta = 13$. To update the robust ASM, a gradient image was calculated based on Gaussian derivatives with a standard deviation of $\sigma_{\mathrm{ASM}} = 4$. The maximum gradient position along the search profile was used to calculate updates for shape points. The model matching was iterated until the average of the shape point movement was below 0.04 mm or at most 100 iterations.

D. Constrained Optimal Surface Finding

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 [25], which allows finding a smooth surface related to a shape prior (ASM segmentation). The algorithm transforms the segmentation problem into a graph optimization problem, which is solved by means of a maximum-flow algorithm [25]. 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 (search profile) are generated. The length l_p of the profile is utilized to constrain the segmentation to the proximity of the initial ASM segmentation. In addition, a surface smoothness constraint Δ is incorporated into the graph as described in [25]. For segmentation we use the following cost function:

$$c_i = \begin{cases} g_{\max} & \text{if } \mathbf{n}_i \cdot \mathbf{g}_{dir_i} < 0 \\ g_{\max} - g_{\mathrm{mag}_i} & \text{otherwise} \end{cases} \quad (8)$$

where c_i represents the cost of the *i*th column element and $g_{\max}$ the maximum gradient magnitude of the volume, similar to the cost function utilized in Section II-C. The gradient calculation is based on Gaussian derivatives with a standard deviation of σ_g. The optimal surface finding is utilized in an iterative coarse to fine fashion using following sequence of values of σ_g and and $\Delta : \{6.0, 3.0, 1.0, 0.5\}$ and $\{10, 8, 5, 2\}$, respectively. For the search profile $l_p = \pm 10$ voxel is used.

III. Evaluation Methodology

A. Image Data and Experimental Setup

For our study, 30 multidetector computed tomography (MDCT) thorax scans of cancer patients with lung tumors were utilized. The images where acquired with several different scanners and imaging protocols. In 26 MDCT scans, the vasculature was contrast enhanced. In each data set, either the left and/or right lung contained one or more lung cancer regions with significant higher density, compared to normal lung tissue. To roughly quantify the size of high density lung pathology, the longest diameter in the axial image plane was measured, similar to the response evaluation criteria in solid tumors (RECIST) [26]. In this context,

note that necrotic lung masses were measured in the same way as solid tumors. The average diameter was 46.28 mm, and the average diameter for right and left lungs was 53.0 mm and 38.8 mm, respectively. The image size varied from 512 × 512 × 424 to 512 × 512 × 642 voxel. The in-plane resolution of the images ranged from 0.58 × 0.58 to 0.82 × 0.82 mm and the slice thickness from 0.6 to 0.7 mm.

With our approach, 60 segmentations of 21 diseased right, 9 normal right, 19 diseased left, and 11 normal left lungs were performed, respectively. All computations were performed on a workstation equipped with a NVIDIA Tesla C1060 Computing Processor which offers 240 thread processors and 4 GB of memory.

In addition to the proposed combination of robust ASM and optimal surface finding (RASM +OSF), segmentations were performed with a standard ASM and robust ASM (RASM) without the subsequent surface finding step. The same automatically generated initialization (Section II-B) was utilized for all three segmentation variants. Further, we applied two different methods for lung segmentation provided by a commercial radiation treatment planning system (Pinnacle^3, Philips, The Netherlands). The first approach which will be denoted as “P1” and utilizes region growing and morphological post-processing steps, similar to many commercially available and clinically used lung segmentation methods. The second will be denoted as “P2” and is based on a deformable template approach. In order to utilize method P2, the user is required to manually place a lung shape template in the volume data and to adapt its scale, before iterative matching is performed. One medical expert processed all test data sets with both methods. In the case of method P2, the expert was asked to repeat the process three times to allow assessing the impact of model initialization on segmentation performance. Performance measures reported for method P2 represent the average over all three repetitions. For both methods, clinically utilized standard parameter settings were used. Method P2 allows the user to manually refine a lung segmentation result. For a fair comparison between methods, the operator was not allowed to use this feature.

B. Independent Reference Standard

For quantitative evaluation of our segmentation method, an independent reference standard was generated by utilizing a commercial lung image analysis software package PW2 from VIDA Diagnostics, Inc., Coralville, IA. First, an automated (conventional) lung segmentation was performed. Second, since the software was not designed to deal with lungs containing large lung cancer regions, two experts inspected all the segmentations slice-by-slice and corrected all segmentation errors manually. In the case of diseased lungs, this process took several hours per lung. Because of this, each case was processed only by one expert.

C. Quantitative Indices

The following quantitative error indices are utilized: Dice coefficient D[27], Hausdorff distance H[27], mean signed border positioning errors (d_s) [27], and mean absolute surface distance (d_a) [28]. In the case of d_s, a negative value indicates that the segmentation boundary is inside and a positive value indicates that the border is outside the reference.

IV. Results

Segmentation performance measures averaged over all left lungs and right lungs are summarized in Table I for the proposed method as well as the ASM and RASM approaches.

Based on the Dice coefficient, the boxplots in Fig. 7 show a comparison among those three methods for left [Fig. 7(a)] and right [Fig. 7(b)] normal and diseased lungs, respectively. The mean and standard deviation of performance measures of our method for left and right lungs with and without disease is presented in Table II.

Table III summarizes the results achieved on the same data with methods P1 and P2 in combination with the results of a statistical comparison with the proposed approach. Corresponding boxplots of the Dice coefficient are depicted in Fig. 9.

Fig. 8(b) depicts a segmentation result which was generated by our method. For comparison, the reference segmentation is shown in Fig. 8(a). A segmentation of the same data set with a conventional approach was previously shown in Fig. 1(b). Additional segmentation examples are depicted in Fig. 10, which also shows corresponding results generated with methods P1 and P2.

On average, 2 min were required for calculating left and right model initialization parameters and another 2 min for segmenting a right or left lung, which resulted in an overall average processing time of 6 min per data set. In this context, the mean time required for the intrinsic robust ASM matching was only 24 s due to code optimization and utilization of GPGPU techniques.

V. Discussion

A. Performance

The robust ASM is an important component of the presented fully automated lung segmentation method. Results presented in Table I and Figs. 7 and 12 demonstrate that our robust ASM matching approach outperforms the standard ASM approach. Even on normal lungs, our robust ASM delivers superior performance. These results are not surprising, because standard ASM matching is a least squares optimization, which is sensitive to outliers. Because the model is only roughly initialized in proximity to the lung, all obstacles between model and target structure like aorta and vessels can cause problems. Our robust matching method can even handle missing data, as shown in Fig. 13.

Since our shape model was built from 41 data sets, smaller local shape variations cannot be explained by the model. The optimal surface finding step allows us to overcome this problem. A good initial match between model and image data is required for this processing step, and the performance of a standard ASM would not be sufficient for this task (Fig. 12). The boxplots in Fig. 7 show that the Dice coefficient is increased significantly by the optimal surface finding step for all constellations of left/right and normal/diseased lungs. This is also clearly reflected in the averaged results for the Dice coefficient D, Hausdorff distance H, and mean absolute surface distance error d_a (Table I). In the case of the mean signed border positioning error (d_s) shown in Table I, the value increases slightly after optimal surface finding, but is within the average dimension of a voxel. Table II shows performance measures for normal left/right and diseased left/right lungs, respectively. The results for the Dice coefficient are in a close range for all possible constellations. Distance-based error metrics for normal left lungs are somewhat lower compared to diseased left lungs. In the case of the right lungs, distance error metrics are higher for normal lungs compared to diseased lungs. The reason for this is that the size of one out of the nine normal right lung is extremely small compared to the corresponding left lung [Fig. 14(a)]. Thus, the shape is significantly deviating from learned shapes. Consequently, the model does not initialize the optimal surface finding in close proximity to the target surface in this region, which results in a larger local distance error. If this case is excluded from the calculation, the following performance measures are obtained for normal right lungs: $D = 0.978 \pm 0.004$, H

= 18.93±4.19 mm, d_a = 0.80±0.15 mm, d_s = 0.63±0.17 mm. These results show a similar pattern to the results for the left lung. The described problem can be addressed by expanding the learning shape set utilized for model generation (Section II-A) such that similar variations are included. Another option would be to allow anisotropic scaling of the model during matching which would give it more flexibility.

In our segmentation results, we observed frequently major deviations from the reference in hilar regions where airways and pulmonary vessels enter/leave the lung. Even for experts, it is hard to segment this area consistently.

Currently, our method requires on average 6 min for the processing of normal or diseased lungs, consisting of initialization and sequential segmentation of the left and right lung. The core component of our approach is a novel robust 3-D ASM matching algorithm which is suitable for large models and can run in parallel. For example, our GPGPU-based implementation required approximately 24 s for matching the RASM to a left or right lung. This demonstrates the feasibility of fast and robust model-based segmentation of large structures. In the current implementation, many processing steps are not optimized nor do they utilize GPGPU approaches. For example, the segmentation of left and right lungs can be done in parallel and parts of the model initialization method can be optimized. Thus, we are optimistic that processing time can be significantly reduced, which is an important issue for routine utilization.

B. Comparison to Other Methods

Segmentation of lungs with large lung cancer regions in chest CT scans is a nontrivial problem. Many of the currently utilized methods are prone to produce incorrect results, as shown in Fig. 1(b). Such methods typically rely on simple strategies (e.g., region growing), that do not incorporate knowledge about the shape of the target object. Problems with standard methods can even occur in the case of normal lungs, as depicted in Fig. 11. This is also clearly demonstrated by our assessment of method P1 (Table III) and corresponding examples depicted in Fig. 10. As a consequence, extensive manual postprocessing of segmentations is necessary. The deformable template approach (method P2) represents an improvement compared to method P1 (Table III and Fig. 9), but still shows local segmentation errors (Fig. 10). The result of method P2 can vary significantly with the initialization [Fig. 10(l)–(n)], which limits reproducibility. In some cases, method P2 provided correct segmentations of lungs with masses [Fig. 10(b)–(d)], while in other cases, it consistently produced segmentation errors [Fig. 10(g)–(i)]. Both methods are outperformed by the proposed approach (RASM+OSF), which shows statistically significant better results for all performance metrics (Table III) and does not require manual initialization. In this context, it also interesting to note that method P2 and our RASM (without OSF) show similar values for the average Dice coefficient (Tables III and I). However, when the Hausdorff distance and the mean signed border positioning error are considered, which are more relevant for refining a segmentation in a subsequent processing step (e.g., OSF), our RASM is the better choice, because it is on average closer to the true lung boundary.

A direct comparison of our work to methods presented in Section I is difficult, because of different objectives and/or data sets utilized for algorithm evaluation. So far, the issue of excessive processing time of more advanced methods was only addressed by one group which proposed to use a hybrid approach that first uses a conventional method, and in the case of segmentation errors, a more time-consuming segmentation by registration approach is utilized [12]. The average time required for segmentation was 18.6 min. In the case of failed segmentations, the segmentation by registration method took 120 min on average, which was required in 22 out of 150 lung scans. Compared to the segmentation by

registration approach, our approach has the advantage that only one model needs to be matched to volumetric image data instead of several data sets, which is faster. For the 22 cases where the segmentation by registration approach was applied, the following performance indices were reported [12]: volumetric overlap fraction $O = 0.94 \pm 0.03$, $H = 23.53 \pm 11.97$ mm, $d_a = 1.90 \pm 2.70$ mm. We also calculated the volumetric overlap fraction for our test data: $O = 0.951 \pm 0.011$ in addition to the results shown in Table I and Table II. Overall, volumetric performance metrics of our approach are in a similar range compared to [12]. In the case of distance-based performance metrics, our method appears to deliver better performance (Table I). However, the different data sets used for evaluation limit the comparability of results.

C. Current Limitations and Future Work

While our robust model matching method successfully deals with outliers and other disturbances (e.g., missing data) (Figs. 12 and 13), it requires learning data which is representative for the target population. If a lung shape is encountered that cannot be explained by the model, the shape needs to be added to the learning set. Such an approach allows to build a more complete model over time and will reduce the likelihood that similar problems are encountered in the future. Also, note that the optimal surface finding step after robust ASM segmentation reduces the need to add new lung shapes to the learning set.

In general, cases of pneumothorax or pleural effusion are difficult to segment automatically, and our model-based approach might require some additional processing steps. For example, Fig. 14(b) depicts a robust ASM matching result of a thorax CT scan with a pleural effusion. The left lung model approximates the “normal” lung location and matches with the diaphragm at the bottom. To segment the left lung, some additional steps are needed and might also allow to quantify the pleural effusion volume, which is of interest to physicians.

For our experiments, we have utilized a simple cost function based on gradient magnitude and direction. Thus, performance can be further improved by utilizing more complex cost functions for model matching and optimal surface finding, which could be based on the relative location of shape points as well as density/gray-value properties and shape features. For example, currently the detected ribs are only utilized for model initialization, but can provide valuable information for cost function design. Also, the proposed work targets larger cancer masses and is not optimized for handling juxtapleural nodules. Again, this problem can be addressed by adapting the cost function as well as the formulation of the smoothness constraint utilized for optimal surface finding.

In the current version of our algorithm, left and right lungs are segmented separately, which can lead to inconsistencies (e.g., overlap). This problem can be solved by utilizing a multiple surface graph search approach as described in [25]. In addition, we plan to extend our approach to 4-D to segment TLC and FRC (functional residual capacity) lung scans simultaneously.

In some cases the optimal surface segmentation has problems in segmenting areas with sharp angles like the area where the diaphragm meets the chest wall (costophrenic angles). To solve this problem, locally more dense mesh vertices in combination with an adaption of search profiles will be required.

Preliminary work investigating the applicability of our lung segmentation method to lungs with other kinds of diseases like idiopathic pulmonary fibrosis is promising and will be subject to future research (Fig. 15).

VI. Conclusion

In this paper, a novel method for the fully automated segmentation of lungs with lung cancer regions was presented. The robustness and effectiveness of our approach was demonstrated on 30 lung scans containing 20 normal lungs and 40 diseased lungs where conventional segmentation methods frequently fail to deliver usable results. Low segmentation errors were achieved in cases with and without high-density pathology, compared to two clinically utilized methods. The presented approach to lung segmentation opens up new avenues for computer-aided lung image analysis. For example, segmentation of diseased lungs and segmentation of the diseased tissue itself are related problems. Thus, we expect that our method will be of significant benefit for the quantification of lung diseases.

A core component of our method is a novel robust ASM matching method. The approach not only allows coping with disturbances (e.g., outliers), 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.

Acknowledgments

The authors would like to thank Dr. E. A. Hoffman and Dr. J. M. Reinhardt at the University of Iowa for providing lung data sets.

This work was supported in part by the Biological Sciences Funding Program of The University of Iowa.

References

[1]. Armato SG, Sensakovic WF. Automated lung segmentation for thoracic CT. Acad. Radiol. 2004; vol. 11(no. 9):1011–1021. [PubMed: 15350582]

[2]. Leader JK, Zheng B, Rogers RM, Sciurba FC, Perez A, Chapman BE, Patel S, Fuhrman CR, Gur D. Automated lung segmentation in X-ray computed tomography. Acad. Radiol. 2003; vol. 10(no. 11):1224–1236. [PubMed: 14626297]

[3]. Silva A, Silva JS, Santos BS, Ferreira C. Fast pulmonary contour extraction in X-ray CT images: A methodology and quality assessment. Proc. SPIE Med. Imag. 2001; vol. 4321:216–224.

[4]. Hu S, Hoffman EA, Reinhardt JM. Automatic lung segmentation for accurate quantitation of volumetric X-ray CT image. IEEE Trans. Med. Imag. Jun.2001 vol. 20(no. 6):490–498.

[5]. Hoffman EA, Ritman EL. Effect of body orientation on regional lung expansion in dog and sloth. J. Appl. Physiol. 1985; vol. 59(no. 2):481–491. [PubMed: 4030600]

[6]. Kitasaka T, Mori K, Hasegawa J, Toriwaki J. Lung area extraction from 3-D chest X-ray CT images using a shape model generated by a variable Bézier surface. Syst. Comput. Jpn. 2003; vol. 34(no. 4):60–71.

[7]. Pu J, Roos J, Yi CA, Napel S, Rubin GD, Paik DS. Adaptive border marching algorithm: Automatic lung segmentation on chest CT images. Computer. Med. Imag. Graph. 2008; vol. 32(no. 6):452–462.

[8]. Pu J, Paik DS, Meng X, Roos JE, Rubin GD. Shape break-and-repair strategy and its application to automated medical image segmentation. IEEE Trans. Vis. Comput. Graphics. Jan.2011 vol. 17(no. 1):115–124.

[9]. Prasad MN, Brown MS, Ahmad S, Abtin F, Allen J, da Costa I, Kim HJ, McNitt-Gray MF, Goldin JG. Automatic segmentation of lung parenchyma in the presence of diseases based on curvature of ribs. Acad. Radiol. 2008; vol. 15(no. 9):1173–1180. [PubMed: 18692759]

[10]. Wang J, Li Q, Li F. Automated segmentation of lungs with severe interstitial lung disease in CT. Med. Phys. 2009; vol. 36(no. 10):4592–4599. [PubMed: 19928090]

[11]. Sluimer I, Prokop M, van Ginneken B. Toward automated segmentation of the pathological lung in CT. IEEE Trans. Med. Imag. Aug.2005 vol. 24(no. 8):1025–1038.

[12]. van Rikxoort EM, de Hoop B, Viergever MA, Prokop M, van Ginneken B. Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection. Med. Phys. 2009; vol. 36:2934–2947. [PubMed: 19673192]

[13]. Sun, S.; McLennan, G.; Hoffman, EA.; Beichel, R. Model-based segmentation of pathological lungs in volumetric CT data. Proc. 3rd Int. Workshop Pulmonary Image Anal.; 2010. p. 967-974.

[14]. Cootes TF, Cooper D, Taylor CJ, Graham J. Active shape models—their training and application. Comput. Vis. Image Understand. 1995; vol. 61(no. 1):38–59.

[15]. Lorensen WE, Cline HE. Marching cubes: A high resolution 3-D surface construction algorithm. ACM SIGGRAPH Comput. Graphics. Jul.1987 vol. 21(no. 4):163–169.

[16]. Heimann T, Wolf I, Williams T, Meinzer H-P. 3-D active shape models using gradient descent optimization of description length. Proc. IPMI, Heidelberg. 2005; vol. 3565:566–577.

[17]. Koenderink JJ, van Doorn AJ. Surface shape and curvature scales. Image Vis. Comput. Oct.1992 vol. 10:557–565.

[18]. Frangi, AF.; Niessen, WJ.; Vincken, KL.; Viergever, MA. Multiscale vessel enhancement filtering. Vol. vol. 1496. Springer; New York: 1998. p. 130-137.MICCAI

[19]. Dryden, IL.; Mardia, KV., editors. Statistical Shape Analysis. Wiley; New York: 1998.

[20]. Bauer C, Pock T, Sorantin E, Bischof H, Beichel R. Segmentation of interwoven 3-D tubular tree structures utilizing shape priors and graph cuts. Med. Image Anal. 2010; vol. 14:172–184. [PubMed: 20060769]

[21]. Comaniciu D, Meer P. Mean shift: A robust approach toward feature space analysis. IEEE Trans. Pattern Anal. Mach. Intell. May; 2002 vol. 24(no. 5):603–619.

[22]. Rogers M, Graham J. Robust active shape model search. Proc. Eur. Conf. Comput. Vis. 2002:517–530.

[23]. Lekadir K, Merrifield R, Yang GZ. Outlier detection and handling for robust 3-D active shape models search. IEEE Trans. Med. Imag. Feb.2007 vol. 26(no. 2):212–222.

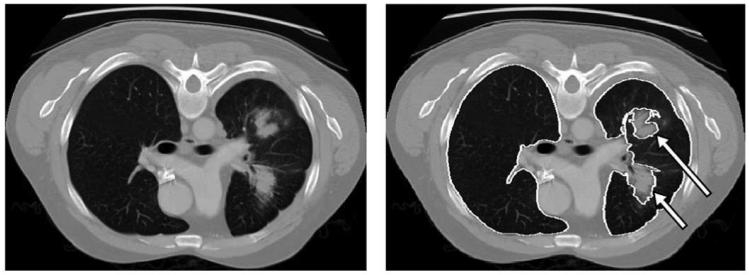
[24]. Storer, M.; Roth, P.; Urschler, M.; Bischof, H. Fast-robust PCA. 16th Scand. Conf. Image Anal.; 2009. p. 430-439.

[25]. Li K, Wu X, Chen D, Sonka M. Optimal surface segmentation in volumetric images—A graph-theoretic approach. IEEE Trans. Pattern Anal. Mach. Intell. Jan.2006 vol. 28(no. 1):119–134. [PubMed: 16402624]

[26]. Therasse P, Arbuck S, Eisenhauer E, Wanders J, Kaplan RS, Rubinstein L, Verweij J, Glabbeke MV, van Oosterom AT, Christian MC, Gwyther SG. New guidelines to evaluate the response to treatment in solid tumors. J. Nat. Cancer Inst. 2000; vol. 92(no. 3):205–216. [PubMed: 10655437]

[27]. Sonka, M.; Hlavac, V.; Boyle, R. Image Processing, Analysis and Machine Vision. CL Engineering; Orlando, FL: 2007.

[28]. Gerig, G.; Jomier, M.; Chakos, M. Vol. vol. 2208. Springer; New York: 2001. Valmet: A new validation tool for assessing and improving 3-D object segmentation; p. 516-523.MICCAI



(a) (b)

Fig. 1.

Segmentation of a lung with cancer using a conventional approach. (a) Axial CT image showing normal right and cancerous left lung tissue. (b) Corresponding segmentation result generated with a conventional lung segmentation method. Segmentation errors are indicated by arrows.

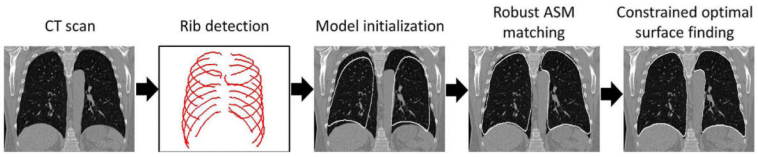


Fig. 2.
Overview of our model-based segmentation approach.

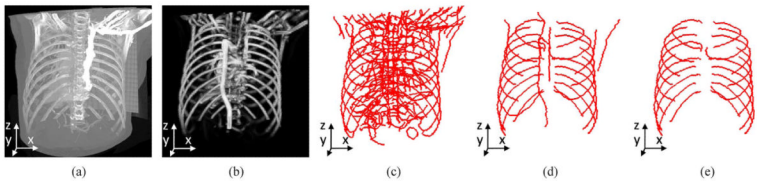


Fig. 3.

Outline of main rib detection processing steps. (a) Volume rendering of the input thorax CT data truncated to a gray-value range between 0 and 500 HU. (b) Volume rendering of Frangi's "vesselness measure" [18] computed at a scale of $\sigma = 5$ mm and (c) corresponding centerlines of rib candidates (Section II-B1). Note that many responses from vessels (e.g., aorta) can be found in (b) and (c), because the CT image is contrast enhanced. Centerlines from vessels and other non-rib structures are removed in subsequent rib detection steps. Output of first (d) and second (e) rib clustering/detection stage (Section II-B2).

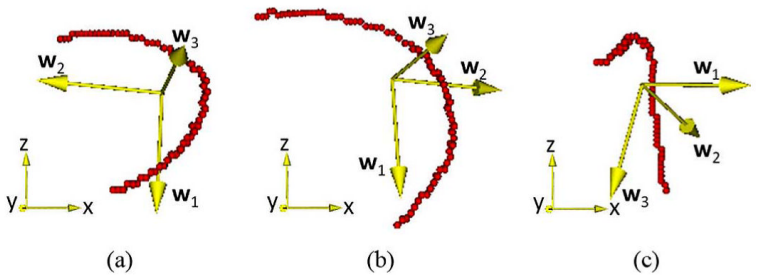


Fig. 4.

Visualization of eigenvector patterns utilized for rib detection. Eigenvectors are shown for ribs [(a), (b)] and the aortic arch (c).

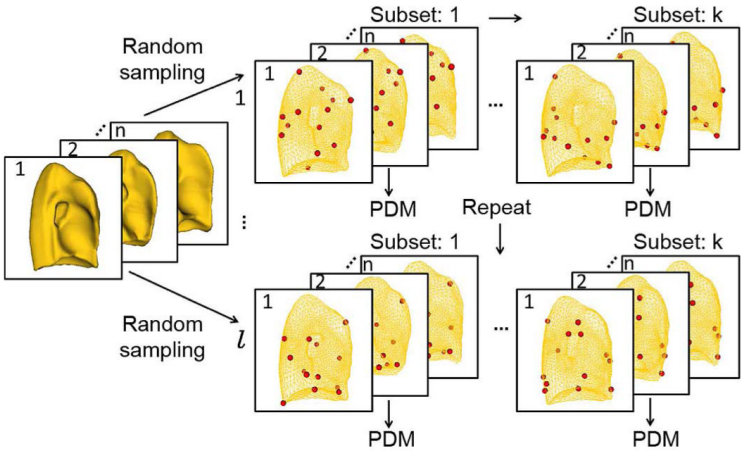


Fig. 5.

Robust shape pattern learning. A random sampling process is repeated *l*-times (rows). In each random sampling process, *k* shape subsets are derived from all *n* training shapes and utilized to generated point subset distribution models.

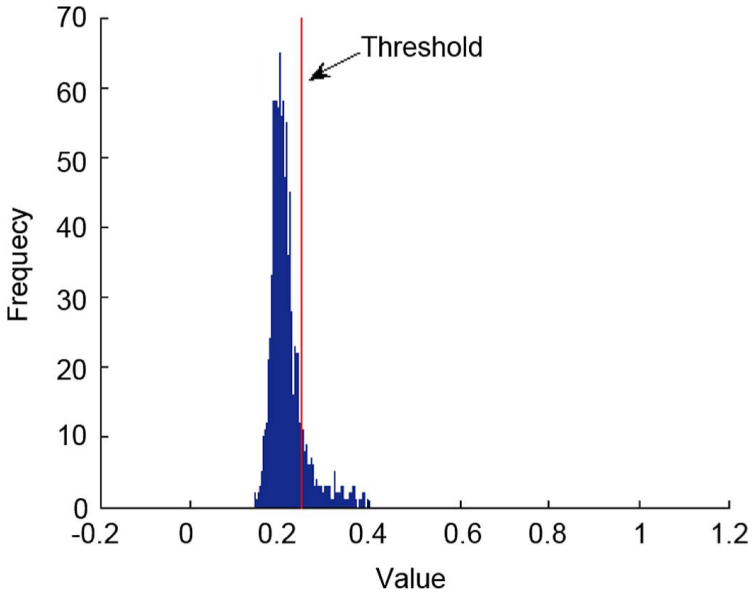


Fig. 6.

Histogram of component values of $\mathbf{v}_{err}$ and automatically derived threshold utilized to detect outlier components.

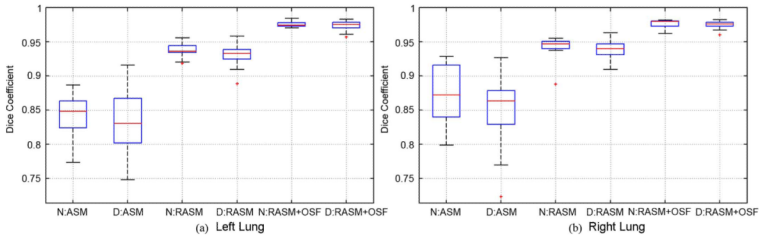
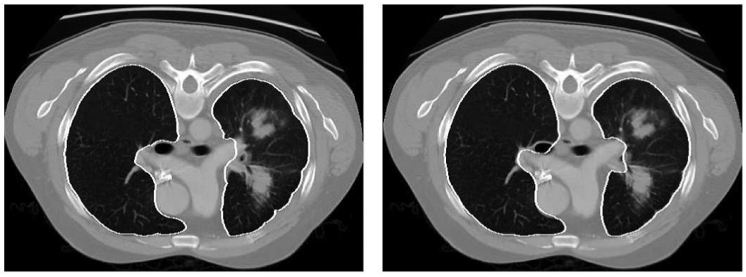


Fig. 7.

Comparison of the Dice coefficient of standard ASM, robust ASM (RASM), and proposed (RASM+OSF) lung segmentation approaches for (a) left and (b) right lungs. Note that boxplots for normal (N) and diseased (D) lungs are shown separately.



(a) (b)

Fig. 8.
Segmentation result for the example shown in Fig. 1(a). (a) Reference segmentation and (b) proposed segmentation approach.

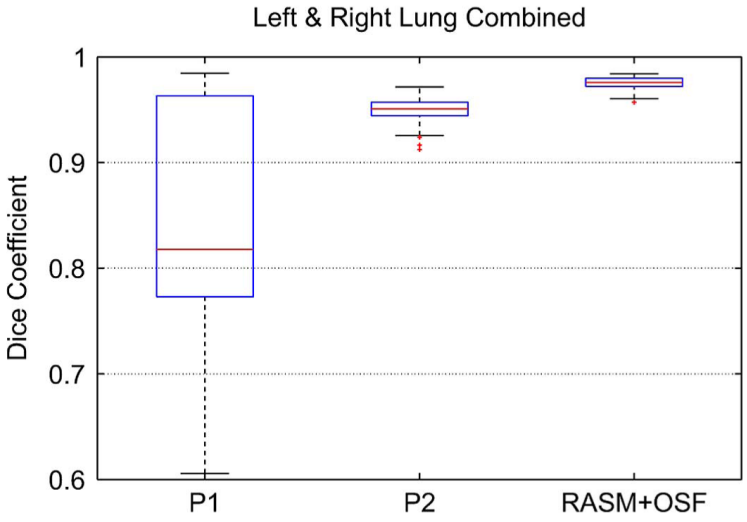


Fig. 9.
Boxplots of the Dice coefficient for P1, P2, and our approach.

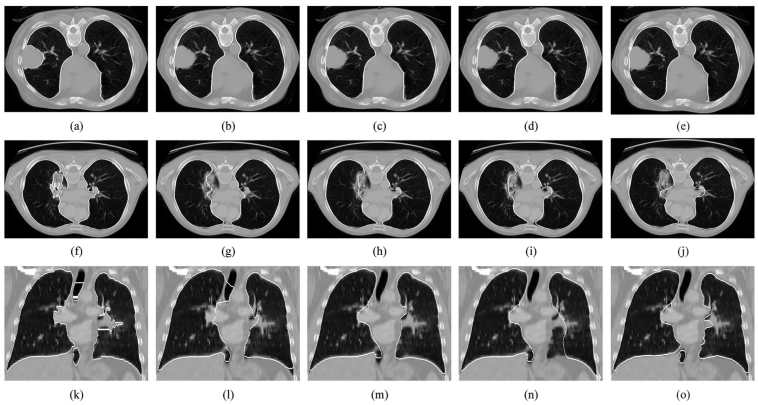
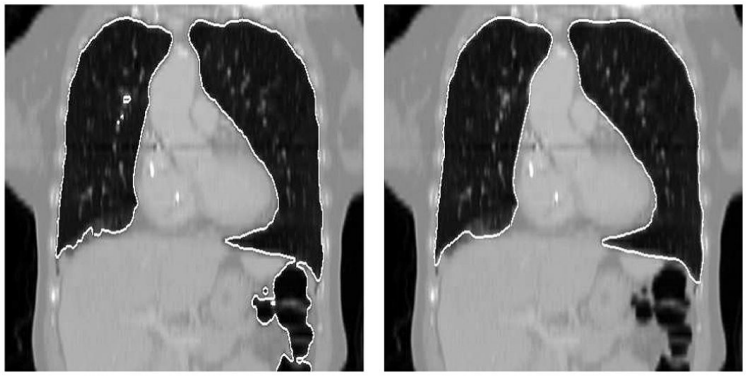


Fig. 10.

Examples of segmentation results on three different data sets (rows). (a) , (f), (k) Results for method P1. (b)–(d), (g)–(i), (l)–(n) Results generated with method P2 based on three different template initializations. (e), (j), (o) Results of the proposed automated approach (RASM+OSF).



(a) (b)

Fig. 11.

Comparison between conventional and proposed lung segmentation methods. (a) The conventional method leaks into the gas filled colon. (b) Our method provides a correct segmentation.

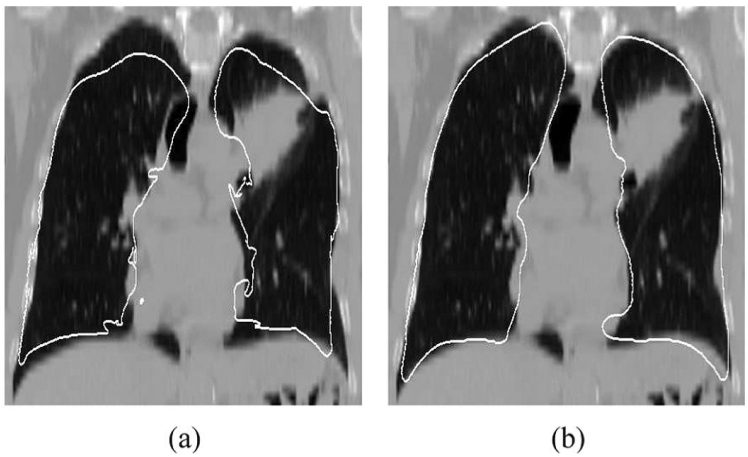
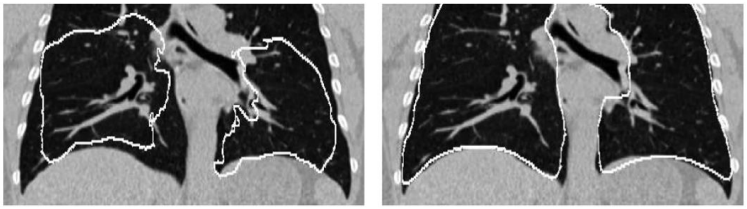


Fig. 12.

Performance comparison between (a) standard ASM and (b) RASM matching (without optimal surface finding step). The RASM delivers a better match for normal and diseased lungs.



(a) (b)

Fig. 13.

Segmentation of an incomplete lung CT data set; the top portion was not scanned. (a) Standard ASM. (b) Robust ASM (without optimal surface finding step). Note that the standard and robust ASM are not aware of the spatial extent of the data set, because of the clamping of gradient values to the boundary (Section II-C1). Surfaces outside of the data set were clipped after the segmentation process was completed.

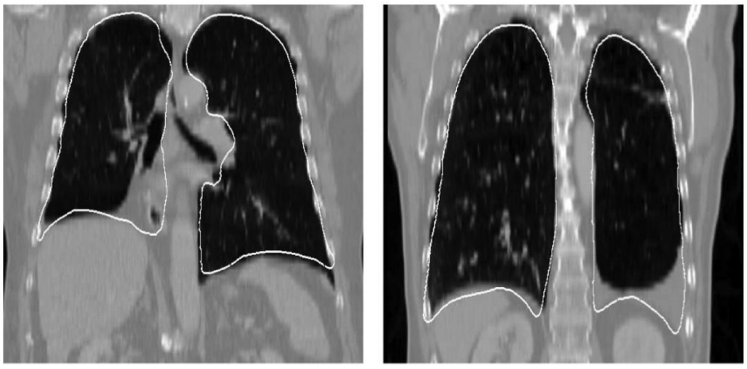


Fig. 14.
Examples of RASM segmentation results (without optimal surface finding step). (a) Case with small right lung and (b) pleural effusion in left lung. See text in Section V for details.

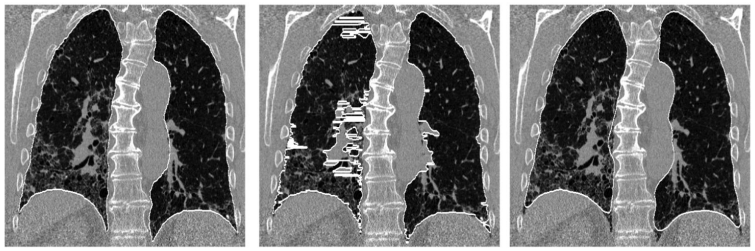


Fig. 15.
Example segmentation results of lung with idiopathic pulmonary fibrosis. (a) A manual reference segmentation. (b) Result of a conventional segmentation method. (c) Preliminary segmentation result of our approach (RASM+OSF).

TABLE I

Comparison of Overall Performance Between Standard ASM, Robust ASM (RASM), and Proposed (RASM +OSF) Lung Segmentation Approaches Averaged Over all Left and Right Lungs Processed. Mean and Standard Deviation (STD) is Given for Each Index

		ASM	RASM	RASM+OSF
D (-)	mean	0.848	0.936	0.975
	std	0.046	0.015	0.006
H (mm)	mean	35.06	22.77	20.13
	std	7.87	5.75	6.17
d_a (mm)	mean	5.47	2.24	0.84
	std	1.37	0.50	0.23
d_s (mm)	mean	−3.02	0.51	0.59
	std	1.14	0.41	0.39

TABLE II

Segmentation Results of the Proposed Method on Normal Left (NL), Diseased Left (DL), Normal Right (NR), and Diseased Right (DR) Lungs. Mean and Standard Deviation (std) is Given for Each Index

		Lung			
		NL	**DL**	**NR**	**DR**
D (-)	mean	0.975	0.974	0.976	0.976
	std	0.004	0.007	0.006	0.005
H (mm)	mean	18.63	19.32	21.64	21.00
	std	5.51	5.82	9.02	5.52
d_a (mm)	mean	0.75	0.84	0.94	0.85
	std	0.10	0.2	0.43	0.18
d_s (mm)	mean	0.44	0.50	0.92	0.60
	std	0.09	0.24	0.87	0.19

TABLE III

Overall Performance Measures for Methods P1 and P2 Averaged Over all Left and Right Test Lungs. Mean, Standard Deviation (std) is Given for Each Index. In Addition, the P-Value of a Paired Wilcoxon Rank Sum Test of the Hypothesis That Method P1 or P2 and Our Approach Come From Distributions With Equal Medians is Reported

		P1	P2
D	mean (-)	0.844	0.949
	std (-)	0.106	0.012
	P-value (-)	2.47e-14	5.92e-20
H	mean (mm)	98.49	33.07
	std (mm)	50.88	7.69
	P-value (-)	1.11e-19	2.05e-15
d_a	mean (mm)	8.56	1.89
	std (mm)	6.14	0.45
	P-value (-)	6.64e-17	4.21e-20
d_s	mean (mm)	15.21	1.25
	std (mm)	12.76	0.72
	P-value (-)	2.31e-05	1.77e-09