Atlas-Driven Lung Lobe Segmentation in Volumetric X-Ray CT Images

Li Zhang, *Member, IEEE*, Eric A. Hoffman, *Member, IEEE*, and Joseph M. Reinhardt*, *Senior Member, IEEE*

Abstract—High-resolution X-ray computed tomography (CT) imaging is routinely used for clinical pulmonary applications. Since lung function varies regionally and because pulmonary disease is usually not uniformly distributed in the lungs, it is useful to study the lungs on a lobe-by-lobe basis. Thus, it is important to segment not only the lungs, but the lobar fissures as well. In this paper, we demonstrate the use of an anatomic pulmonary atlas, encoded with *a priori* information on the pulmonary anatomy, to automatically segment the oblique lobar fissures. Sixteen volumetric CT scans from 16 subjects are used to construct the pulmonary atlas. A ridgeness measure is applied to the original CT images to enhance the fissure contrast. Fissure detection is accomplished in two stages: an initial fissure search and a final fissure search. A fuzzy reasoning system is used in the fissure search to analyze information from three sources: the image intensity, an anatomic smoothness constraint, and the atlas-based search initialization. Our method has been tested on 22 volumetric thin-slice CT scans from 12 subjects, and the results are compared to manual tracings. Averaged across all 22 data sets, the RMS error between the automatically segmented and manually segmented fissures is 1.96 ± 0.71 mm and the mean of the similarity indices between the manually defined and computer-defined lobe regions is 0.988. The results indicate a strong agreement between the automatic and manual lobe segmentations.

Index Terms—Anatomic atlas, lung lobar fissures, pulmonary imaging, segmentation.

I. INTRODUCTION

THE HUMAN lungs are divided into five distinct anatomic compartments or *lobes*. The physical boundaries between the lung lobes are called the *lobar fissures*. The left lung is comprised of two lobes, "upper" and "lower," separated by the left oblique (major) fissure. The right lung contains three lobes: "upper," "middle," and "lower"; the upper and middle lobes are separated by the horizontal (minor) fissure and the middle and lower lobes are separated by the right oblique (major) fissure (see Fig. 1).

The lobar fissures allow lobes to slip and rotate relative to the adjacent lobe(s) [1]–[3]. The lobes provide a means by which the lung can accommodate thoracic shape changes while
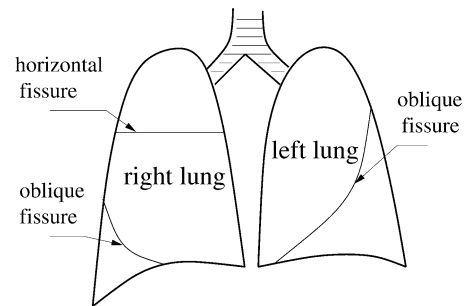
Fig. 1. Lung lobar anatomy. Left lung has two lobes separated by one major fissure; the right lung has three lobes, with the upper and middle lobes separated by a minor fissure and with the middle and lower lobe separated by the major fissure.

maintaining the regional integrity of the alveolar space [4]. It is important to consider interlobar differences in addition to left versus right differences when studying regional lung volumes, the distribution of ventilation and perfusion, analyzing regional lung strain variations, and measuring the variability of parenchymal expansion and deformation [1]–[3].

High-resolution X-ray computed tomography (CT) imaging is routinely used for clinical pulmonary applications and for research studying pulmonary structure–function relationships [5]. To study the lung across volume changes, across time, or across individuals, one must match lung structures between images [6]. Because the lobes deform independently with changes in lung volume, it is critical to locate the fissures prior to the process of matching one lung data set with another. Thus, it is important to segment not only the lungs, but the lobar fissures as well. Furthermore, since disease is usually not uniformly distributed in the lungs, disease evaluation and treatment planning may be improved by lobe-by-lobe assessment of pathology. As an example, it has been found that emphysema patients being evaluated for lung volume reduction surgery are most likely to gain benefit from the surgery if their disease is located predominantly in the upper lobes [7].

The lobar fissures are low contrast surfaces with blurred boundaries when viewed on cross-sectional CT images. Computer-based detection of the fissures is complicated by surrounding vessels and other structures, and noise and artifacts in the images. Automatic or semi-automatic fissure detection methods have been studied by different research groups. Kuhnigk *et al.* used an interactive three-dimensional (3-D) watershed algorithm to detect the lobar fissures on a cost image which was computed from a combination of the original data and the distance map performed on a previously generated vessel mask [8]. The authors showed the low variability of their method with different manual initializations. However, this method is sensitive to the vessel segmentation, which is already a difficult image segmentation task, and their results
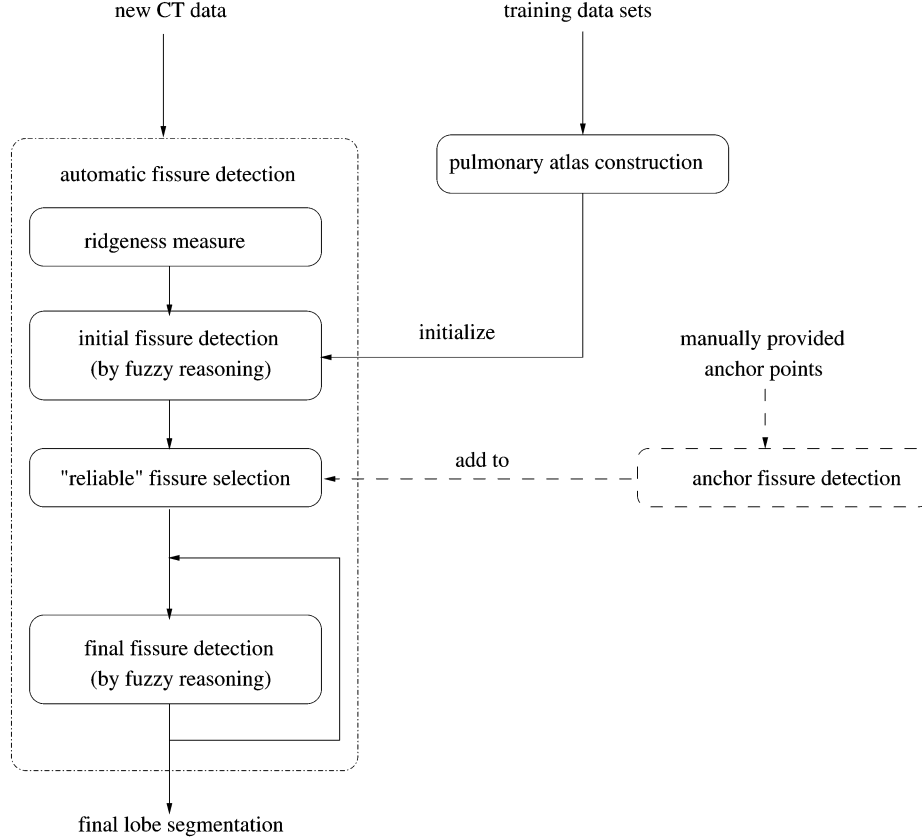
Fig. 2.   Diagram of the lobar fissure detection scheme. Detection is initialized by an anatomic atlas with average fissure and shape variation. The dashed lines indicate optional "anchor points" that may be manually provided to improve the segmentation accuracy.

may be quite different depending on the specific threshold used in vessel segmentation. In addition, their method may not be able to achieve satisfactory results for situations where pulmonary nodules are adjacent to the fissures or when vessels run parallel to and/or near the fissures. Kubo *et al.* extracted lobar fissures by the VanderBurg linear feature detector and morphological operators with directional disk-plane structure elements [9], [10]. This method was improved with intensity surface curvatures to detect fissures near lesions [11]. However, since this method only used local intensity and shape features and did not incorporate any global *a priori* anatomic knowledge, fissures that do not appear with sufficient contrast in CT images could be missed because they were not captured by the linear feature detector. Furthermore, this method did not extract the fissure surface near the lung boundaries. The incomplete segmentation, in addition to the other false negatives and false positives in the fissure extraction, make this method difficult to use in a clinical environment. More recently, Zhou *et al.* and Saita *et al.* independently described methods for lobar fissure detection based on blood vessel or blood vessel plus airway tree analysis [12], [13]. Saita *et al.* developed a method of detecting the "sheet-like" spaces between the pulmonary vasculature—taking advantage of the fact that the fissures divide the vasculature at the lobar boundaries [13]. Zhou *et al.* combined the vessel and airway trees and analyzed these structures with Voronoi analysis to identify the most likely location for the lobar fissures [12].

We have previously reported on a semi-automatic fissure detection method that identifies the oblique fissures, guided by shape constraints [14], [15], and an automatic segmentation method [16]. In this paper, we describe the automatic segmentation method for the oblique lobar fissures by incorporating *a priori* knowledge from an anatomic atlas. We first construct an anatomic atlas from a set of presegmented training data sets. The anatomic atlas is then used to initialize a local fissure detection procedure guided by a fuzzy reasoning system.

II. METHODS

The automatic lobe segmentation procedure consists of two main steps: the pulmonary atlas construction step and the automatic fissure detection step. The atlas construction step provides 3-D fissure surfaces for initialization. In the automatic fissure detection step, the contrast of the original CT image is enhanced by applying a ridgeness operator. Then the fissure detection is performed in the ridgeness map using the anatomic atlas for initialization. The fissure detection is performed based on two-dimensional (2-D) slice information, propagating across slices. There are two main steps in the fissure detection: initial fissure detection and final fissure detection (see Fig. 2). In this section, we first describe the atlas construction from presegmented data, and then we describe fissure detection using the atlas for initialization. In Section II-B5, we describe the use of manually provided "anchor points" to assist and guide the segmentation of difficult or poor quality image data sets.

A. Pulmonary Atlas Construction

1) Lung Surface-Based Registration: The goal of the pulmonary atlas used in this method is not to precisely represent

all of the pulmonary anatomy, but simply to initialize the subsequent fissure detection. Therefore, only two anatomic structures are represented in the atlas: the lung boundaries and the lobar fissures. The lung boundaries are used as features to register the fissure surfaces in the atlas onto a new data set. Our primary reason for choosing the lung boundaries as our anatomic "frame of reference" is that we are seeking a fully automatic lobar segmentation technique and the lung boundaries can be accurately and robustly identified automatically [17], [18]. Our work focuses only on lobar fissure detection, so the lung boundaries are inputs to our lobar segmentation system. Description and validation of methods for lung segmentation have been reported by others [17], [18].

The pulmonary anatomic atlas used in this work is constructed from the predefined fissures in a set of training data. Given n data sets, one data set is randomly selected as the template and the remaining fissures $S_1, S_2, \ldots, S_{n-1}$ are registered onto the template to build the atlas. Once the atlas is established, the location and variation of the average fissure is registered onto the new data set. The registration technique used in this work is a lung surface-based image registration, which has been reported in our previous work [19]. The following paragraphs briefly describe this registration method.

The registration of the fissures in the source images to the target images is accomplished in two steps: 1) a global transformation; and 2) a local deformation. The global transformation scales the bounding boxes of the source lungs and target lungs to the same size and translates them to the same position to reduce the gross differences between the two data sets. The local deformation computes the best match displacement (BMD) fields by matching lung boundaries with similarity measures. After global transformation, each lung is processed separately. For computational efficiency we first down-sample the lung boundaries into sparse sampling points, then for each boundary sample point, a sub-image (cube), centered at the sample point, is defined as edge cube (see Fig. 3). Correspondences between the source and target lung boundary points can be determined by measuring the similarities between the edge cubes at each sampling point. For each sample point in an image (either a source or a target image), we find its BMD vector by searching for the most suitable correspondence in the other image. Two match criteria are used in the search procedure: 1) maximum similarity and 2) minimum magnitude of the BMD vector.

- *Maximum similarity:* We measure the similarity of two edge cubes by counting the voxel differences between the two cubes [20]. Let cube P denote an edge cube in one image, centered at (x_p, y_p, z_p), and cube Q denote an edge cube in the other image, centered at (x_q, y_q, z_q). Then, the similarity between P and Q is measured as

$$s = \frac{1}{\left(\sum_{(i,j,k) \in P} [f(i,j,k) - h(i+dx, j+dy, k+dz)]^2 \right) + 1} \tag{1}$$

where $dx = x_q - x_p$, $dy = y_q - y_p$, $dz = z_q - z_p$. f and h represent the binary values of the source and target images with segmented lungs, respectively. The similarity measure s is in the range of $[0, 1]$, where 1 represents an exact match and a value close to 0 indicates that the two cubes are very different.
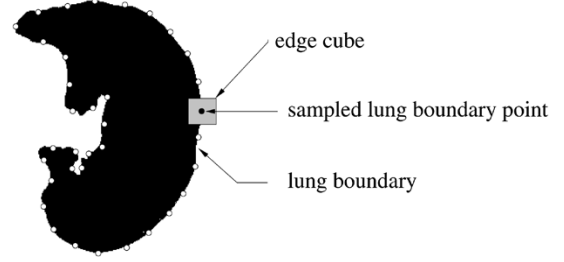

Fig. 3. The edge cube correspondences between the two lung boundaries are defined by measuring the similarities between the edge cubes. These similarities are measured only on the binary masks that define the lung regions in the CT images.

- *Minimum magnitude of BMD vector:* For a cube P in one image, assume we have found its most similar cube Q_0 in the other image, with similarity number $s_{\max}$. However, there may also be cubes $Q_1, Q_2, \ldots, Q_n$ with similarity numbers $s_1, s_2, \ldots, s_n$ that are close to $s_{\max}$, and the best match cube Q' can be selected from $Q_0, Q_1, Q_2, \ldots, Q_n$ by minimizing the magnitude of the match displacement vector. Then the BMD vector for cube P can be calculated as follows:

$$\vec{D}_{\mathrm{BMD}} = \arg\min_i \{ \|\mathbf{c}_{Q_i} - \mathbf{c}_P\| \mid s_{\max} s_i < \epsilon, \\ - i = 0, 1, 2, \ldots, n \} \tag{2}$$

where $\mathbf{c}_{Q_i}$ and $\mathbf{c}_P$ are the centers of cube Q_i and P, respectively. ϵ is a small positive constant. This criterion gives a BMD field that minimizes the distortion energy. Our intent is to find a minimum energy deformation field that preserves the continuity and topology of the geometric shapes in the image.

Two BMD fields are computed: one describes the matching from the source lung to the target lung and the other describes the matching from the target lung to the source lung. Then, the two BMD fields are combined by Burr's dynamic model [21] using Gaussian-weighted distance function to provide a smoothed deformation displacement (SDD) field. The SDD vector for a voxel (x, y, z) in the source image to be deformed can be calculated as follows:

$$\vec{D}_{\mathrm{SDD}}(x,y,z) = \frac{\sum_{m=1}^{M} g_m(x,y,z) \vec{D}_{\mathrm{BMD}}(m)}{\sum_{m=1}^{M} g_m(x,y,z)} \\ - \frac{\sum_{n=1}^{N} g_n(x,y,z) \vec{D}_{\mathrm{BMD}}(n)}{\sum_{n=1}^{N} g_n(x,y,z)} \tag{3}$$

where g_m and g_n are are Gaussian functions, shown in (4) at the top of the next page. In (3) and (4), M and N are the total number of BMD vectors in the source and target image, respectively. (x_m, y_m, z_m) and (x_n, y_n, z_n) indicate the center location of cube m and n. $\vec{D}_{\mathrm{BMD}}(m)$ and $\vec{D}_{\mathrm{BMD}}(n)$ are the BMD vectors of cube m and n. $D_x(n)$, $D_y(n)$, and $D_z(n)$ are x, y, and z components of $\vec{D}_{\mathrm{BMD}}(n)$. The constant σ used to calculate g_m and g_n is the elasticity number of the deformation. The SDD field deforms the lobar fissure surface from the source image onto the target image while preserving the topologies of the fissure surfaces.

2) Anatomic Atlas With Average Fissure and Variation: The anatomic lung atlas for fissure detection initialization is built using the previously identified fissure surfaces from n different

$$g_m(x,y,z) = \exp\left(\frac{-(x-x_m)^2 - (y-y_m)^2 - (z-z_m)^2}{\sigma^2}\right)$$

$$g_n(x,y,z) = \exp\left(\frac{-(x-x_n-D_x(n))^2 - (y-y_n-D_y(n))^2 - (z-z_n-D_z(n))^2}{\sigma^2}\right) \quad (4)$$

subjects. The atlas construction follows the method used by Thompson *et al.* [22] for brain abnormality detection.

After all the fissures from the training data sets are registered onto a template, the 3-D deformed fissures $F_1, F_2, \ldots, F_n$ are parameterized with regular 2-D grid $\Omega(u,v)$ to obtain a parameterized mesh surface set $\mathcal{F} = \{M_i(u,v), 0 \le i < n\}$, where n is the number of the training data sets. The details of the 3-D fissure surface parameterization are given in Appendix A. Let $\mathbf{p}_i(u,v)$ represent the 3-D Cartesian coordinates (x,y,z) of the point located on the ith parameterized mesh surface $M_i(u,v) = \{\mathbf{p}_i(u,v)\}$ for grid node (u,v). Then from the mesh surfaces $\{M_i(u,v), 0 \le i < n\}$, a set of homologous points $\mathbf{p}_i(u,v)$ can be obtained for each mesh grid node (u,v). Then two statistics, the mean, $\mathbf{p}_\mu(u,v)$, and the variance, $\sigma^2(u,v)$, can be calculated for all homologous points at (u,v): $\mathbf{p}_\mu(u,v) = (1/n)\sum_{i=1}^n \mathbf{p}_i(u,v)$, and $\sigma^2(u,v) = (1/n)\sum_{i=1}^n \|\mathbf{p}_i(u,v) - \mathbf{p}_\mu(u,v)\|^2$. The mean positions $\mathbf{p}_\mu(u,v)$ constitute the average fissure surface mesh $M_\mu(u,v)$. Meanwhile, for each grid node (u,v), we can compute the covariance matrix $\mathbf{V}(u,v) = [\sigma^2_{d_m d_n}(u,v)]$, where $d_m, d_n \in \{x,y,z\}$, to represent the directional variation of the displacement map $\{\vec{D}_i(u,v) = \mathbf{p}_i(u,v) - \mathbf{p}_\mu(u,v)\}$, then the mean position $\mathbf{p}_\mu(u,v)$ and the covariance matrix $\mathbf{V}(u,v)$, define a local probability density function (pdf) $g(u,v)$, which varies across different grid nodes. The details of the local PDF calculation and associated probabilistic analysis of the atlas shape accommodation can be found in Appendix B. If we assume that the training data and testing data are all from the same population, the probability analysis of atlas shape accommodation can show whether the atlas is able to accommodate the anatomic variability of different subjects in this population. The experimental results from the probability analysis demonstrated that the average fissure cannot provide close initialization for the fissure detection, and the fissure shape of some testing data sets cannot be completely represented by the atlas and its local probability distributions constructed from the training data sets while the testing and training data sets are both from the same population (see Section III-A and [23]). Therefore, the average fissure surface and its variance calculated in the atlas are only used to provide rough initial estimates for fissure location and shape in the automatic fissure detection, as described in detail in Section II-B.

B. Automatic Fissure Detection

In this paper, we focus only on the lobar fissure detection, so the lung segmentation is considered as a prerequisite. The inputs to the fissure detection system are the original pulmonary CT images and the lung masks from the segmentation, as well as the atlas with an average fissure obtained from different subjects in the same population.

Because the lobar fissure surfaces form locally bright "ridges" in the cross-sectional images, we use a ridgeness operator to enhance the fissure contrast. The fissure detection is carried out in the ridgeness image, with an initialization provided by the atlas. When the average fissure surface in the atlas is deformed to match a new data set, it provides the initial location and shape for the fissure detection in the new data set. Due to the large variation between different subjects and limitations imposed by using only lung surfaces for the image registration, the deformed average fissure may not be very close to the true fissure location in the new data set. As a result, the initial fissure detection, the first step in the two-step fissure detection procedure (see Fig. 2), must be performed over a relatively large search range. Because of noise and artifacts in the image data, the initial detection usually succeeds in regions where the 2-D fissures are clearly visible and continuous, but may fail in regions of low contrast or poor image quality. For these reasons, the results from the initial fissure detection are often unsatisfactory. We follow the initial detection with a reliable fissure selection, then based on the reliable fissure selection results, a final fissure detection step, with more stringent shape constraints, is used to obtain the final segmentation result.

The ridgeness analysis and the algorithm used in the initial fissure and final fissure detection are described in detail in the next parts of this section.

1) Ridgeness Measure: The lobar fissures appear as ridge-like structures in the cross-sectional images (see Fig. 4), thus a ridgeness measure can be used to highlight the fissure and improve overall contrast. Although there are a variety of ridgeness measures in the literature [24], [25], our experience has shown that the 2-D version of the so-called multilocal level set extrinsic curvature measure with enhancement by structure tensor analysis (MLSEC-ST) [24] is a good tradeoff between fissure ridge detection and computation time [23]. A brief description of MLSEC-ST can be found in Appendix C. In Fig. 4, we can see that MLSEC-ST effectively enhances the fissure contrast and reduces noise by identifying local gray level extrema, and will likely produce better results for fissure enhancement than traditional pixel-based edge detection.

Vessel branches, which are sometimes close to the fissures, may also appear as ridge-like structures in the images (see Fig. 4). To eliminate or reduce influence from vessels during fissure detection, an optimal thresholding step, combined with 3-D connected component analysis, is used to identify large vessel branches. The optimal threshold, which maximizes the intensity variance between the object and background [20], is computed from the histogram of voxels in the lung region. The high ridgeness values associated with the bright vessels can thus be removed from the ridgeness image.

2) Initial and Final Fissure Detection by Fuzzy Reasoning: Fuzzy logic theory provides a powerful decision making tool for medical image analysis by its ability to address uncertainties using overlapping fuzzy sets [17], [26]. In a fuzzy reasoning system used for image processing, image features are described by linguistic variables. For example, the ridgeness measure Γ for a image can be described as SMALL, MEDIUM, and LARGE, then a fuzzy set F, $F \in \{$SMALL, MEDIUM,
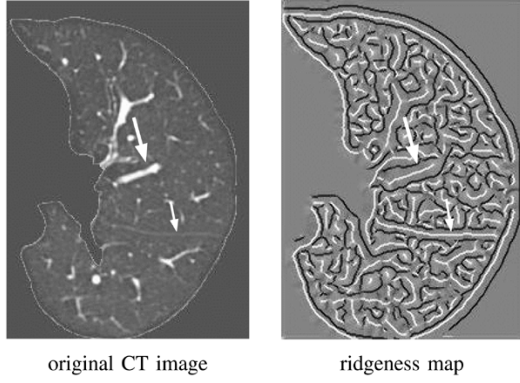
Fig. 4. Fissure contrast enhanced by the ridge detector (2-D MLSEC-ST). The small arrows point to the lobar fissures, and the large arrows point to the vessel branches.

LARGE} in the fuzzy space Γ is defined as a set of ordered pairs

$$F = \{(\gamma, \mu_F(\gamma)) \,|\, \gamma \in \Gamma\} \qquad (5)$$

where $\mu_F(\gamma) : \Gamma \longrightarrow [0, 1]$ is a graded membership function. Once the fuzzy sets are constructed, they interact with each other under fuzzy compositional rules. Fuzzy rules reflect the relationships between the fuzzy sets. Then a de-fuzzification process is applied to the fuzzy solution space to find a numeric value that best represents the information contained in the solution space.

For the initial and final fissure detection, we have three inputs: the ridgeness map from the original CT intensity, the fissure initialization (see Section II-B3), and the segmented lung boundaries. In this context, a fuzzy reasoning system is used to analyze data from the three inputs that are represented by three input fuzzy spaces for the initial and final fissure detection.

- Ridgeness measure Γ. The ridgeness values are derived from the voxel intensity in the CT images, therefore, represents the local image characteristics of the lobar fissures.
- Direction difference Θ. Θ is designed to measure the local 2-D smoothness of the fissure as candidate fissure pixels are evaluated. We use a smoothness constraint based on the anatomical characteristics of the fissures. Θ represents the direction difference between the current advancing direction of the search path and the predicted direction from the smoothness constraint. The predicted direction is first given by the fissure initialization, then, it is adapted gradually by the detection results obtained so far in a graph searching method that will be described later.
- Initialization distance Υ. The location of a fissure initialization provides a force to drive the search toward a preferred location. Υ is the minimum distance between the candidate fissure pixel and the location of the fissure initialization.

The fuzzy reasoning system also has a solution fuzzy space, called the fissure confidence Ξ, which represents the overall confidence that a pixel being located on a fissure.

The fuzzy sets and their corresponding fuzzy membership functions used in this method are shown in Fig. 5. The parameter sets $\{\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4\}, \{\theta_0, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}$, and $\{\upsilon_0, \upsilon_1, \upsilon_2, \upsilon_3, \upsilon_4\}$ have been experimentally selected for this work to be $\{-2.0, 0.5, 0.7, 1.0, 2.0\}$ (ridgeness calculated from grayscale units), $\{0°, 15°, 30°, 35°, 50°, 65°\}$ (angles), and
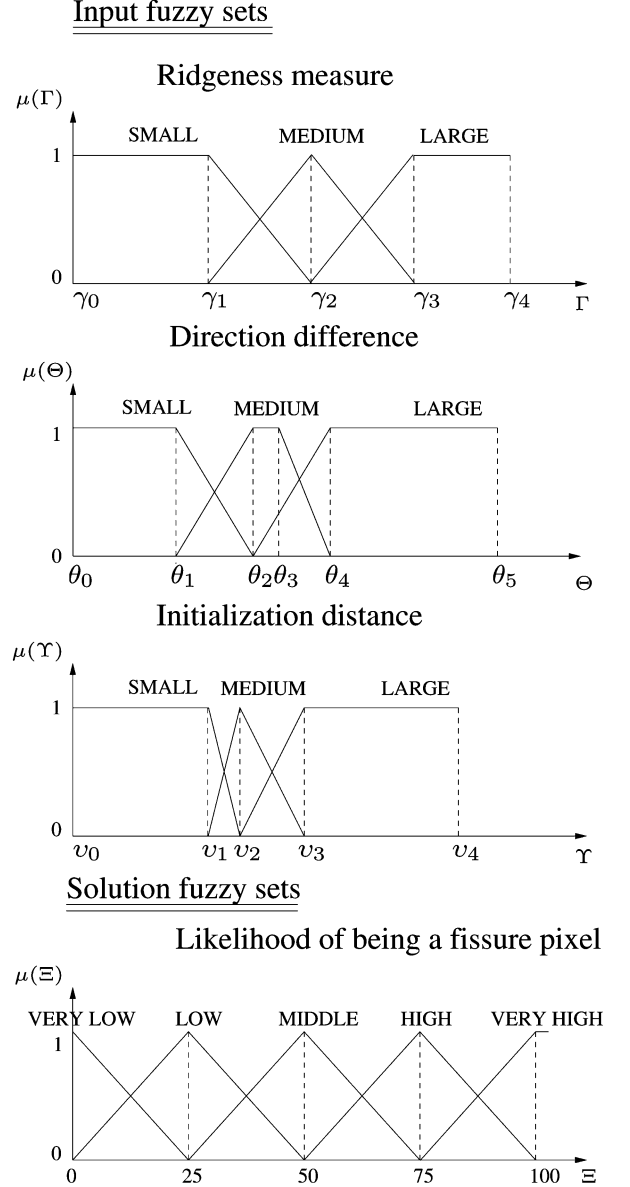

Fig. 5. Fuzzy sets for the automatic fissure search. The values in the input fuzzy spaces Γ, Θ, and Υ are described by the linguistic variables SMALL, MEDIUM, and LARGE. The values in the output fuzzy space Ξ are described by the linguistic variables VERY LOW, LOW, MEDIUM, HIGH, and VERY HIGH.

$\{0, 5, 10, 20, 80\}$ (pixels), respectively. These fuzzy parameter settings could be further tuned by a statistical analysis if a sufficient set of training data with gold standard segmentations were available.

The Γ, Θ, and Υ fuzzy sets are combined by the min-max fuzzy composition [20] with the 27 fuzzy rules given in Table I. The composite moments de-fuzzification [20] is then used to compute the numeric output fissure confidence value c_Ξ.

Once the confidence value for each voxel in the region of interest (ROI) is assigned, a graph searching method, with a cost function that is the complement of the confidence value, is used to find a globally optimal path that is the most likely fissure location in the ROI. The graph search goal nodes are the lung boundary points that are close to the end points of the initial fissure so that the result of 2-D fissure detection on each slice reaches both sides of the lung.
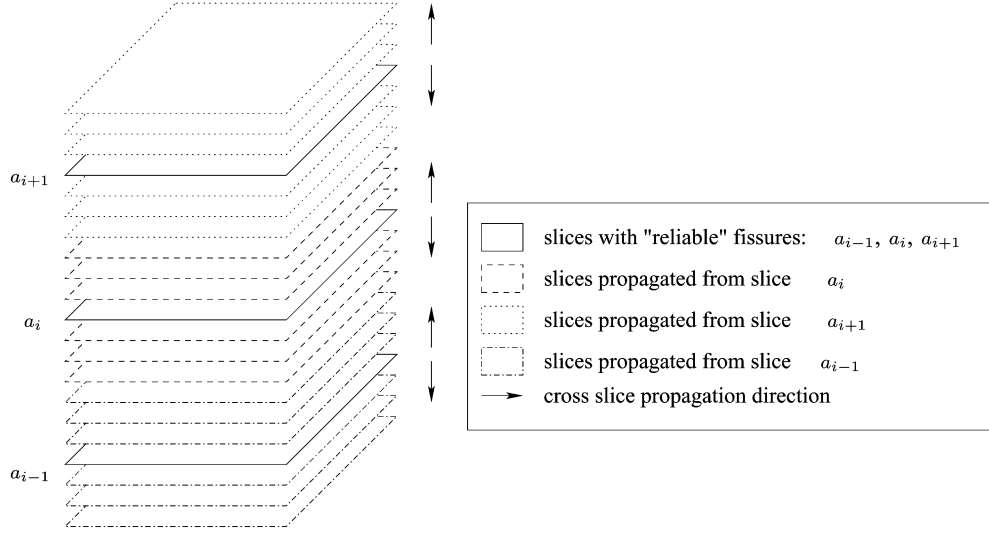
Fig. 6. Final fissure detection across slices. The final fissure detection procedure starts from the neighboring slices of the reliable fissures, and propagates to the slices with the other reliable fissures.

TABLE I
FUZZY RULES TO COMPUTE THE FISSURE LIKELIHOOD Ξ IN THE AUTOMATIC FISSURE SEARCH

Υ: SMALL		Γ		
		LARGE	MEDIUM	SMALL
Θ	SMALL	VERY HIGH	HIGH	MIDDLE
	MEDIUM	HIGH	MIDDLE	LOW
	LARGE	MIDDLE	LOW	VERY LOW

Υ: MEDIUM		Γ		
		LARGE	MEDIUM	SMALL
Θ	SMALL	HIGH	MIDDLE	LOW
	MEDIUM	MIDDLE	LOW	VERY LOW
	LARGE	LOW	VERY LOW	VERY LOW

Υ: LARGE		Γ		
		SMALL	MEDIUM	HIGH
Θ	SMALL	MIDDLE	LOW	VERY LOW
	MEDIUM	LOW	VERY LOW	VERY LOW
	LARGE	VERY LOW	VERY LOW	VERY LOW

3) Differences Between Initial and Final Fissure Detection: As shown in Fig. 2, a two-step fissure detection process is used to segment the fissures. Since the initial fissure initializations provided by the atlas may not be very close to the actual fissures in the new data set, different search strategies are used for the initial fissure detection and final fissure detection. The initial fissure detection selects "reliable" fissure initializations based on only ridgeness values, that is, only the fissures with good contrast are be selected as reliable fissures. The final fissure detection uses a combination of both intensity information and shape information. The key differences between the initial and final fissure detection steps are summarized as follows:

- *Different initialization:* The initial fissure detection uses an atlas initialization by finding the intersection of the 3-D

atlas with each of the 2-D transverse slices in the data set. The final fissure detection uses 2-D fissures propagated from neighboring slices to initialize the detection (see Fig. 6).

- *Different starting node selection for graph searching:* During fissure detection, we do not have any point that is guaranteed to be located on the fissures, so multiple starting nodes are used during the graph search. On the transverse slice, the starting nodes are defined as the points along a ventral–dorsal line segment passing through the middle point of the initial fissure. The size of the line segment reflects the uncertainty: the line segment is longer for the initial fissure detection as the atlas initialization is normally not close; and the segment is shorter for the final fissure detection as the fissures on the neighboring slices should be very close. The position and length of the line segment are determined by the average atlas fissure location and its associated variance (with a statistical significance level of 0.99) in the initial fissure detection, while the neighboring fissures define the starting nodes for the final fissure detection.
- *Usage of fuzzy space initialization distance Υ:* Υ in the fuzzy reasoning system is used only for the final fissure detection.

4) Reliable Fissure Selection: Given a set of candidate fissure locations from the initial search, each consisting of a connected path of pixels on a transverse CT slice, we wish to choose those fissure locations that are most likely to be correct (i.e., "reliable" fissure estimates) and discard those fissure locations that are most likely incorrect. The reliable fissure locations will be used to initialize the subsequent final fissure detection, the second step in the two-step fissure detection procedure (see Fig. 2).

The reliable fissure locations can be selected by evaluating the distribution of ridgeness values on pixels along the fissure. High contrast fissures are bright and continuous in the ridgeness maps. Thus, a path of pixels along an actual fissure will typically satisfy the following two criteria:

- bright in ridge map: voxels on a reliable fissure will have high ridgeness values;

- homogeneous and continuous: the variation in the ridge-ness values for voxels on a reliable fissure will be small.

The mathematical description of the above criteria can be given as follows. Let $\mathcal{P} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ denote a connected subset of n points on the fissure surface $\mathcal{S}$, i.e., $\mathcal{P} \subset \mathcal{S}$. Let the ridge measure for each point $\mathbf{x}_i$ in $\mathcal{P}$ be represented by $\gamma(\mathbf{x}_i)$. The mean and the standard deviation of the ridgeness values in $\mathcal{P}$, can be calculated by

$$\gamma_\mu = \frac{1}{n} \sum_{i=1}^{n} \gamma(\mathbf{x}_i) \tag{6}$$

and

$$\gamma_\sigma^2 = \frac{1}{n} \sum_{i=1}^{n} [\gamma(\mathbf{x}_i) - \gamma_\mu]^2. \tag{7}$$

We can then select the reliable fissures as follows:

$$\begin{cases} \gamma_\mu \geq T_\mu \quad \text{and} \quad \gamma_\sigma \leq T_\sigma \implies \text{fissure is reliable in } \mathcal{P} \\ \text{otherwise} \implies \text{fissure is not reliable in } \mathcal{P} \end{cases} \tag{8}$$

where T_μ and T_σ are predetermined thresholds for the mean and standard deviation of ridgeness values in $\mathcal{P}$. The simplest way to define $\mathcal{P}$ is to directly use the fissures on each transverse slice. In this paper, we use $T_\mu = 0.9$ and $T_\sigma = 0.3$ [23].

5) Initialization With Anchor Points: The atlas-driven segmentation can be also initialized with manually provided anchor points. In the automatic fissure detection, reliable fissures are chosen from the initial detection results by regional ridgeness value analysis. If manually provided anchor points are available for some slices, the fissures on these slices can be detected first using the anchor points as starting and goal nodes for the graph search. These anchor fissures are by definition reliable, and can be combined with or completely replace the automatically selected reliable fissures to initialize the final fissure detection (see Fig. 2), therefore, improve the overall lobe segmentation performance. Furthermore, when scans from patients with irregular fissures or lung parenchymal diseases are to be segmented, and we do not expect satisfied results directly from the automatic fissure detection, manually provided anchor points can be used from the start to guide the detection to avoid wasted automatic detection time.

C. Experimental Methods

Table II summarizes all of the data used for atlas construction and for lobe segmentation validation. All protocols involving the use of humans in research were approved by the University of Iowa Institutional Review Board. A total of 22 multi-detector CT scanner (MDCT) (Philips MX8000) of 12 subjects and seven electron beam CT (EBCT) (Imatron C150L) scans of seven subjects were used to construct atlas and test the segmentation. The MDCT data consisted of 12 scans at total lung capacity (TLC) and 10 scans near functional residual capacity (FRC). The EBCT data was acquired at TLC. The MDCT scans were acquired with 1.2-mm slice thickness; the EBCT images were acquired with contiguous 3-mm-thick slices. Each MDCT scan consists of about 400–580 overlapped slices with 0.6-mm slice spacing. Each EBCT scan consists of 100–140 slices. In both cases, the images were reconstructed onto a 512×512 matrix with pixel dimensions ranging from about 0.6 mm $\times$ 0.6 mm to 0.7 mm $\times$ 0.7 mm.

Nine MDCT TLC scans and the seven EBCT TLC scans were selected for use in the atlas construction step (the "atlas" data

sets). Twenty-two testing cases were used to evaluate segmentation performance, with each of the 22 MDCT data sets playing the role of a "testing" data set. When the testing data set was one of the nine MDCT data sets used for atlas construction, the testing data set was removed from the atlas and a new atlas was constructed using the remaining 15 atlas data sets. The fissures in the atlas data sets were identified by our previously developed semi-automatic fissure detection method [14], and then edited manually to ensure correctness. These semi-automatically defined fissures were used only for atlas construction; they were not used as the gold standard to assess accuracy since they were not independently segmented. After the atlas was constructed, the automatic method was applied to segment the lobar fissures in the testing data set.

To evaluate segmentation accuracy, we used a gold standard obtained by manual analysis. Two trained image analysts each traced the left and right oblique lobar fissures for selected slices for part of the evaluation data sets. While the automatic fissure detection method identified fissures in the entire 3-D lung, it is very time-consuming for the image analysts to do a complete segmentation by tracing all fissure slices (the number of slices on which the fissures are visible averages about 300 per lung on a typical MDCT scan) in all 22 volumetric CT scans for both the left and right fissures. Instead, the analysts only traced fissures on approximately every fifth slice. The analysts were instructed to skip slices on which the fissures were not visible or severely blurred to avoid forcing the analysts to "guess" during segmentation. Accuracy was assessed by measuring the difference between the manual tracing and the computer-generated result using the root mean square (RMS) error and maximum distance. The manually defined fissures, considered as the "gold standard" in this work, are used as the measurement baseline for distance calculation. Let $\mathbf{p}_i$ denote a voxel on a manually defined fissure $F_m = \{\mathbf{p}_i | 1 \leq i \leq n\}$, where n is the total number of the voxels on F_m. Let d_i denote the 3-D Euclidean distance between $\mathbf{p}_i$ and the closest voxel on the automatically detected fissure F_a. Then the RMS error and maximum distance can be calculated as follows:

$$\text{RMS} = \left(\sum_{i=1}^{n} \frac{d_i^2}{n} \right)^{\frac{1}{2}} \tag{9}$$

$$\text{maximum distance} = \max\{d_i | 1 \leq i \leq n\}. \tag{10}$$

In addition, since the purpose of the lobar fissure detection is to divide the lungs into lobes, a similarity measure between the manually defined lobes and the computer identified lung lobes is calculated to evaluate the segmentation agreement between the analysts and the algorithm. While the automatic fissure detection method completely segments lungs into lobes, the 2-D fissures traced by the analysts divided lungs into lobes only for the slices with manual tracings. The sum of the lobe areas on all slices with manual tracings can be considered as a sample of the 3-D complete lobe volume, and is used to represent the lobes for similarity evaluation in this work.

Let M_k and A_k denote the pixels in the manually defined lobe and the computer-defined lobe on a slice k, respectively, where k is a slice with a manual tracing, i.e.,

$$k \in K \tag{11}$$

and

$$K = \{\text{slices with manual tracings}\}. \tag{12}$$

TABLE II
SCAN DATA USED FOR ATLAS CONSTRUCTION AND FOR METHOD TESTING

Subject	Scanner	Condition	Used for Atlas Construction?	Used for Method Evaluation?	Normal/ Abnormal
S1	Spiral	TLC	Yes	Yes	peripheral nodules
		FRC	No	Yes	
S2	Spiral	TLC	Yes	Yes	mediastinal lymph node
S3	Spiral	TLC	Yes	Yes	mediastinal lymph node
		FRC	No	Yes	
S4	Spiral	TLC	Yes	Yes	airway obstruction
		FRC	No	Yes	
S5	Spiral	TLC	Yes	Yes	peripheral nodules
S6	Spiral	TLC	Yes	Yes	normal
		FRC	No	Yes	
S7	Spiral	TLC	Yes	Yes	normal
		FRC	No	Yes	
S8	Spiral	TLC	Yes	Yes	normal
		FRC	No	Yes	
S9	Spiral	TLC	Yes	Yes	normal
		FRC	No	Yes	
S10	Spiral	TLC	No	Yes	normal
		FRC	No	Yes	normal
S11	Spiral	TLC	No	Yes	normal
		FRC	No	Yes	normal
S12	Spiral	TLC	No	Yes	normal
		FRC	No	Yes	normal
E1	EBCT	TLC	Yes	No	normal
E2	EBCT	TLC	Yes	No	normal
E3	EBCT	TLC	Yes	No	normal
E4	EBCT	TLC	Yes	No	normal
E5	EBCT	TLC	Yes	No	normal
E6	EBCT	TLC	Yes	No	normal
E7	EBCT	TLC	Yes	No	normal

TLC: total lung capacity; FRC: functional residual capacity. Spiral: MDCT scanner; EBCT: electron beam CT scanner.

Then the similarity between the manually defined and computer-generated lobes can be measured as follows:

$$S = \frac{2\sum_{k \in K} |M_k \cap A_k|}{\sum_{k \in K} (|M_k| + |A_k|)} \qquad (13)$$

where $|M_k|$, $|A_k|$, and $|M_k \cap A_k|$ represent the number of pixels in sets M_k, A_k, and $M_k \cap A_k$. The similarity index defined in (13) is derived from the kappa coefficient calculation, which is used for reliability studies, and has been introduced in the area of image segmentation to measure the agreement between different classifications [27]–[29]. The similarity index S can also be interpreted as an overlap ratio. If we move the constant 2 in the numerator to the denominator in (13), the numerator becomes $\sum_{k \in K} |M_k \cap A_k|$, which is the overlap part between the manually defined and computer-generated lobes, and the denominator becomes $(1/2)\sum_{k \in K}(|M_k| + |A_k|)$, which is the average volume of the two lobes. The similarity index S ranges from 0 to 1. $S = 0$ indicates no overlap between the two segmentations and $S = 1$ indicates a perfect agreement between the two segmentations.
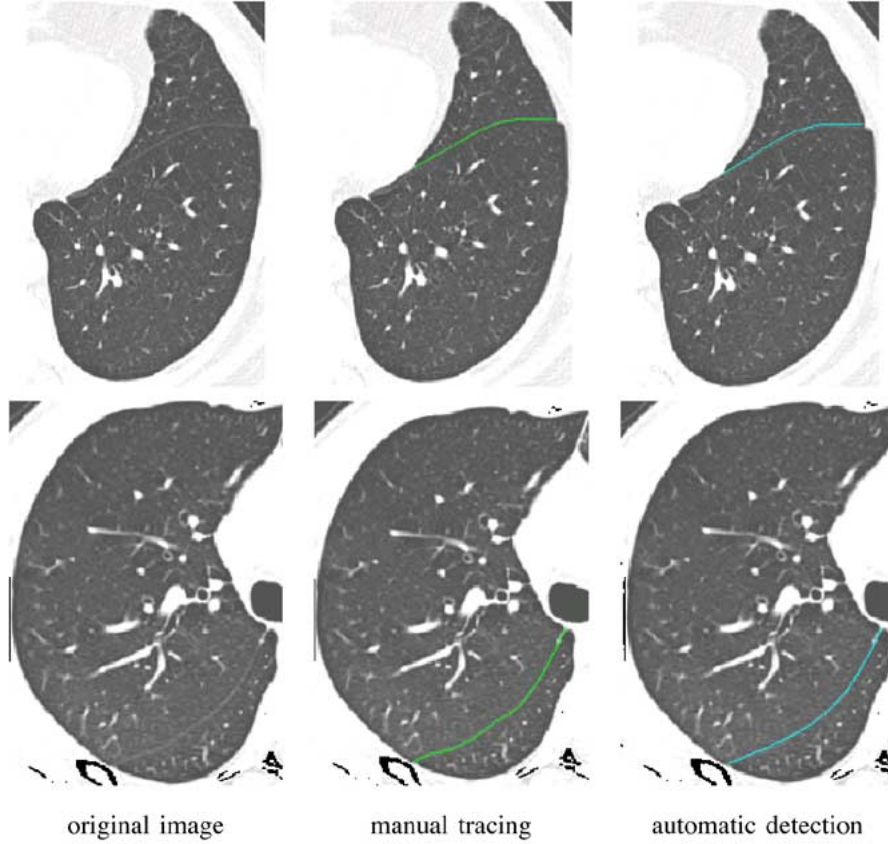
Fig. 7. Comparison between the manual tracings and automatic detection results.

The nonoverlap ratios for the manually defined lobes (R_m) and automatically defined lobes (R_a) are also calculated to indicate the agreement between the image analysts and computer-based algorithm. R_m and R_a are defined as follows:

$$R_m = \frac{\sum_{k \in K}(|M_k| - |M_k \cap A_k|)}{\frac{1}{2}\sum_{k \in K}(|M_k| + |A_k|)}. \quad (14)$$

$$R_a = \frac{\sum_{k \in K}(|A_k| - |M_k \cap A_k|)}{\frac{1}{2}\sum_{k \in K}(|M_k| + |A_k|)}. \quad (15)$$

In this paper, since the upper and lower lobe segmentations are correlated, the similarity index and nonoverlap ratios are calculated only for the lower lobe segmentations to represent the overall classification agreement between the image analysts and the automatic algorithm.

III. RESULTS

A. Atlas Construction

The experimental results from atlas construction show that the root mean square variability σ between different subjects in the normal and slightly diseased population is quite large. That is, after constructing the anatomic atlas using the lung surface-based registration technique with 16 scans from 16 different subjects, the variability $\sigma(u, v)$ at the mesh grid nodes ranges from 9.19 to 24.47 mm along the left oblique fissure and from 10.82 to 30.04 mm along the right oblique fissures.

B. Fissure Segmentation

Fig. 7 shows examples of the manual and automatic delineations for the left and right fissures. Fig. 8 shows the segmenta-
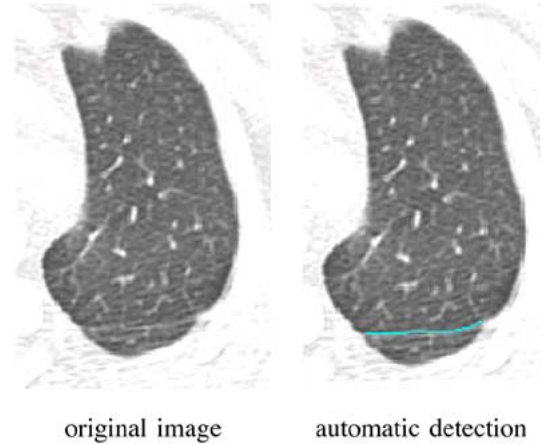

Fig. 8. Detection for a wide and blurred fissure. Streaking artifacts, due to beam hardening during image acquisition, are visible in the images.

tion of a fissure that is very wide and blurred, and almost indistinguishable in some regions. Streaking artifacts, due to beam hardening during image acquisition, are visible in Fig. 8, and make it more difficult to locate the fissures accurately. Fig. 9 shows the automatic detection results for a fissures adjacent to a large nodule. The images in Fig. 9 are from the TLC scan of a cancer patient (subject S1).

Mean and standard deviation, calculated across the two lungs in each of the 22 testing data sets, of the RMS errors (9), maximum distances (10), similarity indices S (13), nonoverlap ratios R_m and R_a [(14) and (15)] are given in Table III. The mean and standard deviation of the RMS errors calculated across 12
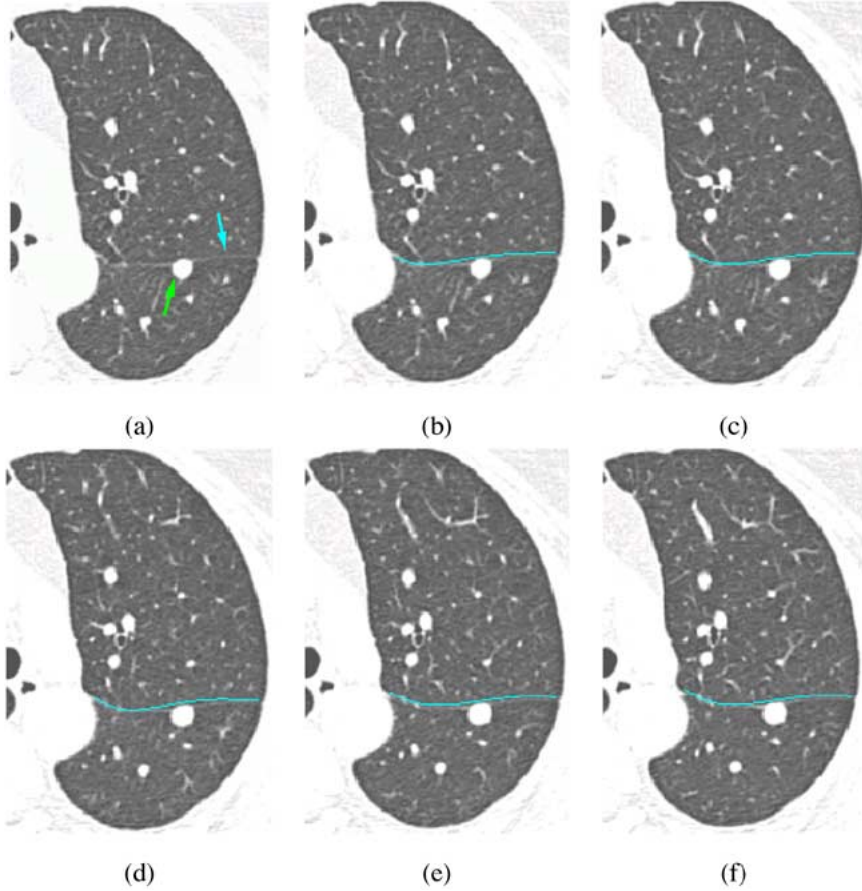
Fig. 9.   Detection for a fissure adjacent to a nodule. A large nodule can be observed from a sequence of contiguous transverse slices. The original image of the first slice in the sequence is shown in (a). The fissure is pointed by the cyan arrow, and the nodule is pointed by the green arrow in (a). (b)–(f) Automatic fissure detection results for the contiguous slices.

TABLE III
RMS ERROR, MAXIMUM DISTANCE, SIMILARITY INDICES S AND THE NONOVERLAP RATIOS R_m AND R_a (MEAN $\pm$ STD.DEV.) CALCULATED FOR THE 22 TESTING DATA SETS FROM 12 SUBJECTS

RMS error (mm)	maximum distance (mm)	similarity index S	non-overlap ratio R_m	non-overlap ratio R_a
1.96 ± 0.71	13.52 ± 4.63	0.988 ± 0.011	0.003 ± 0.004	0.021 ± 0.022

TLC testing data sets is 1.67 ± 0.53 (mm); the mean and standard deviation of the RMS errors calculated across 10 FRC testing data sets is 2.30 ± 0.76 (mm).

IV. DISCUSSION

A. Analysis of Automatic Detection Results

Classification agreement between the manual and automatic results is measured using (13) to evaluate the segmentation accuracy. The similarity index calculated by (13) is derived from the kappa statistic, which indicates the level of the agreement between two classifications. The large value of the similarity index indicates a strong agreement between the manual and automatic lobe segmentation results. The small nonoverlap ratios R_m and R_a also indicate a high level of agreement between the image analysts and the computer-based algorithm. Moreover, our proposed method can successfully detect fissures adjacent to nodules without extra modification.

The RMS errors between the manually defined fissures and the automatic results show that the automatically detected fissures are very close to the manually defined fissures in most cases, although occasional segmentation errors exist, as indicated by the maximum distances. The RMS errors are likely slightly optimistic, however, since we instructed the manual analysts to skip slices that were difficult to segment. Thus, many of the comparisons are made on the highest quality cross-sectional images of the fissures, which in turn are easier to segment both for the human and the computer.

The method segmentation errors are mainly caused by poor image quality. The weak contrast of the fissures, as well as the noise and artifacts, such as streaking artifacts and "double fissure" phenomenon caused by cardiogenic motion, make the fissures difficult to localize even for human experts. A typical situation of detection failure is shown in Fig. 10. However, normally the segmentation errors should not cause significant defects for lobe shape definition because the misclassifications are small as

indicated by the similarity measure. Furthermore, these errors typically occur near the outer lung boundaries, and major pulmonary structures, such as large airways and vessels, are rarely located in these regions. Therefore, although occasional errors exist in the automatic method, the overall performance of the lobe segmentation should be adequate for most quantitative pulmonary structural and functional analysis. In cases where the pulmonary image analysis requires more accurate lobe segmentation, manually provided anchor points can be added into the automatic system to achieve satisfactory results.

It is not uncommon to find lobar fissures that are incomplete in some CT data sets [30]. This may be due to limited scanner resolution or the lack of a fully complete fissure in a particular subject. In this situation, our method extrapolates the fissure to the lung boundary based on shape and initialization information. For some applications this extrapolation might be useful, e.g., for lobar volume analyses, but for other situations this approach might be artificial and misleading.

In terms of processing time, the ridgeness calculation takes about 2 min and the fissure detection takes about 4–6 min, so the total computation time is about 6–8 min for a single lung. The computation time assumes a personal computer with a 3.5 GHz CPU and 2 GB RAM. It takes about 15 min to manually trace the lobar fissures on approximately every fifth slice for a single lung, so a complete segmentation of the major fissure in one lung would take up to 75 min. Therefore, the automatic method delineates fissures faster than the image analysts.

The main result of this work is a method to segment the 3-D lobar fissure surfaces, so it is natural to consider other approaches, such as a deformable model or an active shape model (ASM). For fissure detection it is difficult to find a close initialization of the fissure surfaces, either from a mathematical model or an anatomic atlas. Further, it is well known that the success of deformable models and ASM's generally require fairly close initialization for low-contrast objects. Therefore, although the 3-D model-based segmentation methods may be able to preserve 3-D smoothness and connectivity during lobar fissure detection, the final results may very well be far from the true location due to low contrast of lobar fissures, pulmonary CT image noise and artifacts, and the distraction from other bright structures in lungs, e.g., vessels and scar tissue. In our proposed method, after the initial detection, the reliable fissures are selected to initialize the final detection on neighboring slices. Due to the smoothness of fissure surface, the fissures on neighboring slices are spatially close and similar to each other. Therefore, the fissures from the neighboring slices can provide a very close initialization. Although our proposed method may lose 3-D smoothness occasionally, the overall segmentation results should be close to the true fissure location, providing accurate lobe segmentation, as shown by the validation results.

B. Limitations of the Automatic Method

The subjects listed in Table II have normal or nearly normal anatomy. The automatic method should be suitable for screening applications for detecting early lung disease in patients with nearly normal anatomy, but it has not been extensively tested on subjects with abnormal anatomy or severe parenchymal disease. The method is based on the following assumptions: 1) the input CT images are complete scans covering the entire lung with contiguous slices; 2) the fissure search with the atlas initialization can always find enough reliable results so that the final

fissure search succeeds; and 3) the fissure contrast is sufficient for the graph search to succeed. Incomplete scans and scans with noncontiguous slices cannot be easily initialized by the atlas because the registration from the atlas to new data sets is based on the lung boundaries. Initialization may also fail on subjects with very unusual anatomy or with disease that disrupts the normal lobar architecture. In images for which the fissure contrast is very low, for example, in very fibrotic tissue or for normal tissue in thick slice data, the fissure search for the atlas initialization may not be able find even one reliable initialization result, or the detected reliable fissure initializations may be insufficient to make the subsequent fissure segmentation robust and accurate. In Fig. 11, a typical transverse slice from a thick-slice (3-mm slice thickness) CT image is shown on the left, compared with a slice from a thin-slice CT image (1.2-mm slice thickness) on the right. In Fig. 11, the fissure on the 3-mm-thick slice is much less distinguishable and, therefore, much more difficult to be recognized by fully automated algorithms.

C. Performance Improved by Anchor Points

Occasional segmentation errors exist in several of the automatic results, as indicated by the relatively large RMS error. Anchor points were provided manually for these test cases on two or four slices on which the segmentation errors are large. It takes about 2–3 min to put anchor points on two or four slices (two or three anchors per slice) for one lung. The computer then detects fissures with the manually provided anchor points. The results are compared with manual tracings and fully automatic results in Fig. 12. The total number of the slices with fissures in one lung for the test case shown in Fig. 12 ranges from about 160 to 282 slices. From Fig. 12, we can see that the occasional segmentation errors can be easily corrected by minor manual interaction.

D. Horizontal Fissure Detection by Anchor Points

The atlas-driven automatic method cannot detect the horizontal fissure. For this reason, we did not construct a horizontal fissure atlas, but instead use anchor points are provided on some key slices to detect the horizontal fissures for cases where the horizontal fissure is visible. The horizontal fissures are more distinguishable on sagittal views than transverse views since the horizontal fissures are normally oriented parallel or nearly parallel to the transverse slices. Therefore, unlike the oblique fissure detection, which are oriented more perpendicular to the transverse slices and thus more distinguishable on transverse views, the delineation of the horizontal fissures is performed on sagittal views rather than transverse views.

Fig. 13 shows horizontal fissure delineations obtained using the anchor point initializations. Note that part of the oblique fissure is also detected in Fig. 13 to completely segment the lung. Fig. 14 gives four different views of the 3-D visualization for a complete lung lobe segmentation, with both the oblique fissures and the horizontal fissure.

E. Initialization With Other Atlases

For high-resolution and high-quality CT images, the quality of the atlas used for initialization should not affect the final segmentation results appreciably. However, for the images in which the boundaries of anatomic structures are not so well defined, a better atlas is expected to give more robust and accurate results.
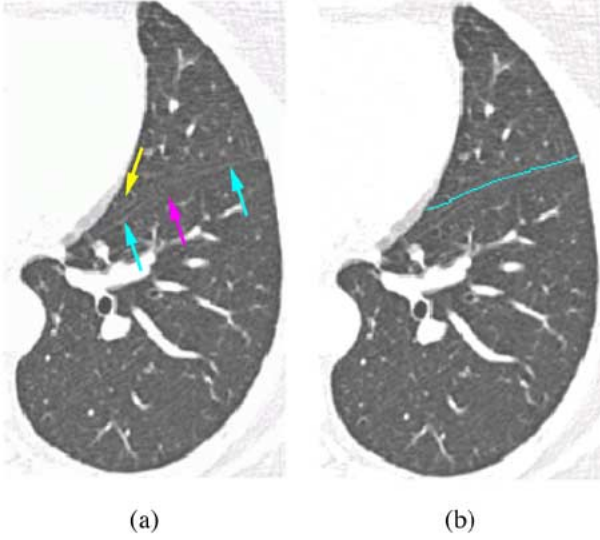
(a)   (b)

Fig. 10. A typical fissure detection error. (a) a original CT image with a blurred fissure (pointed by the cyan arrow) close to the lung boundary. The fissure is too blurred to distinguish in the middle part (pointed by the arrow magenta). Meanwhile, the nearby ridge-like influential structures are presented above the fissure (pointed by the yellow arrow); (b) automatic fissure detection result.

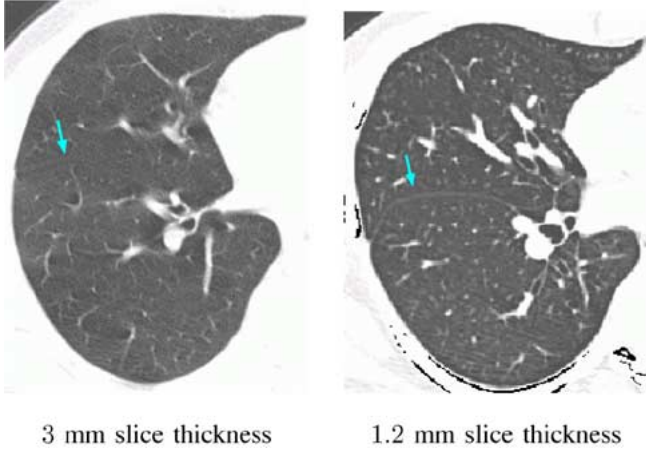3 mm slice thickness   1.2 mm slice thickness

Fig. 11. Comparison of fissure contrast. The fissures are marked by the arrows. The fissures on 3-mm-thick slices are much less distinguishable.

For example, as new fully automatic, robust, and accurate techniques for airway tree and vessel tree segmentation and registration continue to be developed, the additional structural information from the airways and vessels can be utilized to build a better anatomic atlas that more accurately represents the relative position of the fissures within the surrounding lung anatomy. Moreover, if additional subjects can be investigated and classified into groups, for example, based on disease type or body geometry, the group average atlas should be able to provide closer initial fissure positions than the generalized average atlas used in the current implementation.

In our work, the only link between the atlas construction and fissure detection is the deformed average fissure surface and shape variation. Therefore, if an improved atlas can provide the average fissure location, the associated shape variation, and an appropriate registration technique to deform the average atlas fissure onto new data sets, it would be easy to substitute an improved atlas into the current fissure detection system (see Fig. 2).
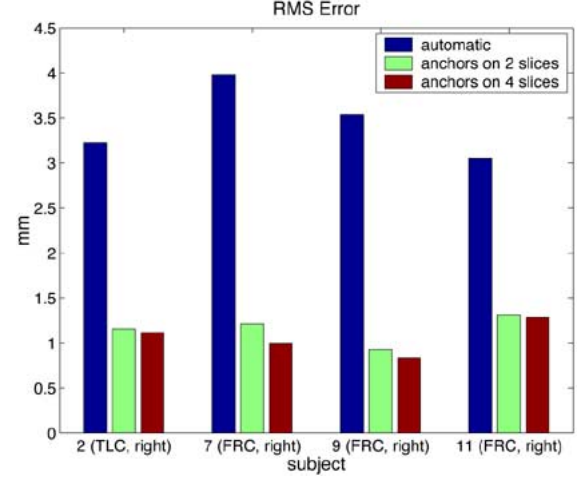

Fig. 12. Performance improved by anchor points: RMS error is reduced by introducing subject-specific anatomical information provided by a human expert.

(a)   (b)
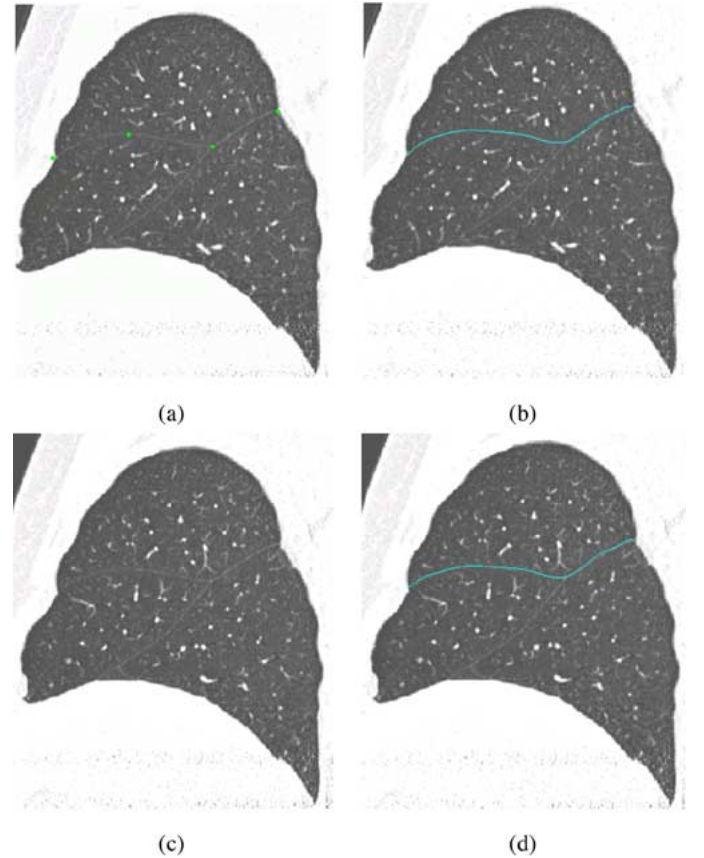
(c)   (d)

Fig. 13. Horizontal fissure delineation (combined with the oblique fissure). (a) Original CT image with anchor points (marked by green dots) on sagittal views; (b) Horizontal fissure delineation for (a); (c) Original CT image without anchor points; (d) Horizontal fissure delineation for (c).

V. CONCLUSION

An automatic lobe segmentation framework with atlas initialization has been proposed in this work. The validation results show that some occasional errors exist due to poor image quality, but in most cases the automatic method can delineate fissures close to those defined by manual tracings. Occasional errors occur near lung boundaries, but their influence on the
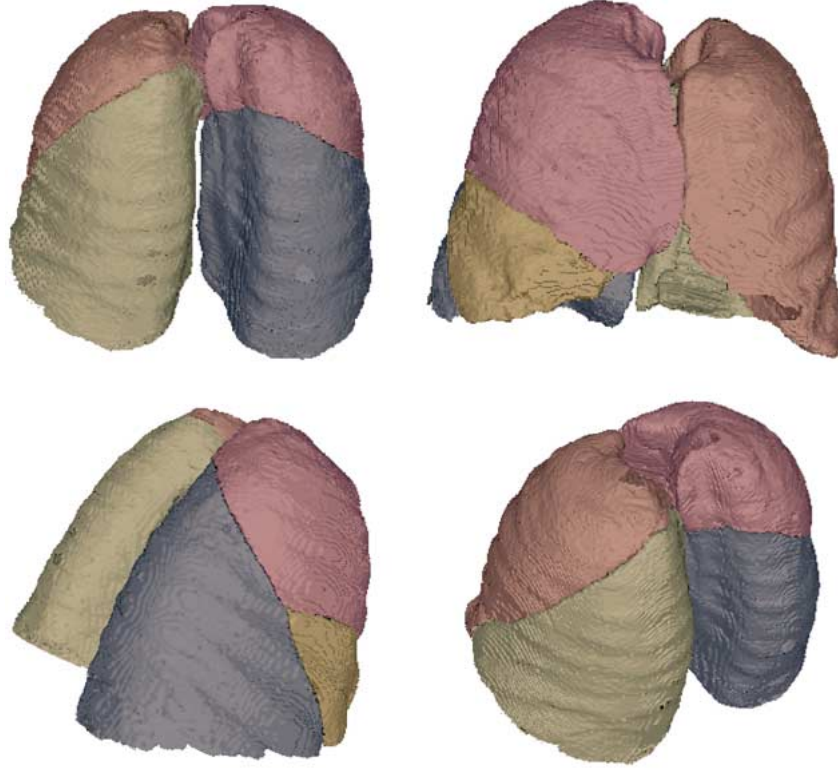
Fig. 14.   Four different views of the 3-D visualization for a complete lung lobe segmentation, with both the oblique fissures and the horizontal fissure.
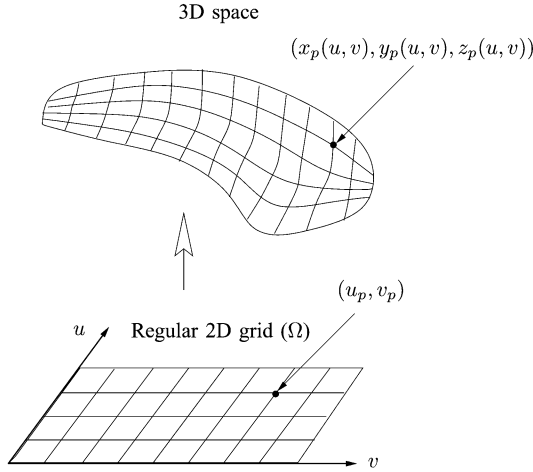

Fig. 15.   Fissure surface parameterization. Parameterization of a 3-D surface is analogous to stretching and distorting a rubber sheet with a plane regular grid $\Omega(u, v)$ onto the 3-D surface. Once the rubber sheet has been exactly fit onto the 3-D surface, the distorted 2-D grid provides a different indexing system (u, v) other than the original Cartesian coordinate indexing (x, y, z).

overall lobe segmentation is small. The agreement between the automatic and manual lobe segmentations is strong, as suggested by a similarity measure. Thus the automatic results should be accurate enough for most computer-based pulmonary quantitative analysis. Further, the lobe segmentation system described here is flexible, and minor manual intervention can be introduced to improve the method performance for cases where the application requires more accurate lobe segmentations.

The lobar segmentation can be the first step in a complete decomposition of the lung into its major anatomic components, and can be used to build a normal lung atlas, which can then be used to report CT image-based quantitative measurements on a whole lung or regional basis.

APPENDIX

A. Three-Dimensional Fissure Surface Parameterization

A (u, v) parameterization of a 3-D surface is analogous to stretching and distorting a planar rubber sheet with a regular grid onto the 3-D surface. Once the rubber sheet has been exactly fit onto the 3-D surface, the distorted 2-D grid provides a different indexing system than the original Cartesian coordinate indexing (see Fig. 15).

To do mesh parameterization, we re-arrange the 3-D lobar fissure surface $F(x, y, z)$ into stacked parallel open in-plane curves C_z indexed by z coordinates. Then, the generation of a mesh $M = \{\mathbf{p}(u, v)\}$, where $\mathbf{p}(u, v)$ is a mesh grid point, of size $I \times J$ for the 3-D fissure surface $F(x, y, z)$, can be done in two main steps. First, the curve C_z is re-sampled into I points with equal arc lengths so that C_z is indexed by the first grid parameter u, where $0 \leq u < I$. Points on each in-plane curve C_z with the same index u consist a new family of curves $\{D_u\}$. Similarly, curve D_u is re-sampled into J points spacing with equal arc length (see Fig. 16). The x, y, z position of a grid point is calculated from the two original points whose arc lengths are closest to the required arc length of the grid point by linear interpolation. For example, in Fig. 16, the position of the grid point $\mathbf{p}_z(u)$ can be calculated from the original points $\mathbf{q}_a$ and $\mathbf{q}_b$ by

$$\mathbf{p}_z(u) = \mathbf{q}_a + \mu_C(\mathbf{q}_b - \mathbf{q}_b), \tag{A-1}$$

where the linear factor μ_C is assigned based on the arc length. Let $l(C_z)$ denote the total arc length of the curve C_z, $l(\mathbf{q}_a)$ and
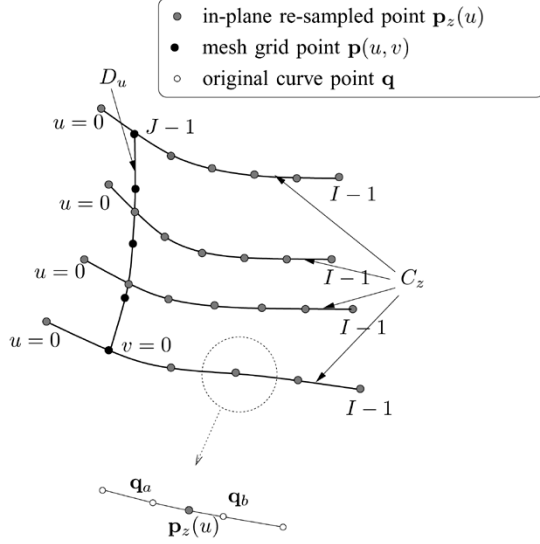
Fig. 16. Calculation of surface parameterization. Surface parameterization with regular grid (u, v) of size $(I \times J)$ can be accomplished with two curve arc length parameterization. The in-plane curve C_z is first re-sampled into I points with equal arc lengths so that C_z is indexed by the first grid parameter u, where $0 \leq u < I$. Points on each in-plane curve C_z with the same index u consist a new family of curves $\{D_u\}$. Similarly, curve D_u is re-sampled into J points spacing with equal arc length. The position of the grid point $\mathbf{p}_z(u)$ can be calculated from the original points $\mathbf{q}_a$ and $\mathbf{q}_b$ by arc length-based linear interpolation.

$l(\mathbf{q}_b)$ denote the arc lengths of $\mathbf{q}_a$ and $\mathbf{q}_b$, respectively. Then, we have

$$\mu_C = \frac{(u/I)l(C_z) - l(\mathbf{q}_a)}{l(\mathbf{q}_b) - l(\mathbf{q}_a)}. \quad (A-2)$$

B. Probabilistic Analysis for Atlas Accommodation

The directional variation of the displacement map $\{\vec{D}_i(u, v) = \mathbf{p}_i(\mathbf{u}, \mathbf{v}) - \mathbf{p}_\mu(\mathbf{u}, \mathbf{v})\}$ on x, y, and z axis can be represented by its local covariance matrix $\mathbf{V}(u, v) = [\sigma_{mn}^2(u, v)]$, where m and n are matrix indices and $1 \leq m, n \leq 3$, that is

$$\mathbf{V}(u, v) = \begin{bmatrix} \sigma_{xx}^2(u, v) & \sigma_{xy}^2(u, v) & \sigma_{xz}^2(u, v) \\ \sigma_{xy}^2(u, v) & \sigma_{yy}^2(u, v) & \sigma_{yz}^2(u, v) \\ \sigma_{xz}^2(u, v) & \sigma_{yz}^2(u, v) & \sigma_{zz}^2(u, v) \end{bmatrix}. \quad (B-1)$$

$\vec{D}_i(u, v)$ can be considered as a random variable with zero mean in the local probability distribution, but the x, y, and z components of $\vec{D}_i(u, v)$ do not necessarily have the same variability. In general, x, y, and z are weakly correlated since the nondiagonal elements in the covariance matrix $\mathbf{V}(u, v)$ are normally far less than the diagonal elements [31].

We assume that a point $\mathbf{p}(u, v)$ on a mesh $M = \{\mathbf{p}(u, v)\}$, which is chosen from the same population as the fissure surfaces in the family $\mathcal{F}$, is a 3-element random variable in a multivariate

normal distribution with mean vector $\mathbf{p}_\mu(u, v)$ and covariance matrix $\mathbf{V}(u, v)$, i.e., $\mathbf{p}(u, v) \sim N_3(\mathbf{p}_\mu(u, v), \mathbf{V}(u, v))$. Therefore, the local PDF for the displacement $\vec{D}(u, v)$ at a grid node (u, v) can be written as shown in (B-2) at the bottom of the page [32]. Equation (B-2) shows that the local pdf $g(\vec{D}(u, v))$ depends on the local covariance matrix $\mathbf{V}(u, v)$. As the covariance matrix $\mathbf{V}(u, v)$ differs for each grid node (u, v), the local distribution also differs for individual grid nodes on the surface meshes.

Using the local PDF $g(\vec{D}(u, v))$, a 3-D spatial anisotropic confidence region of the displacement vector $\vec{D}(u, v)$ can be calculated. A $q\%$ confidence region for a grid node (u, v) means that $q\%$ of the homologous points $\mathbf{p}_i(u, v)$ in the family $\mathcal{F} = \{M_i(u, v), 0 \leq i < n\}$ fall in the confidence region, which is symmetrically defined around the average position $\mathbf{p}_\mu(u, v)$.

Assume that the x, y, and z components of the displacement vectors $\vec{D}(u, v)$ are mutually independent, then the quadratic form in the exponent in (B-2) $Q = [\vec{D}(u, v)]^T[\mathbf{V}(u, v)]^{-1}[\vec{D}(u, v)]$ is a three-degree of freedom chi-square distribution $\chi^2(3)$ [32]. The quadratic form Q is the square of Mahalanobis distance from a point $\mathbf{p}(u, v)$ on a general mesh to a point $\mathbf{p}_\mu(u, v)$ on the average mesh, representing the local variation. Then, the $q\%$ confidence region, or equivalently the $100(1 - \alpha)\%$ confidence region, where α denotes the significance level, for the 3-D position of a surface mesh node (u, v) can be given by the equation

$$\Pr\{[\vec{D}(u, v)]^T[\mathbf{V}(u, v)]^{-1}[\vec{D}(u, v)] \leq \xi_\alpha\} = 1 - \alpha, \quad (B-3)$$

where ξ_α is the critical value of the $\chi^2(3)$, and defined as $\Pr\{X_{\chi^2(3)} \geq \xi_\alpha\} = \alpha$, and where $X_{\chi^2(3)}$ is a random variable in the $\chi^2(3)$ distribution. A complete mathematical proof of (B-3) can be found in [31].

For a new data set, a probabilistic analysis from the atlas assigns a probability value to points on the fissure surface $M_{\text{new}} = \{\mathbf{p}_{\text{new}}(u, v)\}$ in the new data. Let $\vec{D}_{\text{new}}(u, v)$ denote the displacement vector from a point on the average mesh to the corresponding point on the new subject fissure. Then the probability for $\mathbf{p}_{\text{new}}(u, v)$ is [31]

$$\text{probability}[\mathbf{p}_{\text{new}}(u, v)]$$
$$= \Pr\left\{ X_{\chi^2(3)} \geq [\vec{D}_{\text{new}}(u, v)]^T[\mathbf{V}(u, v)]^{-1}[\vec{D}_{\text{new}}(u, v)] \right\}$$
$$(B-4)$$

where $X_{\chi^2(3)}$ is the random variable in $\chi^2(3)$. The probability of $\mathbf{p}_{\text{new}}(u, v)$ can be interpreted in the context of the anatomic atlas. For example, in Fig. 17, the probability of the point $\mathbf{p}_A(u, v)$, which is from subject A, is large because of the small Mahalanobis distance, meaning the atlas represents the fissure of subject A well locally at the node (u, v); whereas the probability of the point $\mathbf{p}_B(u, v)$, which is from subject B, is small, so the dispersion of the fissure B is large at the node

$$g[\vec{D}(u, v)] = \frac{1}{(2\pi)^{\frac{3}{2}}|\mathbf{V}|^{\frac{1}{2}}} \exp\left\{ -\frac{[\mathbf{p}(u, v) - \mathbf{p}_\mu(u, v)]^T[\mathbf{V}(u, v)]^{-1}[\mathbf{p}(u, v) - \mathbf{p}_\mu(u, v)]}{2} \right\}$$

$$= \frac{1}{(2\pi)^{\frac{3}{2}}|\mathbf{V}|^{\frac{1}{2}}} \exp\left\{ -\frac{[\vec{D}(u, v)]^T[\mathbf{V}(u, v)]^{-1}[\vec{D}(u, v)]}{2} \right\} \quad (B-2)$$

local PDF of Q at node (u, v) in the atlas

$$\text{square of Mahalanobis distance}$$
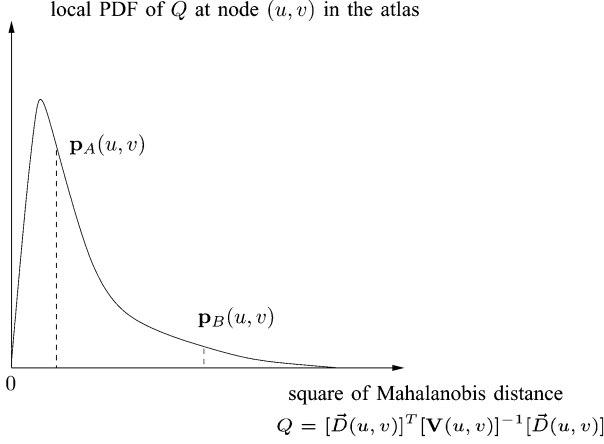$$Q = [\vec{D}(u,v)]^T [\mathbf{V}(u,v)]^{-1} [\vec{D}(u,v)]$$

Fig. 17. Local probability analysis in the atlas. The probability of the point $\mathbf{p}_A(u,v)$, which is from subject A, is large because of the small Mahalanobis distance, meaning the atlas represents the fissure of subject A well locally at the node (u,v); whereas the probability of the point $\mathbf{p}_B(u,v)$, which is from subject B, is small, so the dispersion of the fissure B is large at the node (u,v), and, therefore, is unlikely to be well represented by the atlas.

(u,v) and, therefore, is unlikely to be well represented by the atlas.

An analysis of the mesh position probabilities can be used to determine the likelihood that a new subject is from the same population on which the atlas is based on, or stated differently, whether the shape information encoded in the atlas represents the shape variation present in the new data. In our work, we assume that the training data and testing data are all from the same population, and the atlas, built from the training data, accommodates the anatomic variability of different subjects in this population. The probabilistic analysis between the atlas and new data sets provides a way to test how well this assumption fits reality.

C. MLSEC-ST-Based Ridgeness Measure

The MLSEC operators obtain the measure of curvatures through the relationship between the local curvature and the parallelism of the normal vectors of the curve. The parallelism of a vector field can be measured by the divergence operator $\text{div}(\cdot)$, therefore, the curvature κ_d of a d-dimensional level set can be computed by the divergence of the gradient vector field $\mathbf{w}$ as follows (proven by López *et al.* in [24]). Let $\mathbf{0}_d$ denote the d-dimensional zeros vector, then the gradient vector field $\mathbf{w} = [L_{x^1}, \ldots, L_{x^d}]^t$ of a d-dimensional image L can be normalized as

$$\bar{\mathbf{w}} = \begin{cases} \mathbf{w}/\|\mathbf{w}\| & \text{if } \|\mathbf{w}\| > 0 \\ \mathbf{0}_d & \text{if } \|\mathbf{w}\| = 0 \end{cases} \qquad \text{(C-1)}$$

and

$$\kappa_d = -\text{div}(\bar{\mathbf{w}}). \qquad \text{(C-2)}$$

Calculation of $\text{div}(\bar{\mathbf{w}})$ is digitalized in [24] with 4-connectivity neighborhoods for 2-D images. The 2-D MLSEC $\bar{\kappa}_2$ can be given as follows:

$$\bar{\kappa}_2[i,j] = -\frac{1}{2}(\bar{w}^1[i+1,j] \\ -\bar{w}^1[i-1,j] + \bar{w}^2[i,j+1] - \bar{w}^2[i,j-1]), \qquad \text{(C-3)}$$

where $\bar{w}^1$ and $\bar{w}^2$ are the first and second components of $\bar{\mathbf{w}}$ in 2-D.

The MLSEC operators can be further improved by modifying the gradient vector $\bar{\mathbf{w}}$ to the dominant gradient vector $\tilde{\mathbf{w}}$ with the *structure tensor* before the calculation of the divergence. The structure tensor enhanced operator is called MLSEC-ST in [24], [33], and can be defined by

$$\tilde{\kappa}_d = -\text{div}(\tilde{\mathbf{w}}). \qquad \text{(C-4)}$$

The dominant gradient vector $\tilde{\mathbf{w}}$ at point $\mathbf{x}$ can be calculated by

$$\tilde{\mathbf{w}}(\mathbf{x}; \sigma) = \text{sign}\left[\mathbf{w}_\Lambda^t(\mathbf{x}; \sigma) \cdot \mathbf{w}(\mathbf{x})\right] \mathbf{w}_\Lambda(\mathbf{x}; \sigma). \qquad \text{(C-5)}$$

where $\mathbf{w}(\mathbf{x})$ is the gradient vector at $\mathbf{x}$, and $\mathbf{w}_\Lambda(\mathbf{x}; \sigma)$ denotes the eigenvector corresponding to the largest eigenvalue of the structure tensor $\mathbf{S}(\mathbf{x}; \sigma)$. The structure tensor $\mathbf{S}(\mathbf{x}; \sigma)$ are defined as

$$\mathbf{S}(\mathbf{x}; \sigma) = G(\mathbf{x}; \sigma) * [\mathbf{w}(\mathbf{x}) \cdot \mathbf{w}^t(\mathbf{x})], \qquad \text{(C-6)}$$

where $G(\mathbf{x}; \sigma)$ is a d-dimensional Gaussian function with standard deviation σ and "$\cdot$" indicates matrix multiplication. The variable σ is the integration scale and controls the size of the neighborhood in which the structure tensor is defined. The convolution "$*$" is element-wise. $\tilde{\kappa}_d$ can then be calculated by (C-3) by substituting $\bar{\mathbf{w}}$ with $\tilde{\mathbf{w}}$.

A multiplication factor $\mathsf{C} \in [0,1]$ is introduced to suppress the curvature measure in isotropic regions [24], [33]. C, which is defined through the eigenvalues of $\mathbf{S}(\mathbf{x}; \sigma)$ and approaches 0 in homogeneous regions, modifies the MLSEC-ST operator to $\mathsf{C}\tilde{\kappa}_d$. The calculation of C can be found in [24] and [33].

REFERENCES

[1] R. D. Hubmayr and S. S. Margulies, "Regional ventilation in statically and dynamically hyperinflated dogs," *J. Appl. Physiol.*, vol. 81, no. 4, pp. 1815–1821, Oct. 1996.

[2] J. R. Rodarte, R. D. Hubmayr, D. Stamenovic, and B. J. Walters, "Regional lung strain in dogs during deflation from total lung capacity," *J. Appl. Physiol.*, vol. 58, no. 1, pp. 164–172, Jan. 1985.

[3] J. R. Rodarte, M. Chaniotakis, and T. A. Wilson, "Variability of parenchymal expansion measured by computed tomography," *J. Appl. Physiol.*, vol. 67, no. 1, pp. 226–231, Jul. 1989.

[4] E. A. Hoffman and E. L. Ritman, "Effect of body orientation on regional lung expansion in dog and sloth," *J. Appl. Physiol.*, vol. 59, no. 2, pp. 481–491, 1985.

[5] E. A. Hoffman and G. McLennan, "Assessment of the pulmonary structure—Function relationship and clinical outcomes measures: Quantitative volumetric CT of the lung," *Acad. Radiol.*, vol. 4, no. 11, pp. 758–776, 1997.

[6] B. Li, G. E. Christensen, G. McLennan, E. A. Hoffman, and J. M. Reinhardt, "Establishing a normative atlas of the human lung: Inter-subject warping and registration of volumetric CT," *Acad. Radiol.*, vol. 10, no. 3, pp. 255–265, Mar. 2003.

[7] National Emphysema Treatment Trial Research Group, "A randomized trial comparing lung-volume reduction surgery with medical therapy for severe emphysema," *New Engl. J. Med.*, vol. 348, no. 21, pp. 2059–2073, May 22, 2003.

[8] J.-M. Kuhnigk, H. K. Hahn, M. Hindennach, V. Dicken, S. Krass, and H.-O. Peitgen, "Lung lobe segmentation by anatomy-guided 3-D watershed transform," *Proc. SPIE (Medical Imaging)*, vol. 5032, pp. 1482–1490, 2003.

[9] M. Kubo, N. Niki, S. Nakagawa, K. Eguchi, M. Kaneko, N. Moriyama, H. Omatsu, R. Kakinuma, and N. Yamaguchi, "Extraction algorithm of pulmonary fissures from thin-section CT images based on linear feature detector method," *IEEE Trans. Nucl. Sci.*, vol. 46, no. 6, pp. 2128–2133, Dec. 1999.

[10] M. Kubo, Y. Kawata, N. Niki, K. Eguchi, H. Ohmatsu, R. Kakinuma, M. Kaneko, M. Kusumoto, N. Moriyama, K. Mori, and H. Nishiyama, "Automatic extraction of pulmonary fissures from multidetector-row CT images," in *Proc. IEEE Int. Conf. Image Processing (ICIP'01)*, 2001, pp. 1091–1094.

[11] M. Kubo, N. Niki, K. Eguchi, M. Kaneko, and M. Kusumoto, "Extraction of pulmonary fissures from thin-section CT images using calculation of surface-curvatures and morphology filters," in *Proc. IEEE Int. Conf. Image Processing (ICIP'00)*, 2000, pp. 637–640.

[12] X. Zhou, T. Hayashi, T. Hara, H. Fujita, R. Yokoyama, T. Kiryu, and H. Hoshi, "Automatic recognition of lung lobes and fissures from multi-slice CT images," *Proc. SPIE (Medical Imaging)*, vol. 5370, pp. 1629–1633, 2004.

[13] S. Saita, M. Yasutomo, M. Kubo, Y. Kawata, N. Niki, K. Eguchi, H. Ohmatsu, R. Kakinuma, M. Kaneko, M. Kusumoto, N. Moriyama, and M. Sasagawa, "An extraction algorithm of pulmonary fissures from multi-slice CT image," *Proc. SPIE (Medical Imaging)*, vol. 5370, pp. 1590–1597, 2004.

[14] L. Zhang, E. A. Hoffman, and J. M. Reinhardt, "Lung lobe segmentation by graph search with 3-D shape constraints," in *Proc. SPIE (Medical Imaging)*, C.-T. Chen and A. V. Clough, Eds., 2001, vol. 4321, pp. 204–215.

[15] L. Zhang and J. M. Reinhardt, "Detection of lung lobar fissures using fuzzy logic," in *Proc. SPIE (Medical Imaging)*, C.-T. Chen and A. V. Clough, Eds., 1999, vol. 3660, pp. 188–199.

[16] L. Zhang, E. A. Hoffman, and J. M. Reinhardt, "Atlas-driven lung lobe segmentation in volumetric X-ray CT images," *Proc. SPIE (Medical Imaging)*, vol. 5032, pp. 309–319, 2003.

[17] M. S. Brown, M. F. McNitt-Gray, N. J. Mankovich, J. G. Goldin, J. Hiller, L. S. Wilson, and D. R. Aberle, "Method for segmenting chest CT image data using an anatomical model: Preliminary results," *IEEE Trans. Med. Imag.*, vol. 16, no. 6, pp. 828–839, Dec. 1997.

[18] S. Hu, E. A. Hoffman, and J. M. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images," *IEEE Trans. Med. Imag.*, vol. 20, no. 6, pp. 490–498, Jun. 2001.

[19] L. Zhang and J. M. Reinhardt, "3-D pulmonary CT image registration with a standard lung atlas," in *Proc. SPIE (Medical Imaging)*, C.-T. Chen and A. V. Clough, Eds., 2000, vol. 3978, pp. 67–77.

[20] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis and Machine Vision*. Pacific Grove, CA: Brooks/Cole, 1998.

[21] D. J. Burr, "A dynamic model for image registration," *Comput. Graphics Image Process.*, vol. 15, pp. 102–112, 1981.

[22] P. M. Thompson, J. Moussai, S. Zohoori, A. Goldkorn, A. A. Khan, M. S. Mega, G. W. S. J. L. Cummings, and A. W. Toga, "Cortical variability and asymmetry in normal aging and Alzheimer's disease," *Cereb. Cortex*, vol. 8, no. 6, pp. 492–509, 1998.

[23] L. Zhang, "Atlas-Driven Lung Lobe Segmentation in volumetric X-Ray CT images," Ph.D. thesis, Univ. Iowa, Iowa City, Dec. 2002.

[24] A. M. López, D. Lloret, J. Serrat, and J. J. Villanueva, "Multilocal creaseness based on the level-set extrinsic curvature," *Comput. Vis. Image Understanding*, vol. 77, pp. 111–144, 2000.

[25] J. A. Maintz, P. A. van den Elsen, and M. A. Viergever, "Evaluation of ridge seeking operators for multimodality medical image matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 4, pp. 353–365, Apr. 1996.

[26] W. Park, E. A. Hoffman, and M. Sonka, "Segmentation of intrathoracic airway trees: A fuzzy logic approach," *IEEE Trans. Med. Imag.*, vol. 17, no. 4, pp. 489–497, Aug. 1998.

[27] B. M. Dawant, S. L. Hartmann, J.-P. Thirion, F. Maes, D. Vandermeulen, and P. Demaerel, "Automatic 3-D segmentation of internal structures of the head in MR images using a combination of similarity and free-form transformation: Part I, methodology and validation on normal subjects," *IEEE Trans. Med. Imag.*, vol. 18, no. 10, pp. 909–916, Oct. 1999.

[28] M. S. Atkins and B. T. Mackiewich, "Fully automatic segmentation of the brain in MRI," *IEEE Trans. Med. Imag.*, vol. 17, no. 1, pp. 98–107, Feb. 1998.

[29] A. P. Zijdenbos, B. M. Dawant, R. A. Margolin, and A. C. Palmer, "Morphometric analysis of white matter lesions in MR images: Method and validation," *IEEE Trans. Med. Imag.*, vol. 13, no. 4, pp. 716–724, Dec. 1994.

[30] K. Hayashi, A. Aziz, K. Ashizawa, H. Hayashi, K. Nagaoki, and H. Otsuji, "Radiographic and CT appearances of the major fissures," *Radiographics*, vol. 21, no. 4, pp. 861–874, Jul.–Aug. 2001.

[31] P. M. Thompson, C. Schwartz, and A. W. Toga, "High-resolution random mesh algorithms for creating a probabilistic 3-D surface atlas of the human brain," *NeuroImage*, vol. 3, pp. 19–34, 1996.

[32] R. V. Hogg and A. T. Craig, *Introduction to Mathematical Statistics*. Upper Saddle River, NJ: Prentice-Hall, 1995.

[33] A. M. López, F. Lumbreras, J. Serrat, and J. J. Villanueva, "Evaluation of methods for ridge and valley detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 4, pp. 327–334, Apr. 1999.