Pulmonary lobe separation in expiration chest CT scans based on subject-specific priors derived from inspiration scans

Christian Bauer
Michael Eberlein
Reinhard R. Beichel

Pulmonary lobe separation in expiration chest CT scans based on subject-specific priors derived from inspiration scans

Christian Bauer,^a Michael Eberlein,^b and Reinhard R. Beichel^{a,b,*}

^aUniversity of Iowa, Department of Electrical and Computer Engineering, Iowa City, Iowa, United States

^bUniversity of Iowa, Carver College of Medicine, Iowa City, Iowa, United States

Abstract. Segmentation of pulmonary lobes in inspiration and expiration chest CT scan pairs is an important prerequisite for lobe-based quantitative disease assessment. Conventional methods process each CT scan independently, resulting typically in lower segmentation performance at expiration compared to inspiration. To address this issue, we present an approach, which utilizes CT scans at both respiratory states. It consists of two main parts: a base method that processes a single CT scan and an extended method that utilizes the segmentation result obtained on the inspiration scan as a subject-specific prior for segmentation of the expiration scan. We evaluated the methods on a diverse set of 40 CT scan pairs. In addition, we compare the performance of our method to a registration-based approach. On inspiration scans, the base method achieved an average distance error of 0.59, 0.64, and 0.91 mm for the left oblique, right oblique, and right horizontal fissures, respectively, when compared with expert-based reference tracings. On expiration scans, the base method's errors were 1.54, 3.24, and 3.34 mm, respectively. In comparison, utilizing proposed subject-specific priors for segmentation of expiration scans allowed decreasing average distance errors to 0.82, 0.79, and 1.04 mm, which represents a significant improvement ($p < 0.05$) compared with all other methods investigated. © 2018 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.5.1.014003](https://doi.org/10.1117/1.JMI.5.1.014003)]

Keywords: pulmonary lobe segmentation; inspiration; expiration; chest CT.

Paper 17208RR received Jul. 7, 2017; accepted for publication Jan. 11, 2018; published online Feb. 9, 2018.

1 Introduction

The human lungs are divided into five lobes by pulmonary fissures (Fig. 1). The left lung consists of an upper and a lower lobe, and the right lung of an upper, a lower, and a middle lobe. Each lobe forms an independent functional unit with its own vasculature and airways.

For a number of clinical investigations, only a single lung CT scan acquired at full inspiration is sufficient. However, acquisition of inspiration and expiration scans pairs (Fig. 2) enables the analysis of regional, lobe-based ventilation, and perfusion patterns, which facilitates the treatment of lung diseases such as chronic obstructive pulmonary disease (COPD) and enables the subsequent assessment of treatment response. For example, in the case of COPD, assessment of air trapping in relation to lobar structures plays an important role in determining the suitability of interventions, such as bronchoscopic lung volume reduction by insertion of one-way endobronchial valves.¹

Segmentation of lung lobes in CT scans is a nontrivial task because lobar fissures are often incomplete and do not completely separate lungs into lobes [Fig. 3(a)]. The incompleteness of fissures may either be anatomically founded or be the result of insufficient image resolution or other imaging-related factors. In the literature, a number of approaches for lung lobe segmentation have been reported. A review of state-of-the-art methods can be found in Doel et al.² Typically, lobe segmentation methods consist of a lung segmentation step, and utilize various methods

to enhance/segment fissures, vessels, and/or airways, providing additional input for the actual lobe segmentation process. Some approaches mainly rely on an identification of fissure surfaces in scans (e.g., Ref. 3). Such an approach may become challenging if fissures are incomplete [Fig. 3(a)] or pathology has affected the appearance of the lung parenchyma [Fig. 3(b)]. To deal with this issue, other information such as absence of vessels and airways^{4,5} near fissures or other anatomical knowledge have been utilized as alternatives. For example, shape priors were utilized and are typically derived from atlases to model the average fissure location and its variation.^{6,7} A different approach was used by van Rikxoort et al.⁸ From an atlas, they propose to select a reference dataset, which best describes the dataset to be segmented. Bragman et al.⁹ presented a method to derive an average fissure prior in form of a confidence map from unlabeled training data. However, the shape of individual lobar boundaries can show large anatomical variation. In addition, lung pathology may further deform lobes [Fig. 3(c)]. Some previously presented methods^{3–5,8} obtain binary segmentations of fissure surfaces, vasculature, or the airway tree, which are then utilized as hard constraints or to obtain distance measures for lobe estimation. However, an incomplete segmentation of one of these structures may negatively affect the segmentation process.

The majority of existing lobe segmentation methods were only developed for and evaluated on inspiration scans. The reliable detection of fissures in expiration scans is more challenging because the density of the lung parenchyma in expiration scans

*Address all correspondence to: Reinhard R. Beichel, E-mail: reinhard-beichel@uiowa.edu

Fig. 1 Pulmonary lobes and fissures. The left oblique (LO) fissure (red) separates the left lung into left upper lobe (LUL) and left lower lobe (LLL). The right lung is separated by the right oblique (RO) fissure (green) and right horizontal (RH) fissure (blue) into right upper lobe (RUL), right lower lobe (RLL), and right middle lobe (RML). (a) 3-D Fissure model and (b) sagittal CT slice through right lung. Arrows indicate the location of the fissures.

Fig. 2 A pair of (a) inspiration and (b) expiration lung CT scans. Both images show coronal slices with the same gray-value window.

Fig. 3 Challenges for lung lobe segmentation methods. All images show a sagittal slice of the right lung of an expiration CT scan. (a) Incomplete RH fissure (blue) and accessory fissure (yellow). (b) Pathologic parenchyma containing structures, which are similar to fissures. (c) Deformed lung lobes showing some air trapping. Dashed lines indicate oblique (green) and horizontal (blue) fissures.

increases, which reduces the visibility of fissures (Fig. 2). In addition, in some clinical protocols the expiration scans are acquired at lower radiation dose compared with inspiration scans to reduce x-ray exposure, which negatively affects image quality. So far, only a few authors report evaluation results on expiration scans. An overview of these methods and a performance comparison between inspiration and expiration scans based on the distance error metric d_{RMS} is provided in Table 1. The error d_{RMS} is defined as the root mean square error of three-dimensional (3-D) Euclidean distances between all reference points and their closest point in the algorithm's segmentation result. For all approaches, segmentation performance on expiration scans is worse when compared with results on inspiration scans.

All methods listed in Table 1 process inspiration and expiration scans independently. However, in clinical practice, an expiration scan is typically acquired with a corresponding inspiration scan. Therefore, we present a method for separating the pulmonary lobes in expiration CT scans with a subject-specific prior derived from the corresponding inspiration scan to improve the quality of lobe segmentation in the expiration scan and increase consistency of results between respiratory states. As fissures may be incomplete, we will use the term interlobar boundaries in the remainder of this paper to refer to the complete boundaries that are necessary to separate the lungs into lobes. We evaluate our method on CT scans from four cohorts and compare results to another approach that is based on image registration between inspiration and expiration CT scans.

2 Image Data

For this work, we utilized pairs of inspiration and expiration chest CT scans from four different cohorts; scans from subjects with no abnormalities (normals), subjects with COPD, asbestosis, and asthma. For each cohort, 20 scan pairs were available, which were split into two datasets using stratified random sampling. The development dataset consisted of 40 scan pairs and was utilized for method development and parameter selection (Sec. 3). The test dataset consisted of the remaining 40 scan pairs and was exclusively utilized for performance evaluation as described in Sec. 4.

All inspiration scans were acquired at total lung capacity, and all expiration scans at functional residual capacity. The mean and standard deviation of CT scan in-plane resolution were 0.62 ± 0.06 mm, and the slice thickness was 0.71 ± 0.08 mm. Scans from cohorts normal, COPD, and asbestosis were acquired with Siemens scanners (SOMATOM Definition Flash or Sensation 64) and reconstructed with $B35f$ kernels. The tube current was 120 kV_p. For cohorts normal and asbestosis, inspiration and corresponding expiration scans were obtained with matching tube currents, ranging from 220 to 320 mA. For cohort COPD, tube currents were 400 or 440 mA for inspiration scans and 100 or 110 mA for expiration scans. Scan parameters for the asthma cohort were not available.

3 Method

An overview of our approach for separation of the pulmonary lobes in pairs of inspiration and expiration scans is given in Fig. 4. First, the lungs in inspiration and expiration scan pairs are segmented using a four-dimensional (4-D) lung segmentation method.^{10,11} Second, we identify the interlobar surfaces separating the lungs into lobes based on an optimal surface

Table 1 Comparison of lung lobe segmentation performance between inspiration and expiration scans for three published methods. Note that the reported indices were produced on different datasets; see cited papers for more details.

Method	Fissure surface	Inspiration	Expiration	Evaluation dataset
		d_{RMS} (mm)	d_{RMS} (mm)	
Ukil and Reinhardt ⁵	Left oblique	2.31 ± 1.72	2.70 ± 1.19	12 inspiration/expiration scan pairs of healthy subjects.
	Right oblique	1.57 ± 0.17	2.41 ± 0.58	
	Right horizontal	1.63 ± 0.99	2.11 ± 1.94	
Ross et al. ⁶	Left oblique	1.45 ± 1.04	2.88 ± 5.23	50 high-dose inspirations scans and 50 low-dose expiration scans from subjects with COPD. Inspiration and expiration scans were not from the same individuals.
	Right oblique	1.60 ± 1.51	2.25 ± 2.85	
	Right horizontal	3.47 ± 2.06	5.29 ± 16.32	
Zhang et al. ⁷	Obliques combined	1.67 ± 0.53	2.30 ± 0.76	10 inspiration/expiration scan pairs and two extra inspiration scans.

Fig. 4 Method overview. The result of the inspiration scan is utilized as a subject-specific prior for segmentation in the expiration scan.

segmentation (OSS)¹² approach, for which we propose a suitable cost function that enables the adaption of a prior surface to the relevant features of the CT scan. For the inspiration scan, we utilize the OSS with a rough generic surface prior for the three major fissures to obtain the interlobar boundaries. Subsequently, for the corresponding expiration scan, the results from the inspiration scan are utilized as a subject-specific surface prior for segmentation. All parameters listed in the following sections were selected using the development dataset (Sec. 2).

3.1 Cost Function Design

In addition to utilizing CT density information for segmentation of fissures, other relevant anatomical information (e.g., airways and vessels) is also frequently considered to increase robustness, and different approaches have been proposed.^{2,4,9} We follow such an approach, but this type of information alone is typically not sufficient to achieve good performance on expiration scans. Because we address this issue by utilizing a subject-specific prior (Sec. 3.2.2), we can use the same cost function for both respiratory states (Fig. 5).

3.1.1 Vessel likelihood $V(x)$

We apply the method of Frangi et al.¹³ to calculate a vesselness likelihood [Fig. 5(b)] on four scales $\sigma_1 \in \{1, 2, 3, 4\}$ mm with the following set of parameters: $\alpha = 0.5$, $\beta = 0.5$, and $c = 500$. Frangi's method calculates the eigenvalues $|e_1| \leq |e_2| \leq |e_3|$ of the Hessian matrix at each scale σ_1 and derives measures R_A , R_B , and N that are combined using Gaussian weighting.¹³ The term R_A allows distinguishing between plate-like and line-like structures, R_B indicates blob-like structures, and N is a noise measure.

3.1.2 Fissure likelihood $F(x)$

A fissureness likelihood can be obtained in an analogous fashion (Fig. 6).^{9,14} For example, Bragman et al.⁹ uses the following approach: $B_{\sigma_1}(\mathbf{x}) = H(e_3)e^{\frac{R_A^2}{2\sigma^2}}e^{\frac{R_B^2}{2\sigma^2}}$ [Fig. 6(b)], where H represents the Heaviside step function. To suppress false responses, we add an additional term: $S_{\sigma_1, \sigma_2}(\mathbf{x}) = B_{\sigma_1}(\mathbf{x})(1 - e^{-\frac{N^2}{2\sigma^2}})e^{\frac{|f'_{\sigma_2}(\mathbf{x})|^2}{2\sigma^2}}$. The first part is identical to Frangi's noise suppression term, whereas the second part is

Fig. 5 CT scan and corresponding derived feature likelihood volumes and final cost volume. (a) Sagittal slice of CT scan, (b) MIP of vessel likelihood, (c) MIP of fissure likelihood, (d) tissue density likelihood corresponding to CT image shown in (a), and (e) combined cost volume.

Fig. 6 Fissureness calculation. (a) CT slice, (b) fissure enhancement $B_{\sigma_1}(\mathbf{x})$ following Bragman et al.⁹ with recommended parameters, (c) initial sheetness measure $S_{\sigma_1, \sigma_2}(\mathbf{x})$, and (d) final fissureness $F(\mathbf{x})$ after tensor voting on $S_{\sigma_1, \sigma_2}(\mathbf{x})$.

based on the image gradient $I'_{\sigma_2}(\mathbf{x})$ at scale σ_2 and suppresses responses near the surface of vessels [Fig. 6(c)]. The scale σ_2 for gradients suppression should be selected slightly larger than the scale utilized for calculation of the Hessian matrix, and we use $\sigma_2 = \eta\sigma_1$ to fulfill this requirement. For calculation, we utilize the following parameters: $\alpha = 0.25$, $\beta = 0.5$, $c = 20$, $\delta = 5$, $\eta = 1.5$, and scale $\sigma_1 = 0.75$ mm. To attenuate isolated, spurious responses and enhance responses that are part of larger sheetlike structures, we apply tensor voting¹⁵ on $S_{\sigma_1, \sigma_2}(\mathbf{x})$ with $\sigma_3 = 5.0$ mm and $\phi = 0.05$ resulting in the fissureness likelihood volume $F(\mathbf{x})$ [Figs. 5(c) and 6(d)].

3.1.3 Tissue density likelihood $T(\mathbf{x})$

To deal with pathologically thick fissures, we utilize the approach of Lassen et al.⁴ and incorporate a tissue density likelihood $T(\mathbf{x}) = \min\{\max\{I(\mathbf{x})/1000 + 1.0, 1\}M_V(\mathbf{x}), 1\}$, where I is the image intensity and M_V represents a rough vessel mask obtained by thresholding $V(\mathbf{x})$ at 0.1 followed by morphological dilation by one voxel.

3.1.4 Cost function $C(\mathbf{x})$

A high fissure or tissue density likelihood should result in low costs, whereas a high vesselness likelihood should result in high costs. Therefore, we define the cost function by $C(\mathbf{x}) = -F(\mathbf{x}) + V(\mathbf{x}) - 0.25T(\mathbf{x})$ weighting the individual terms by their importance. To avoid that the OSS is distracted by features outside of the lung or the airway walls, which also appear as bright sheet-like structures, we set the costs outside of the

lungs or within a 2 mm range of an airway segmentation mask¹⁶ to the median cost value inside of the lung.

3.2 Lobe Separation

After obtaining cost images for the left and right lungs, we divide each lung into two parts by finding the optimal separating surface. Subsequently, the right lung is further divided to obtain the third lung lobe. This process is performed for the inspiration and expiration scan pair. Note that different priors are utilized for the two respiratory states, as described in detail in the following sections.

3.2.1 Lobe separation at inspiration utilizing a generic prior

For inspiration scans, planar meshes are utilized as rough generic priors for the left oblique (LO) fissure, right oblique (RO) fissure, and right horizontal (RH) fissure (Fig. 7). The

Fig. 7 Initial planar prior surfaces for (a) LO, (b) RO, and (c) RH fissure at inspiration.

location and orientation of each of these prior surfaces relative to the bounding box of the right and left lungs, respectively, were obtained as follows. Construction of the generic surface priors (Fig. 7) is based on the reference tracings of visible fissures in a subset of 10 CT scan pairs (subjects with no significant abnormalities) of the development dataset. Thus, the location and shape of fissures in these datasets have not been altered by disease. For these datasets, a set of point correspondences were established by fitting our robust 4-D lung active shape model (ASM)¹⁰ to the CT volumes. Utilizing these point correspondences, the reference tracings of all 20 datasets (10 scan pairs) were warped into the coordinate space of one of the CT scans using a thin plate spline (TPS) transformation. Then, for each of the three fissure surfaces, the center location and orientation relative to the bounding box of the lung were obtained as follows. To find the center, the average of all points of the warped reference tracing was calculated. By applying principal components analysis (PCA) to these points after mean normalization, the corresponding three eigenvectors with associated eigenvalues were obtained. The eigenvector associated with the smallest eigenvalue was used as normal direction for the planar surface mesh prior. These three prior meshes are fitted to the cost volume in two steps: a coarse and subsequent fine surface finding step.

Coarse surface finding. For each of the three surfaces, the prior surface mesh is adapted to the volume using OSS in combination with the previously defined cost volume. OSS enables transforming the segmentation problem into a graph-optimization problem and allows obtaining a globally optimal solution in low-order polynomial computation time. For details, we refer the reader to the work of Li et al.¹² For this purpose, a search profile is constructed for each mesh vertex of the prior surface, following a straight line normal to the initial plane [Fig. 8(a)]. Each search profile has several possible discrete positions through which the final surface can pass with associated costs. These costs are defined by $C(\mathbf{x})$ (Sec. 3.1). In addition, neighboring search profiles are coupled with hard¹² and soft¹⁷ smoothness constraints for regularization. The hard constraint¹² ensures that the final surface's mesh vertices on neighboring search profiles stay within a maximum given distance. The soft smoothness constraint¹⁷ penalizes any discontinuity

Fig. 8 Illustration of OSS graph structure used for fissure segmentation. Based on an initial surface (green) search profiles are constructed (blue) with graph nodes (red and green dots). Note that graph edges for soft and hard smoothness constraints are not shown. (a) Structure for planar prior surface and (b) structure for non-planar prior surface with parallel search profiles.

between neighboring surface points (i.e., increasing costs with increasing discontinuity). Because location and shape of the lobe-separating fissure surfaces vary between subjects and our initial planar meshes provide only a rough prior, a large search range combined with relaxed smoothness constraints is utilized. The vertices in the planar mesh are organized in a rectangular grid with a spacing of 3 mm. For each mesh vertex, we construct search profiles with a length of ± 50 mm from the initial mesh and use a sampling interval of 1 mm on the search profiles. Parameters for hard and soft smoothness constraints were set to ± 4 mm and 0.005, respectively. An example showing the result of this initial surface adaption step is shown in Fig. 9(c). The obtained surface follows the visible fissure surface in the dataset with roughly the same orientation as the initial prior and—due to the definition of the cost function—the OSS tries to avoid cutting through vessels. However, one can note that in areas without visible fissures (purple), the continuation of the surface is not very smooth.

Fine-scale surface finding. To ensure a plausible smooth continuation of the interlobar surface in areas without visible fissures, we obtain a more accurate dataset specific surface prior for the inspiration scan. Figure 9(c) shows the fissureness $F(\mathbf{x})$ (Sec. 3.1.2) color-coded at the determined surface location. From this result, the visible fissure surface portion is obtained based on thresholding [$F(\mathbf{x}) \geq 0.2$] and identification of the largest connected component. Figure 9(d) shows an example of a so-obtained fissure surface. Then, a TPS is fitted through these fissure surface points providing smooth inter- and extrapolation in areas without visible fissures [Fig. 9(e)]. For rare cases where the method does not find any fissure surface points, the initial surface finding result is used as the final result. The TPS extrapolated surface is again refined utilizing OSS and the previously calculated cost volume. Similar to the first OSS step, parallel search profiles are constructed starting at the mesh vertex points of the prior surface, but the direction of the search profiles is obtained based on a PCA of the identified fissure surface points [Fig. 8(b)] as described previously. In contrast to the first OSS step, the search range is only ± 10 mm, with a finer sampling distance of 0.25 mm along the search profiles and more restrictive hard (± 1 mm) and soft smoothness constraints (0.00125). Note that search profiles and corresponding graph nodes on them are shifted relative to each other compared with the graph given in Fig. 8(a). The definition of the smoothness constraints encourages surface smoothness and penalizes major deviations from the given surface shape prior.

Separating the third lobe of the right lung. We utilize the aforementioned surface finding steps with oblique fissure priors, to split the right as well as left lung into two parts. While the left lung has only two lobes, the right lung needs to be further subdivided to obtain the third lobe. Because of anatomical variations of fissure locations, the first lobe-separating surface found for the right lung (Fig. 10) can either separate (a) the right lower lobe (RLL) or (b) the right upper lobe (RUL) from the rest of the lung. Depending on the found solution, the remaining lobe-separating surface can either be in the upper or lower part of the lung. To address this issue, we find the best lobe-separating surface by evaluating the two options and selecting the solution, which maximizes the average fissureness measure at the identified lobe surface. To identify the lobe-separating surfaces in each of the two subparts of the lung, we

Fig. 9 Interlobar surface segmentation steps at inspiration. (a) Lung mesh (top) and CT cross-sections (middle and bottom), (b) initial planar surface prior, (c) initial OSS result with color-coded fissureness, (d) initial fissure segmentation, (e) TPS inter- and extrapolation, and (f) final interlobar boundary.

Fig. 10 Lobe separation in the right lung. The first found interlobar surface might separate (a) the lower right lobe or (b) the upper right lobe from the rest of the lung.

first modify their cost image. For each subpart, all costs outside of the subpart or within a 2-mm range of the first lobe-separating surface are set to the median cost value inside of the subpart. For identification of a lobe-separating surface in the upper part of the lung, we apply the two surface finding steps with the planar RH fissure prior surface [Fig. 7(c)], and for the lower part of the lung with the planar RO fissure prior [Fig. 7(b)].

3.2.2 Lobe separation at expiration utilizing a subject-specific prior

The segmentation results of the inspiration scan (Sec. 3.2.1) provide subject-specific information about the patient's lung lobe anatomy. We utilize this information as a prior for segmentation in the expiration scan to obtain segmentation results, which are

Fig. 11 Warping of fissure surfaces from inspiration scan to expiration scan. (a) ASM lung meshes matched to inspiration and expiration scan provide landmarks for warping the inspiration scan results to the expiration scan. Sagittal view of warped surfaces at expiration of (b) right lung and (c) left lung, respectively.

more accurate and more consistent with results of the inspiration scan. The approach is shown in Fig. 11. First, we warp the interlobar surfaces from the inspiration scan to the expiration scan and then adapt these prior surfaces to the cost image using OSS.

We obtain a sparse set of point correspondences between the two datasets using the ASM-based segmentation of the lungs. These ASMs are fitted to the inspiration and expiration CT scan pair simultaneously using the method proposed by Gill and Beichel.¹⁰ The method is computationally efficient, robust against gross lung pathology, and provides a set of correspondences between the two datasets, which are approximately evenly distributed on the lung surface. Using these points, we warp the inspiration scan results to the expiration scan using a TPS-based transformation (Fig. 11). On the development dataset, we assessed the quality of the registration to facilitate parameter selection. The warped oblique interlobar surfaces showed a mean unsigned distance error to the actual fissures of ~ 4 mm without any bias (signed error). The warped RH interlobar surfaces showed a bias of about 5 mm. After correction for this bias, the average remaining mean unsigned distance error was found to be ~ 4 mm. Thus, the approach preserves the shape of the warped interlobar surfaces and obtains the approximate location within the lungs, which makes the warped surfaces well suited as priors for a subsequent OSS-based refinement step.

To adapt the warped interlobar boundaries to the expiration scan, the final OSS step follows similar ideas as presented for the fine-scale surface finding at inspiration (Sec. 3.2.1). Starting from the warped interlobar surfaces, parallel search profiles are constructed. Their directions are obtained based on a PCA of the warped surfaces. The search range is ± 15 mm, the sampling distance along the search profiles is 0.25 mm, and hard and soft smoothness constraints are ± 1 mm and 0.00125, respectively. To compensate for the systematic error of the warped RH surface, the warped surface is shifted by 5 mm in the inferior direction, before OSS graph construction.

4 Experimental Setup

We assess the performance of the proposed method on pairs of inspiration and expiration lung CT scans (test dataset). In addition, we compare the results on the expiration scans to two other approaches. The first alternative approach processes the expiration scan in the same manner as the inspiration scan (Sec. 3.2.1) instead of utilizing the subject-specific prior. The second alternative approach utilizes the publicly available elastix registration framework¹⁸ with parameters suggested by the authors for pulmonary image registration¹⁹ to warp the segmentation result from the inspiration scan to the expiration scan. A summary of all assessed approaches with short descriptions is given in Table 2.

4.1 Reference Standard

For each evaluation datasets, we established an independent reference standard for LO, RO, and RH interlobar boundaries. For this purpose, a pulmonologist with several years of experience in analyzing lung CT scans manually traced the three boundaries in every 10th coronal slice of each CT scan. To facilitate tracing in the coronal slices, the observer was able to draw auxiliary contours in axial and sagittal image slices. In addition, the pulmonologist processed each expiration scan immediately after the corresponding inspiration scan, and therefore was able to utilize information from the inspiration scan for tracing in the

Table 2 Evaluated segmentation approaches.

Approach	Description
Gen _{Insp}	Inspiration scan segmented using generic surface priors (Sec. 3.2.1).
Gen _{Exp}	Expiration scan segmented using generic surface priors (Sec. 3.2.1) and no information from the inspiration scan.
Reg _{Exp}	Inspiration scan results Gen _{Insp} warped to expiration scan based on a pulmonary image registration. ^{18,19}
Sub _{Exp}	Expiration scan segmented using inspiration scan result Gen _{Insp} as subject-specific surface prior (Sec. 3.2.2).

expiration scan. Because fissures can be incomplete, the pulmonologist was instructed to (a) provide accurate tracings of visible fissures with one label and (b) use another label to provide an estimate of the remaining interlobar boundaries using other anatomical clues and expert knowledge (Fig. 12). This approach enables evaluating segmentation results against visible fissures as well as the complete interlobar boundaries. In a few cases, the pulmonologist was not able to identify any fissures, and thus did not generate a reference. This resulted in two LO and two RO fissures without reference tracings, and therefore were excluded from statistical analysis.

4.2 Performance Metrics

Based on the distance of the reference tracings to the algorithm-identified interlobar surfaces, we calculated the mean d_{mean} , root mean square d_{RMS} , and maximum d_{max} distances for each fissure surface similarly as described by other authors.⁴⁻⁸ We obtained these metrics for all visible fissures and complete

Fig. 12 Reference standard for the three interlobar boundaries. Different colors indicate whether a fissure was visible in the dataset (red, green, and blue) or not (cyan, magenta, and yellow). (a) 3-D rendering of tracings, (b) coronal slice of a CT scan, and (c) corresponding tracing.

interlobar boundaries. In addition, based on this reference, a fissure completeness score^{6,8} was calculated for each fissure. It is defined as a percentage of visible fissure tracings versus tracings of the complete interlobar boundary.

4.3 Statistical Significance Tests

Performance differences between approaches were tested for statistical significance. Due to the highly skewed nature of the observed measurements, the log of the distance at each reference tracing point was taken, which was then averaged across all reference tracing points within a subject. This resulted in an overall discrepancy measure for each subject d_{ML} suitable for statistical significance tests. This discrepancy was determined for each of the four investigated approaches, and separately for visible fissures and complete interlobar boundaries. For analytical purposes, a repeated measures framework was used since we have (nonindependent) response measurements on each subject. As such a subject's d_{ML} will be correlated across methods. To account for the dependent nature of the response measurements, a multivariate analysis of variance (MANOVA) was conducted. The MANOVA allows for a completely unspecified covariance/correlation matrix to account for the interdependence of these repeated measures.

5 Results

Performance metrics d_{mean} , d_{RMS} , and d_{max} for each of the approaches Gen_{Insp} , Gen_{Exp} , Reg_{Exp} , and Sub_{Exp} are summarized in Tables 3 and 4. Both tables list results for individual fissure surfaces as well as all fissures combined. Table 3 shows results obtained when comparing segmentations to tracings of visible fissures only, whereas Table 4 lists results for tracings of the complete interlobar boundaries. Note that larger

local errors (outliers) are quite common in fissure segmentation, and mean values can be considerably influenced by outliers. Thus, the median value should be considered too when comparing methods. For visible fissure surfaces, methods Gen_{Insp} , Gen_{Exp} , Reg_{Exp} , and Sub_{Exp} produced an overall (ALL) median d_{mean} value of 0.62, 0.83, 1.53, and 0.68 mm, respectively. For interlobar surfaces, the same methods resulted in an overall (ALL) median d_{mean} value of 1.05, 1.54, 1.91, and 1.14 mm, respectively. For example, while Gen_{Exp} can produce larger errors in a few cases [larger average error than Reg_{Exp} ; see example in Fig. 18(d)], Reg_{Exp} is more prone to bias (larger median error than Gen_{Exp}). In contrast, Sub_{Exp} and Gen_{Insp} show better and more balanced performance than Gen_{Exp} and Reg_{Exp} .

The exponentiated average d_{ML} across subjects is presented in Tables 5 and 6. The conducted MANOVA on d_{ML} resulted in the p -values for the pairwise comparisons, which are summarized in Tables 5 and 6. The joint test was highly significant for the visible fissure data ($p < 0.0001$, Wilks Λ) and also the interlobar boundary data ($p < 0.0001$, Wilks Λ), indicating significant differences among groups. Pairwise followup tests indicated a significant difference between all pairings of approaches when $\alpha = 0.05$. In conclusion, we have strong evidence to suggest that there is a difference among the four approaches. The order of overall d_{ML} performance, from best to worst, is Gen_{Insp} , Sub_{Exp} , Gen_{Exp} , and finally Reg_{Exp} .

Examples of segmentation results for subjects with no abnormalities, COPD, asbestosis, and asthma are shown in Figs. 13–16, respectively. For method Gen_{Insp} , the average value for d_{mean} over these cohorts when considering visible fissures was 0.59, 0.63, 0.74, and 0.77 mm, respectively. On the same cohorts, approach Sub_{Exp} resulted in d_{mean} values of 0.58, 0.67, 0.99, and 1.00 mm, showing that our approach

Table 3 Mean, RMS, and maximum distance errors between reference tracings of visible fissures and segmentation results. Results are provided separately for each fissure surface as well as combined (ALL).

Approach	Measure (mm)	Visible fissures			
		LO	RO	RH	ALL
Gen_{Insp}	d_{mean}	0.59 ± 0.18	0.64 ± 0.29	0.91 ± 0.72	0.68 ± 0.25
Gen_{Insp}	d_{RMS}	1.02 ± 0.43	1.18 ± 0.85	1.72 ± 1.78	1.43 ± 0.94
Gen_{Insp}	d_{max}	7.13 ± 3.31	8.79 ± 5.16	9.64 ± 10.77	13.25 ± 9.95
Gen_{Exp}	d_{mean}	1.54 ± 2.22	3.24 ± 6.05	3.34 ± 9.97	2.19 ± 3.39
Gen_{Exp}	d_{RMS}	2.54 ± 3.46	4.63 ± 7.16	4.24 ± 10.52	4.23 ± 5.12
Gen_{Exp}	d_{max}	10.39 ± 9.49	15.65 ± 14.82	11.64 ± 15.24	22.10 ± 17.14
Reg_{Exp}	d_{mean}	2.08 ± 2.15	2.16 ± 1.70	1.96 ± 1.23	2.04 ± 1.40
Reg_{Exp}	d_{RMS}	2.97 ± 2.73	3.07 ± 2.22	3.10 ± 2.15	3.23 ± 1.90
Reg_{Exp}	d_{max}	10.66 ± 6.82	11.65 ± 6.68	12.03 ± 10.11	16.74 ± 10.08
Sub_{Exp}	d_{mean}	0.82 ± 0.71	0.79 ± 0.54	1.04 ± 0.92	0.82 ± 0.44
Sub_{Exp}	d_{RMS}	1.35 ± 1.35	1.34 ± 1.14	1.94 ± 2.18	1.63 ± 1.26
Sub_{Exp}	d_{max}	6.89 ± 6.04	7.60 ± 5.56	9.74 ± 10.49	12.95 ± 10.85

Table 4 Mean, RMS, and maximum distance errors between reference tracings of the complete interlobar boundaries and segmentation results. Results are provided separately for each interlobar surface as well as combined (ALL).

Approach	Measure (mm)	Complete interlobar boundaries			
		LO	RO	RH	ALL
Gen _{Insp}	d_{mean}	0.92 ± 0.81	1.02 ± 0.61	2.12 ± 1.56	1.18 ± 0.54
Gen _{Insp}	d_{RMS}	1.73 ± 1.42	2.06 ± 1.45	3.91 ± 2.66	2.71 ± 1.38
Gen _{Insp}	d_{max}	10.82 ± 6.78	13.31 ± 7.51	19.54 ± 12.06	22.85 ± 11.55
Gen _{Exp}	d_{mean}	1.86 ± 2.12	3.48 ± 5.63	3.79 ± 7.35	2.65 ± 3.14
Gen _{Exp}	d_{RMS}	3.30 ± 3.29	5.32 ± 6.95	5.35 ± 8.06	5.28 ± 4.63
Gen _{Exp}	d_{max}	14.84 ± 9.24	18.96 ± 14.90	17.96 ± 14.31	27.87 ± 14.90
Reg _{Exp}	d_{mean}	2.36 ± 2.05	2.39 ± 1.72	2.59 ± 1.45	2.40 ± 1.40
Reg _{Exp}	d_{RMS}	3.62 ± 2.69	3.50 ± 2.46	4.14 ± 2.18	4.02 ± 1.88
Reg _{Exp}	d_{max}	14.96 ± 8.85	14.44 ± 9.36	17.85 ± 9.57	23.13 ± 10.41
Sub _{Exp}	d_{mean}	1.12 ± 0.75	1.12 ± 0.78	2.06 ± 1.44	1.34 ± 0.58
Sub _{Exp}	d_{RMS}	2.10 ± 1.48	2.03 ± 1.57	3.72 ± 2.54	2.89 ± 1.36
Sub _{Exp}	d_{max}	11.68 ± 7.34	11.07 ± 7.33	17.81 ± 10.73	21.93 ± 9.89

Table 5 Tests for statistically significant performance differences between approaches for visible fissures.

Approach	$\exp(\overline{d_{\text{ML}}})$	Pairwise comparison p -value		
		Sub _{Exp}	Gen _{Exp}	Reg _{Exp}
Gen _{Insp}	0.315	0.0009	0.0007	<0.0001
Sub _{Exp}	0.364	—	0.0122	<0.0001
Gen _{Exp}	0.467	0.0122	—	<0.0001
Reg _{Exp}	0.777	<0.0001	<0.0001	—

Table 6 Tests for statistically significant performance differences between approaches for the complete interlobar boundaries.

Approach	$\exp(\overline{d_{\text{ML}}})$	Pairwise comparison p -value		
		Sub _{Exp}	Gen _{Exp}	Reg _{Exp}
Gen _{Insp}	0.407	<0.0001	0.0003	<0.0001
Sub _{Exp}	0.478	—	0.0107	<0.0001
Gen _{Exp}	0.613	0.0107	—	0.0003
Reg _{Exp}	0.912	<0.0001	0.0003	—

closely tracks the performance of the generic approach on inspiration volumes with largest differences occurring for asbestosis and asthma data sets. In contrast, methods Gen_{Exp} and Reg_{Exp} resulted in average error values that were larger by 50% or more.

Figures 17(a) and 17(b) compare fissure completeness scores between inspiration and expiration scans, and Fig. 17(c) provides a comparison between the three fissures. The average fissure completeness of inspiration scans was 84.46% compared with 80.88% on expiration scans. The average fissure completeness on inspiration and expiration scans combined was for the RH fissures 62.48%, compared with 82.72% and 85.78% on LO and RO fissures, respectively.

The presented lung lobe segmentation approach represents a part of a complete lung CT analysis system, which segments the lungs (robust 4-D ASM matching) and airways in prior processing steps. The overhead introduced by lobe segmentation of an inspiration and expiration scan pair was ~17 min. In comparison, computation times of the lung registration-based approach (Reg_{Exp}) required about 106 min per dataset.¹⁹

6 Discussion

6.1 Performance

In experiments, we compared our method with subject-specific prior for segmentation of the expiration scan (Sub_{Exp}) with two alternative approaches (Table 2)—one processing the expiration scan in the same fashion as the inspiration scan (Gen_{Exp}), and another one utilizing a pulmonary image registration method (Reg_{Exp}). As can be seen from the results presented in Tables 3 and 4 as well as our statistical analysis of performance differences (Tables 5 and 6), results with the generic prior on expiration scans Gen_{Exp} were significantly worse than those obtained with the same method on inspiration scans Gen_{Insp}. While Gen_{Exp} can also produce good quality segmentation results, the method frequently produces larger segmentation errors. Utilizing the proposed Sub_{Exp} approach with a subject-specific prior increased segmentation performance significantly. Figure 18 depicts an example, demonstrating the achievable improvement in segmentation quality. Overall,

Fig. 13 Results on three inspiration/expiration scan pairs of subjects with no significant abnormalities.

Fig. 14 Results on three inspiration/expiration scan pairs of subjects with COPD.

Fig. 15 Results on three inspiration/expiration scan pairs of subjects with asbestosis.

Fig. 16 Results on three inspiration/expiration scan pairs of subjects with asthma.

Fig. 17 Fissure completeness. (a) Boxplot of fissure completeness in inspiration scans. (b) Boxplot of fissure completeness in expiration scans. (c) Cumulative histograms of fissure completeness for the three surfaces.

Fig. 18 Example of improvements achievable with the proposed approach. (a) Inspiration scan, (b) corresponding Gen_{Insp} result, (c) expiration scan with barely visible RH fissure (visibility score = 0.17), (d) corresponding Gen_{Exp} result with RH interlobar $d_{mean} = 33.01$ mm, and (e) Sub_{Exp} result with RH interlobar $d_{mean} = 1.47$ mm.

results of the registration-based method (Reg_{Exp}) were found to be less accurate, when compared with Gen_{Insp} and Sub_{Exp} . Figure 19 shows an example where the registration-based method had problems to deliver a good result. In this context, more advanced registration methods are emerging, which aim to take sliding of adjacent lobes into account or focus specifically

on better alignment of fissures between respiratory states (e.g., Refs. 20 and 21). However, while these methods are promising, trade-offs between performance and computing time need to be considered. In comparison, utilizing Sub_{Exp} produces significantly better results with a performance close to that achieved on inspiration scans Gen_{Insp} . However, the remaining

Fig. 19 Example of a case where the registration-based approach shows local errors. (a) Expiration scan and (b) Reg_{Exp} result. The sliding of lung lobes creates discontinuities, which is challenging for image registration.

Fig. 20 Example of an inspiration scan result with RH fissure distance error $d_{\text{mean}} = 2.33$ mm. The method followed an accessory fissure in the dataset instead of the actual lobar fissure. (a) Inspiration scan, (b) Gen_{Insp} result, and (c) independent reference standard with RH fissure location shown in pink color.

performance difference between inspiration scan results Gen_{Insp} and expiration scan results Sub_{Exp} was found to be statistically significant.

6.2 Comparison with Other Approaches

Obtaining accurate segmentations in expiration scans is more challenging compared with inspiration scans. This can be clearly seen when comparing results between inspiration and expiration scans reported by other authors (Table 1). Also, we found the fissures in expiration scans appear less complete than in inspiration scans (Fig. 17). To address this issue, we presented

a method for lung lobe segmentation in pairs of inspiration and expiration CT scans, which utilizes a subject-specific shape prior for segmenting the expiration volume. Although other lobe segmentation approaches with similar cost function components exist,^{4,9} they process an expiration CT scan without considering information of the corresponding inspiration scan. Our experiments have demonstrated the relative performance gain that can be achieved by utilizing a subject-specific prior. Thus, we expect that by incorporating such information into these methods, a similar relative performance gain can be achieved. In this context, the utilized OSS represents a well-suited framework to include such mesh-based surface priors. Also, segmentation of distal airways is frequently utilized by other approaches to initialize the process of determining the lung lobes. However, finding distal airways is typically more challenging in expiration scans. Again, we avoid such issues by utilizing a subject-specific prior for OSS. Compared with results report by other authors who evaluated their approach on inspiration and expiration scans (Table 1), methods Gen_{Insp} and Sub_{Exp} (Table 3) show less error for inspiration and expiration scans in our evaluation. However, note that a direct comparison is not possible due to the different set of evaluation datasets and reference tracings.

6.3 Current Limitations

For the majority of datasets, results of proposed method Gen_{Insp} and Sub_{Exp} show only small segmentation errors. Examples of cases with larger segmentation errors of Gen_{Insp} and Sub_{Exp} are given in Figs. 20 and 21. Such cases are usually related to ambiguous image information. On typical cases, segmentation results for LO and RO fissures are both approximately in a range of about two voxels, on average. Distance errors for RH fissures are slightly larger. A similar pattern has been reported by other authors (e.g., Ref. 6) and is probably caused by the lower fissure visibility of the RH fissure compared with oblique fissures (Fig. 17). When comparing metrics for visible fissures versus complete interlobar boundaries, one can note that the distance error increases. Arguably, neither a human expert nor an automated method can provide accurate delineations of the lung lobes in such areas. However, the resulting distance error still stays relatively low, indicating a good level of agreement between proposed method and the human expert.

7 Conclusion

Obtaining accurate fissure segmentations of expiration scans is typically more challenging than segmentation of inspiration

Fig. 21 Example of an expiration scan result with RO fissure distance error $d_{\text{mean}} = 2.77$ mm. In this case, the RO fissure contains a discontinuity. (a) Inspiration scan, (b) Gen_{Insp} result, (c) expiration scan, (d) Sub_{Exp} result, and (e) independent reference standard with RO fissure location shown in pink color.

scans, as demonstrated by our analysis as well as other publications. To address this issue, we have presented a method that utilizes information from the inspiration scan as a prior for segmentation of the expiration scan of the same subject. This approach enabled increasing robustness and achieving better consistency between inspiration and expiration results. In our evaluation on four cohorts, which included cases with severe disease, the method was able to produce good results in inspiration as well as expiration scans, even in challenging cases.

Disclosures

The authors have no conflict of interest to report.

Acknowledgments

This work was supported in part by NIH/NHLBI Grant No. R01HL111453 as well as a PILOT grant from the Institute for Clinical and Translational Science at the University of Iowa via the NIH/NCATS Grant No. UL1TR000442. The authors thank Knute D. Carter and Cody Hansen at the University of Iowa for their support in statistical analysis.

References

1. T. D. Koster and D.-J. Slebos, "The fissure: interlobar collateral ventilation and implications for endoscopic therapy in emphysema," *Int. J. Chronic Obstruct. Pulm. Dis.* **11**, 765–773 (2016).
2. T. Doel, D. J. Gavaghan, and V. Grau, "Review of automatic pulmonary lobe segmentation methods from CT," *Comput. Med. Imaging Graphics* **40**, 13–29 (2015).
3. J. Pu et al., "Pulmonary lobe segmentation in CT examinations using implicit surface fitting," *IEEE Trans. Med. Imaging* **28**(12), 1986–1996 (2009).
4. B. Lassen et al., "Automatic segmentation of the pulmonary lobes from chest CT scans based on fissures, vessels, and bronchi," *IEEE Trans. Med. Imaging* **32**(2), 210–222 (2013).
5. S. Ukil and J. M. Reinhardt, "Anatomy-guided lung lobe segmentation in x-ray CT images," *IEEE Trans. Med. Imaging* **28**(2), 202–214 (2009).
6. J. C. Ross et al., "Pulmonary lobe segmentation based on ridge surface sampling and shape model fitting," *Med. Phys.* **40**(12), 121903 (2013).
7. L. Zhang, E. A. Hoffman, and J. M. Reinhardt, "Atlas-driven lung lobe segmentation in volumetric x-ray CT images," *IEEE Trans. Med. Imaging* **25**(1), 1–16 (2006).
8. E. M. van Rikxoort et al., "Automatic segmentation of pulmonary lobes robust against incomplete fissures," *IEEE Trans. Med. Imaging* **29**(6), 1286–1296 (2010).
9. F. J. S. Bragman et al., "Pulmonary lobe segmentation with probabilistic segmentation of the fissures and a groupwise fissure prior," *IEEE Trans. Med. Imaging* **36**(8), 1650–1663 (2017).
10. G. Gill and R. R. Beichel, "Lung segmentation in 4D CT volumes based on robust active shape model matching," *Int. J. Biomed. Imaging* **2015**, 10 (2015).
11. G. Gill and R. R. Beichel, "An approach for reducing the error rate in automated lung segmentation," *Comput. Biol. Med.* **76**, 143–153 (2016).
12. K. Li et al., "Optimal surface segmentation in volumetric images—a graph-theoretic approach," *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(1), 119–134 (2006).
13. A. F. Frangi et al., "Multiscale vessel enhancement filtering," *Lect. Notes Comput. Sci.* **1496**, 130–137 (1998).
14. T. Doel et al., "Pulmonary lobe segmentation from CT images using fissureness, airways, vessels and multilevel B-splines," in *9th IEEE Int. Symp. on Biomedical Imaging (ISBI 2012)*, pp. 1491–1494 (2012).
15. C.-K. Tang, G. Medioni, and M.-S. Lee, "N-dimensional tensor voting and application to epipolar geometry estimation," *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(8), 829–844 (2001).
16. C. Bauer, M. Eberlein, and R. R. Beichel, "Airway tree reconstruction in expiration chest CT scans facilitated by information transfer from corresponding inspiration scans," *Med. Phys.* **43**(3), 1312–1323 (2016).
17. Q. Song et al., "Optimal graph search segmentation using arc-weighted graph for simultaneous surface detection of bladder and prostate," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2009)*, pp. 827–835, Springer, Berlin, Heidelberg (2009).
18. S. Klein et al., "elastix: a toolbox for intensity-based medical image registration," *IEEE Trans. Med. Imaging* **29**(1), 196–205 (2010).
19. M. Staring et al., "Pulmonary image registration with elastix using a standard intensity-based algorithm 2," Technical Report, University of Leiden Medical Center, Department of Radiology, Leiden, The Netherlands (2010).
20. V. Delmon et al., "Registration of sliding objects using direction dependent b-splines decomposition," *Phys. Med. Biol.* **58**(5), 1303–1314 (2013).
21. A. Schmidt-Richberg et al., "Lung registration with improved fissure alignment by integration of pulmonary lobe segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2012)*, pp. 74–81, Springer, Berlin, Heidelberg (2012).

Biographies for the authors are not available.