Published in final edited form as:

Phys Med Biol. 2012 August 07; 57(15): 4905–4930. doi:10.1088/0031-9155/57/15/4905.

Tracking the Motion Trajectories of Junction Structures in 4D CT Images of the Lung

Guanglei Xiong[1,2], **Chuangzhen Chen**[2,3], **Jianzhou Chen**[2,3], **Yaoqin Xie**[4], **Lei Xing**[1,2]

[1]Biomedical Informatics Program, Stanford University, Stanford, CA 94305, USA

[2]Department of Radiation Oncology, Stanford University, Stanford, CA 94305, USA

[3]Department of Radiation Oncology, Tumor Hospital, Shantou University Medical College, Guangdong 515031, China

[4]Key Laboratory for Health Informatics, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China

Abstract

Respiratory motion poses a major challenge in lung radiotherapy. Based on 4D CT images, a variety of intensity-based deformable registration techniques have been proposed to study the pulmonary motion. However, the accuracy achievable with these approaches can be sub-optimal because the deformation is defined globally in space. Therefore, accuracy of the alignment of local structures may be compromised. In this work, we propose a novel method to detect a large collection of natural junction structures in the lung and use them as the reliable markers to track the lung motion. Specifically, detection of the junction centers and sizes is achieved by analysis of local shape profiles on one segmented image. To track the temporal trajectory of a junction, the image intensities within a small region of interest surrounding the center is selected as its signature. Under the assumption of the cyclic motion, we describe the trajectory by a closed B-spline curve and search for the control points by maximizing a metric of combined correlation coefficients. Local extrema are suppressed by improving the initial conditions using random walks from pair-wise optimizations. Several descriptors are introduced to analyze the motion trajectories. Our method was applied to thirteen real 4D CT images. More than 700 junctions in each case are detected with an average positive predictive value of greater than 90%. The average tracking error between automated and manual tracking is sub-voxel and smaller than the published results using the same set of data.

Keywords

Lung; 4D CT; radiotherapy; respiratory motion; junction detection; motion trajectory; motion tracking; registration

I. INTRODUCTION

Lung respiratory motion is a complex biomechanical process (Tustison et al. 2011). The lung is a soft and elastic organ, consisting of bronchi, bronchioles and alveolar ducts, surrounded by alveoli and blood vessels. The movement of pulmonary tissue is driven by

contraction and expansion of muscles in the diaphragm and chest wall. The motion is heterogeneous, i.e. the lung deforms spatially in different magnitude and direction. The deformation is also discontinuous, manifested as relative sliding between the lung and the chest wall, as well as between lobes. In addition, the motion is nonlinear and hysteretic, leading to distinct paths during inhalation and exhalation. Furthermore, physiologic motion patterns are patient-specific and even non-stationary.

In recent years, a number of 4D CT procedures have been developed by sorting the CT data to corresponding phases or amplitudes from multiple respiratory cycles according to simultaneously recorded internal or external respiratory signals (Ford et al. 2003; Low et al. 2003; Rietzel et al. 2005). The acquired images encode motion information with high spatial resolution and moderate temporal resolution. Respiratory motion is closely related to the function of the lung and a detailed knowledge of lung motion is thus valuable in diagnosis and prognosis by itself. Clinically, accurate motion information is also vital in radiotherapy to reduce healthy tissue irradiation while allowing target dose escalation (Xing et al. 2006). In radiation therapy treatment planning, for example, it is crucial to determine the motion extent and define appropriate margins to ensure adequate dose coverage. In the delivery stage, radiation beam is ideally gated or tracked with respective to the motion information derived from 4D CT data and online internal or external respiratory surrogates.

Deformable registration (Maintz and Viergever 1998) is the most common way to align deforming structures in 4D images and to derive the motion (Li et al. 2008; Xing et al. 2009; Xing et al. 2007). In this approach, homologous correspondences are established by satisfying optimal similarity metrics under proper regularization through explicit iterative processes or by analytical constructions. For deformable lung registration, there are two major categories in terms of the definition of the similarity: One is intensity-based, which aims at diminishing the mismatch of dense image intensities. The metrics widely used are the mean squared difference (MSD), the correlation coefficient (CC), and the mutual information (MI). A metric of sum squared tissue volume difference was also developed (Yin et al. 2011; Yin et al. 2009) to improve lung registration in regions lack of structural landmarks by particularly considering intensity changes during breathing. Depending on the applications, the metric can be defined over the whole image domain or in a region of interest (Chao et al. 2008; Chao et al. 2007). Lung deformation can be defined by either parametric or nonparametric transformations. The former approach parameterizes the spatial transform with a set of parameters and casts registration as a problem of optimization of the chosen metric. Parameterization with B-splines is common (McClelland et al. 2006; Rietzel and Chen 2006; Schreibmann and Xing 2006), while others, e.g. Fourier series (Amit 1994), piecewise affine (Collins and Evans 1997; Heath et al. 2007), and polyaffine (Arsigny et al. 2005), are also proposed. On the other hand, the nonparametric approach defines the transform as a dense vector field and search the solution of the registration problem by solving a physical process, such as elastic solid (Davatzikos 1997), fluid flow (Foskey et al. 2005), optical flow (Dougherty et al. 2003; Guerrero et al. 2004; Horn and Schunck 1981), demons (Sarrut et al. 2006; Thirion 1998), biomechanics (Sundaram and Gee 2005) and several others.

Alternatively, landmark-based registration attempts to align the sparse paired landmarks, mostly points or sometimes curves or surfaces. Landmarks can be either feature-based or anatomy-based. The former identifies landmarks by searching for salient local intensity patterns, which may not be anatomically meaningful and thus hard to be interpreted and verified. Therefore, anatomy-based landmarks are often preferred when they are available and not difficult to be detected. Besides detection, a matching process must be followed to establish correspondences between landmarks on one image and another. Traditionally, this task is largely manual and thus is very subjective and error prone. In order to reduce manual procedures, computer-assisted and semi-automatic tools have been developed recently to successfully generate large numbers of paired landmarks (Murphy et al. 2011) for the sake of evaluation of registration performance. However, they are still relatively time-consuming and thus impractical for direct use in landmark-based registration, especially when considering many patients and each with 4D data. Finally, the deformations at the landmarks are extrapolated using thin plate spline (Bookstein 1989; Coselmon et al. 2004), elastic body spline (Davis et al. 1997; Kohlrausch et al. 2005) or moving least squares (Schaefer et al. 2006) to obtain a dense deformation field. In both categories of registration, extra regularization constraints are proposed to favor *a priori* properties of the deformation, e.g. inverse consistency and diffeomorphism (Beg et al. 2005; Vercauteren et al. 2009).

To date, a majority of previous studies on quantification and analysis of pulmonary motion from 4D image data adopted a pair-wise registration paradigm. Some only considered the deformation between two distinct phases, e.g. the EMPIRE10 study (Murphy et al. in press), in which many state-of-the-art registration algorithms were evaluated using two phases at opposing ends of the breathing cycle. Others (e.g. (Boldea et al. 2008; Xie et al. 2009)) considered a series of deformations among all phases by registering either each phase to a reference phase, between consecutive phase pairs, or more expensively among any combination of phase pairs. Although these approaches provide an immediate extension from addressing paired data to 4D data, they do not take into account the order of the images in time and the temporal continuity is not imposed. Followed with the successful efforts in cardiac motion analysis (e.g. (Ledesma-Carbayo et al. 2005; McEachen et al. 2000; Perperidis et al. 2005; Peyrat et al. 2010; Sundar et al. 2009)), a few recent studies on pulmonary motion analysis have shifted towards a new and more natural paradigm of simultaneously registering images from all phases by incorporating the temporal dimension into the optimization metric and deformation representation. Metz *et al.* (Metz et al. 2011) proposed a generic free-form deformation model with B-spline parameterizations for both spatial and temporal dimensions. Optionally, cyclic motion can be smoothly imposed on the temporal dimension. Similar to the group-wise registration techniques, the MSD-based similarity metric is defined with respect to an implicit reference image instead of a reference image. Parallel to their work, Vandemeulebroucke *et al.* (Vandemeulebroucke et al. 2011) proposed a similar technique but the cyclic motion is either continuously or smoothly imposed. The MSD-based metric is defined with respect to a chosen reference image, in which the deformation is set to be zero. In both studies, the user-selected landmarks and those extracted from the resulted 4D deformation model were compared for the purpose of evaluation. Compared to aforementioned approaches based on dense image intensities, Castillo *et al.* proposed to separately register individual voxels from a subset chosen on a

coarse grid. The trajectory of a voxel is parameterized by a cubic polynomial. It is then calculated by performing nonlinear least squares fit to a local compressible flow equation. Cyclic motion is not imposed since only the expiratory trajectories are considered. Finally, the dense deformation field is generated by moving least squares. Again, the user-selected landmarks were utilized for evaluation.

In this paper, we propose a novel method to detect a large collection of natural junction structures in the lung and use them as the reliable markers to track the lung motion. The current work represents an substantial extension (improved method, more data, and further validation) to our preliminary results in Xiong et al. 2011. Junction structures, such as vessel and bronchial bifurcations, are abundantly distributed in the lung and visually apparent in 4D CT images even without any contrast enhancement. They are the natural anatomy-based markers that can be reliably detected and verified. They may also serve as the noninvasive counterparts of the implanted fiducial markers (e.g. metal beads) for tumor tracking in image guided radiotherapy (Seppenwoolde et al. 2002). From the perspective of continuum mechanics, the junctions, as the material particles, follow the motion of the whole lung as a continuum, which realizes a Lagrangian description of the motion kinematics. We choose the landmark-based approach as opposed to the intensity-based approach partly because intensity information from both local and distant regions affects the local deformation in the latter approach, which may compromise the accuracy of aligning local structures. In addition, the common assumption of spatial smoothness regularizes the deformation but it contradicts the fact that the deformation is discontinuous close to the pleural surfaces and between lobes as mentioned. Instead of using the polynomial, we describe the motion trajectory of a junction by a B-spline space curve to avoid the oscillation effects due to the Runge's phenomenon (Runge 1901) if the order is too high. The cyclic motion is explicitly imposed. Specifically, detection of the junction centers and sizes is achieved by analysis of local shape profiles on one segmented image. To track the temporal trajectory of a junction, the image intensities within a small region of interest surrounding the center is selected as its signature. Under the assumption of the cyclic motion, we describe the trajectory by a closed B-spline curve and search for the control points by maximizing a metric of combined correlation coefficients. Local extrema are suppressed by improving the initial conditions using random walks from pair-wise optimizations. To analyze the motion, several descriptors are introduced to characterize a trajectory.

II. Methods

We start off by briefly describing our method for lung and vessel segmentation in Section II–A.1. Next, an insight on local shape profiles of junctions is made and utilized to construct an automated junction detection algorithm in Section II–A.2. In Section II–B.1, we present how to parameterize an individual trajectory and mention its properties relevant to tracking. Then in Section II–B.2, we describe the similarity metric that can be used to measure how well the intensities along the trajectory match. Next in Section II–B.3, we introduce the optimization strategy and a special technique to suppress local minima, if found although rarely. Finally in Section II–B.4, quantitative descriptors of tracked trajectories are introduced.

A. Lung Junction detection

A.1 Lung and Vessel Segmentation—This step is only performed on one reference image (e.g. at the maximal inhale phase) of a 4D CT dataset. Lung region is segmented using thresholding and morphological operations. The image is cropped to cover only the body region, which is denoted as I_1^0. Denoting Hounsfield unit as [H.U.], it is then thresholded between low −900 [H.U.] and high −500 [H.U.] intensity values, yielding I_1^1, which mainly contains the lung and part of the trachea close to its wall. To remove the trachea, I_1^0 is thresholded below −900 [H.U.] with a seeding point, yielding I_1^2. The seeding point is automatically detected by searching for the near-circular, air-filled region in the few slices of I_1^0 by noting that it is sufficient as long as it is inside the trachea. I_1^2 is then dilated using a sphere structuring element (SE) with a radius of 4 voxels, yielding I_1^3, which will include the whole trachea including the part close to the wall. To remove the trachea, I_1^1 is subtracted by I_1^3, yielding I_1^4, which contains the lung and a few small unwanted regions. To select the lung, I_1^4 is decomposed by an analysis of connected components. Usually, the two components with the largest volumes correspond to the left and right lungs, denoted as I_1^5. Sometimes, both lungs are connected because there is no clear boundary in certain spots between them. This case can be easily detected by comparing the ratio of the first and the second largest volume. No further separation is needed in our study. Since I_1^5 has some large concavities due to bronchi penetrating into the lung around the helium and holes due to lower intensities of bronchioles and higher intensities of vessels inside the lung, it is closed using a sphere SE with a radius of 20 voxels, yielding I_1^6. There may be small holes remaining in I_1^6, which are filled iteratively (Soille 2003), yielding I_1^7. Finally, I_1^7 is eroded using a sphere SE with a radius of 2 voxels to remove the banding region close to the chest wall, yielding I_1^8, which serves as the lung mask.

Using I_1^8, I_1^0 is masked to include only the lung region, yielding I_1^9. To segment the pulmonary blood vessels, we propose a simple yet effective algorithm. The major challenge for this task is the nonuniform background caused by the partial volume effects. It manifests as higher intensities of big vessels around the helium than those of distal smaller ones. Therefore, a global thresholding is not applicable. Instead, we choose an adaptive thresholding strategy within a range of ambiguous intensities. First, a baseline threshold is determined from the intensities of I_1^0 within I_1^8 using Otsu's method (Otsu 1979). That is, the voxels above baseline+25 [H.U.] are set to belong to vessels, while those below baseline-75 [H.U.] are set to belong to background. Any voxel in this range is decided based on the difference between its intensity and the Gaussian blurred intensity calculated in the neighboring region. If the difference is greater than 10 [H.U.], the voxel is considered to belong to a vessel. To address the issue of various sizes of the vessels in the lung, we test for multiple widths of the Gaussian kernel: 3, 4, and 5 voxels. If any test is positive, the voxel of interest is considered to belong to a vessel. The resultant binary image from adaptive thresholding is denoted as I_1^{10}, from which any disconnected component with less than 10

voxels is removed from further consideration, yielding I_1^{11}. Finally, the well-composed property (Latecki et al. 1995) is enforced to ensure that there is no critical and nonrealistic configuration in the vessel segmentation, yielding I_1^{12}. For a 3-D binary image, it is well-composed if and only if the set of points in the pixel boundaries between the foreground and the background forms a 2-D manifold. Certain critical and nonrealistic configurations circumvent well-composedness whenever there are two neighboring background or foreground pixels that are 8-connected but not 4-connected.

Note that all the parameters mentioned in this section are set empirically and fixed in all of our experiments. Our lung and vessel segmentation method may be replaced by other methods (e.g. those reviewed in (Sluimer et al. 2006)). But one must be careful to use Hessian-based vessel enhancement methods (Frangi et al. 1998). If not designed properly, they tend to alter vessel sizes and more importantly they may weaken junction structures of interest while enhancing tubular structures. In addition, junctions are not isolated spheres and cannot be simply treated as blob structures and detected using the existing blob detection methods.

A.2 Detection by Analyzing Local Shape Profile—Our goal is to search for the center and size (defined by the minimal distance from the center to the boundary) for a large collection of junctions in the segmented reference image. The basic idea for junction detection comes from an insight of the unique characteristics of the local shape profile of a junction. There are three or more disconnected components inside the spherical shell between two concentric spheres of appropriate sizes at the center of a junction, whereas two or less at other locations (e.g. near/off the center of a tube, near/off the center of an endpoint, and off the center of a junction). Fig. 1(a) illustrates six typical cases in a synthetic 2D example.

To analyze the local shape profiles, the maximal distance d of a ray traveling within the object in the direction θ at different locations is plotted in Fig. 1(b). The portion of the object is separated into disconnected regions by the corresponding inner r_{in} and outer r_{out} radii of the two concentric spheres in Fig. 1(a). For a variety of scales of pulmonary vessels, it is important to determine the appropriate r_{in} and r_{out} that robustly differentiate junctions from other structures. Let $d_{\text{min}} = \min_{\theta \in [0,2\pi)} d(\theta)$ and it indicates the scale of the structure. Then, $r_{\text{in}} >= d_{\text{min}}$, because there is always only one disconnected region otherwise. To adapt to different scales, simple linear relations of $a_{\text{in}} = \alpha_{\text{in}} d_{\text{min}}$ and $r_{\text{out}} = \alpha_{\text{out}} d_{\text{min}}$ are chosen to make the spherical shell capture the size of a local shape, where a_{in} and a_{out} are user-defined constants. Furthermore, a_{in} should be larger than 1 by an amount related to the tolerance of commonly observed small bulges. In addition, a_{out} should be larger than a_{in} to make each disconnected component a sufficient size but not too large to save the computation cost. We empirically use $a_{\text{in}} = 3.0$ and $a_{\text{out}} = 4.0$ for all experiments.

Based on the insight of local shape profile, each candidate voxel within the object of interest in the image is checked whether it belongs to a junction. The distance d_{min} for all candidates can be obtained in one pass using a fast distance transform (Maurer et al. 2003). In practice, the approach to use Euclidean distance to find voxels between concentric spheres may result

in a wrong detection by generating three disconnected components when there is another non-contacting branch nearby (see e.g. Fig. 2(a)). Instead, the minimal distance from the candidate is employed. It is computed using the fast marching method seeded at the candidate (Sethian 1999), and modified to keep track of voxels inside the spherical shell. Then, these voxels are assigned to different components based on the connectedness. As mentioned before, the candidate with three or more disconnected components is considered as a junction point.

It should be noted that multiple candidates can belong to a particular junction due to similar profiles among voxels close to the center of a junction. Fortunately, those for a given junction are in the neighborhood of each other and can be easily clustered. For two candidates 1 and 2 as shown in Fig. 2(b), they are considered to be directly related and belong to the same junction if $d^{1,2} < \max\left\{d^1_{\min}, d^2_{\min}\right\}$. where $d^{1,2}$ is the distance between two candidates. A cluster consists of all candidates that are directly and indirectly related. The candidate with the largest $d_{\min}$ in each cluster is chosen as the center of the junction and its $d_{\min}$ is the size of the junction.

To speed up the detection, it is sufficient to consider only the voxels close to the centerline of the object. Thus, the set of candidates is narrowed down to be $C = \{\upsilon \in S | D_x(\upsilon) \vee D_y(\upsilon) \vee D_z(\upsilon)\}$, where S is the set of all voxels within the object in the image and $D_x(\upsilon) := D_{-x}(\upsilon) < D(\upsilon) \wedge D_{+x}(\upsilon) < D(\upsilon)$ with $D(\upsilon)$, $D_{-x}(\upsilon)$, and $D_{+x}(\upsilon)$is the distance transform at the voxel υ and the adjacent voxels in both directions along x axis, respectively. $D_y(\upsilon)$ and $D_z(\upsilon)$ are defined similarly. Furthermore, all tests of candidates in C are independent with each other. Therefore, the algorithm for junction detection is parallelizable.

B. Junction Trajectory Tracking

B.1 Trajectory Parameterization—A B-spline curve is a sequence of lower-order polynomial curve segments that joins continuously (De Boor 2000). It is used here to represent the trajectory that fits the motion of the junction under smoothness regularization to suppress noises and artifacts. Thanks to the convex hull property of the B-spline curve, the oscillation effect by a single higher-order polynomial is avoided and the trajectory is completely bounded by the control polygon. Another advantage is that each control vertex only influences locally as a consequence of the compact support of B-spline basis functions. Without loss of generality, the fourth-order B-splines with periodic basis functions and uniform knot vectors are considered here. In addition, the assumption of cyclic motion from all images of the 4D data is explicitly imposed; i.e. the last image is temporally next to the first image. Note three repeated control vertices are placed at the end of the control polygon to close the curve and ensure the same order of smoothness at the joint of any two curve segments. Mathematically, the k-th curve segment ($k = 1 \ldots m$) of the B-spline curve is

$$\boldsymbol{C}_k(s) = [b_0(s)\ b_1(s)\ b_2(s)\ b_3(s)] \begin{bmatrix} \mathbf{P}_{(k+0) \bmod m} \\ \mathbf{P}_{(k+1) \bmod m} \\ \mathbf{P}_{(k+2) \bmod m} \\ \mathbf{P}_{(k+3) \bmod m} \end{bmatrix} \quad (1)$$

$$[b_0(s)\ b_1(s)\ b_2(s)\ b_3(s)]^T = \frac{1}{6}\begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 0 & 4 \\ -3 & 3 & 3 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} s^3 \\ s^2 \\ s \\ 1 \end{bmatrix} \quad (2)$$

where, b_h, $h = 0\ldots3$ are the basis functions with $\Sigma_h\, b_h = 1$, mod is the modulo, $s \in [0,1)$ is the time fraction, and m is the number of curve segments, or equally, control vertices $\boldsymbol{P}_l$, $l = 1...m$. It determines the flexibility of the curve. Notice that $m \in [4, n]$ is required, where n is the number of images in the 4D data. The choice of m depends on the complexity of the actual trajectory and the temporal resolution of the data. In the presence of noises and artifacts, smaller m is generally preferred when sufficient in order to avoid over-fitting. Fig. 3 shows an example B-spline curve with $m = 5$ control vertices and $n = 10$ images.

Let $I_i(\boldsymbol{x})$, $i = 1\ldots n$ denote the images in the 4D data. Assuming the images are evenly distributed in time and the interval is unity, the i-th image is at time $t = i$. For any intermediate time $t \in [1, n + 1)$, the corresponding curve segment is $k = \left\lfloor \frac{m(t-1)}{n} + 1 \right\rfloor$ and the time fraction within the segment is $s = k - \left\lfloor \frac{m(t-1)}{n} + 1 \right\rfloor$ The derivatives of the position at s with respect to the support control vertices are the basis functions and zero for non-support control vertices, i.e.

$$\frac{\partial \mathbf{C}_k(s)}{\partial \mathbf{P}_l} = \begin{cases} b_h(s) \text{ if } l = (k+h) \bmod m, h = 0\ldots3 \\ 0 \qquad\qquad \text{otherwise} \end{cases} \quad (3)$$

B.2 Tracking by Optimizing Combined Correlation Coefficients—To find the corresponding points in other phased images, let the center of a junction detected in the reference image ($t = r$) be $\boldsymbol{x}_c$ and its size be d_c. Its trajectory due to respiratory motion is modeled as $\boldsymbol{T}(\boldsymbol{x}_c, t) = \boldsymbol{x}_c + \boldsymbol{C}_k(s)$, where $\boldsymbol{C}_k(s)$ is the displacement from $\boldsymbol{x}_c$ as defined in Eq. (1). An observation that we utilize to simplify the tracking problem is the shape of an individual junction, and the statistics of intensity values over the neighborhood region will not significantly change during the respiration. Therefore, the movement of any voxel $\boldsymbol{x}_j$ in the neighborhood region $N(\boldsymbol{x}_c)$ is assumed to be constant with respect to that of $\boldsymbol{x}_c$, i.e. the deviation between them stays constant as in the reference image

$$\boldsymbol{T}(\boldsymbol{x}_j, t) - \boldsymbol{T}(\boldsymbol{x}_c, t) = \boldsymbol{x}_j - \boldsymbol{x}_c, \boldsymbol{x}_j \in N(\boldsymbol{x}_c) \quad (4)$$

where, the size of the region $N(\boldsymbol{x}_c)$ is junction-specific and is chosen to be α_{neighbor}, where $\alpha_{\text{neighbor}} = 4.0$ empirically. In addition, the voxels outside the lung mask are excluded from $N(\boldsymbol{x}_c)$ since they may move quite differently from those inside due to the sliding motion.

We need to define a metric that measures the similarity among the intensity values $I_i(\boldsymbol{T}(\boldsymbol{x}_j, i))$ on the trajectory. Unlike the popular choice of mean squared intensity differences, we propose to use the sum of normalized correlation coefficients with subtracting means. There are two advantages compared to the former. One is the metric is bounded and suitable to

indicate the reliability of a tracked trajectory. The other is that it is insensitive to additive intensity fluctuations among the images, which are very common because of the density change of lung tissue due to ventilation, especially considering the neighborhood region at the junction center in our case. Mathematically, the metric with respect to the reference image I_r is

$$M(\boldsymbol{p}) = \sum_{i=1}^{n} M_i(\boldsymbol{p}) \tag{5}$$

$$M_i(\boldsymbol{p}) = \frac{\sum_{x_j \in N(x_c)} (I_r(\boldsymbol{x}_j) - \bar{I}_r)(I_i(\boldsymbol{T}(\boldsymbol{x}_j, i)) - \bar{I}_i)}{\sqrt{\sum_{x_j \in N(x_c)} (I_r(\boldsymbol{x}_j) - \bar{I}_r)^2 \sum_{x_j \in N(x_c)} (I_i(\boldsymbol{T}(\boldsymbol{x}_j, i)) - \bar{I}_i)^2}} \tag{6}$$

where, $\boldsymbol{p}$ is a vector of the coordinates of all control vertices $\boldsymbol{P}_l$, $l = 1 \ldots m$. The mean intensities $\bar{I}_r$ and $\bar{I}_i$ are computed in the reference image as $\bar{I}_r = \frac{1}{|N(x_c)|} \sum_{x_j \in N(x_c)} I_r(\boldsymbol{x}_j)$ and in the i-th image as $\bar{I}_i = \frac{1}{|N(x_c)|} \sum_{x_j \in N(x_c)} I_i(\boldsymbol{T}(\boldsymbol{x}_j, i))$. It is straightforward to show that $-n \leq M(\boldsymbol{p}) \leq n$. Ideally, the relation $M(\boldsymbol{p}) = n$ holds if the intensities on the trajectory perfectly correlate. In practice, two problems can cause the value of $M(\boldsymbol{p})$ far away from n. One is that the original observation does not hold, i.e. the shape and intensities values of a junction change significantly. The other is that the tracking algorithm converges to local extrema. In fact, each individual $M_i(\boldsymbol{p})$ (ideally equal to 1) also indicates the reliability of a tracked trajectory passing through the i-th image. There are possibilities when some $M_i(\boldsymbol{p})$ close to one but others not.

Notice that in Eq. (5), the summation actually includes the contribution from the reference image $M_r(\boldsymbol{p})$. In other words, the position of the junction at the reference image is not strictly constrained as $\boldsymbol{T}(\boldsymbol{x}_c, r) = \boldsymbol{x}_c$ (as shown in Fig. 3 with $r = 1$), but leaves these particular degrees of freedom to the optimization. It is a deliberate choice to account for the fact that the trajectory may need to deviate slightly from $\boldsymbol{x}_c$ in order to achieve a better overall metric while satisfying the smoothness set forth by the B-spline curve.

Solving the trajectory tracking problem then becomes the estimation of the optimal parameters $\boldsymbol{p}^*$ that maximizes the metric $M(\boldsymbol{p})$

$$\boldsymbol{p}^* = \arg\max_{\boldsymbol{p}} M(\boldsymbol{p}) \tag{7}$$

Again, the optimization problem of each junction is independent with each other. Therefore, the algorithm for junction tracking is also parallelizable. In our landmark-based approach, spatial regularization between junctions is not introduced, as mentioned in the introduction.

B.3 Optimization Strategy—Optimization of Eq. (7) is handled by a Quasi-Newton approach in the form of the limited memory BFGS (L-BFGS) method (Nocedal 1980). The main advantage is due to its high precision and improved convergence rate compared to

simple gradient descent algorithms. Another feature of L-BFGS is that simple bounds can be placed to constrain the range of parameters (Byrd et al. 1995). Thanks to the convex hull property of B-spline curves, the feature can be used to explicitly incorporate a prior knowledge of the scope of the trajectory (e.g. in each of the three dimensions).

We start the optimization with zero initial conditions, i.e. the whole trajectory is simply a single point at $\boldsymbol{x}_c$. In most cases, this trivial choice leads to converge to the actual trajectory. However, in some cases, especially for the large movement or poor image quality, the optimization converges to a local extremum, which can be easily detected if the value of the metric $M(\boldsymbol{p})$ is far away from n. It is necessary to search for an improved initial condition, which is closer to the actual trajectory.

As illustrated in Fig. 4, we propose to use random walks from pair-wise optimizations. The iterative procedure is as follows: starting from $\boldsymbol{x}_c$ at $t = r$, suppose a good match ($M_i(\boldsymbol{p})$ is close to 1) is found at $t = i$ and the corresponding position is $\boldsymbol{T}'(\boldsymbol{x}_c, i)$. To search for the position at the next time step $t = i + 1$, a pair-wise optimization using L-BFGS can be performed to maximize $M_{i+1}(\boldsymbol{p})$ in Eq. (6), which involves only $I_t(\boldsymbol{x}_j)$ and $I_{i+1}(\boldsymbol{T}(\boldsymbol{x}_j, i + 1))$, $\boldsymbol{x}_j \in N(\boldsymbol{x}_c)$. It is simpler to optimize in the pair-wise fashion than globally since the movement of the junction between the consecutive steps is smaller. To further improve the coverage, a number of random positions in a neighborhood centered at $\boldsymbol{T}'(\boldsymbol{x}_c, i)$ are sampled and used to initialize a series of optimizations. The position with the largest $M_{i+1}(\boldsymbol{p})$ is then chosen to be $\boldsymbol{T}'(\boldsymbol{x}_c, i + 1)$. The iteration continues until it returns to r. However, error in each step may accumulate in this manner. We instead walk halfway in both the forward and backward directions. Finally, the control vertices $\boldsymbol{P}'_l, l = 1 \ldots m$ of the B-spline curve that best fit $\boldsymbol{T}'(\boldsymbol{x}_c, i)$, $i = 1 \ldots n$ are computed by solving a least squares problem based on Eq. (1). The parameters $\boldsymbol{p}'$ from this set of control vertices serves as the new initial conditions to repeat the optimization.

L-BFGS requires the derivatives of the metric $M(\boldsymbol{p})$ with respect to $\boldsymbol{p}$ are explicitly defined, which can be obtained by differentiating Eq. (5) as

$$\begin{aligned}\frac{\partial M_{(\boldsymbol{p})}}{\partial \boldsymbol{p}} = \sum_{i=1}^{n} & \frac{1}{\sqrt{\sum_{x_j \in N(x_c)} (I_r(\boldsymbol{x}_j) - \bar{I}_r)^2 \sum_{x_j \in N(x_c)} (I_i(\boldsymbol{T}(\boldsymbol{x}_j, i)) - \bar{I}_i)^2}} \\ & \cdot \left[\sum_{x_j \in N(x_c)} (I_r(\boldsymbol{x}_j) - \bar{I}_r) \frac{\partial I_i(\boldsymbol{T}(\boldsymbol{x}_j, i))}{\partial \mathbf{p}} \right. \\ & - \frac{\sum_{x_j \in N(x_c)} (I_r(\boldsymbol{x}_j) - \bar{I}_r)(I_i(\boldsymbol{T}(\boldsymbol{x}_j, i)) - \bar{I}_i)}{\sum_{x_j \in N(x_c)} (I_i(\boldsymbol{T}(\boldsymbol{x}_j, i)) - \bar{I}_i)^2} \\ & \left. \cdot \sum_{x_j \in N(x_c)} (I_i(\boldsymbol{T}(\boldsymbol{x}_j, i) - \bar{I}_i) \frac{\partial I_i(\boldsymbol{T}(\boldsymbol{x}_j, i))}{\partial \boldsymbol{p}} \right]\end{aligned} \tag{8}$$

where, $\bar{I}_r$ and $\bar{I}_i$, are defined as before. The derivatives $\frac{\partial I_i(\boldsymbol{T}(\boldsymbol{x}_j, i))}{\partial \boldsymbol{p}}$ represent the change of the intensity at position $\boldsymbol{T}(\boldsymbol{x}_j, i)$ on the trajectory with respect to the change of the parameters $\boldsymbol{p}$, which are computed using the chain rule as

$$\frac{\partial I_i(\boldsymbol{T}(\boldsymbol{x}_j, i))}{\partial \boldsymbol{p}} = \frac{\partial I_i(\boldsymbol{T}(\boldsymbol{x}_j, i))}{\partial \boldsymbol{T}(\boldsymbol{x}_j, i)} \frac{\partial \boldsymbol{T}(\boldsymbol{x}_j, i)}{\partial \boldsymbol{p}} \quad (9)$$

where, the first term is the intensity gradient. Evaluating the intensity and its gradient at $\boldsymbol{T}(\boldsymbol{x}_j, i)$, usually a non-grid position, requires a continuous representation of the image. It is obtained by interpolation using fourth-order B-splines. The second term is readily available in Eq. (3).

B.4 Trajectory Descriptors—The qualitative assessment of the tracked trajectory through direct visualization is helpful to reveal the complexity of the motion pattern, such as motion heterogeneity and hysteresis. On the other hand, the quantitative analysis of the trajectory is very important to measure the extents of such complexities. For this purpose, we define the following descriptors to characterize a trajectory:

a. *Long axis length* Q_L: it is the length of the long axis, which is defined as the displacement from the maximal inhale to the maximal exhale.

b. Short *axis length* Q_S: it is the length of the short axis, which is defined as the displacement from a point to the point half cycle later on the trajectory with the smallest distance.

c. *Hysteresis* Q_H: it measures the deviation between the inhale and exhale paths and is defined as Q_S/Q_L. $Q_H \geq 0$ and $Q_H = 0$ if there is no hysteresis.

d. Max *curvature* Q_C: it measures the degree of maximal turning along the trajectory. Thanks to the B-spline representation, the curvature of any point on the trajectory can be analytically computed as $\frac{|C'(t) \times C''(t)|}{|C''(t)|^3}$.

e. Max *torsion* Q_T: it measures the degree of maximal twisting. The torsion of any point on the trajectory can be analytically computed as $\frac{|(C'(t) \times C''(t)) \cdot C'''(t)|}{|C'(t) \times C''(t)|^2}$.

f. *Uniformity* Q_U: it measures how points at different phases are distributed temporally along the trajectory and is defined as $-\sum_{i=1}^{n} \tilde{l}_i \ln \tilde{l}_i$, where $\tilde{l}_i$ is the length of the trajectory between the i-th and $(i+1)$-th phases, normalized by the perimeter $Q_U \leq \ln n$ and $Q_U = \ln n$ if uniformly distributed.

III. Experiments and Results

A. Algorithm Implementation

The algorithms were implemented in C++, with the support of the ITK library (Ibanez et al. 2005). All experiments were carried out on a computer with a 3.0 GHz quad-core processor and 8 GB memory. Both junction detection and tracking were implemented in parallel taking advantage of the multiple cores in the processor, whenever possible.

B. Datasets

We employ 4D CT datasets of the lungs from two sources in our experiments. One is publicly available from DIR-Lab (http://www.dir-lab.com). It has ten cases (referred as Cases 1-10); each composed of ten CT images corresponding to respiratory phases. The first image is at maximal inhale and the sixth image is at maximal exhale. There are two sets of manually identified landmarks (mostly vessel bifurcations) for each case. One is 300 landmarks in the maximal inhale and maximal exhale phases. The other is 75 landmarks in the six expiratory phases. The other source is from CREATIS (Vandemeulebroucke et al. 2011). It has three cases (referred as Cases 11-13). Each case has 100 landmarks identified in all ten phases. For more details of image acquisition protocols, image specifications, and landmark extraction procedures, we refer the readers to their original papers and websites.

C. Junction Detection

The first image at the maximal inhale phase is selected as the reference image, in which the junction detection algorithm is performed. In Fig. 5, an example of lung segmentation, vessel segmentation, and junction detection in one axial slice is shown. Fig. 6 shows the detected junctions in Case 1, depicted as spheres of corresponding sizes, in volume-rendered reference image masked by the lung region. Visually, they are in a large number and spread quite uniformly across both the left and right lungs. Note that the sufficiency and proper distribution of the landmarks are necessary for adequately sampling the motion through trajectory tracking. For a more quantitative assessment in Fig. 7, we show the histogram-based spatial distributions in three main directions for all ten cases. General trends are observed as follows: in inferior-superior direction, the distribution does not change much except at apex and base. However, the number ofjunctions clearly increases from anterior to posterior, which is primarily related to the spatially-varying distribution of the lung volume in this direction. The gap in the right-left direction is caused by the separation of the left and right lungs. Furthermore, the number of junctions in the right lung is greater than that of the left lung, also attributing to the difference in lung volumes.

Table I summarizes the results of junction detection. In each case, we have automatically detected a range of 700-1100 junctions in less than 5 minutes of the segmentation and detection time. This is significantly more than what manual or semi-automatic approaches can achieve in the same time frame. Certainly, such a large number of junctions may not be always necessary, especially when motion in only a small region is of interest. In order to evaluate the accuracy, two radiation oncologists (observer 1 and 2), who are expert in lung anatomy, independently inspected 1/4 (random selected) of all detected junctions for each case. This was done using MimVista (MIM Software Inc., Cleveland, OH) by overlaying the detected junctions (spheres in different colors) on the original CT image and displaying in three orthogonal views. After seeing every detected junction, the observers have three judgments to make: “correct”, “uncertain”, and “incorrect”, indicating the existence of a true junction. The positive predictive value (PPV) is then computed as the percentage of true positives (the number of “correct” and half of the number of “uncertain”) of the total number of positives. The average PPV for observer 1, observer 2, and both observers are 92.0%, 90.3%, and 91.2%, respectively. Notice that we do not attempt to compute the sensitivity for three reasons. First, the problem of correctly detecting all junctions in the lung is still

challenging and we admit that there are a considerable number of the false negatives led by our method. Second, manually generating such a ground truth data is intimidating if not impossible. Moreover, the number of junctions we are able to detect is considered to be sufficient for our purpose. The computation of specificity is also not applicable since true negatives are not defined. To assess the inter-observer difference, we define agreement and disagreement values as the percentages of same judgments and different judgments. On average the agreement and disagreement are 94.8% and 5.2%, respectively. In regard to the junction sizes, we plot the histograms using seven bins in Fig. 8. A clear exponential trend exhibits in all cases, which agrees with the exponential branching pattern of the pulmonary vessels.

D. Tracking using Manually Identified Landmarks

We test our tracking algorithm using the manually identified landmarks from DIR-Lab as the ground-truth data. The landmarks on the first phase are supplied as inputs and they are tracked in the following phases. To evaluate the accuracy, the tracking error is defined by the Euclidean distance between manual and tracked landmark positions, or commonly known as the target registration error (TRE). It should be noted that manual landmarks are not necessarily all junctions and their locations are delineated only in the voxel without sub-voxel accuracy. Furthermore, we have assigned the size of all landmarks to 2.0 mm since no sizing information is available. In addition, tracking is performed in all images but accuracy is assessed only on $I_1 \ldots I_6$, or I_1 and I_6.

In order to determine the appropriate number of control vertices m that is sufficient to capture the motion, we track 75 landmarks for Case 1 using different numbers of control vertices $m = 4 \ldots 10$. In the following, we employ box plots to show the distribution of the tracking errors. On each box, the central line is the median, the edges of the box are the 25th $p_{25\%}$ and 75th percentile $p_{75\%}$, and the whiskers extend to $p_{75\%} + 1.5(p_{75\%} - p_{25\%})$ and $p_{25\%} - 1.5(p_{75\%} - p_{25\%})$. Remaining outliers are not plotted for clarity. Box plots of the tracking errors at Phases 1-6 among are shown in Fig. 9. There are no significant differences of the tracking errors for Phases 2-6 (p-value=0.9983, 0.9994, 0.9973, 0.9902, and 0.9771, respectively using ANOVA) except for Phase 1 (p-value=4.13×10^{-39}). Recall that we do not explicitly impose zero displacements at Phase 1 (the reference). The results suggest that fewer numbers of control vertices are already sufficient to represent the motion of the lung due to the limited temporal resolution. Increasing the control vertices allows the trajectory to be dragged more closely to the reference phase. In respect of the metric of the combined correlation coefficients $M(\boldsymbol{p}^*)$, there is also no significant difference (p-value=0.9999) among $m = 4 \ldots 10$ (box plots in Fig. 10). Therefore, we choose $m = 7$ in our further evaluations for a compromise between the simplicity of the trajectory and the additional flexibility of the represented motion.

Next, we track 75 landmarks in each Case 1-10 and show the box plots of the tracking errors at Phases 1-6 in Fig. 11. In addition, we track 300 landmarks and show the box plots of the tracking errors at Phases 1 and 6 in Fig. 12. In both figures, the box plots of the tracking errors of all ten cases are shown on the right. In comparison, we define the baseline tracking errors as the distances between the manual and stationary (i.e. without trajectory tracking)

reference landmark positions, which indicate the extents of the motion. It is clear that the trajectory tracking dramatically reduces the tracking errors.

Table II summarizes the mean and standard deviations of the tracking errors in Fig. 11 and Fig. 12. It suggests that the proposed tracking algorithm can achieve sub-voxel accuracy. The mean TRE values of combined Phases 1-6 in the experiment with 75 landmarks and those of combined Phase 1 and 6 in the experiment with 300 landmarks are smaller than the published results in (Metz et al. 2011) using the same data (first five cases) and the same landmark sets. However, it should be mentioned that the algorithm in (Metz et al. 2011) are not designed to specifically optimize for alignment at the landmarks. Therefore, optimization in a global manner may compromise the alignment of local structures. In fact, the purpose of our algorithm is to explicitly align the local junction structures. Compared to the results in (Castillo et al. 2010), the median of TRE values per phase in combined Cases 1-10 is smaller in the experiment with 75 landmarks (from reading Figure 6 of (Castillo et al. 2010), though the exact value is not available). In comparison of the mean TRE values in our experiment with 300 landmarks and theirs with larger number of landmarks, 6 out of 10 are smaller, whereas our standard deviations are consistently smaller. Unfortunately, we have no access to this larger set of landmarks for a more fair comparison. In addition, the performance of our algorithm from case to case is found to be loosely related to the extent of the lung motion. But in each case as shown in Fig. 11, the performance per phase does not depend on the extent of the motion at a given phase number (excluding Phase 1).

Similarly, we track 100 landmarks in each Case 11-13 from CREATIS. Note that Phase 6 at maximal exhale is selected as the reference in order to compare with (Vandemeulebroucke et al. 2011). Fig. 13 shows the box plots of the tracking errors and Table III summarizes the mean and standard deviations. Again, the proposed tracking algorithm can achieve sub-voxel accuracy. Compared to the results in (Vandemeulebroucke et al. 2011), the mean TRE values of Phase 1 and Phases 1-5,7-10 in each case is smaller using the same data and the same landmark sets.

The computational time depends on the size of the local region, the optimization convergence rate, and the necessity to search for an improved initial condition. On average, the time for tracking one trajectory is about 2-3 seconds.

E. Tracking using Automatically Detected Junctions

Finally, we track the junctions automatically detected in Section III–C. Fig. 14 shows the optimized metrics $M(\boldsymbol{p}^*)$ for all cases, of which a majority are greater than 7.0. Note that the tracked trajectories with metrics less than 8.0 are considered to be unreliable and thus discarded for further analysis. Fig. 15 shows some examples of tracked trajectories in Case 1, colored by their metric values. The trajectories are actually space curves in 3D and can be quite complex. From maximal inhale at Phase 1 (marked by a sphere), a trajectory moves forward through the successive phases (marked by cones with tips pointing to the direction), until reaching maximal exhale at Phase 6. Inversely, the trajectory clearly has a different return path, i.e. hysteresis. To characterize the hysteresis, we draw their major and minor axes (marked by arrows). As mentioned before, the ratio of the lengths between the minor and major axes is used to indicate the degree of hysteresis (1 for full hysteresis, 0 for no

hysteresis). Fig 16 displays an example tracked trajectory with the volume-rendered junction structure in each phase.

To analyze the trajectories quantitatively, we compute the descriptors defined in Section II–B.4 for all tracked trajectories with $M(\boldsymbol{p}^*) > 8.0$. Fig. 17 demonstrates the results overlaid on the trajectories and colored by their corresponding values for Case 1. From the figure, we can see the long axis length is greater at the bottom portion of the lung, which is expected because the diaphragm pushes the tissue up during respiration. For the short axis length, there is a large region with higher values in the bottom of left lung than that of the right. But, this region does not have high hysteresis values, which are essentially short axis lengths normalized by long axis lengths. Conversely, the hysteresis values are larger in the upper lung, especially in the left. For the max curvature, there is a similar region with higher short axis length but lower curvature value, which suggests the trajectories turn less sharply than elsewhere. On the other hand, the max torsion in the upper lung is clearly higher than that in the bottom. In addition, the uniformity values in the left lung are roughly larger than those in the right. Finally, we should emphasize that the above analysis is only made on this particular case. Therefore, it may not be generalizable to other cases for the obvious patient-specific nature of lung motion. But similar analysis can be performed in those cases.

IV. Conclusion

We have proposed a framework to both detect junction structures and track their trajectories in 4D CT images. In the problem of detection, we identified more than 700 junctions with an average accuracy of greater than 90%. We believe the automation of detecting junctions as landmarks should be useful in various applications of lung analysis, where manual or semi-automatic annotation is still a common practice. In principle, any landmarks can be used to study the motion of the lung as long as they possess the intensity profiles that can be reliably tracked. However, it is difficult, if not impossible, to detect a large number of other types of landmarks throughout the lung other than the vessel junctions. It should be emphasized that our junction detection algorithm is only performed in one reference phase. Therefore, no pair-wise matching is necessary in our approach, compared to common landmark-based approaches, e.g. in (Tashiro et al. 2006). For a particular junction, its locations in the following phases are determined simultaneously by the optimization in the trajectory tracking step.

In the problem of tracking, we do not pursue the common intensity-based approach which first registers in the whole image space and then extracts the motion trajectories. In this approach, one assumes the spatial smoothness of transformation and may compromise the accuracy of alignment of local structures. Another drawback is that one has to optimize a large number of transform parameters together and can be extremely time-consuming. In our approach, trajectories of the landmarks are separately tracked by optimizing a combination of correlation coefficients of intensities on the trajectory from all images. Our choice is more apparent when only local deformation is of interest, e.g. in the application of tracking local structures around a tumor. We have achieved sub-voxel accuracy and obtained better results than other published work using the same data and the same landmark sets. Our approach to generate better initialization suppresses local extrema. Although worst cases may still

happen rarely, they can be detected by extremely low correlation coefficients and disregarded for further consideration. But we should mention that distorted junctions due to imaging noises or artifacts may also lead to low correlation coefficients even the alignment has been achieved. Another potential application of our tracking algorithm is to ease the tedious task of generating large validation data sets for 4D image registration. Given the automated or manual landmarks on one phase, the corresponding landmark positions on successive phases are automatically detected. Once examined and confirmed by experts, they can be used to evaluate other registration methods. There is one problem we have not addressed in this paper: how to generate a full 4D deformation field from a large number of trajectories. It is only necessary if the goal is to obtain motion information in the whole image space. Although we could simply produce a series of 3D deformation fields using thin plate spline (Bookstein 1989), this approach does not account for the temporal consistency. Future work will be directed to tackle this problem more thoroughly.

Acknowledgment

We would like to thank Drs. R. Castillo and T. Guerrero from MD Anderson Cancer Center for making the DIR-Lab data publicly available. We also acknowledge Drs. D. Sarrut, J. Vandemeulebroucke, and V. Delmon from Léon Bérard Cancer Center, University of Lyon, France for sharing the CREATIS data. We thank the reviewers for their constructive comments. This work was supported by the National Cancer Institute under Grants (1R01 CA133474 and 1R21 CA153587), and National Science Foundation of China (NSFC) Joint Research Fund for Overseas Research Chinese, Hong Kong and Macao Young Scholars (30928030).

References

Amit Y 1994 A Nonlinear Variational Problem for Image Matching. SIAM J Sci Comput 15(1):207–224.

Arsigny V, Pennec X, and Ayache N. 2005 Polyrigid and polyaffine transformations: A novel geometrical tool to deal with non-rigid deformations - Application to the registration of histological slices. Med Image Anal 9(6):507–523. [PubMed: 15948656]

Beg MF, Miller MI, Trouve A, and Younes L. 2005 Computing large deformation metric mappings via geodesic flows of diffeomorphisms. Int J Comput Vision 61(2): 139–157.

Boldea V, Sharp GC, Jiang SB, and Sarrut D. 2008 4D-CT lung motion estimation with deformable registration: quantification of motion nonlinearity and hysteresis. Med Phys 35(3):1008–1018. [PubMed: 18404936]

Bookstein FL. 1989 Principal Warps: Thin-Plate Splines and the Decomposition of Deformations. IEEE T Pattern Anal 11(6):567–585.

Byrd RH, Lu PH, Nocedal J, and Zhu CY. 1995 A Limited Memory Algorithm for Bound Constrained Optimization. SIAM J Sci Comput 16(5):1190–1208.

Castillo E, Castillo R, Martinez J, Shenoy M, and Guerrero T. 2010 Four-dimensional deformable image registration using trajectory modeling. Phys Med Biol 55(1):305–327. [PubMed: 20009196]

Chao M, Li T, Schreibmann E, Koong A, and Xing L. 2008 Automated contour mapping with a regional deformable model. Int J Radiat Oncol 70(2):599–608.

Chao M, Schreibmann E, Li T, Wink N, and Xing L. 2007 Automated contour mapping using sparse volume sampling for 4D radiation therapy. Med Phys 34(10):4023–4029. [PubMed: 17985648]

Collins DL, and Evans AC. 1997 Animal: Validation and applications of nonlinear registration-based segmentation. Int J Pattern Recogn 11(8):1271–1294.

Coselmon MM, Balter JM, McShan DL, and Kessler ML. 2004 Mutual information based CT registration of the lung at exhale and inhale breathing states using thin-plate splines. Med Phys 31(11):2942–2948. [PubMed: 15587645]

Davatzikos C 1997 Spatial transformation and registration of brain images using elastically deformable models. Comput Vis Image Und 66(2):207–222.

Davis MH, Khotanzad A, Flamig DP, and Harms SE. 1997 A physics-based coordinate transformation for 3-D image matching. IEEE T Med Imaging 16(3):317–328.

De Boor C 2000 A practical guide to splines. New York: Springer.

Dougherty L, Asmuth JC, and Gefter WB. 2003 Alignment of CT lung volumes with an optical flow method. Acad Radiol 10(3):249–254. [PubMed: 12643551]

Ford EC, Mageras GS, Yorke E, and Ling CC. 2003 Respiration-correlated spiral CT: A method of measuring respiratory-induced anatomic motion for radiation treatment planning. Med Phys 30(1):88–97. [PubMed: 12557983]

Foskey M, Davis B, Goyal L, Chang S, Chaney E, Strehl N, Tomei S, Rosenman J, and Joshi S. 2005 Large deformation three-dimensional image registration in image-guided radiation therapy. Phys Med Biol 50(24):5869–5892. [PubMed: 16333161]

Frangi A, Niessen W, Vincken K, and Viergever M. 1998 Multiscale vessel enhancement filtering. Medical Image Computing and Computer-Assisted Intervention. p 130–137.

Guerrero T, Zhang G, Huang TC, and Lin KP. 2004 Intrathoracic tumour motion estimation from CT imaging using the 3D optical flow method. Phys Med Biol 49(17):4147–4161. [PubMed: 15470929]

Heath E, Collins DL, Keall PJ, Dong L, and Seuntjens J. 2007 Quantification of accuracy of the automated nonlinear image matching and anatomical labeling (ANIMAL) nonlinear registration algorithm for 4D CT images of lung. Med Phys 34(11):4409–4421. [PubMed: 18072506]

Horn BKP, and Schunck BG. 1981 Determining Optical-Flow. Artif Intell 17(1–3):185–203.

Ibanez L, Schroeder W, Ng L, and Cates J. 2005 The ITK Software Guide. Clifton Park, NY: Kitware Inc.

Kohlrausch J, Rohr K, and Stiehl H. 2005 A new class of elastic body splines for nonrigid registration of medical images. J Math Imaging Vis 23(3):253–280.

Latecki L, Eckhardt U, and Rosenfeld A. 1995 Well-Composed Sets. Comput Vis Image Und 61(1):70–83.

Ledesma-Carbayo MJ, Kybic J, Desco M, Santos A, Suhling M, Hunziker P, and Unser M. 2005 Spatio-temporal nonrigid registration for ultrasound cardiac motion estimation. IEEE T Med Imaging 24(9): 1113–1126.

Li G, Citrin D, Camphausen K, Mueller B, Burman C, Mychalczak B, Miller RW, and Song Y. 2008 Advances in 4D medical imaging and 4D radiation therapy. Technol Cancer Res Treat 7(1):67–81. [PubMed: 18198927]

Low DA, Nystrom M, Kalinin E, Parikh P, Dempsey JF, Bradley JD, Mutic S, Wahab SH, Islam T, Christensen G et al. . 2003 A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing. Med Phys 30(6):1254–1263. [PubMed: 12852551]

Maintz JB, and Viergever MA. 1998 A survey of medical image registration. Med Image Anal 2(1):1–36. [PubMed: 10638851]

Maurer CR, Qi RS, and Raghavan V. 2003 A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions. IEEE T Pattern Anal 25(2):265–270.

McClelland JR, Blackall JM, Tarte S, Chandler AC, Hughes S, Ahmad S, Landau DB, and Hawkes DJ. 2006 A continuous 4D motion model from multiple respiratory cycles for use in lung radiotherapy. Med Phys 33(9):3348–3358. [PubMed: 17022231]

McEachen JC, Nehorai A, and Duncan JS. 2000 Multiframe temporal estimation of cardiac nonrigid motion. IEEE T Image Process 9(4):651–665.

Metz CT, Klein S, Schaap M, van Walsum T, and Niessen WJ. 2011 Nonrigid registration of dynamic medical imaging data using nD+t B-splines and a groupwise optimization approach. Med Image Anal 15(2):238–249. [PubMed: 21075672]

Murphy K, van Ginneken B, Klein S, Staring M, de Hoop BJ, Viergever MA, and Pluim JP. 2011 Semi-automatic construction of reference standards for evaluation of image registration. Med Image Anal 15(1):71–84. [PubMed: 20709592]

Murphy K, van Ginneken B, Reinhardt J, Kabus S, Ding K, Deng X, Cao K, Du K, Christensen G, Garcia V et al. . in press. Evaluation of Registration Methods on Thoracic CT: The EMPIRE10 Challenge. IEEE Trans Med Imaging.

Nocedal J 1980 Updating Quasi-Newton Matrices with Limited Storage. Math Comput 35(151):773–782.

Otsu N 1979 Threshold Selection Method from Gray-Level Histograms. IEEE T Syst Man Cyb 9(1):62–66.

Perperidis D, Mohiaddin RH, and Rueckert D. 2005 Spatio-temporal free-form registration of cardiac MR image sequences. Med Image Anal 9(5):441–456. [PubMed: 16029955]

Peyrat JM, Delingette H, Sermesant M, Xu CY, and Ayache N. 2010 Registration of 4D Cardiac CT Sequences Under Trajectory Constraints With Multichannel Diffeomorphic Demons. IEEE T Med Imaging 29(7):1351–1368.

Rietzel E, and Chen GTY. 2006 Deformable registration of 4D computed tomography data. Med Phys 33(11):4423–4430. [PubMed: 17153421]

Rietzel E, Pan T, and Chen GTY. 2005 Four-dimensional computed tomography: Image formation and clinical protocol. Med Phys 32(4):874. [PubMed: 15895570]

Runge C 1901 Über empirische Funktionen und die Interpolation zwischen äquidistanten Ordinaten. Zeitschrift für Mathematik und Physik 46(224–243):20.

Sarrut D, Boldea V, Miguet S, and Ginestet C. 2006 Simulation of four-dimensional CT images from deformable registration between inhale and exhale breath-hold CT scans. Med Phys 33(3):605–617. [PubMed: 16878564]

Schaefer S, McPhail T, and Warren J. 2006 Image deformation using moving least squares. ACM T Graphic 25(3):533–540.

Schreibmann E, and Xing L. 2006 Image registration with auto-mapped control volumes. Med Phys 33(4):1165–1179. [PubMed: 16696494]

Seppenwoolde Y, Shirato H, Kitamura K, Shimizu S, van Herk M, Lebesque JV, and Miyasaka K. 2002 Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. Int J Radiat Oncol 53(4):822–834.

Sethian JA. 1999 Level set methods and fast marching methods. Cambridge, U.K. ; New York: Cambridge University Press xx, 378 p. p.

Sluimer I, Schilham A, Prokop M, and van Ginneken B. 2006 Computer analysis of computed tomography scans of the lung: A survey. IEEE T Med Imaging 25(4):385–405.

Soille P 2003 Morphological image analysis: principles and applications. Berlin; New York: Springer xvi, 391 p.p.

Sundar H, Litt H, and Shen DG. 2009 Estimating myocardial motion by 4D image warping. Pattern Recogn 42(11):2514–2526.

Sundaram TA, and Gee JC. 2005 Towards a model of lung biomechanics: pulmonary kinematics via registration of serial lung images. Med Image Anal 9(6):524–537. [PubMed: 15896996]

Tashiro M, Minohara S, Kanai T, Yusa K, Sakurai H, and Nakano T. 2006 Three-dimensional velocity mapping of lung motion using vessel bifurcation pattern matching. Med Phys 33(6):1747–1757. [PubMed: 16872082]

Thirion JP. 1998 Image matching as a diffusion process: an analogy with Maxwell's demons. Med Image Anal 2(3):243–260. [PubMed: 9873902]

Tustison NJ, Cook TS, Song G, and Gee JC. 2011 Pulmonary kinematics from image data a review. Acad Radiol 18(4):402–417. [PubMed: 21377592]

Vandemeulebroucke J, Rit S, Kybic J, Clarysse P, and Sarrut D. 2011 Spatiotemporal motion estimation for respiratory-correlated imaging of the lungs. Med Phys 38(1):166–178. [PubMed: 21361185]

Vercauteren T, Pennec X, Perchant A, and Ayache N. 2009 Diffeomorphic demons: Efficient non-parametric image registration. Neuroimage 45(1):S61–S72. [PubMed: 19041946]

Xie YQ, Chao M, and Xing L. 2009 Tissue Feature-Based and Segmented Deformable Image Registration for Improved Modeling of Shear Movement of Lungs. Int J Radiat Oncol 74(4):1256–1265.

Xing L, Lee L, and Timmerman R. 2009 Adaptive Radiation Therapy and Clincial Perspectives In: Timmerman R, and Xing L, editors. Image Guided and Adaptive Radiation Therapy. Baltimore Lippincott Williams & Wilkins p 16–40.

Xing L, Siebers J, and Keall P. 2007 Computational challenges for image-guided radiation therapy: Framework and current research. Semin Radiat Oncol 17(4):245–257. [PubMed: 17903702]

Xing L, Thorndyke B, Schreibmann E, Yang Y, Li TF, Kim GY, Luxton G, and Koong A. 2006 Overview of image-guided radiation therapy. Med Dosim 31(2):91–112. [PubMed: 16690451]

Xiong G, and Xing L. 2011 Automated detection of junctions structures and tracking of their trajectories in 4D images Information Processing in Medical Imaging, Monastery Irsee, Germany.

Yin Y, Hoffman EA, Ding K, Reinhardt JM, and Lin CL. 2011 A cubic B-spline-based hybrid registration of lung CT images for a dynamic airway geometric model with large deformation. Phys Med Biol 56(1):203–218. [PubMed: 21149947]

Yin Y, Hoffman EA, and Lin CL. 2009 Mass preserving nonrigid registration of CT lung images using cubic B-spline. Med Phys 36(9):4213–4222. [PubMed: 19810495]

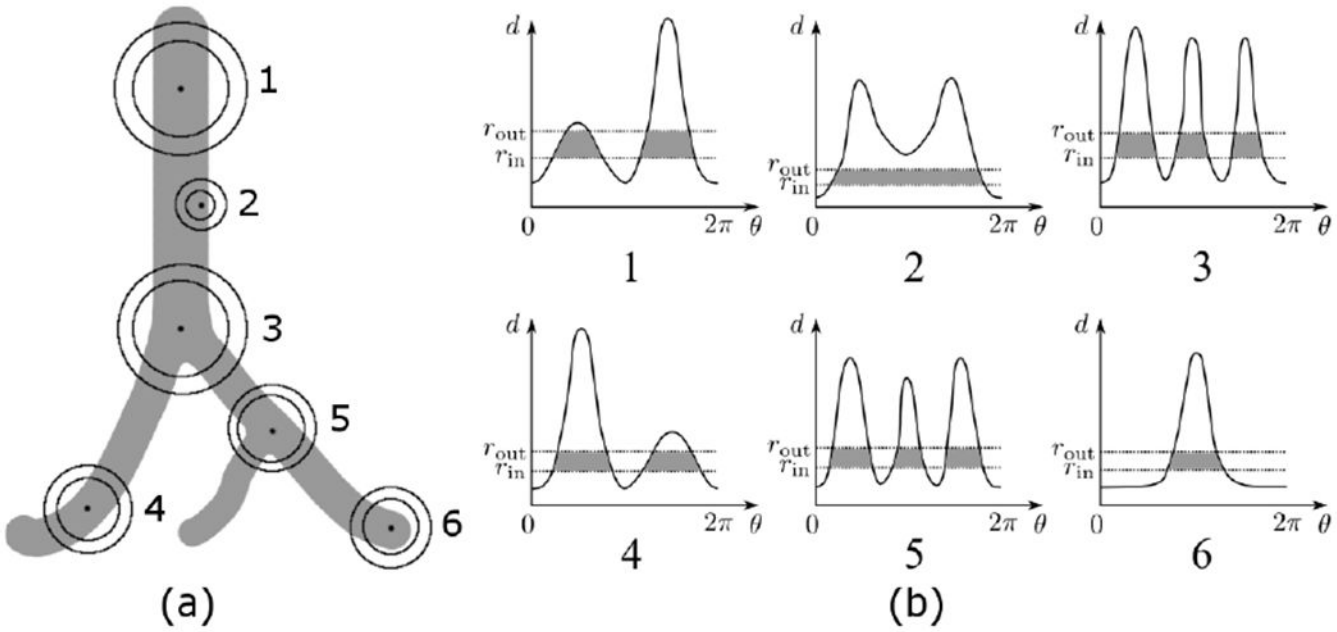

Fig. 1.

(a) Six typical cases of concentric spheres at different locations of a synthetic 2D object; (b) their corresponding local shape profiles: the maximal distance d of a ray traveling with the object in the direction θ. The number of disconnected components between inner r_{in} and outer r_{out} radii can be used to differentiate junctions from other structures, such as tubular structure and endpoint.

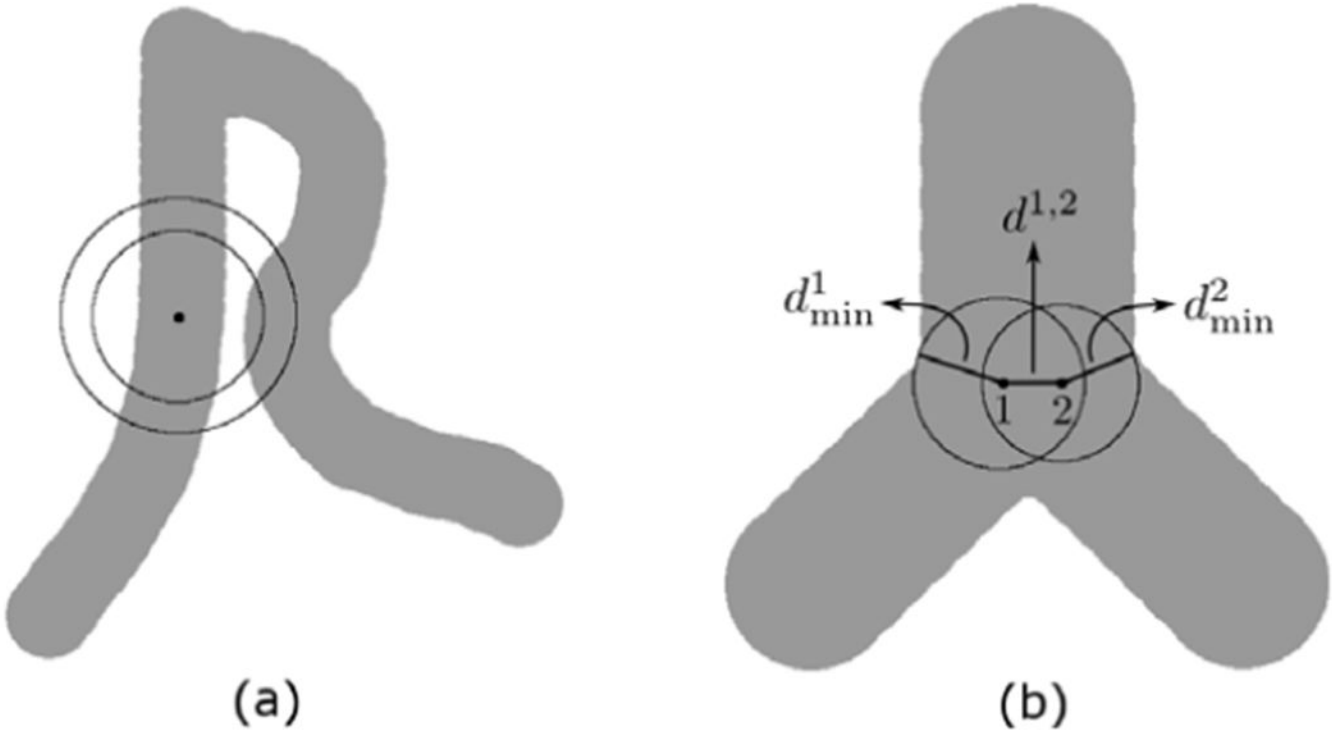

Fig. 2.

(a) An example shows that Euclidean distance may lead to three disconnected components when there is another non-contacting branch nearby; (b) The candidates 1 and 2 are clustered together, i.e. belong to the same junction, if $d^{1,2} < \max\left\{d^1_{\min}, d^2_{\min}\right\}$.

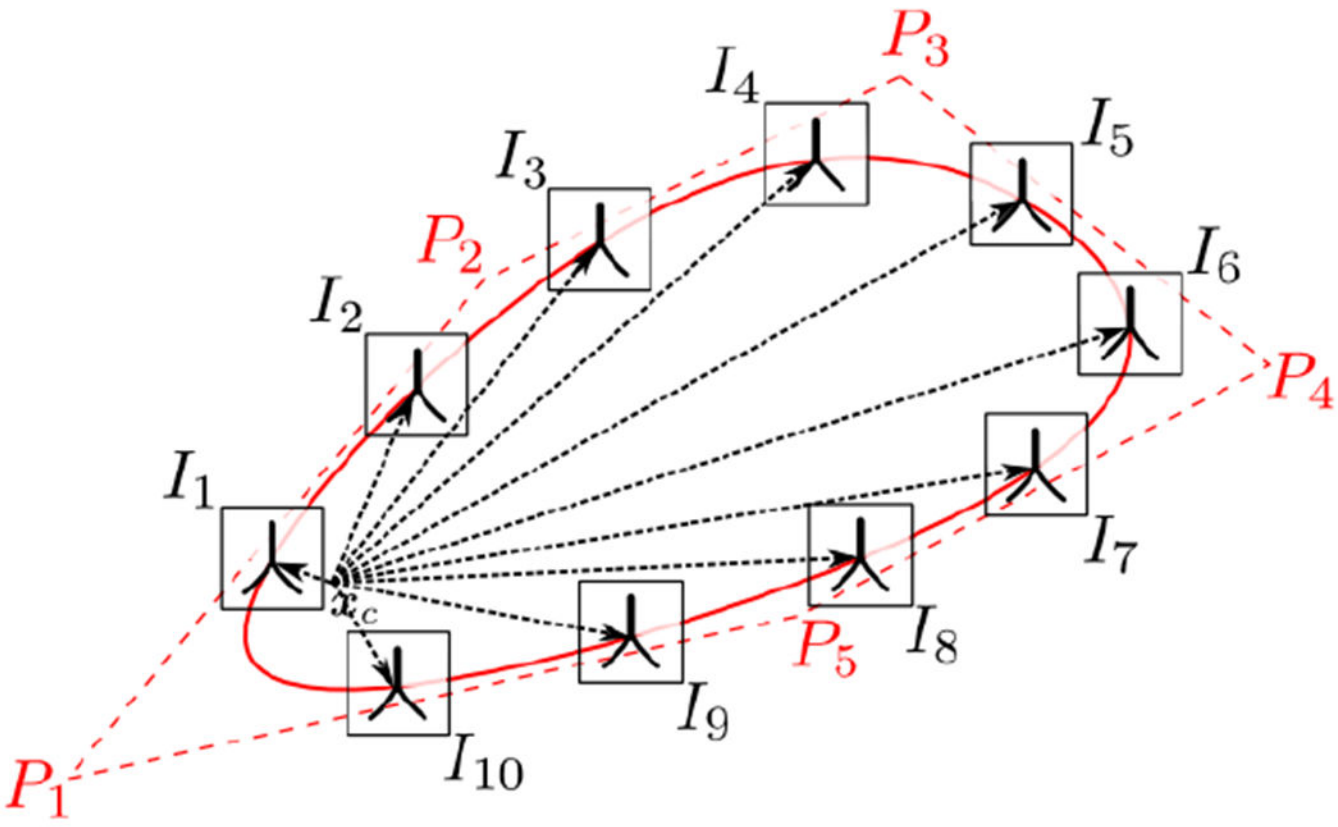

Fig. 3.

An example of trajectory parameterization using B-spline with m=5 control vertices (P_1…P_5) and tracking using n=10 images (I_1…I_{10}). The B-spline curve and its control polygon are in solid and dashed lines. A tracked trajectory should pass through corresponding junction centers in all images. The dotted lines depict the displacements at different time points.

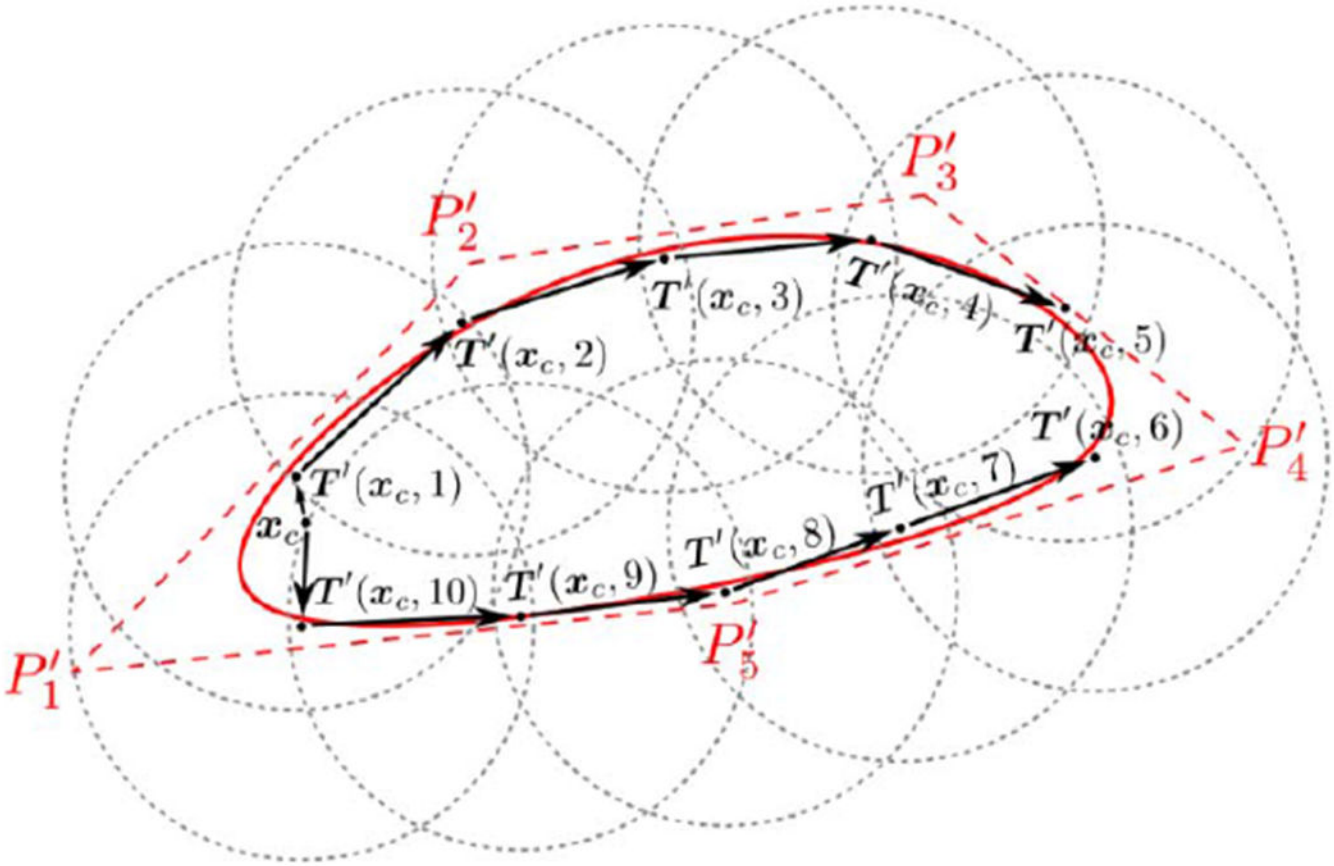

Fig. 4.

An illustration of searching for an improved initial condition by random walks. Random positions in the neighborhood (gray dashed circle) of the previous step $\boldsymbol{T}'(\boldsymbol{x}_c, i)$ are used to initialize pair-wise optimizations to determine the next step $\boldsymbol{T}'(\boldsymbol{x}_c, i+1)$. Both forward and backward directions are searched halfway to avoid accumulating error. The new initial conditions, i.e. $\boldsymbol{P}'_l, l = 1 \ldots m$, are obtained by a least square fitting.

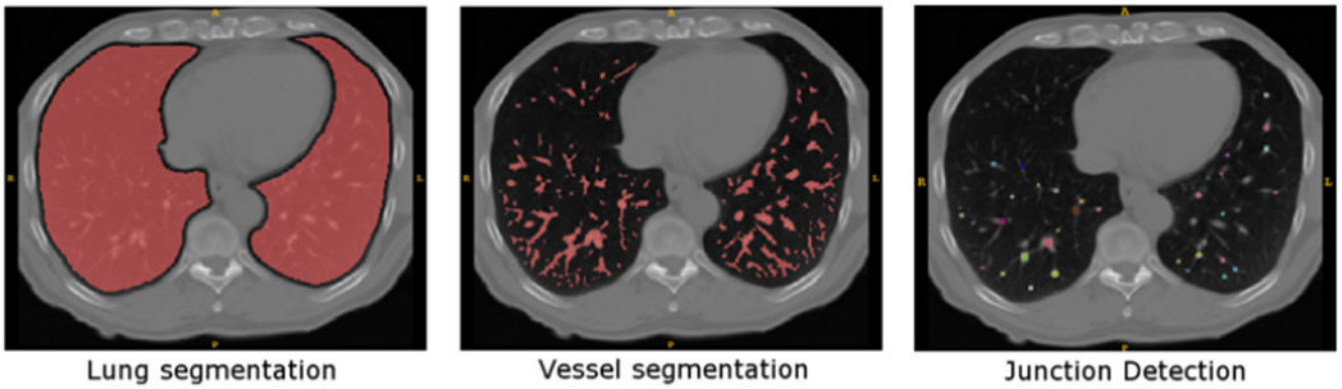

Fig. 5.

Results of lung segmentation, vessel segmentation, and junction detection in one axial slice. The segmented lung region and vessels are marked in red. The detected junctions are marked in different colors.

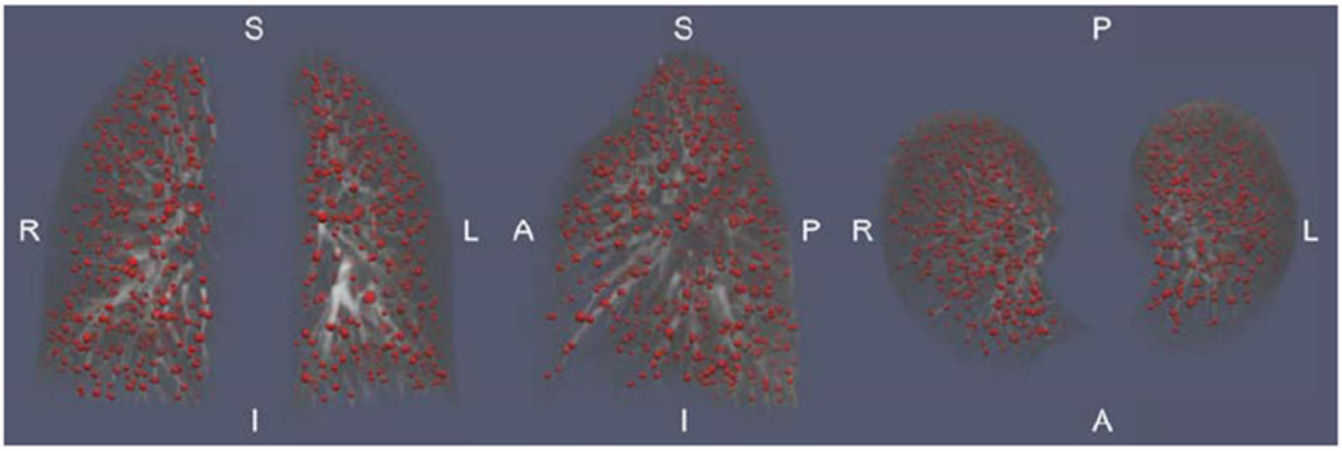

Fig. 6.

Detected Junctions in Case 1. The reference image masked by the lung region is volume-rendered in coronal, sagittal, and transverse views. Junctions are depicted as red spheres of radii corresponding to junction sizes. Letters indicate anatomical directions.

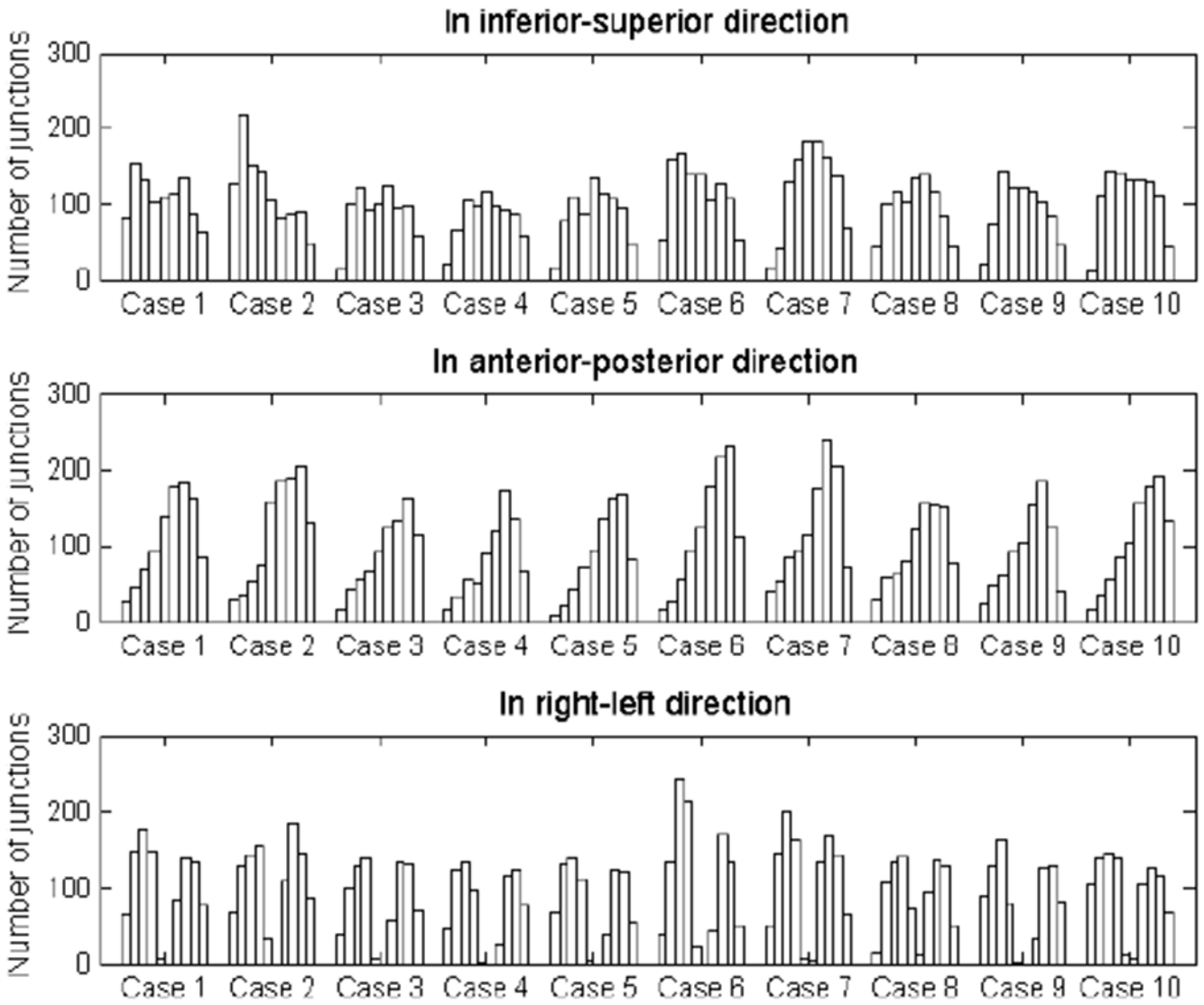

Fig. 7.
Spatial distributions of the detected junctions in three main directions for Cases 1-10.

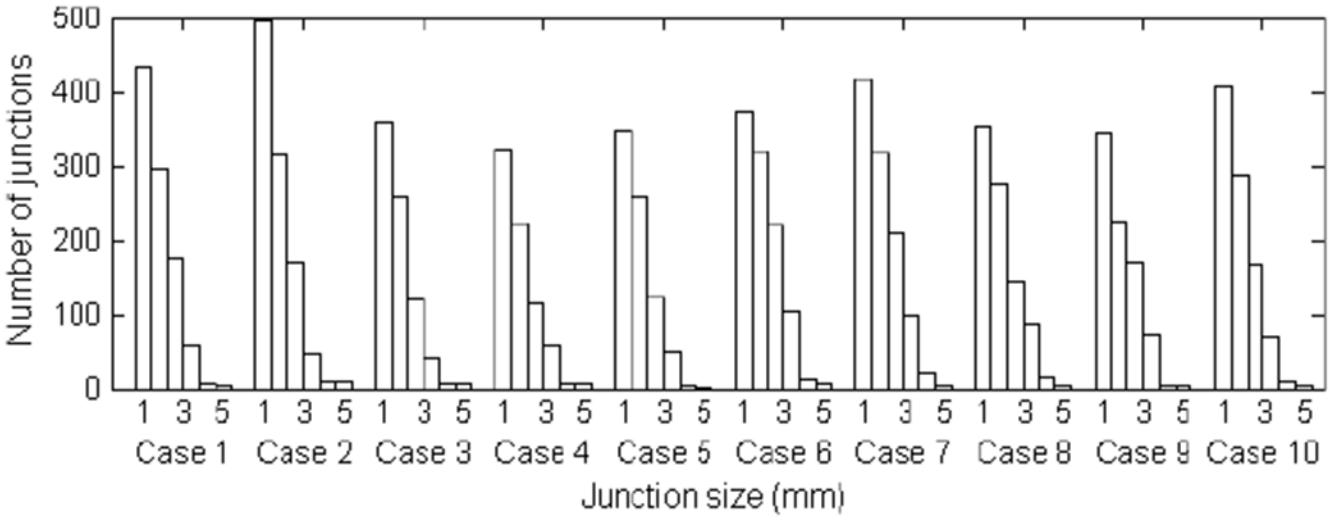

Fig. 8.
Sizes of the detected junctions for Cases 1-10.

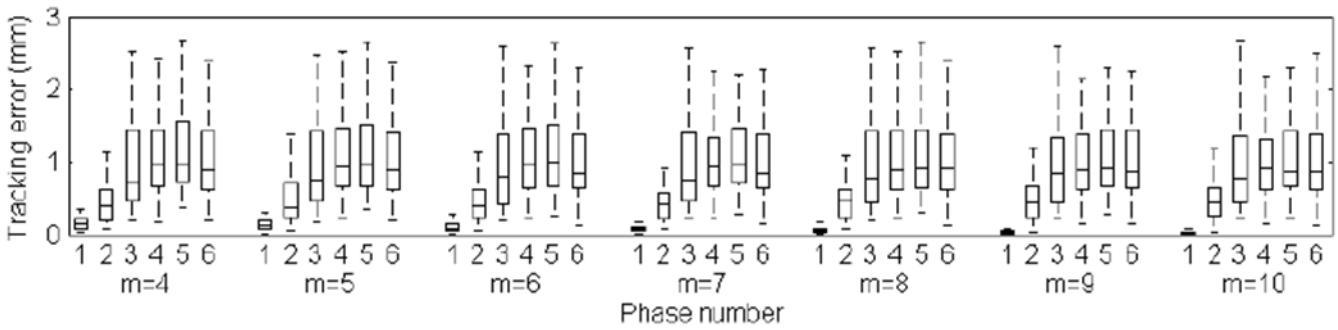

Fig. 9.

Tracking errors for Case 1 using different numbers of control vertices $m = 4…10$.

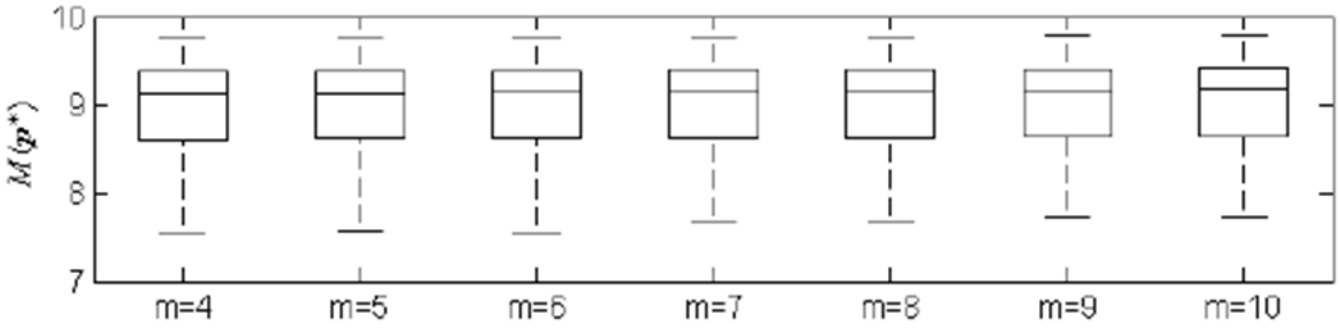

Fig. 10.
The Metric $M(\boldsymbol{p}^*)$ for $m = 4\ldots10$.

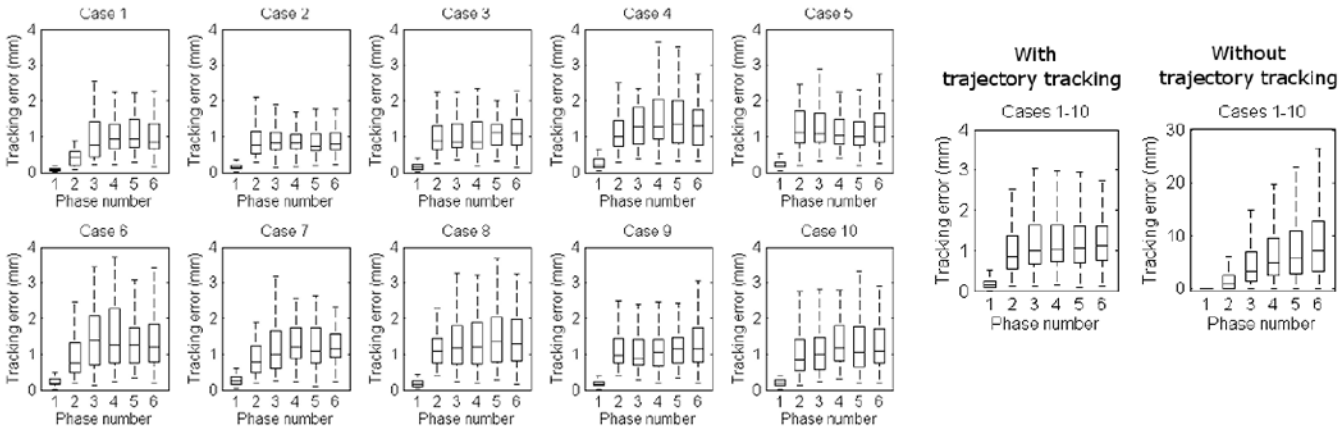

Fig. 11.

Tracking errors for each of Cases 1-10 based on 75 landmarks at Phases 1-6. The right two box plots (note that they are in different scales) compare the combined and baseline tracking errors of all ten cases with and without trajectory tracking.

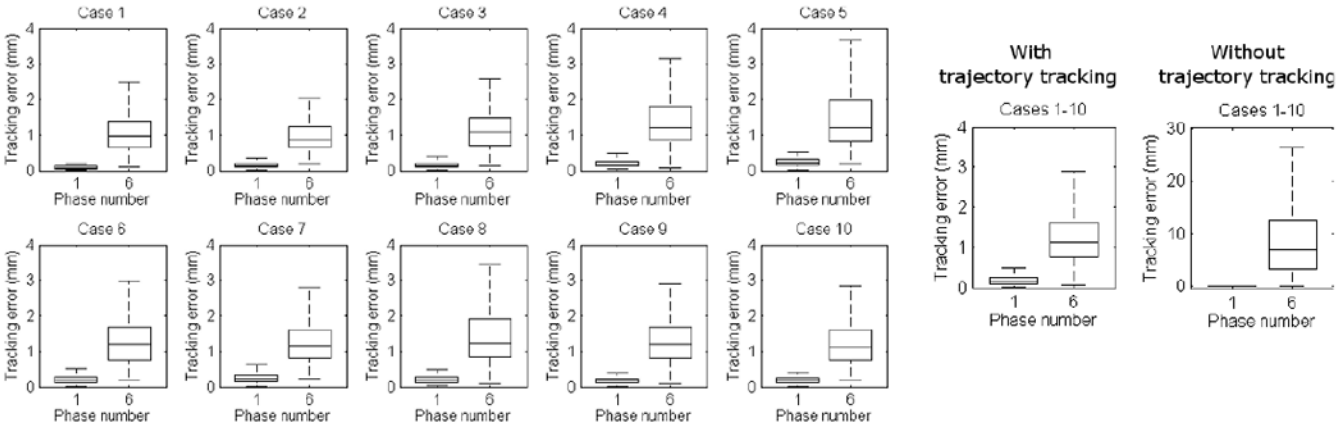

Fig. 12.

Tracking errors for each of Cases 1-10 based on 300 landmarks at Phases 1 and 6. The right two box plots (note that they are in different scales) compare the combined and baseline tracking errors of all ten cases with and without trajectory tracking.

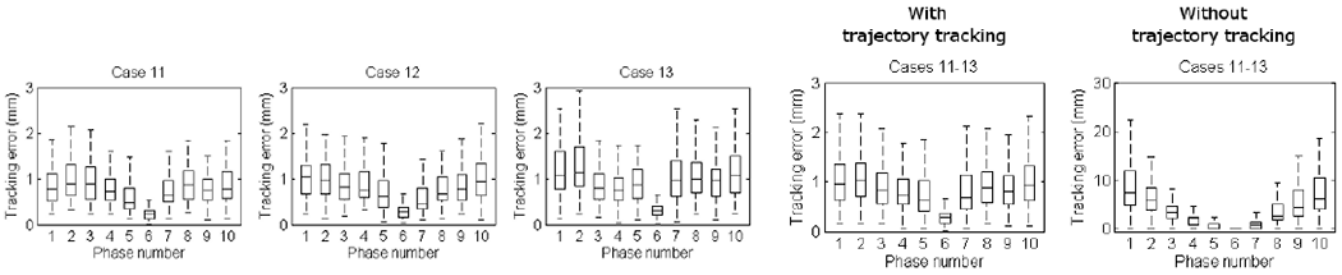

Fig. 13.

Tracking errors for each of Cases 11-13 based on 100 landmarks at all ten phases. The right two box plots (note that they are in different scales) compare the combined and baseline tracking errors of all three cases with and without trajectory tracking.

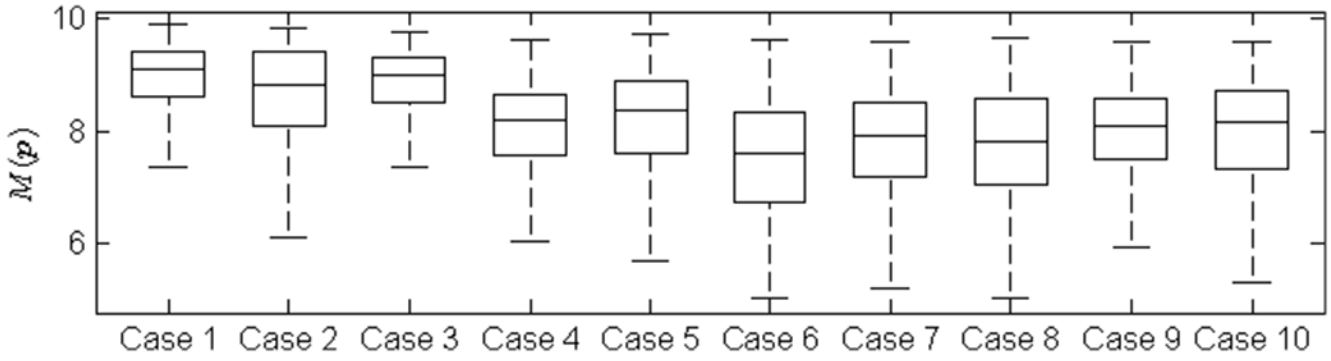

Fig. 14.
The metrics $M(\boldsymbol{p}^*)$ for Cases 1-10.

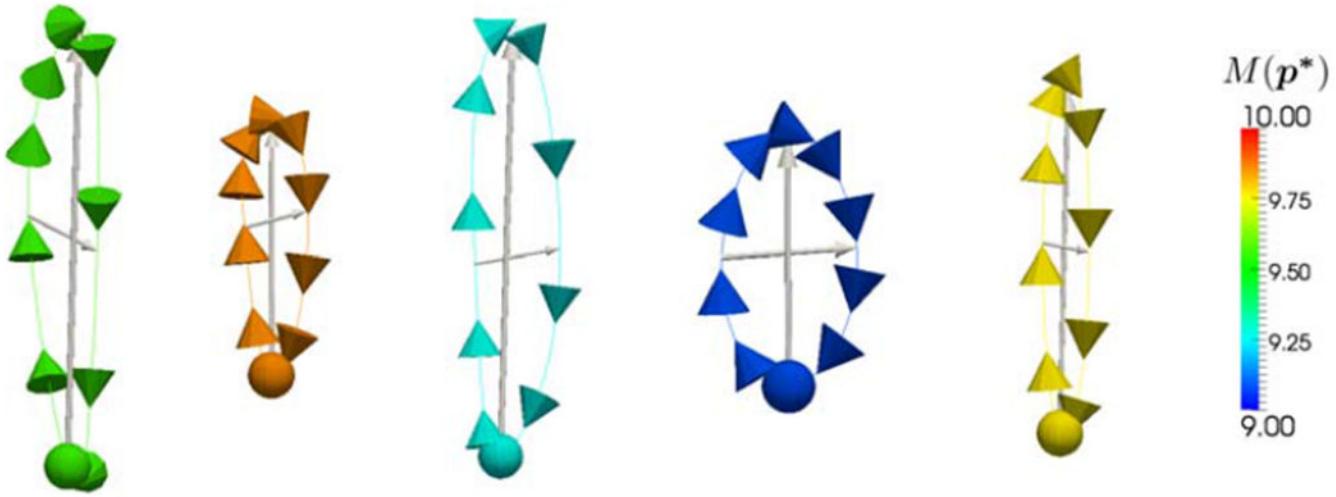

Fig. 15.

Example trajectories tracked in Case 1 with color-coded metric values $M(\boldsymbol{p}^*)$. The spheres mark the trajectory at Phase 1 and cones at successive phases with tips pointing to the direction.

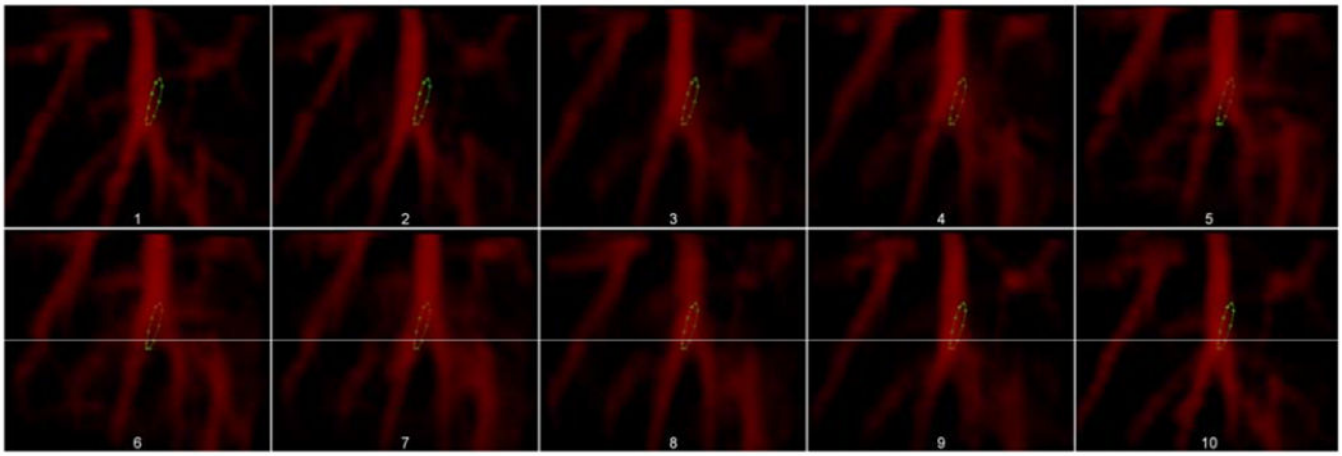

Fig. 16.

An example trajectory (in green) displayed with the volume-rendered junction structure (in red) in each phase. Note that the trajectory tracks the movement of the junction center.

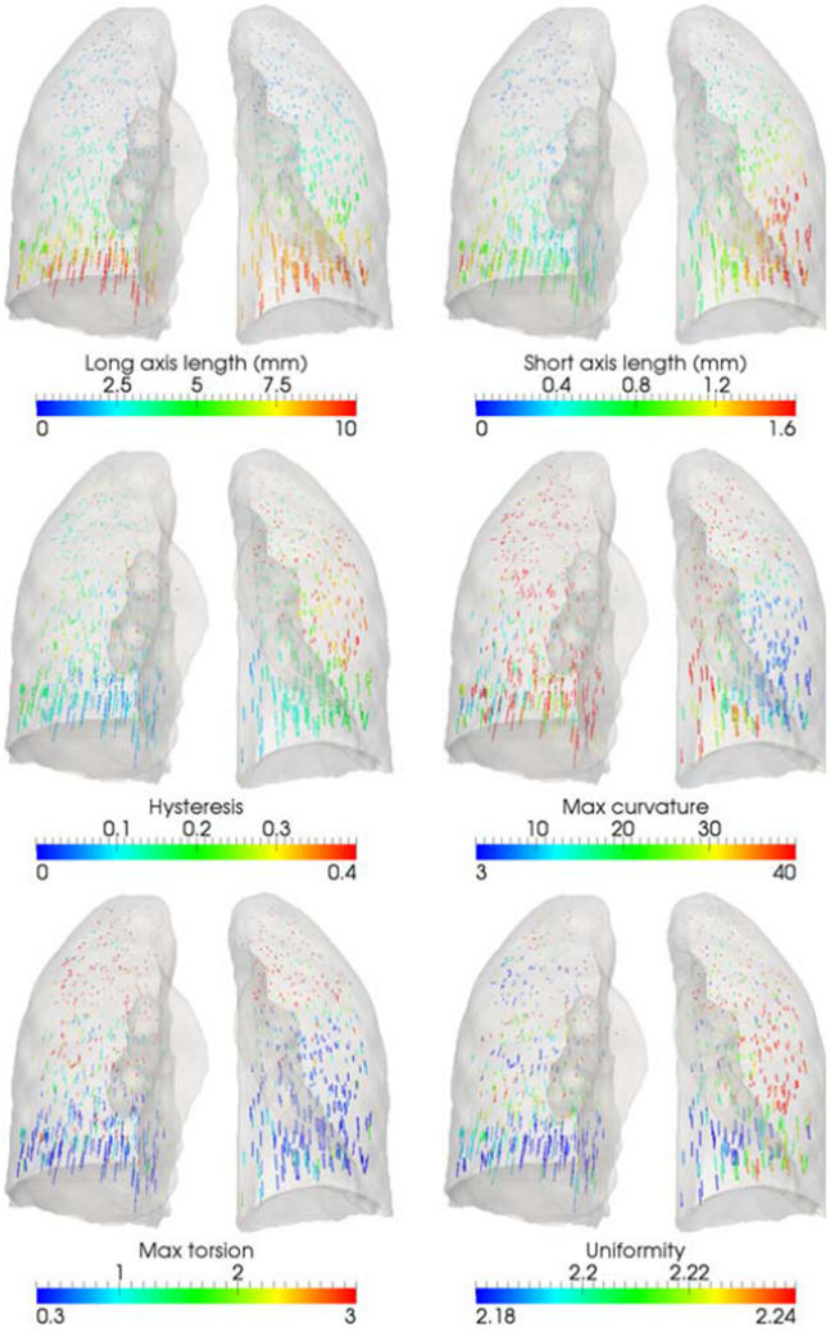

Fig. 17.
Descriptors of trajectories for Case 1.

TABLE I

A Summary of Junction Detection

Case	Number of Junctions	Segmentation Time (s)	Detection Time (s)	PPV by observer 1	PPV by observer 2	Agreement	Disagreement
1	977	73	45	91.6%	91.4%	98.8%	1.2%
2	1050	83	53	88.3%	83.0%	85.0%	15.0%
3	800	55	48	93.5%	93.1%	97.5%	2.5%
4	739	72	88	91.6%	88.9%	93.4%	6.6%
5	788	49	45	91.0%	89.2%	93.0%	7.0%
6	1043	98	106	87.6%	84.6%	93.3%	6.7%
7	1071	135	129	93.5%	93.7%	97.0%	3.0%
8	883	75	72	92.4%	91.0%	95.3%	4.7%
9	830	131	160	92.4%	91.1%	96.2%	3.8%
10	953	88	63	98.2%	97.3%	98.2%	1.8%

TABLE II

MEAN AND STANDARD DEVIATION OF THE TRACKING ERRORS (MM) USING DIR-LAB DATA

Case	Experiment with 75 landmarks								Experiment with 300 landmarks				
	Phase 1	**Phase 2**	**Phase 3**	**Phase 4**	**Phase 5**	**Phase 6**	**Phases 1-6**	**Phases 1-6***	**Phase 1**	**Phase 6**	**Phase 1,6**	**Phase 1,6***	**Phase 6****
1	0.10 (0.08)	0.50 (0.38)	0.99 (0.70)	1.13 (0.77)	1.17 (0.77)	1.05 (0.61)	0.82 (0.72)	0.93 (0.65)	0.09 (0.08)	1.05 (0.52)	0.57 (0.61)	0.99 (0.48)	0.97 (1.02)
2	0.16 (0.12)	0.95 (0.62)	0.97 (0.62)	0.89 (0.43)	0.91 (0.54)	0.92 (0.49)	0.80 (0.57)	0.89 (0.51)	0.15 (0.10)	1.00 (0.51)	0.58 (0.56)	0.96 (0.49)	0.86 (1.08)
3	0.17 (0.13)	1.08 (0.74)	1.05 (0.60)	1.03 (0.57)	1.15 (0.55)	1.14 (0.51)	0.94 (0.64)	1.05 (0.56)	0.17 (0.13)	1.15 (0.62)	0.66 (0.66)	1.11 (0.62)	1.01 (1.17)
4	0.24 (0.16)	1.16 (0.64)	1.44 (0.85)	1.60 (0.97)	1.48 (0.91)	1.36 (0.68)	1.21 (0.88)	1.40 (1.10)	0.20 (0.14)	1.39 (0.72)	0.79 (0.79)	1.49 (1.08)	1.40 (1.57)
5	0.23 (0.13)	1.53 (1.20)	1.32 (0.75)	1.35 (0.96)	1.19 (0.67)	1.38 (0.77)	1.16 (0.92)	1.27 (1.10)	0.24 (0.13)	1.43 (0.89)	0.83 (0.87)	1.37 (1.21)	1.67 (1.79)
6	0.23 (0.12)	1.05 (0.90)	1.45 (0.91)	1.55 (1.10)	1.30 (0.72)	1.45 (0.96)	1.16 (0.95)	-	0.22 (0.14)	1.32 (0.80)	0.76 (0.79)	-	1.58 (1.65)
7	0.24 (0.13)	0.95 (0.68)	1.21 (0.83)	1.44 (1.00)	1.39 (0.96)	1.32 (0.64)	1.08 (0.86)	-	0.24 (0.14)	1.29 (0.70)	0.76 (0.73)	-	1.46 (1.29)
8	0.20 (0.14)	1.17 (0.68)	1.44 (1.11)	1.47 (1.09)	1.44 (1.01)	1.48 (1.02)	1.20 (1.01)	-	0.21 (0.14)	1.39 (0.82)	0.79 (0.83)	-	1.77 (2.12)
9	0.17 (0.09)	1.16 (0.65)	1.21 (0.79)	1.13 (0.58)	1.24 (0.72)	1.31 (0.74)	1.04 (0.75)	-	0.17 (0.09)	1.32 (0.68)	0.75 (0.75)	-	1.19 (1.12)
10	0.20 (0.14)	1.12 (0.88)	1.24 (0.96)	1.52 (1.08)	1.20 (0.66)	1.26 (0.74)	1.09 (0.90)	-	0.18 (0.10)	1.24 (0.68)	0.71 (0.72)	-	1.59 (1.87)
1-10	0.19 (0.13)	1.07 (0.80)	1.23 (0.84)	1.31 (0.91)	1.24 (0.78)	1.26 (0.75)	1.05 (0.84)	-	0.19 (0.13)	1.26 (0.71)	0.72 (0.74)	-	1.25 (1.43)

*Results from Metz et al. 2011.

**Results from Castillo et al. 2010, which were obtained using more landmarks.

TABLE III

MEAN AND STANDARD DEVIATION OF THE TRACKING ERRORS (MM) USING CREATIS DATA

	Experiment with 100 landmarks												
Case	**Phase 1**	**Phase 2**	**Phase 3**	**Phase 4**	**Phase 5**	**Phase 6**	**Phase 7**	**Phase 8**	**Phase 9**	**Phase 10**	**Phases 1-5,7-10**	**Phase 1-5,7-10***	**Phase 1***
11	0.86 (0.44)	1.00 (0.44)	0.96 (0.54)	0.89 (0.94)	0.60 (0.35)	0.24 (0.13)	0.81 (0.44)	0.96 (0.54)	0.83 (0.39)	0.89 (0.45)	0.87 (0.54)	1.00 (0.69)	0.96 (0.57)
12	1.30 (1.01)	1.46 (1.16)	1.01 (0.84)	0.80 (0.36)	0.88 (0.40)	0.33 (0.16)	1.16 (0.79)	1.15 (0.91)	1.03 (0.69)	1.22 (0.74)	1.11 (0.83)	1.27 (1.09)	1.56 (1.34)
13	1.13 (0.83)	1.18 (0.97)	0.97 (0.72)	0.87 (0.39)	0.71 (0.41)	0.31 (0.17)	0.65 (0.47)	0.83 (0.50)	0.88 (0.48)	1.12 (0.78)	0.93 (0.67)	1.16 (1.15)	1.53 (1.70)
11-13	1.10 (0.81)	1.21 (0.93)	0.98 (0.71)	0.85 (0.62)	0.73 (0.40)	0.29 (0.16)	0.87 (0.63)	0.98 (0.69)	0.91 (0.54)	1.08 (0.69)	0.97 (0.70)	1.14 (1.00)	-

*Results from Vandemeulebroucke et al. 2011.