Location registration and recognition (LRR) for serial analysis of nodules in lung CT scans

Michal Sofka *, Charles V. Stewart

Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY 12180–3590, United States

ARTICLE INFO

ABSTRACT

In the clinical workflow for lung cancer management, the comparison of nodules between CT scans from subsequent visits by a patient is necessary for timely classification of pulmonary nodules into benign and malignant and for analyzing nodule growth and response to therapy. The algorithm described in this paper takes (a) two temporally-separated CT scans, I_1 and I_2, and (b) a series of nodule locations in I_1, and for each location it produces an affine transformation that maps the locations and their immediate neighborhoods from I_1 to I_2. It does this without deformable registration and without initialization by global affine registration. Requiring the nodule locations to be specified in only one volume provides the clinician more flexibility in investigating the condition of the lung. The algorithm uses a combination of feature extraction, indexing, refinement, and decision processes. Together, these processes essentially "recognize" the neighborhoods. We show on lung CT scans that our technique works at near interactive speed and that the median alignment error of 134 nodules is 1.70 mm compared to the error 2.14 mm of the Diffeomorphic Demons algorithm, and to the error 3.57 mm of the global nodule registration with local refinement. We demonstrate on the alignment of 250 nodules, that the algorithm is robust to changes caused by cancer progression and differences in breathing states, scanning procedures, and patient positioning. Our algorithm may be used both for diagnosis and treatment monitoring of lung cancer. Because of the generic design of the algorithm, it might also be used in other applications that require fast and accurate mapping of regions.

1. Introduction

Lung cancer is the deadliest of all cancers worldwide with an estimated 1.3 million deaths and 1.5 million new cases in 2007 (Garcia et al., 2007). This high prevalence makes systems for diagnosis and treatment monitoring particularly important, as reflected in the number of techniques being proposed by the computer vision and medical imaging communities (Ginneken et al., 2001; Sluimer et al., 2006). In the clinical workflow for lung cancer management, follow-up scans are necessary for timely classification of pulmonary nodules into benign and malignant (Fig. 1). Central to this workflow is the comparison of nodules between scans from subsequent visits by a patient.

In our work, we are concerned with the question of volumetric registration tools to assist in longitudinal monitoring of pre-cancerous and cancerous locations in the body. The idea is that nodule locations, having been selected within a CT scan using a combination of CAD and manual screening techniques, are automatically located in a scan taken at a different time. The regions surrounding the nodule and its corresponding location in the second scan are aligned (rigid or affine) with the first scan in preparation for display to the radiologist, or for subsequent quantitative analysis such as comparative measurement of nodule volumes (Fig. 2). Stating the problem more precisely: given are image volumes I_1 and I_2, a set of locations $\mathcal{L} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ from I_1, and a set of surrounding 3D neighborhoods, $\{\mathcal{N}(\mathbf{x}_k)\}$. The goal is to find, for each $\mathbf{x}_k$, the affine transformation $\mathbf{T}_k : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ that best aligns $\mathcal{N}(\mathbf{x}_k)$ with a region of I_2. Importantly, implicit in this problem statement is a *decision* about the neighborhood of I_2 that forms the co-domain of the transformation. In essence, the problem requires *recognizing* $\mathcal{N}(\mathbf{x}_k)$ as part of the registration process. We refer to this problem as the *Location Registration and Recognition* or LRR problem. A solution to this problem will become important with continuing increases in the number and size of CT scans and with increasing use of longitudinal studies that require alignment of scans (Sluimer et al., 2006).

Previous techniques for serial analysis of the lung nodules assume that the nodule locations are specified in *both* volumes in advance (Betke et al., 2003; Kawata et al., 2001; Reeves et al., 2006).

* Corresponding author.
E-mail addresses: sofka@cs.rpi.edu (M. Sofka), stewart@cs.rpi.edu (C.V. Stewart).
URLs: http://www.cs.rpi.edu/~sofka/ (M. Sofka), http://www.cs.rpi.edu/~stewart/ (C.V. Stewart).
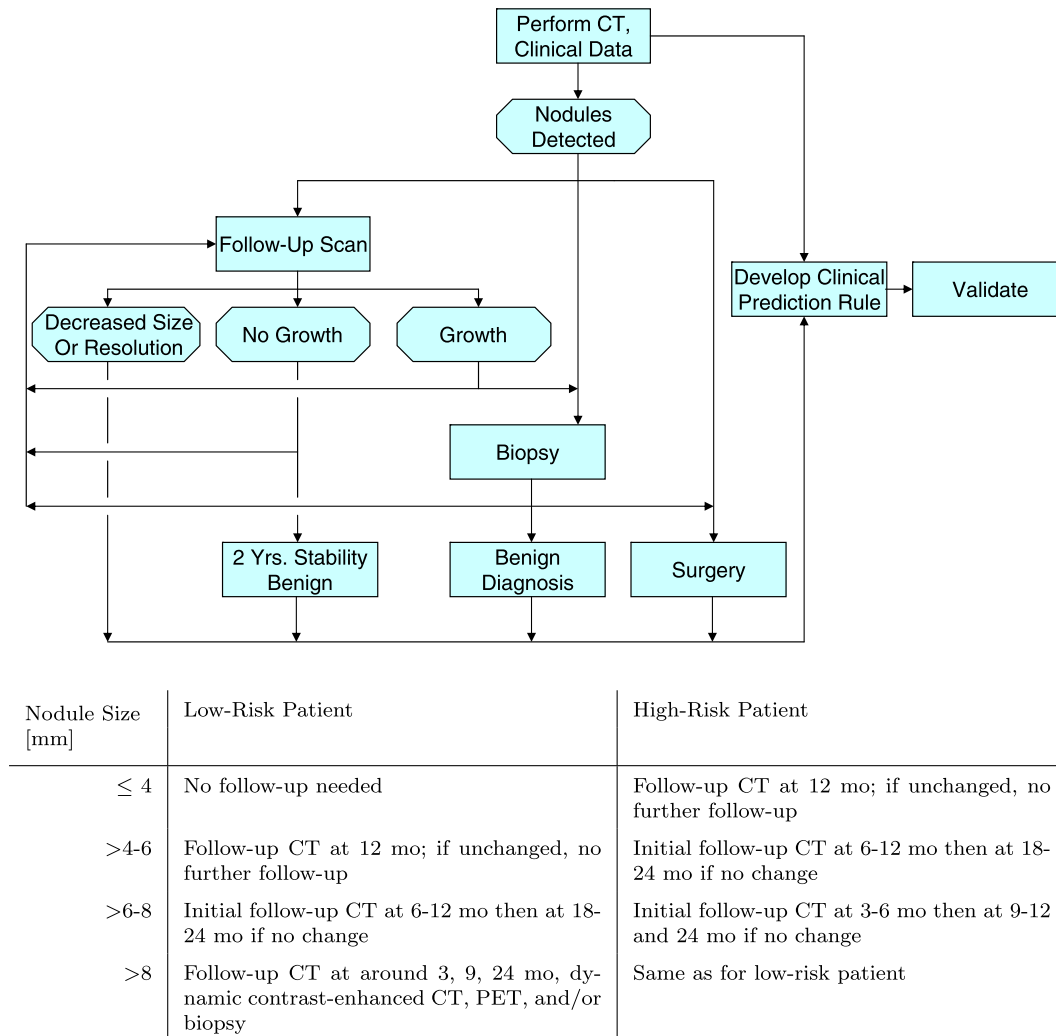
Fig. 1. Timely differentiation of pulmonary nodules into benign and malignant is important for diagnosis and treatment planning. The diagram (top) shows a clinical workflow for lung cancer management. Follow-up scans are necessary to determine treatment according to the growth of the nodules. The table (bottom) shows recommendations for follow-up and management of nodules detected incidentally at non-screening CT (MacMahon et al., 2005). Surgery (the treatment of choice for stages I and II non-small cell lung cancer Walter et al., 2007) may directly follow the initial scan and nodule discovery.

The LRR problem is more challenging since the nodules are indicated only in one of the volumes. This gives the clinician greater flexibility in analyzing the condition of the lung. The immediately-obvious solution to the LRR problem is application of global deformable registration techniques, followed by local affine registration to produce the final local transformations. We choose a different approach for several important reasons: (1) deformable registration tends to be slow, whereas a useful solution to the LRR problem will run at interactive speeds for many scans in succession, (2) deformable registration is not necessary because our interest is a sparse set of local alignments, and (3) a solution to our LRR problem could provide highly-reliable interest-point correspondences for deformable registration. This is not to say, however, that advances in the speed, reliability and accuracy of deformable registration will not eventually be useful in solving the LRR problem. Instead, in this paper, we propose and explore an alternative approach. An important part of our approach is a decision component to ensure with high-confidence that particular neighborhoods are correctly matched. This verification has been left to manual assessment in both the global deformable techniques (Matsopoulos et al., 2005; Sluimer et al., 2006) and nodule matching and registration algorithms (Shi et al., 2007; Wiemker et al., 2008).

Our method is feature-based, involving a preprocessing step to extract a variety of features distributed throughout the scan volumes. The main algorithm is applied separately for each location $\mathbf{x}_k$ and its surrounding neighborhood $\mathcal{N}(\mathbf{x}_k)$. The processing at each location consists of steps of hypothesis generation, hypothesis refinement, and decision. Hypothesis generation involves matching keypoint locations taken from inside $\mathcal{N}(\mathbf{x}_k)$ using 3D Shape-Context methods (Belongie et al., 2002; Mori et al., 2005). Hypotheses are rank-ordered and then refined using the Iterative Closest Point (ICP) algorithm (Besl and McKay, 1992; Chen and Medioni, 1992). The decision step computes a vector of measures on the estimated transformations and on the ICP correspondences and combines them to form a decision function based on a Support Vector Machine (SVM). All steps of the algorithm are designed to be robust to differences between scans caused by changes in scanning procedures, patient positioning, and physical changes in the lung caused by breathing state differences and progression of the cancer. This includes robustness to intensity differences, that may be caused by the introduction of a contrast agent, although we did not explicitly test this here.

Two types of experiments are presented here, one to guide specific design decisions and the second to validate the overall resulting algorithm. Comparisons are made to a global registration
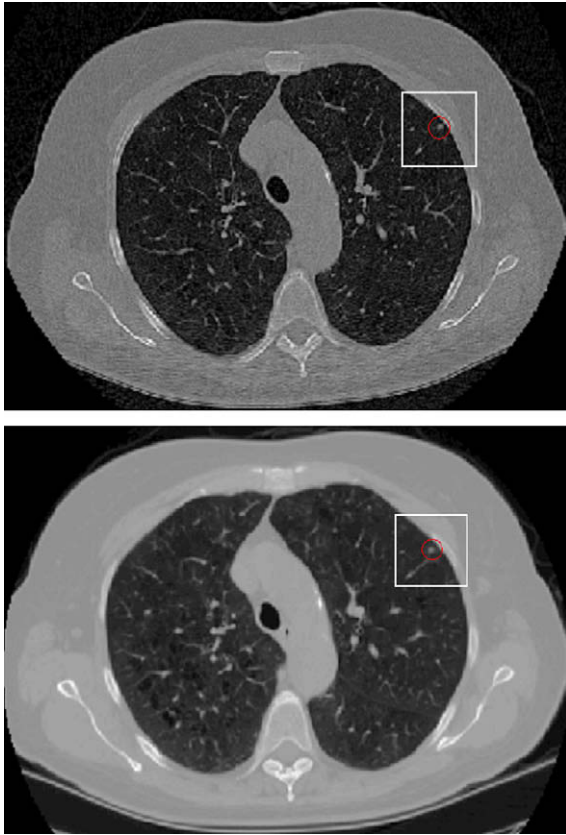
(2) Evaluation of shape-context methods for indexing and for generating an initial transformation estimate in CT scans.
(3) Development of a decision criteria for determining when a region of one scan has been correctly recognized and aligned against a second scan.
(4) Open source implementation of a clinically-relevant method for aligning regions containing pulmonary nodules.

Part of this work and an earlier version of the system have been reported in our conference publication (Sofka and Stewart, 2008). In this paper, we include experiments with pulmonary nodules, show how to handle multiple independent motions inside the alignment regions, add more evaluation of the decision step, present analysis of the affine transformation and the local deformation, compare against global nodule registration followed by local refinement, add multiresolution hierarchy to the keypoint indexing, and provide an open source implementation of the entire algorithm.

The paper is organized as follows. The relevant literature is reviewed in Section 2. The Location Registration and Recognition algorithm is proposed in Section 3. The experimental results are presented in Section 4. The paper concludes in Section 5.

2. Background

Our background overview focuses on feature-based registration methods related to the proposed LRR algorithm (Section 2.1), on nodule matching and registration approaches (Section 2.2), and on deformable registration algorithms (Section 2.3).

2.1. Feature-based registration

Feature-based registration starts with a preprocessing step to extract various image quantities, such as keypoints, descriptors, and features. Keypoint extraction algorithms have been widely studied in both medical imaging and computer vision (Hartkens et al., 2002; Lowe, 2004). Typically, keypoints are distributed throughout a volume at distinctive locations such that they can be detected at the same distinctive locations in another volume (repeatability). Perhaps the most widely-used 2D keypoints are detected as local extrema in the Laplacian-of-Gaussian pyramid in both spatial and scale dimensions (Yang et al., 2007; Lowe, 2004). Complementary to these keypoints are Harris corners (Mikolajczyk and Schmid, 2004). We use similar keypoints in our work.

Recent work in computer vision has emphasized the computation of descriptor vectors at keypoint locations. The vectors summarize the local image content and serve as signatures at keypoint locations. The descriptors are compared between images using indexing methods (Lowe, 2004). Given a descriptor vector of a keypoint detected in one volume, the goal is to find the most similar descriptor of a keypoint in another volume. The corresponding keypoints can then be used for computing an approximate local transformation at various keypoint locations throughout the volume, not just at predetermined anatomical points (Wörz and Rohr, 2006). Two sets of descriptor vectors are the SIFT (Lowe, 2004; Lai and Hua, 2008) and Shape-Context descriptors (Belongie et al., 2002; Frome et al., 2004; Mori et al., 2005; Liu and Chen, 2004), both of which emphasize the distribution of points and gradients. The primary differences between these descriptors, once gradient information is added to Shape-Contexts, are the spatial organization of the bins and the choice of points — all points in a region or just the edge points. In Kelman et al. (2007), we showed that it is possible to construct repeatable descriptors from points found by a feature extraction algorithm. In Section 3.2, we will show how

Fig. 2. Given a location with a surrounding neighborhood in one CT scan (top), the goal of the Location Registration and Recognition (LRR) algorithm is to find the corresponding location in another scan (bottom) and align the local regions surrounding the locations. The location in the bottom volume is not known in advance. Locations of non-calcified solid nodules are shown in the axial slices without aligning the volumes.

technique with local refinement (Betke et al., 2003; Kawata et al., 2001; Shi et al., 2007; Wiemker et al., 2008) and to a state-of-the-art deformable registration algorithm, Diffeomorphic Demons (Vercauteren et al., 2007). The overall results show that LRR is in most cases in agreement with the resulting deformation fields and it provides better alignment in several cases involving changes to the neighborhood surrounding $\mathbf{x}_k$. The algorithm finishes on average in 6 s per initial location and is able to report whether the registration produced an accurate alignment or not.

The new Location Registration and Recognition (LRR) algorithm may be used to align regions surrounding lung nodules for subsequent analysis of nodule growth (Betke et al., 2003; El-Baz et al., 2009; Kostis et al., 2003; Reeves et al., 2006). The nodule locations can be obtained from an automatic nodule detection algorithm (Agam et al., 2005; Kostis et al., 2003) applied to time-separated CT scans. Initializing LRR at nodules found only in one of the volumes can help discover misdetections. Alternatively, the LRR algorithm might be used in an interactive registration algorithm, in which user selects points (or nodules) in one of the volumes with a mouse click. This would greatly improve workflow in those systems where the user is currently required to find two corresponding points (Azar et al., 2006).

Our main contributions are as follows:

(1) The overall algorithm that maps a local image region from one scan to a corresponding region in another scan and aligns the two regions.

to use Shape-Contexts computed from generic features to generate an initial transform between two 3D image regions in lung CT scans.

In comparison to keypoints, features are much more dense and less distinctive and often include other attributes computed from the image intensities (Shen and Davatzikos, 2002; Xue et al., 2004). We use feature correspondences to drive the estimation process of image registration. To detect features in medical images, Rohr (1997) extended four 2D differential operators to 3D. Hartkens et al. (2002) presented a careful evaluation of these operators for the detection of anatomical landmarks in Magnetic Resonance (MR) and Computed Tomography (CT) images. In our previous work on registration of 2D images (Sofka et al., 2007; Yang et al., 2007), we successfully used generic edge-like and corner-like features (similar to the differential operators of Rohr (1997)). Here, we extend our generic features to 3D and use them to drive the registration.

One of the most widely-adopted approaches for correspondence matching and transformation estimation is the Iterative Closest Point (ICP) algorithm (Besl and McKay, 1992; Chen and Medioni, 1992). Many ICP enhancements have been proposed to improve convergence properties, increase efficiency, and enhance robustness (Rusinkiewicz and Levoy, 2001). In our algorithm, constraints generated by matching generic features are used in a robust ICP objective function to refine a transformation estimate between two local regions. Final decisions on the alignment accuracy are often left to manual assessment. Since this is not always practical, there is a need for reliable automatic methods. The technique in Brown et al. (2007) for aligning 2D images applies statistical tests of randomness based on the number of overall and consistent keypoint matches. The algorithm in Yang et al. (2007) combines measures of accuracy, consistency, and stability. Similar methods have not yet been applied to 3D medical imaging. In our approach, the measurement vectors of successful and unsuccessful alignments are used in a Support Vector Machine (SVM) to construct a separating hyperplane in the measurement space Duda et al., 2001, Chapter 5. SVMs have been used for feature selection and feature-based classification of normal and abnormal regions by Fan et al. (2005) and for pooling results of multiple registrations of 2D endoscopy images by Seshamani et al. (2009).

2.2. Nodule matching and registration

In serial data analysis — the application of the LRR algorithm — the goal is to locally align regions surrounding two corresponding nodules. Some techniques assume that the corresponding nodule pairs are known (Reeves et al., 2006), others determine the pairing by nodule matching (Betke et al., 2003) or interactively (Kawata et al., 2001). The primary disadvantage of these techniques is that they require the nodule locations to be known in *both* volumes in advance. Conversely, LRR starts from nodules specified in only *one* of the volumes and the algorithm automatically finds the corresponding nodules in the other volume.

In the work by Reeves et al. (2006), the alignment of two pre-determined corresponding nodules is refined by minimizing mean-squared difference of the intensity volumes of interest. The paper proposes a nodule segmentation method for nodule growth analysis and techniques for removing pleural surfaces and vessels attached to nodules. These steps could be applied after the LRR algorithm to quantify nodule growth.

The algorithms for automatically pairing nodules begin by global registration of lung volumes (Kawata et al., 2001) or segmented structures (Betke et al., 2003; Shen et al., 2002; Shi et al., 2007). In Kawata et al. (2001), a technique for analyzing evolution of pulmonary nodules over time starts by manually specifying two corresponding regions of interest. The initial global rigid registration

is followed by local refinement and displacement field computation. In the algorithm of Betke et al. the global transform is used to find nodule correspondences by mapping each nodule from one volume and finding the closest nodule in the other volume. The approach in Okada and Huang (2007) finds a single point-wise correspondence to a given point by relative configurations to precomputed stable features. In contrast to LRR, neither of the two algorithms (Betke et al., 2003; Okada and Huang, 2007) aligns the local regions.

Very few techniques attempt to find and align the corresponding nodule regions automatically using nodules specified only in one of the scans. In Shi et al. (2007), the alignment of a volume of interest around a candidate nodule is initialized by registration of the ribs and refined by template matching. The technique might fail due to poor template matching when a nodule is missing in one of the scans or due to inconsistent rib segmentation when one of the scans does not cover the entire lung. In Wiemker et al. (2008), the corresponding nodule locations are found by a coordinate transform derived from the relative positions of segmented lung volumes. Only translation parameters are computed (similarly to Shen et al. (2002)) and the approach fails in the presence of large cardiac motion or in neighborhoods with changes due to lung motion relative to the rib cage. The final decision on alignment accuracy in both of these algorithms (Shi et al., 2007; Wiemker et al., 2008) is done manually.

The major advantage of the proposed Location Registration and Recognition (LRR) algorithm is that it starts from nodule locations in one of the volumes without pairing the nodules first. It does not rely on global registration of the lung volumes or segmented structures, which might fail when one of the scans does not cover the entire lung. The proposed solution to the LRR problem also includes an automatic decision of the final alignment accuracy which has been previously left to manual assessment.

2.3. Deformable registration

Global deformable registration algorithms provide a mapping at each image location. This way, these algorithms solve the Location Registration and Recognition (LRR) problem for all locations in a given volume. Such a solution is expensive and not necessary, since we are only interested in a sparse set of locations. In this paper, we use a result of a deformable registration technique, Diffeomorphic Demons, for an experimental analysis of the LRR algorithm. Diffeomorphic Demons gave the most accurate results in a comparative evaluation on the inter-subject thoracic registration (Urschler et al., 2007). The algorithm is four times more computationally expensive than the original Demons algorithm (Thirion, 1998), which was second in this evaluation.

Deformable registration techniques have been widely used for the alignment of lung CT data. Boldea et al. (2003) ran experiments with the Demons algorithm using breath holding techniques. Matsopoulos et al. (2005) use correspondences between points around bone structures to fit radial basis functions to the entire scans of patients with non-small cell lung cancer. Other techniques have been based on volumetric B-splines (Camara et al., 2007; Gorbunova et al., 2008; Kaus et al., 2004; Murphy et al., 2008; Vik et al., 2008; Yin et al., 2009), hybrid of intensity and feature-based approaches (Urschler et al., 2006b), alignment of surfaces (El-Baz et al., 2009; Kaus et al., 2004; Vik et al., 2008), and thin-plate spline (TPS) mappings (Urschler et al., 2006a). Some of these algorithms are initialized by matching pulmonary (Li et al., 2003) or bone (Matsopoulos et al., 2005) structures. Others use point correspondences during the registration to compute the final alignment (Urschler et al., 2006a). Many of these techniques could benefit from the LRR algorithm that we

propose. One possibility is to use LRR to initialize a global deformable registration.

The image-wide mapping provided by the global deformable registration is not necessary to solve the LRR problem. One difficulty of the deformable techniques is that local constraints, such as those that ensure diffeomorphism (Vercauteren et al., 2007; Cootes et al., 2004), rigidity, smoothness (Vercauteren et al., 2007; Cahill et al., 2009; Gorbunova et al., 2008; Pock et al., 2007; Yin et al., 2009), and mass and weight preservation (Gorbunova et al., 2008; Yin et al., 2009) need to be included. Since the LRR algorithm provides a local affine mapping, it is possible to directly evaluate the local regions by a quantitative comparison. A set of local alignments provided by LRR might be used to initialize deformable registration.

3. Location registration and recognition

The LRR algorithm finds the alignment of regions surrounding the locations $\mathcal{L} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ in I_1 with the corresponding regions in I_2 using correspondences between image features. Denoting $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ as the image-wide mapping function, the following distance-based energy functional is at the minimum when all regions are correctly aligned:

$$\mathcal{F}_{def} = \sum_{(\mathbf{p}_i, \mathbf{q}_j) \in \mathcal{C}} d(f(\mathbf{p}_i), \mathbf{q}_j), \tag{1}$$

where $\mathcal{C}$ is a set of all point correspondences, d is the distance between a fixed feature point $\mathbf{q}_j$ from I_2 and a moving feature point $\mathbf{p}_i$ from I_1, mapped with the function f. Abusing the notation, $\mathbf{p}_i$ and $\mathbf{q}_j$ represent both the feature points and their locations. The correspondences are found by matching moving feature points $\{\mathbf{p}_i\}$ to the corresponding fixed feature points $\{\mathbf{q}_j\}$ while rejecting outliers (Yang et al., 2007; Sofka et al., 2007; Chui et al., 2004). Finding the function f, which is generally non-linear, is an ill-posed problem. Moreover, the image-wide mapping is not necessary for solving the LRR problem. The LRR algorithm uses the first-order approximation at $\mathbf{x}_k$:

$$f(\mathbf{p}_i) \approx f(\mathbf{x}_k) + \mathbf{J}(\mathbf{x}_k)(\mathbf{p}_i - \mathbf{x}_k) \tag{2}$$
$$= \mathbf{b}_k + \mathbf{A}_k(\mathbf{p}_i - \mathbf{x}_k)$$
$$= \mathbf{T}_k(\mathbf{p}_i; \theta_k). \tag{3}$$

The Jacobian $\mathbf{J}$ of the function f is evaluated at the location $\mathbf{x}_k$ and becomes locally the affine component $\mathbf{A}_k$ of the transformation $\mathbf{T}_k$. The parameters θ_k are the parameters of the affine transformation $\mathbf{T}_k$, i.e. elements of the matrix $\mathbf{A}_k$ and the vector $\mathbf{b}_k$, mapping the region $\mathcal{N}(\mathbf{x}_k)$ surrounding the location $\mathbf{x}_k$. The region size is fixed for all locations and its size is determined experimentally for the particular application. The energy (1) can now be rewritten:

$$\mathcal{F}(\mathbf{x}_1, \ldots, \mathbf{x}_N; \theta_1, \ldots, \theta_N) = \sum_{k=1}^{N} \sum_{(\mathbf{p}_i, \mathbf{q}_j) \in \mathcal{C}_k} d(\mathbf{T}_k(\mathbf{p}_i; \theta_k), \mathbf{q}_j)$$
$$= \sum_{k=1}^{N} \mathcal{E}(\theta_k), \tag{4}$$

where $\mathcal{C}_k$ is a set of correspondences between feature points $\{\mathbf{p}_i\}$ in the neighborhood of $\mathbf{x}_k$ and feature points $\{\mathbf{q}_j\}$ in the neighborhood of $\mathbf{x}'_k = \mathbf{T}_k(\mathbf{x}_k)$. For the purposes of the current work, each local affine transform is independent of the others. Therefore, each $\mathcal{E}(\theta_k)$ is applied separately. The overall system diagram is depicted in Fig. 3. The initialization estimation step provides an approximate transform $\mathbf{T}_k$ which is valid in the neighborhood of $\mathbf{x}_k$. Subsequently, this estimation transformation is refined into an accurate alignment using correspondences $\mathcal{C}_k$ between image points. The decision step determines whether the final transform accurately aligns the two local regions. The algorithm is outlined in Fig. 4, with the details described below.

3.1. Feature extraction

Features are detected at the surface boundaries of anatomical regions, intersections of organs, and at the centers of tubular structures (e.g. airways and vessels). This way, the features represent the structural content of the image volumes and are used in all aspects of the LRR process. In order to ensure that features are available to drive the LRR algorithm anywhere in the volume, the emphasis is to extract multiscale features distributed throughout the volumes (coverage) without missing anatomical structures (completeness). During the alignment process, we rely on robust statistical estimation to ensure that correspondences generated when a feature in one volume has no analog in the other are ignored.

The datasets are often anisotropic, i.e. they have sample spacing in the X and Y dimensions different from the spacing in the Z dimension. Therefore, they are processed in physical coordinates and resampled to be isotropic when necessary (using implementation provided in ITK Ibáñez et al., 2003). At each voxel location $\mathbf{x}$, the intensity gradient $\nabla I(\mathbf{x})$ is computed. The outer product (auto-correlation) matrix $\mathbf{M}(\mathbf{x})$, also known as the structure tensor (Jähne, 1993), is then computed as a Gaussian-weighted average of $(\nabla I(\mathbf{x}))(\nabla I(\mathbf{x}))^{\top}$ over the neighborhood of $\mathbf{x}$. A strength is assigned to each point: $m(\mathbf{x}) = \text{trace}(\mathbf{M}(\mathbf{x}))$ (Fig. 5(a)). Other measures could be used as well (Hartkens et al., 2002). The features are locally pruned by discarding those that have strength smaller than $\mu_m + k\sigma_m$, $k = 1$, where μ_m and σ_m are median and std. dev. strength values computed in overlapping volume neighborhoods ($30 \times 30 \times 30$ mm). Fig. 5(b) shows features obtained after this step.
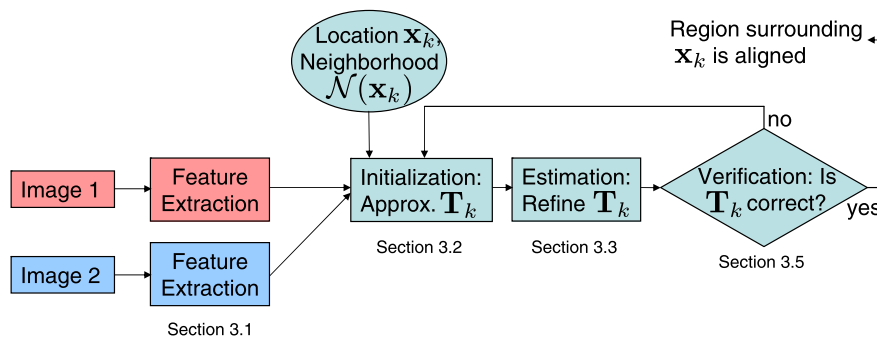

Fig. 3. Diagram of the Location Registration and Recognition system. The initial transform $\mathbf{T}_k$ maps the region $\mathcal{N}(\mathbf{x}_k)$ surrounding the location $\mathbf{x}_k$ from image I_1 into image I_2. The transform $\mathbf{T}_k$ is refined into accurate alignment in the estimation stage using correspondences between image features. If the decision step determines that the alignment is correct, the algorithm finishes. Otherwise, a new initialization is generated.

Input: Image volumes I_1 and I_2 and a set of locations $\mathcal{L} = \{\mathbf{x}_1, ..., \mathbf{x}_N\}$, pre-selected manually or semi-automatically from I_1.

Output: Set $\{\hat{\boldsymbol{\theta}}_1, ..., \hat{\boldsymbol{\theta}}_N\}$, where each $\hat{\boldsymbol{\theta}}_k$ is a vector of affine parameters or the zero-vector $\mathbf{0}$ indicating that no transformation was found.

Preprocessing: Extract features (Sec. 3.1), extract keypoints, and compute descriptor vectors (Sec. 3.2) in both volumes. Compute over-segmentations of the volumes. Form a k-d tree storing all descriptors from I_2. These steps may be done offline. (A) **Repeat** for each pre-selected location $\mathbf{x}_k$:

(1) Establish a local 3D neighborhood $\mathcal{N}(\mathbf{x}_k)$ surrounding location $\mathbf{x}_k$.

(2) Gather keypoints from I_1 within the neighborhood $\mathcal{N}(\mathbf{x}_k)$ into a set $\mathcal{U}_k$ (Sec. 3.2).

(3) Indexing: for each keypoint in $\mathcal{U}_k$, search the k-d tree to find the keypoint in I_2 with the closest descriptor. The two keypoints form a hypothesized match (Sec. 3.2).

(4) Order the hypothesized matches by increasing descriptor distance. Consider the best $M = 20$.

(5) Compute an initial rigid transform for each of the M hypotheses by aligning local keypoint coordinate axes.

(6) Reorder the list of hypotheses to put the initializations that provide the best initial alignments at the top.

(7) **Repeat**

    (a) Pick the next match from the list and its computed initial rigid transform. Initialize the affine mapping function $\mathbf{T}_{k,r}$, $r = 1, ..., M$ and establish 3D feature neighborhoods $\mathcal{M}(\mathbf{x}_k)$ and $\mathcal{M}(\mathbf{x}'_k)$, $\mathbf{x}'_k = \mathbf{T}_{k,r}(\mathbf{x}_k)$.

    (b) Gather features in the neighborhoods $\mathcal{M}(\mathbf{x}_k)$ and $\mathcal{M}(\mathbf{x}'_k)$ belonging to the same anatomical regions as determined by the over-segmentations.

    (c) Estimate affine transformation parameters $\hat{\boldsymbol{\theta}}_k$ of the mapping function $\mathbf{T}_k$ using feature correspondences in a robust generalization of ICP (Sec. 3.4).

    (d) Apply the decision classifier (Sec. 3.6): If $\hat{\boldsymbol{\theta}}_k$ is *correct*, save it and proceed to the next pre-selected location (A). Otherwise continue to the next match (Step 7a).

    Until all matches from the list have been processed.

(8) No transformation has been found for $\mathbf{x}_k$, so set $\hat{\boldsymbol{\theta}}_k = \mathbf{0}$.

Fig. 4. LRR algorithm outline.

Features are classified as 1D (sheet-like), 2D (tube-like) or 3D (corner-like) according to the number of maxima of the strength measures along each of the feature orientation axes. The local orientation is assigned to each feature-based on the eigen-decomposition of $\mathbf{M}(\mathbf{x})$: $\mathbf{M}(\mathbf{x}) = \sum_{d=1}^{3} \lambda_d(\mathbf{x}) \boldsymbol{\Gamma}_d(\mathbf{x}) \boldsymbol{\Gamma}_d(\mathbf{x})^\top$, with $\lambda_1(\mathbf{x}) \leqslant \lambda_2(\mathbf{x}) \leqslant \lambda_3(\mathbf{x})$. The decomposition will produce eigenvectors $\{\boldsymbol{\Gamma}_d(\mathbf{x})\}$ which are defined up to a sign. In order to assign an orientation to each feature, sign ambiguities are resolved by choosing the sign that makes the derivative in direction $\boldsymbol{\Gamma}_d(\mathbf{x})$ positive. The resulting vectors $\{\widetilde{\boldsymbol{\Gamma}}_d(\mathbf{x})\}$ define the local orientation axes of each feature.

According to the feature type, non-maximum suppression is applied in all three dimensions when the feature is corner-like, in 2D (in a plane spanned by $\widetilde{\boldsymbol{\Gamma}}_2$ and $\widetilde{\boldsymbol{\Gamma}}_3$) when the feature is tube-like, and in 1D (along direction of the eigenvector $\widetilde{\boldsymbol{\Gamma}}_3$ when the feature is sheet-like). Surviving locations are interpolated to subvoxel accuracy by finding a peak of the parabolic surface fitted to the local intensity neighborhood. Therefore, features can be located between voxels, rather than strictly at voxel coordinates. See Fig. 5(c) for an example.

One more filtering step is applied to select the final set of features. Candidate features are sorted by $m(\mathbf{x})$ values. Then, they are accepted one-by-one starting with the highest strength and continuing until the list is exhausted. When a feature (at location $\mathbf{x}$) is accepted, each remaining point whose strength is less than $\tau m(\mathbf{x})$ ($\tau = 0.8$) and whose position is within distance Ω of $\mathbf{x}$ ($\Omega = 4$ mm, currently) is eliminated. This provides a semi-dense set of features (31 per cm³ on average). The filtering parameters were found experimentally with the emphasis on coverage and completeness (see the discussion at the beginning of this section). An example of a feature extraction result is shown in Fig. 5(d). Fig. 16 shows features as part of a checkerboard mosaic of *X–Y* slices taken from aligned volumetric neighborhoods.

3.2. Keypoints and indexing

Corner-like features are used to form keypoints and compute descriptors for matching and indexing. This section gives an overview of this process. Section 3.3 shows how keypoint indexing is used to generate initial matches for a nodule (or any other location of interest).

Keypoints are found by repeating the foregoing process with $\Omega = 6$ mm. The larger spacing for filtering the points reduces the
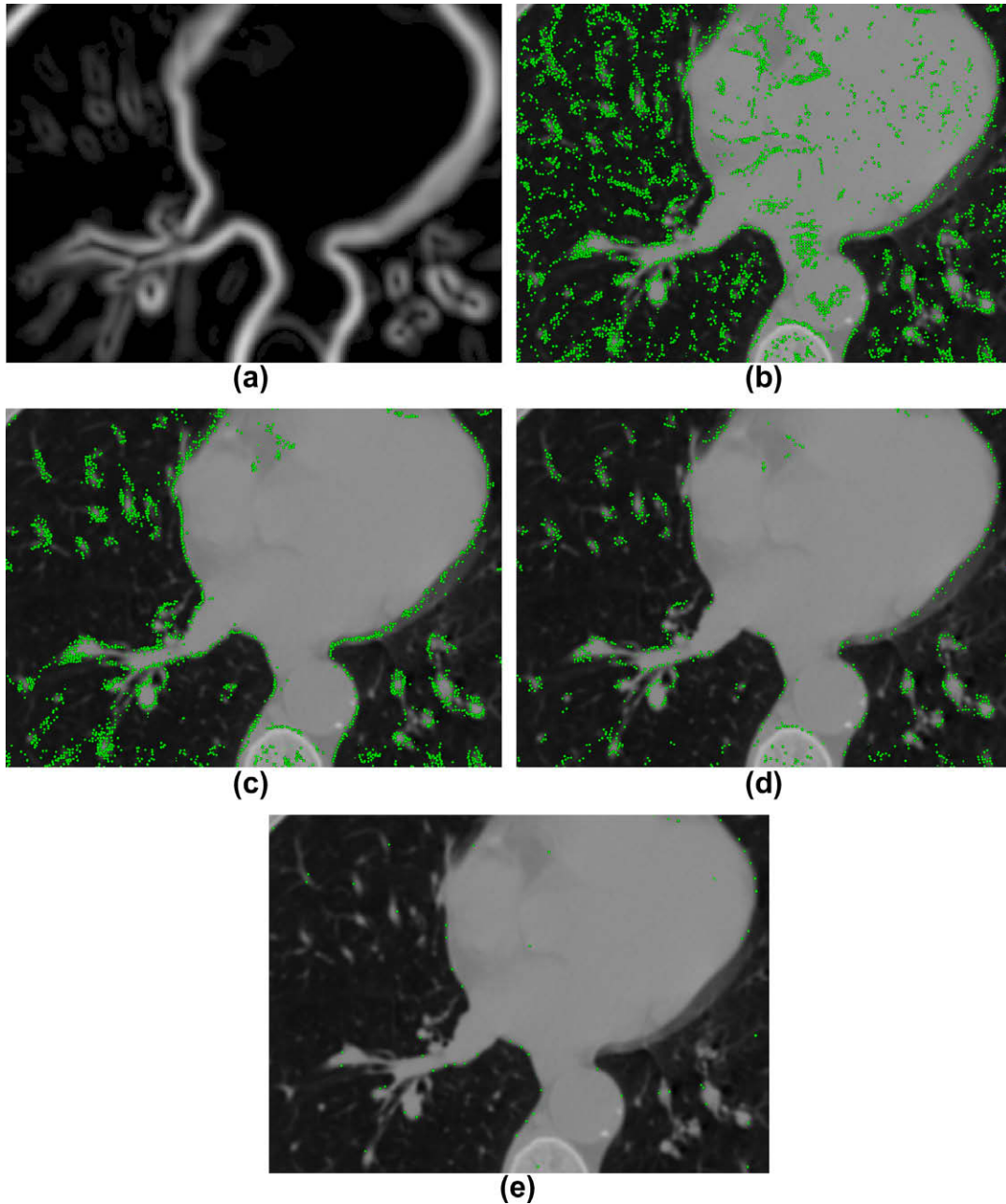
Fig. 5. An example showing the feature and keypoint extraction steps on an axial slice of a source volume centered on the lung. First, the strength is computed as the trace of the outer product of the intensity gradient at each voxel location (a). The locations with the highest strength after local contrast filtering are shown in (b). Second, non-maximum suppression is applied and feature locations are refined to subvoxel accuracy (c). Finally, the feature extraction result is obtained after spatial filtering (d). The features are widely distributed throughout the image. Occasionally, features are missing along the boundaries of structures because they are detected in 3D and might get suppressed in this slice while being present in the neighboring slices. Corner-like features are spatially filtered once more with a larger filtering radius to produce keypoints (e). On average, 800,000 features and 50,000 keypoints are extracted per volume.

size of the feature set, making matching more efficient. Keeping only corner-like features (Fig. 5(e)), an associated descriptor vector is computed for each feature to form a keypoint.

Descriptors are computed by adapting the Shape-Contexts (Belongie et al., 2002; Mori et al., 2005) approach, previously used for 2D images and range data (Frome et al., 2004). Specifically, at each keypoint location $\mathbf{x}_m$, we form a local spherical coordinate system centered at $\mathbf{x}_m$. The region radius is $r = 30$ mm, which we found to be a good balance among cost, distinctiveness, and location accuracy. Each feature with the location $\mathbf{x}_i$ within r from $\mathbf{x}_m$ has a strength $m(\mathbf{x}_i)$ and an orientation given by the basis vectors $\{\widetilde{\boldsymbol{\Gamma}}_d(\mathbf{x}_i)\}$ (Section 3.1). The coordinate system of a keypoint at location $\mathbf{x}_m$ is formed as the weighted average of the feature orientations: $\boldsymbol{\Theta}_d(\mathbf{x}_m) = \sum_i m(\mathbf{x}_i)\widetilde{\boldsymbol{\Gamma}}_d(\mathbf{x}_i)/\sum_i m(\mathbf{x}_i)$, for each dimension d (Fig. 6).

The descriptor is represented relative to the keypoint orientation (Lowe, 2004). In the local spherical coordinate system, four angular bins and four radial bins are formed, together with an additional central bin (Fig. 7). Major orientation inside each bin is computed as the weighted average of feature orientations (Fig. 8). The major orientations are concatenated into a 195-component descriptor vector ([4 elevation $\times$ 4 azimuth $\times$ 4 radial + 1 central] $\times$ 3 dim. normal = 195 bins). This is nearly the size of 2D image descriptors (Lowe, 2004) and much smaller than descriptors
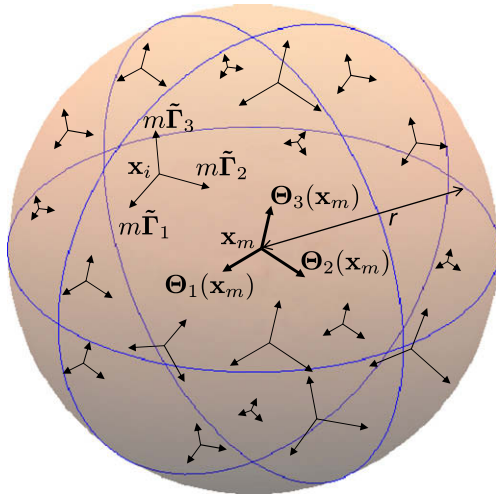
Fig. 6. Local coordinate system centered at a keypoint location $\mathbf{x}_m$ is formed from features with locations $\{\mathbf{x}_i\}$ within the distance r from $\mathbf{x}_m$. The coordinate system $\Theta(\mathbf{x}_m)$ is established as the average of feature orientations $\{\tilde{\Gamma}(\mathbf{x}_i)\}$ weighted by the feature strengths $\{m(\mathbf{x}_i)\}$.

Fig. 8. The feature orientations in the rotated coordinate system, i.e. relative to the keypoint orientation, are denoted $\{\tilde{\Gamma}'_d(\mathbf{x}_i)\}$. Features that fall into each bin and the neighboring bins along each dimension are gathered into a set $\mathcal{B}_k$. The weighted average of the feature directions is then computed as $\eta = \frac{1}{|\mathcal{B}_k|}\sum_{\mathbf{x}_i \in \mathcal{B}_k} b_i \tilde{\Gamma}'_3(\mathbf{x}_i)/\sum_{\mathbf{x}_i \in \mathcal{B}_k} b_i$. The direction $\tilde{\Gamma}'_3$ was derived from the eigenvector corresponding to the largest eigenvalue of the intensity outer product matrix at the feature location (Section 3.1). The weight $b_i = 1$ at the center of the bin and linearly decreases to 0 at a distance equal to the bin size along each dimension. This reduces the boundary effects since each feature can contribute to multiple bins with the weight distributed by b_i. The resulting orientation η is recorded as the major orientation for each bin.

used for range data (Frome et al., 2004). This smaller descriptor results in faster matching and is sufficient for 3D volumes due to more distinctive spatial distribution of points than in range data. The range points are located only on surfaces visible to the scanner, so many 3D bins will be empty, whereas in a CT volume, the features are detected in all directions, resulting in a richer descriptor.

Keypoints — we are not yet discussing nodule locations — are matched by finding the keypoint in the fixed image I_2 that minimizes the Euclidean distance between descriptor vectors. The search is accelerated by storing the descriptor vectors in k-d trees, one for each image I_2.

3.3. Initial matching for a nodule

For a given nodule location $\mathbf{x}_k$ in image I_1, we gather the keypoints and their descriptors in a surrounding region $\mathcal{N}(\mathbf{x}_k)$ and match them against the keypoints and descriptors from image I_2, and from these matches we generate a series of initial transformation estimates between I_1 and I_2. We use all the keypoints in the

region because (a) indexing is fast and (b) it is difficult to predict in advance which keypoint will produce a good match. The region size was chosen to produce a sufficient number of matches without sacrificing the accuracy of the initial estimates.

When two keypoints are matched, the rigid transform $\mathbf{T}_k$ between the two keypoint neighborhoods is obtained by aligning the keypoint coordinate axes as follows. Let the location and orientation of a matched keypoint from I_1 be $\mathbf{x}_m$ and $\mathbf{R}_{\mathbf{x}_m}(\Theta_1(\mathbf{x}_m), \Theta_2(\mathbf{x}_m), \Theta_3(\mathbf{x}_m))$, respectively. Let the location and orientation of a keypoint from I_2 be $\mathbf{y}_n$ and $\mathbf{R}_{\mathbf{y}_n}(\Theta_1(\mathbf{y}_n), \Theta_2(\mathbf{y}_n), \Theta_3(\mathbf{y}_n))$, respectively. Then, the initial rigid transform $\mathbf{T}_k(\mathbf{R}_k, \mathbf{t}_k)$ is obtained as: $\mathbf{R}_k = \mathbf{R}_{\mathbf{y}_n} \mathbf{R}_{\mathbf{x}_m}^{\top}$ and $\mathbf{t}_k = \mathbf{y}_n - \mathbf{x}_m$.

3.4. Transformation estimation

The indexing of keypoints produces a list of matches ordered based on the descriptor distances, which means that the most similar descriptors are near the top of the list. Each match is used to generate an initial rigid transform by aligning local keypoint coordinate axes. The degree to which the structures are aligned using the initial transform varies. The ordering of matches is changed
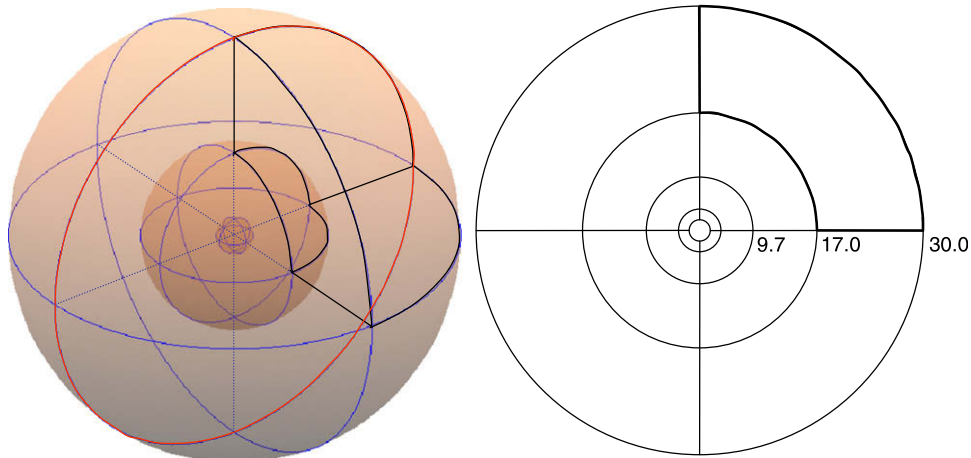

Fig. 7. The keypoint descriptor is computed by forming a local spherical coordinate system oriented along the keypoint's axes. A cross-section through the descriptor along the plane of the red circle is shown on the right. Four angular bins and four radial bins are formed, (up to radius $r = 30$ mm), together with an additional central bin. One such bin is highlighted with a black line. Following Belongie et al. (2002), the partitioning is uniform in the log-polar space, which corresponds to a linearly increasing positional uncertainty with distance from the keypoint location. (For interpretation of references to color in this figure legend, the reader is referred to the web version of this article.)
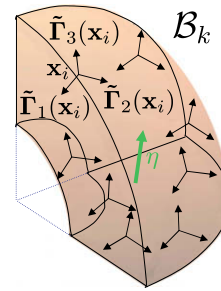
to place those that provide the best initial alignments near the top. The reordering is based on the initial alignment error of features from the initial matches.

Starting from the initializations, the estimation procedure finds the transformation $\mathbf{T}_k$ with parameters $\hat{\theta}_k$ that best aligns the neighborhoods $\mathcal{M}(\mathbf{x}_k)$ and $\mathcal{M}(\mathbf{x}_k')$, $\mathbf{x}_k' = \mathbf{T}_k(\mathbf{x}_k)$, by applying a robust form of ICP:

$$\mathcal{E}_\rho(\boldsymbol{\theta}_k) = \sum_{(\mathbf{p}_i,\mathbf{q}_j)\in\mathcal{C}} v_j \rho(d(\mathbf{p}_i,\mathbf{q}_j;\theta_k)/\sigma), \tag{5}$$

where $\mathcal{C}$ is a set of correspondences, each established between a moving point $\mathbf{p}_i$ from $\mathcal{M}(\mathbf{x}_k)$ and a fixed point $\mathbf{q}_j$ from $\mathcal{M}(\mathbf{x}_k')$. Function $\rho(\cdot)$ is the Beaton–Tukey M-estimator robust loss function (Stewart, 1999):

$$\rho(u) = \begin{cases} \frac{a^2}{6}\left[1 - \left(1 - \left(\frac{u}{a}\right)^2\right)^3\right], & |u| \leqslant a \\ \frac{a^2}{6}, & |u| > a. \end{cases} \tag{6}$$

The constant a is set to 4, which means that normalized alignment error distances beyond 4σ have a fixed cost. Spatial weight $v_j = 1/\left(1 + c\left(\mathbf{x}_k' - \mathbf{q}_j\right)^2\right)$, with $c = 5$, increases the influence of the matches close to the region center and decreases the influence when approaching the region boundaries. This improves the accuracy of the estimate at $\mathbf{x}_k$.

The error distance d is computed as follows. The residual vector is $\mathbf{e} = \mathbf{T}_k(\mathbf{p}_i; \theta_k) - \mathbf{q}_j$. The squared error is computed from $\mathbf{e}$ as $d^2 = \mathbf{e}^\top \mathbf{M} \mathbf{e}$, where $\mathbf{M} = \mathbf{I}$ for point-to-point, $\mathbf{M} = \mathbf{I} - \mathbf{t}\mathbf{t}^\top$ for point-to-tube, and $\mathbf{M} = \mathbf{n}\mathbf{n}^\top$ for point-to-sheet distance. Here, $\mathbf{I}$ is the identity matrix and $\mathbf{t}$ and $\mathbf{n}$ are the unit tangent and the unit normal at the fixed point location $\mathbf{q}_j$. The distance d is normalized by error scale σ.

The objective function (5) is minimized using Iteratively Reweighted Least-Squares (IRLS) (Stewart, 1999; Meer, 2004) which alternates two steps: (1) establishing correspondences $\mathcal{C}$ and recomputing weights $\{w_{ij}\}$ and (2) finding parameters $\hat{\theta}_k$ by minimizing the squared errors after applying robust function and weights $\{w_{ij}\}$:

$$\mathcal{F}(\boldsymbol{\theta}_k) = \sum_{(\mathbf{p}_i,\mathbf{q}_j)\in\mathcal{C}} v_j w_{ij} d^2(\mathbf{p}_i,\mathbf{q}_j;\theta_k). \tag{7}$$

The distance-based robust weight w_{ij} is computed as:

$$w_{ij} = \sum_{(\mathbf{p}_i,\mathbf{q}_j)\in\mathcal{C}} w(d^2(\mathbf{p}_i,\mathbf{q}_j;\theta_k)/\sigma)/\sigma^2, \tag{8}$$

where $w(\cdot)$ is the robust M-estimator weight function (Stewart, 1999):

$$w(u) = \frac{\rho'(u)}{u} = \begin{cases} 1 - \left(\frac{u}{a}\right)^2, & |u| \leqslant a \\ 0, & |u| > a. \end{cases} \tag{9}$$

At the first iteration, the scale σ is estimated using the MUSE algorithm (Miller and Stewart, 1996). The MUSE algorithm extracts the best fit from the data by minimizing objective function of unbiased scale estimates. After the first iteration, the scale is computed as:

$$\sigma^2 = \sum_{(\mathbf{p}_i,\mathbf{q}_j)\in\mathcal{C}} v_j w_{ij} d^2(\mathbf{p}_i,\mathbf{q}_j;\hat{\theta}_k) / \sum_{(\mathbf{p}_i,\mathbf{q}_j)\in\mathcal{C}} v_j w_{ij}. \tag{10}$$

The estimation is symmetric, which means that constraints from forward and backward directions are used together to estimate $\hat{\theta}_k$. This increases the stability of the estimation.

Parameter estimation is the most computationally expensive part of the algorithm. Speedups are achieved by efficient matching and by a multi-resolution scheme. The feature matching is accelerated by using Voronoi maps Duda et al., 2001, Chapter 4 for efficient nearest neighbor search. The multi-resolution starts by randomly sampling the moving points in the neighborhood $\mathcal{M}(\mathbf{x}_k)$ to obtain at most $F = 1500$ initial features – enough to compute an accurate affine transformation estimate. The estimation is then run until convergence while randomly selecting F features at each iteration. The final transformation is used to initialize the estimation at a finer resolution. The computation at the finer level starts by setting $F \leftarrow 2F$ and resampling points in $\mathcal{M}(\mathbf{x}_k)$. The multi-resolution scheme stops when the finest resolution is reached, i.e. all available points from $\mathcal{M}(\mathbf{x}_k)$ are used. The transformation estimate at the finest resolution is the final estimate $\hat{\theta}_k$.

3.5. Independent motions

The affine transformation model used by LRR is not accurate in regions, such as near bones, where different tissues that appear in the region may be moving independently during the breathing cycle. In this case matching all features in the region can lead to misalignments. Filtering feature matches by finding independent motions or using robust estimation (Čech et al., 2008; Ferrari et al., 2004; Lhuillier and Quan, 2005; Yang et al., 2007) is difficult because the error distances (Eq. (5)) for separately moving features can be quite small (smaller than four standard deviations, see Eq. (6)). Instead, we use an oversegmentation of the lung CT volume computed using the watershed segmentation algorithm (Ibáñez et al., 2003) to select features from I_1 to match during transformation estimation. An example of two segmented volumes is on the bottom of Fig. 9.

Using the segmentations, LRR starts by collecting all moving features in the segment at $\mathbf{x}_k$ into $\mathcal{P}$ and all fixed features at $\mathbf{x}_k'$ into $\mathcal{Q}$. If the sets $\mathcal{P}$ and $\mathcal{Q}$ contain enough features ($|P| \geqslant 1500, |Q| \geqslant 1500$), the estimation starts according to the multi-resolution scheme (Section 3.4) using the features in $\mathcal{P}$ and $\mathcal{Q}$. Otherwise, the segments in the moving and fixed images are sorted according to the Euclidean distance of the nearest segment boundary point from $\mathbf{x}_k$ and $\mathbf{x}_k'$, respectively. The features from the first closest segment in each sorting are included in $\mathcal{P}$ and $\mathcal{Q}$. More features are added from farther segments until the desired size (1500 features) is reached. Once the estimation starts, no additional features are added. Therefore, the foregoing process indirectly selects which features are used at the finest resolution.

In the example in Fig. 10, the lung volume moves differently from the surrounding tissue and bone. LRR is initialized at a nodule location close to the lung wall. Only features that are within the same anatomical region surrounding the query point $\mathbf{x}_k$ and the mapped query point $\mathbf{x}_k'$ are used during estimation. The accuracy of the final estimate is improved because the feature matches outside the lung volume are not included. Deformable registration (Section 2.3) uses a smoothness constraint (Vercauteren et al., 2007), essentially producing an average of the lung motion and the motion of the surrounding tissue and bone. This causes inaccuracies near the boundaries of the two motions. The problem has been addressed by using an L1 regularizer term instead of the L2 term (Pock et al., 2007) but the evaluation on the clinical datasets is limited (only four examples studied). In Cahill et al. (2009) an improvement of 10% in the median alignment error is achieved by using a locally adaptive regularizer.

3.6. Alignment decision process

A crucial step in the LRR algorithm is the decision process that determines whether the final alignment is sufficiently reliable to be called "correct". This decision process discards alignments that were produced from incorrect initializations (using non-corresponding neighborhoods) and rejects estimation results with misaligned regions. By relying on this decision criteria, we can truly recognize the nodule location $\mathbf{x}_k$ and its surrounding neighborhood
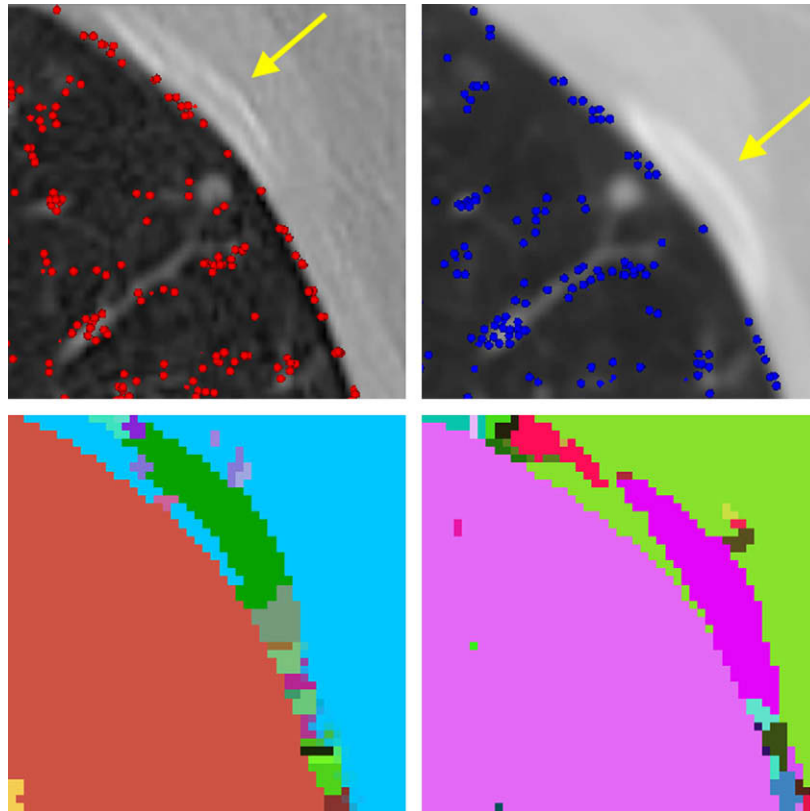
Fig. 9. Top: The lung is expanding due to breathing which results in a different relative position between the nodule and the neighboring rib (indicated by an arrow) in the images. LRR estimates the motion at the nodule location inside the lung volume while ignoring the other motions. Accuracy of the estimate is improved since only feature matches from one motion are used. Bottom: Over-segmented volumes are used to add only features from the moving (left) and fixed (right) regions of the same anatomy. Each segment produced by the watershed algorithm is shown with a different color. (For interpretation of references to color in this figure legend, the reader is referred to the web version of this article.)

$\mathcal{N}(\mathbf{x}_k)$ in the entire volume I_2. When the decision step does not accept the alignment, the next match from the rank-ordered list is used to initialize a new estimation process. The algorithm continues until an alignment is accepted or until the list is exhausted. In the latter case, LRR exits with a statement that no transformation has been found for the location $\mathbf{x}_k$.

The decision process computes an eight-dimensional vector $\mathbf{\Phi}_k = \{\phi_1, \ldots, \phi_8\}$ from the forward ($\mathbf{T}_k$) and backward $\left(\mathbf{T}_k^{-1}\right)$ transformations and the final ICP feature correspondences. In the following, we show how to compute $\phi_1, \ldots, \phi_4$ using the forward transform ($\mathbf{T}_k$). The measures $\phi_5, \ldots \phi_8$ are computed similarly using the backward transform $\left(\mathbf{T}_k^{-1}\right)$. Each of these measures is intended to be small when the alignment is correct and large when it is incorrect.

The first measure is the average mapping error between fixed and mapped moving points. It is calculated from the final ICP feature correspondences $\mathcal{C}$. Using the notation above, the mapping errors are weighted by the IRLS robust weights and averaged:

$$\phi_1 = \sum_{(\mathbf{p}_i,\mathbf{q}_j)\in\mathcal{C}} w_{ij} d(\mathbf{p}_i, \mathbf{q}_j; \hat{\theta}_k) \Bigg/ \sum_{(\mathbf{p}_i,\mathbf{q}_j)\in\mathcal{C}} w_{ij}. \tag{11}$$

The second measure uses only the subset $\mathcal{C}_s$ of correspondences involving sheet features. For correspondence $(\mathbf{p}_i, \mathbf{q}_j)$, let $\mathbf{m}_i$ and $\mathbf{n}_j$ be vectors normal to the local sheet structure at $\mathbf{p}_i$ and $\mathbf{q}_j$, respectively. Each normal vector $\mathbf{m}_i$ at a moving point location is mapped with the transform $\mathbf{T}_k$ producing a vector $\mathbf{m}_i'$. The average of angles between each $\mathbf{m}_i'$ and $\mathbf{n}_i$ is then evaluated:

$$\phi_2 = \sum_{(\mathbf{p}_i,\mathbf{q}_j)\in\mathcal{C}_s} w_{ij} \cos^{-1}\left(|\mathbf{m}_i' \cdot \mathbf{n}_j|\right) \Bigg/ \sum_{(\mathbf{p}_i,\mathbf{q}_j)\in\mathcal{C}_s} w_{ij}. \tag{12}$$

Random (incorrect) alignments produce large values of these normal angles.

Using a subset $\mathcal{C}_t$ with correspondences $(\mathbf{p}_i, \mathbf{q}_j)$ formed from tubular features and replacing the normal vectors $\{\mathbf{m}_i\}$ and $\{\mathbf{n}_j\}$ by vectors $\{\mathbf{s}_i\}$, $\{\mathbf{t}_j\}$ tangent to the tubular structures, the previous steps are modified to compute the third measure as:

$$\phi_3 = \sum_{(\mathbf{p}_i,\mathbf{q}_j)\in\mathcal{C}_t} w_{ij} \cos^{-1}\left(|\mathbf{s}_i' \cdot \mathbf{t}_j|\right) \Bigg/ \sum_{(\mathbf{p}_i,\mathbf{q}_j)\in\mathcal{C}_t} w_{ij}. \tag{13}$$

The fourth and final measure is based on the transfer error covariance of the transformation parameters Hartley and Zisserman, 2000, Chapter 4. The neighborhood $\mathcal{M}(\mathbf{x}_k)$ is sampled in 3D at regular intervals to obtain the set of locations $\{\mathbf{l}_i\}$. The measure, which characterizes the stability of the parameter estimate, is then calculated as:

$$\phi_4 = \max_{\mathbf{l}_i} \mathrm{trace}(\mathbf{J}_{\theta_k}(\mathbf{l}_i)\mathbf{S}_{\theta_k}\mathbf{J}_{\theta_k}(\mathbf{l}_i)^\top), \tag{14}$$

where $\mathbf{J}_{\theta_k}(\mathbf{l}_i) = \partial \mathbf{T}_k / \partial \hat{\theta}_k$ evaluated at $\mathbf{l}_i$.

The vector $\mathbf{\Phi}_k$ is used to train a soft-margin SVM classifier with a radial basis kernel (Vapnik, 1998). The classifier has a binary output: aligned or misaligned. It is trained only once for each application using a training set of neighborhood alignments. The vector is computed for all initializations from the training set, even those that did not provide a correct final alignment. Alignments with changes are left out. The decision boundary is shifted (Chang and
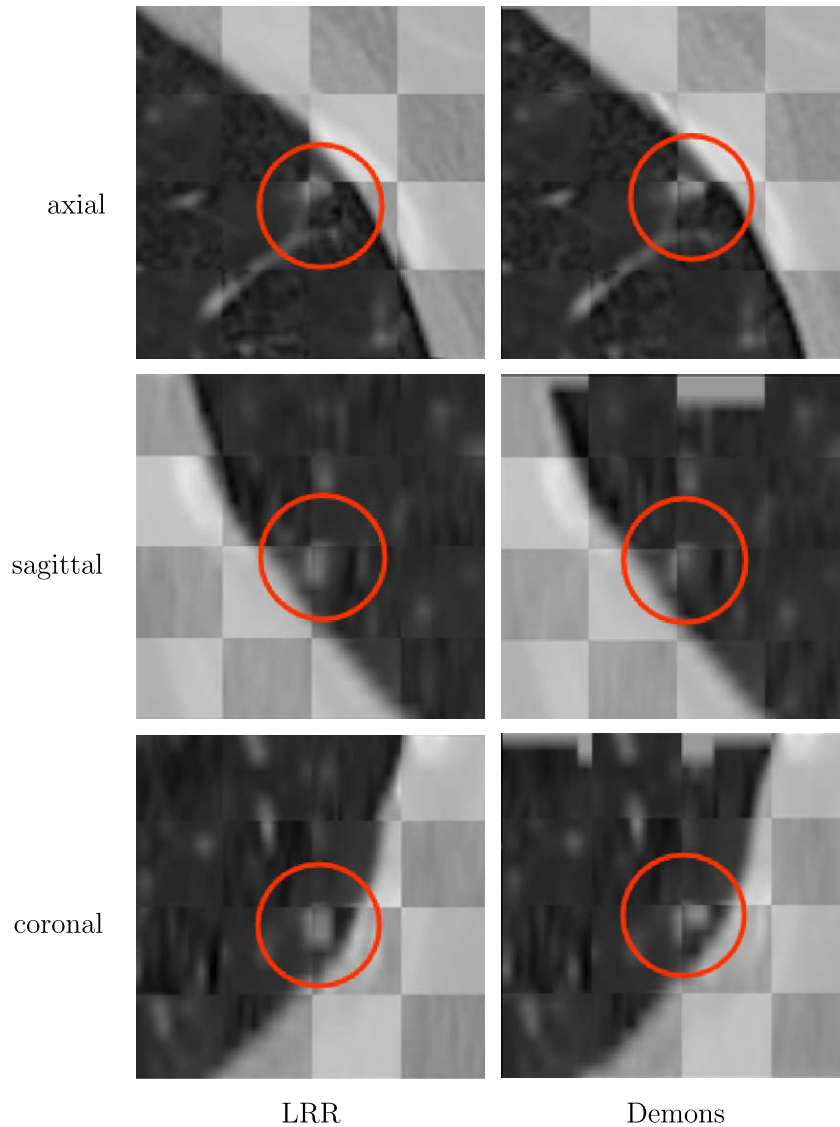
Fig. 10. By using features only within the same anatomical regions, LRR alignment near the boundaries of two independent motions is more accurate than the alignment by the Diffeomorphic Demons algorithm. The smoothness constraint in the deformable registration algorithms essentially averages the two motions.

Lin, 2001) to bias the classifier towards the vectors of the misaligned results. This causes the decision process to limit the cases when an incorrect result would be accepted. On the other hand, correct alignments are occasionally rejected. This is less of a problem because the next match is tried from the rank-ordered list. The algorithm can still produce a correct and successfully verified alignment. See Section 4.3 for the experimental analysis of the decision algorithm.

3.7. Speedup by multiresolution indexing

Features, keypoints, and descriptors are precomputed for each image so that LRR processing at different locations $\{\mathbf{x}_k\}$ is fast in an interactive application. In the case when the number of locations $\{\mathbf{x}_k\}$ is small and an additional computational cost when applying LRR at each location can be afforded, most of the precomputation time can be saved. To achieve this efficiency, we experimented with a simple multiresolution indexing scheme motivated by a recent technique applied to 2D images (Liu et al., 2009).

A multiresolution hierarchy of image volumes I_1 and I_2 is created by subsampling the original volumes. At resolution i, only the descriptors in I_2 within region R_i of size r_i^3, centered at $\mathbf{c}_i$ are used for matching (Fig. 11). When matching a keypoint from I_1, we find the keypoint from I_2 within R_i that minimizes the inter-descriptor distance. The best match of keypoints from $\mathcal{N}(\mathbf{x}_k)$ produces an initial transform $\mathbf{T}_{k,1}$. At the next resolution, $i+1$, the matching is repeated but this time only the descriptors in I_2 within the region of size r_{i+1}, centered at $\mathbf{c}_{i+1} = \mathbf{T}_{k,1}(\mathbf{x}_k)$ are used ($r_{i+1} < r_i$). This produces a new match with refined location and orientation. At the coarsest resolution, the size r_0 equals the size of the volume I_2 and the center $\mathbf{c}_0$ corresponds to the volume center. We used two resolutions in our tests and we set the size $r_1 = 30$ mm.

The multiresolution indexing achieves computation savings in applications where only few locations $\{\mathbf{x}_k\}$ are processed in *batch* processing. Features, keypoints, and descriptors need to be computed within $\mathcal{N}(\mathbf{x}_k)$ in I_1 and only within the region R_i of size r_i in I_2 at each resolution i, which is much smaller than the whole volume size I_2 at the finest resolution. However, to obtain the
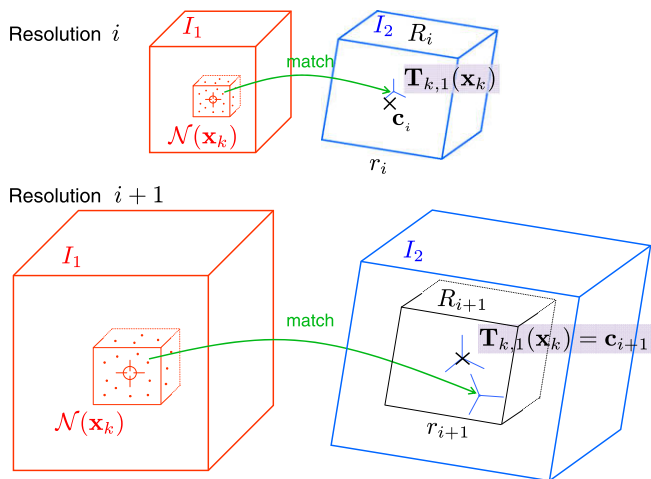
Fig. 11. Multiresolution indexing scheme. At resolution i, only descriptors within region R_i of size r_i^3, centered at $\mathbf{c}_i$ are used for matching. Features, keypoints, and descriptors only need to be computed within the neighborhood $\mathcal{N}(\mathbf{x}_k)$ in I_1 and within the region R_i which is smaller than the size of the volume I_2 at the finest resolution.

highest *interactive* speed, precomputing the features, keypoints, and descriptors is desirable.

4. Experiments

Our experimental evaluation starts by testing each of the LRR stages: the assessment of keypoint indexing is in Section 4.2 and the estimation and the decision steps are tested in Section 4.3. We will then turn to studying the overall algorithm performance. In Section 4.4, LRR is initialized at nodule locations and the final alignments are evaluated qualitatively and quantitatively. Section 4.5 analyzes the local affine alignments of regions with respect to the deformation fields obtained by the Diffeomorphic Demons algorithm. In Section 4.6, LRR is compared to a global nodule registration technique with local refinement. LRR alignments on a database to evaluate response to therapy in lung cancer are evaluated in Section 4.7. Timing results are presented in Section 4.8, and the experimental results are summarized in Section 4.9.

4.1. Data sets

The data for our first set of experiments in Sections 4.2–4.6 are screening and diagnostic lung CT scans, with two scans for each patient taken approximately one year apart. The scans were obtained with a GE Healthcare LightSpeed CT scanner with a tube current from 30 mA to 60 mA (screening scans) and from 250 mA to 440 mA (diagnostic scans), peak voltage from 120 kV to 140 kV, reconstruction diameter from 180 mm to 441 mm, and exposure time from 0.6 s and 1.1 s. After masking out the background, the size of the volumes ranges from $478 \times 382 \times 106$ voxels to $511 \times 422 \times 122$ voxels. Their slice spacing ranges from 2.5 to 5 mm and the voxel width ranges from 0.63 to 0.91 mm. Most of our thoracic scans were reconstructed using the lung kernel, but several of them using the soft kernel as well. Soft kernels produce images with lower spatial resolution, but higher contrast (Prokop et al., 2000).

An expert annotated 12 volume pairs (lung kernel) containing 35 nodule pairs. Using additional 10 volumes of the soft kernel CT reconstruction yields 22 volume pairs with a total of 67 nodule pairs available for experiments. Throughout the experiments, one nodule from each nodule pair is used for the initialization and the result of an experiment is compared to the corresponding

nodule from the pair. The reference nodule annotations are accurate for comparative evaluation of several methods for nodule alignment since any annotation errors will affect the accuracy of all techniques. The nodules are non-calcified and have different aspect (i.e. solid, partial solid, and non-solid). Of the 67 nodules, 19 have diameter from 1 to 4 mm, 22 have diameter from 4 to 8 mm, 16 have diameter from 8 to 12 mm, and 10 have diameter from 12 to 20 mm. Several examples are in Fig. 12. Notice the variability in their shape and location and the different amounts of noise in the images.

The experiments in Section 4.7 are on The Reference Image Database to Evaluate Response (RIDER) to therapy in lung cancer (Armato et al., 2008). The scans were obtained with a GE Healthcare Systems LightSpeed CT scanner with a tube current from 180 mA to 441 mA, peak voltage of 120 kV, reconstruction diameter from 260 mm to 466 mm, and exposure time from 0.5 s to 1.1 s. After masking out the background, the size of the volumes ranges from $510 \times 318 \times 46$ to $511 \times 511 \times 321$. Their slice spacing is 1.25 and 5 mm and the voxel width ranges from 0.49 to 0.91 mm. Of the 52 nodules in the pilot dataset (RIDER1), 4 have diameter from 2 to 10 mm, 23 have diameter from 2 to 15 mm, and 25 have diameter from 16 to 31 mm. Of the 32 nodules in the second dataset (RIDER2), the tumor sizes ranges from 11 to 93 mm, with a mean of 38 mm.

4.2. Keypoint indexing

Since our approach hinges on effective keypoint matching, it is important to examine some empirical tests that motivate the approach and guide several design decisions. These experiments are based on a small set of CT volume pairs. For each pair, the Diffeomorphic Demons algorithm (Vercauteren et al., 2007), currently the best algorithm for global lung CT registration (Urschler et al., 2007), is applied to produce a transformation $\mathbf{T}_G$ (Fig. 14). For any keypoint match, if, after application of $\mathbf{T}_G$ to the keypoint from I_1, the keypoints are within 9 mm of each other and their axes are within 20°, then this is considered a "good" candidate match. This distance and orientation tolerance is defined to categorize as "good" those matches, that can provide initial transform accurate enough to be successfully refined into the final estimate. The tolerance therefore defines the acceptable level of inaccuracy for the estimation stage rather than characterizing the quality of the matches in the absolute sense. We use the rule to examine the results of keypoint matching at a random sampling of 1000 locations in and around the lung in each scan. Here is a summary of the results:

- Consider the *closest* keypoint $\mathbf{u}_i$ to location $\mathbf{x}_k$. On average, $\mathbf{x}_k$ has a good match, independent of the descriptor distance, 72% of the time. Also on average, the best descriptor match for $\mathbf{x}_k$ is "good" 40% of the time. Together, these show the promise of the indexing approach, but also clearly indicate that more than just matching the closest keypoint is needed.
- The next issue is how far we need to look before finding a good keypoint match. On average, a good match exists within 10 mm of $\mathbf{x}_k$ for 20% of the locations, within 20 mm for 65%, and within 30 mm for 98%. When requiring the good match to be the best descriptor match, these values are 2%, 18%, and 96% for 10 mm, 20 mm, and 30 mm, respectively. This shows much greater promise.
- Since there may be many keypoints in $\mathcal{N}(\mathbf{x}_k)$ (on average 160 within 30 mm), knowing that a good one exists is not sufficient. Many incorrect initializations might be generated before the good keypoint match is encountered. We must therefore consider the effects of rank-ordering of
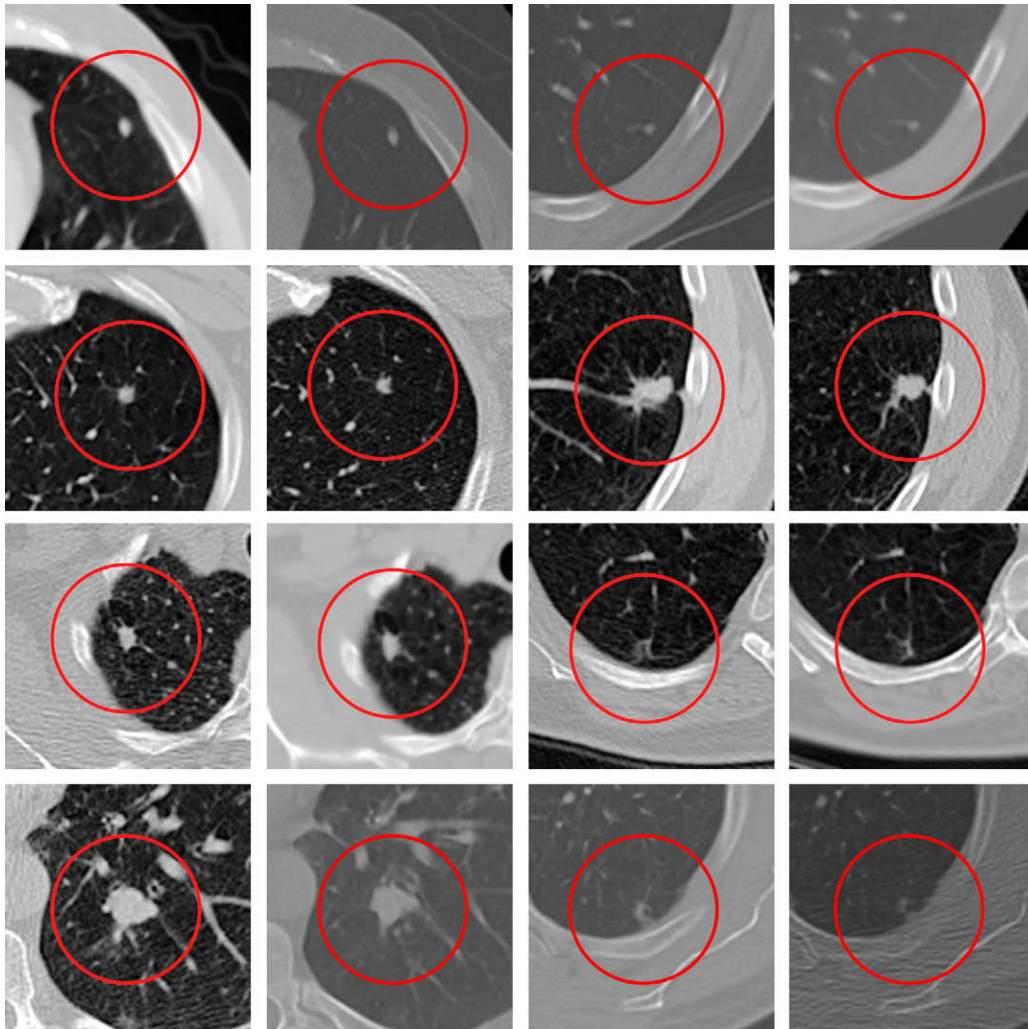
Fig. 12. Examples of nodule pairs in neighborhoods of size 100 × 100 mm. The red circle with 30 mm radius illustrates the size of the region used to compute the keypoint descriptors. The keypoint locations, however, might not be exactly at the nodule centers and in the same slice. Notice the variability in the nodule shape and location within the thoracic cavity. Also notice the different amounts of noise in the images. Independent motions of the lung volume and the bone (brightest) are clearly visible in several cases (1st row, 3rd and 4th column, and 2nd row, all columns). (For interpretation of references to color in this figure legend, the reader is referred to the web version of this article.)

the matches. Indexing followed by rank-ordering based on descriptor distance must put good matches near the top. Experiments show that when the top 20 matches are considered for all keypoints within 30 mm, 65% are good. More importantly, 95% of the $\mathbf{x}_k$ have at least one good match within the top 20. These encouraging results justify the design path taken here. It is crucial to realize that this 95% figure does not represent an upper bound on the overall success rate of the algorithm since the alignment process can often overcome initial misalignments from keypoint matches that are not considered "good". As a final note, if we replace our ordering by descriptor distance with Lowe's widely-used distance ratio (first to second best match for a keypoint), the previous percentages drop slightly to 62% and 93%.

4.3. Estimation and decision

Testing of the estimation and decision components of the algorithm is done by a visual assessment of the algorithm output. The schematic diagram of the experimental setup is in Fig. 15. Estimation is initialized by keypoint descriptor matches and is run until

convergence (or stopped when an invalid transformation is produced). A set of 600 initializations at random locations in six volumes is tested and the result of every estimation presented as a panel of nine images (Fig. 13). Each panel contains an axial, sagittal, and coronal slice through the fixed and mapped moving volume neighborhood. It also contains a checkerboard image with alternating fixed and mapped moving slices. An independent observer visually examined each panel and marked each alignment as successful, unsuccessful, or undefined. The successful and unsuccessful estimation results were divided based on the initialization being good (using a good match) or bad. In this evaluation, the categorization of matches is determined by the Diffeomorphic Demons deformable registration algorithm. As a reminder a "good" match is defined by applying the deformable registration to the keypoint from I_1 and requiring the mapped keypoint to be within 9 mm and its axes within 20° of the matched keypoint (Section 3.2). The decision algorithm is trained as described in Section 3.6 and the classification results are evaluated using the successful/unsuccessful labels of the estimation results.

Several examples of final alignments are shown in Fig. 16. They represent both cases where LRR is in agreement with the deformable registration and cases where the location registration
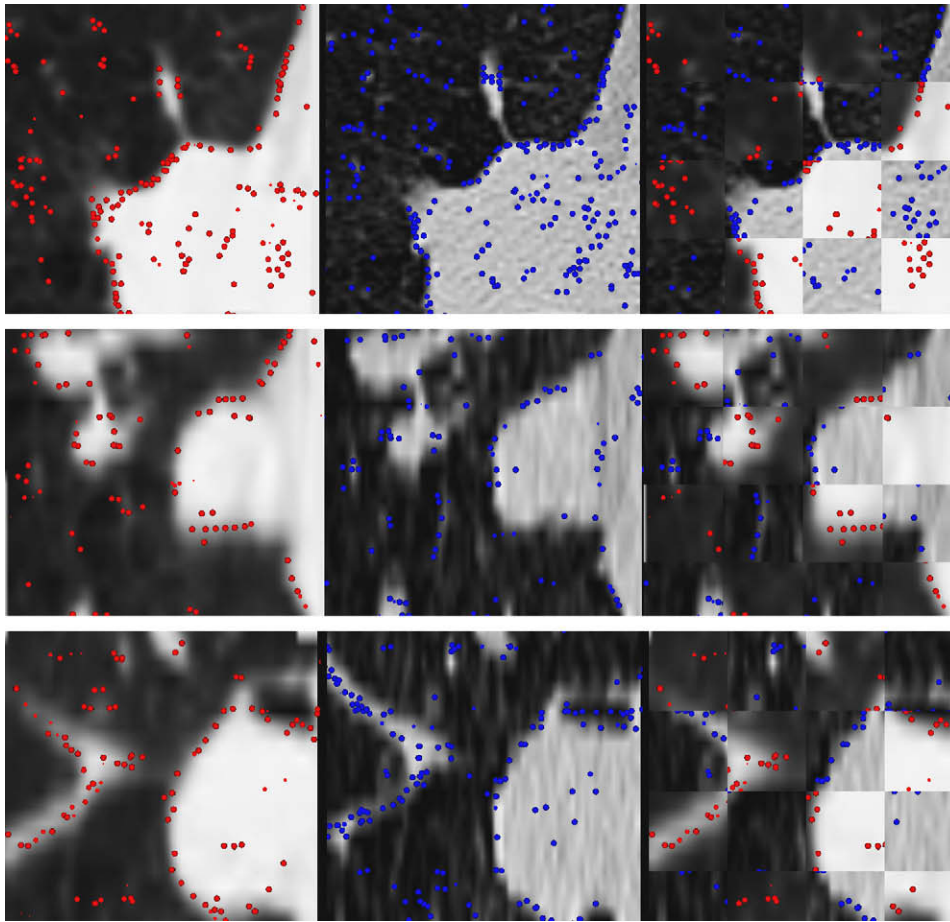
Fig. 13. A panel of nine images for each result presented to an independent observer for the alignment evaluation. The rows show axial, coronal, and sagittal slices. The columns show mapped moving slices, fixed slices, and checkerboards alternating mapped moving and fixed slices. The features are superimposed onto the images.
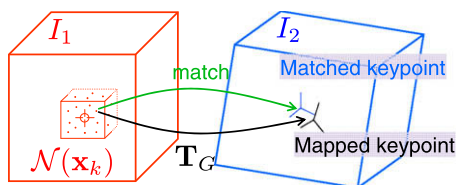

Fig. 14. Empirical analysis of keypoint matching using the Diffeomorphic Demons mapping (transformation $\mathbf{T}_G$). The location and orientation of a corresponding match in I_2 for a keypoint from the moving image I_1 is compared to the location and orientation after mapping the keypoint from I_1 with $\mathbf{T}_G$. See text for the results.

alignment appears to be more accurate. The algorithm correctly handled breathing state differences close to the lung wall which tend to cause misalignments in the deformable registration. After counting the successful and unsuccessful alignments, we have learned that for two volume pairs, all good (and several bad) initializations produced successful alignment (Table 1). Corresponding neighborhoods with partial alignment but visible changes (see Fig. 16) were marked as undefined. The estimation did not produce a successful alignment in seven cases for one volume pair and in two cases in each of the remaining three volume pairs. Such cases are automatically identified and a different initialization is tried. The overall algorithm succeeds on 97.3% of all initial locations.

The decision algorithm is tested for each volume pair using the leave-one-out procedure. The decision classifier is first trained using all but one volume pair using the steps described in Section

3.6. The testing on the left-out pair then produces counts for two types of errors: false positives (FP) and false negatives (FN). False positives are estimation results deemed aligned by the decision algorithm but manually graded as misaligned. False negatives are results deemed misaligned by the algorithm but classified as aligned by the grader. On the rest of the alignments, the grader and the algorithm are in agreement, i.e. they both identify the same results as correctly aligned (true positives) or misaligned (true negatives).

Out of 574 neighborhood pairs, the total number of false positives was 2, while the number of false negatives was 29. This shows that the algorithm is not likely to accept an incorrect alignment. Rejection simply means that another initialization is tried. The bias towards misalignments was introduced during training (Section 3.6) by shifting the decision boundary of the SVM classifier. The most ambiguous cases for the decision algorithm (and manual grading) are those with local regions containing changes between them (marked as undefined).

We have also compared the classification capability of each of the six decision components. The results using the full 8-component measurement vector are used as a reference. The soft-margin SVM classifier trained with different parameters, scale of the radial basis kernel and a penalty cost for mislabeled examples, produces different false positive and true positive rates for each parameter setting. The training and testing procedure is repeated with a 5-component measurement vector leaving out each of the component in turn. The true positive and false positive rates from each experiment and parameter setting are used to compare the classifiers in an ROC analysis.
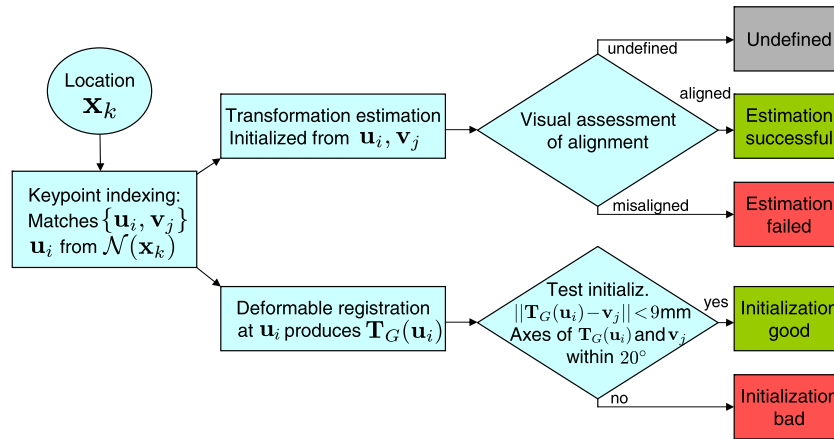
Fig. 15. Schematic diagram of the experimental evaluation of the estimation. The match $(\mathbf{u}_i, \mathbf{v}_j)$ is used to initialize the estimation and the result is visually assessed as aligned, misaligned, or undefined. The keypoint $\mathbf{u}_i$ from the neighborhood of the query location $\mathbf{x}_k$ is mapped using the deformable registration $\mathbf{T}_G$. The match $(\mathbf{u}_i, \mathbf{v}_j)$ provides a "good" initialization if the matched keypoint $\mathbf{v}_j$ and the mapped keypoint $\mathbf{T}_G(\mathbf{u}_i)$ are within 9 mm and their axes within 20°. The results in Table 1 are grouped based on the initialization being good or bad.
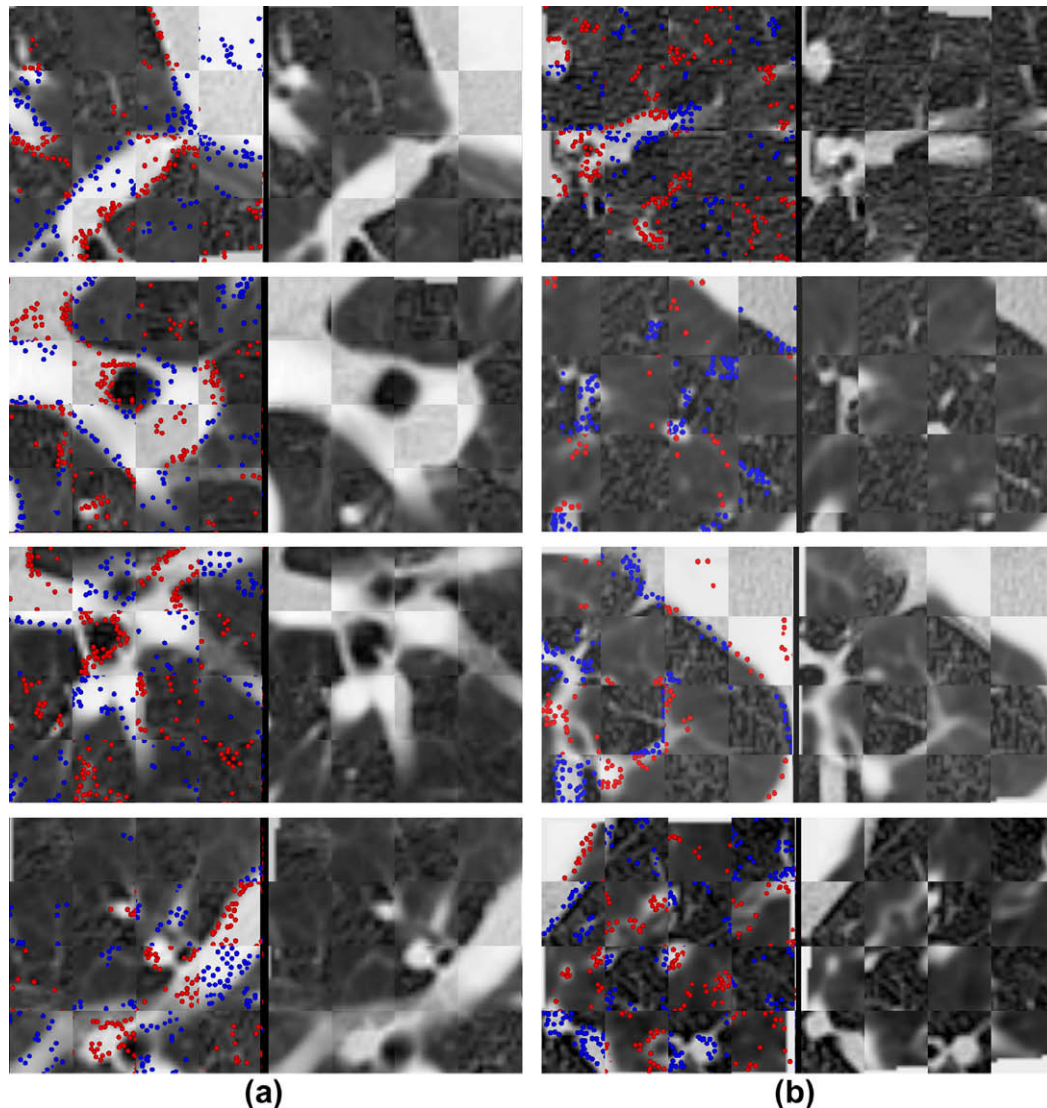

Fig. 16. Examples of LRR (1st and 3rd column) vs. deformable registration (2nd and 4th column). Agreement of both results (a) and examples where LRR alignment is better (b). Features detected in fixed (blue) and moving (red) images drive the registration. The robust estimation ensures that the outliers (i.e. feature points that do not have a direct match in the other volume) are ignored. The outliers might be caused by a difference in the breathing state as in the bottom right example (upper half of the volume). The regions are still well aligned in the lower half. (For interpretation of references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Estimation results using 100 random initializations per volume pair grouped based on whether the initialization was good or bad (see schematic diagram of the experimental setup in Fig. 15). Numbers indicate successful/unsuccessful alignments that were initialized with a good (2nd/4th column) or a bad (3rd/5th column) match. Estimation succeeds in many cases where the initialization is poor (3rd column) and the number of estimation failures is small (4th column). The volume pairs for this experiment were randomly selected.

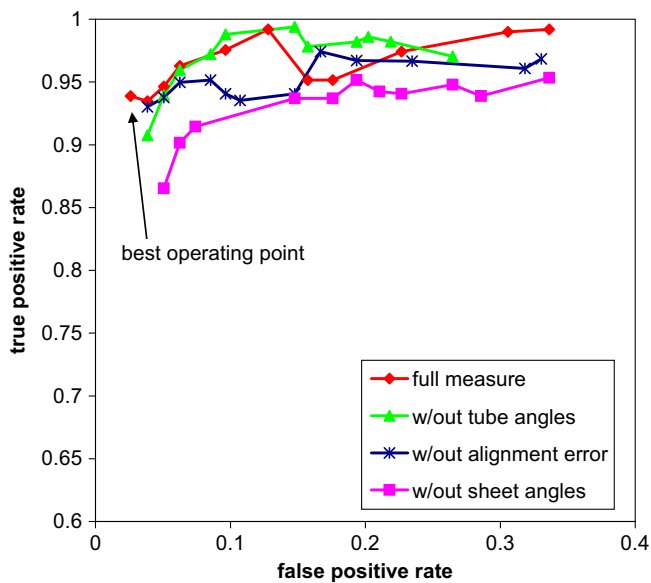| Initialization Estimation	Good Success	Bad Success	Good Failure	Bad Failure	Undef.
Pair 1	63	10	2	16	9
Pair 2	81	5	2	9	3
Pair 3	86	3	0	11	0
Pair 4	78	6	2	11	3
Pair 5	92	6	0	2	0
Pair 6	43	17	7	22	11

Fig. 17. Testing the classification power of each of the decision measure components. Varying the two SVM parameters, each curve is plotted by averaging all true positive rates for a given false positive rate. (This averaging causes the curves to be non-monotonic.) The classifier with the full 8-component measurement vector gives the best results. The highest decrease in performance was in a classifier without the measure of angles between sheet structures. The trace of the transfer error covariance is the least indicative so the ROC curve was left out of the plot.

ROC curves are typically obtained by varying a single discrimination threshold. Since the soft-margin SVM has two parameters, we plot each curve by averaging all true positive rates for a given false positive rate. The results in Fig. 17 show that the full 8-component measure achieves the smallest false positive rate at a high true positive rate. The measure based on the trace of the transfer error covariance is the least indicative so the ROC curve was left out of the plot. Considering the remaining components of the measurement vector, the most decision power is in the measure of angles between sheet structures. The classifier without this measure performs the worst. Less substantial decrease can be seen by removing the alignment error component. The classifier without the measure of angles between tubular structures produces higher true positive rates than the full 8-component classifier if we tolerate higher false positive rates.

4.4. Nodule locations

Our next set of experiment uses the manually-matched locations of nodules in two time-separated CT volumes to evaluate the complete LRR algorithm. The LRR algorithm was initialized at

Table 2
Quantitative analysis of mapping errors computed according to the figure above. Given is a nodule location in one volume (Nodule A). The error between the corresponding nodule location (Nodule B) and the location of the Nodule A aligned with Diffeomorphic Demons (d_{dem}, 2nd column) and LRR (d_{LRR}, 3rd column) was computed. Median, 25th, and 75th percentile errors computed for all nodules are lower for the LRR algorithm. The last column shows errors between the nodule locations mapped using the Demons algorithm and the locations mapped with LRR (d_{com}). On average, the differences agree with the amount of improvement by LRR.
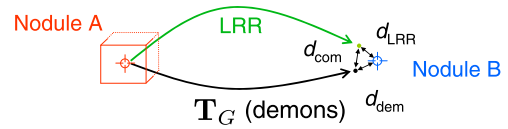
Error [mm]/algorithm	d_{dem}	d_{LRR}	d_{com}
25th percentile	1.43	1.25	0.33
Median	2.14	1.70	0.55
75th percentile	3.40	2.94	1.63

nodule locations in a scan from one of the time points (e.g. the first volume from the pair). The algorithm was run without any user intervention and the final transformation was saved. This was repeated for nodule locations from the other time point (the second volume from the pair). This produced a total of 134 alignments using both time points of our 67 nodule pairs for the initialization. The locations from one volume were mapped using the final transforms and also using the Diffeomorphic Demons deformation field. The mapped locations were then compared to the locations of the corresponding nodules in the other volume.

Table 2 reports 25th percentile, median, and 75th percentile of the nodule location mapping error for LRR and the Diffeomorphic Demons algorithm as compared to the nodule annotations. The mapping error of nodule locations for the Demons algorithm is small (the median error is 2.14 mm), but the mapping error for the LRR algorithm is smaller (the median error is 1.70 mm). The low mapping error is achieved through robust estimation in a feature-based registration algorithm with features detected to subvoxel accuracy (Section 3.1). The accuracy is especially important for small nodules (1–4 mm), where high mapping errors could result in the estimated location being outside of the nodule volume. Table 2 also compares the error distances between the nodule locations mapped using both algorithms relative to each other.

Generally, the differences between the algorithms correspond with the amount of improvement by LRR. This means that compared to the Demons-mapped locations, the LRR algorithm maps the locations closer to the locations of the corresponding nodules. One cause of the deformation field inaccuracies is the globally enforced smoothness constraint (Vercauteren et al., 2007). In the presence of two independent motions, such as the motions of bone and the lung volume (Section 3.5), the smoothness constraint tends to average the two motions (Fig. 10). This can also happen when the nodule shape or the neighboring structure changes. In other cases, the deformable registration can fail because of an insufficient common surfaces between the images or because the structures are very subtle (Fig. 18). Several nodule alignments are shown in Fig. 19. The images show a variety of the nodule shapes and sizes. The nodule neighborhoods have a different structural complexity. Despite these challenges, all regions are correctly aligned and recognized by the LRR algorithm.

The LRR algorithm aligned all 134 nodule regions. In one case, the error between the LRR-mapped location and the annotated nodule location was 12.3 mm. The alignment was verified as correct by the algorithm but should have been rejected and the next initialization tried. The failure was caused by insufficient structure present in the neighborhoods, which resulted in poor feature extraction. This example can be seen in Fig. 20.
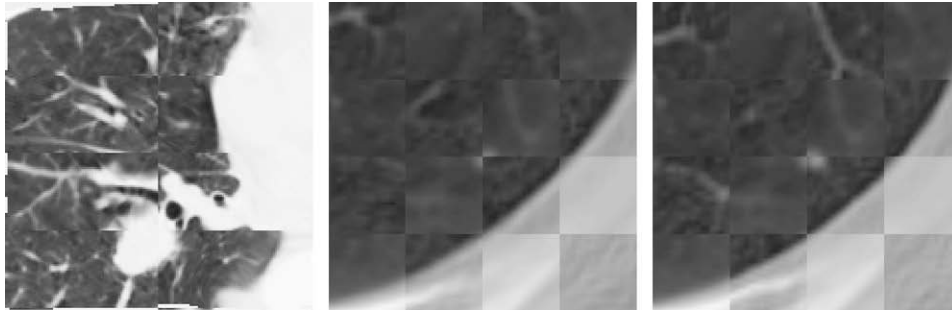
Fig. 18. The deformable registration can fail when the initialization or large changes due to breathing state differences causes the corresponding structures to be far apart (left). Another failure might happen when the structural content in the two images is too subtle to produce sufficient constraints for the deformable registration (middle). LRR deals with these problems by robust estimation and by feature extraction which is adaptive to regions with small intensity gradients. The region neighborhoods around the given nodule location are correctly aligned by LRR (right).
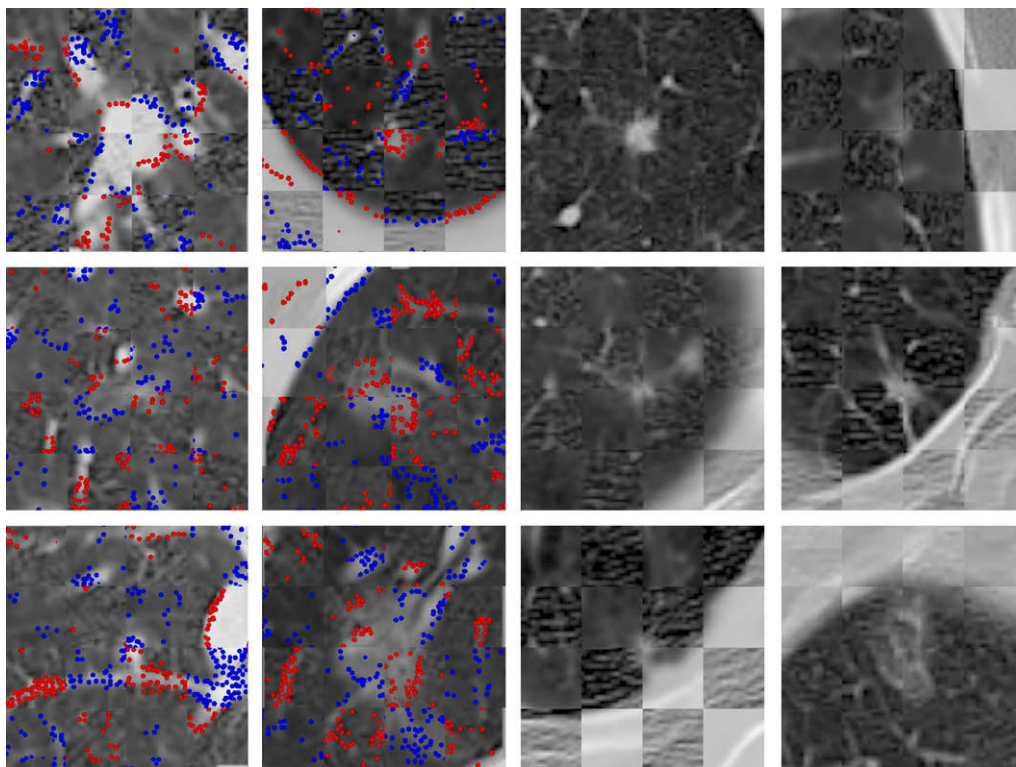

Fig. 19. Examples of nodule alignments shown as a checkerboard image alternating fixed and mapped moving axial slices. Images in the two left columns have superimposed fixed features (blue) and mapped moving features (red). LRR correctly aligns nodules of various shapes and sizes in neighborhoods with different structural complexity. (For interpretation of references to color in this figure legend, the reader is referred to the web version of this article.)
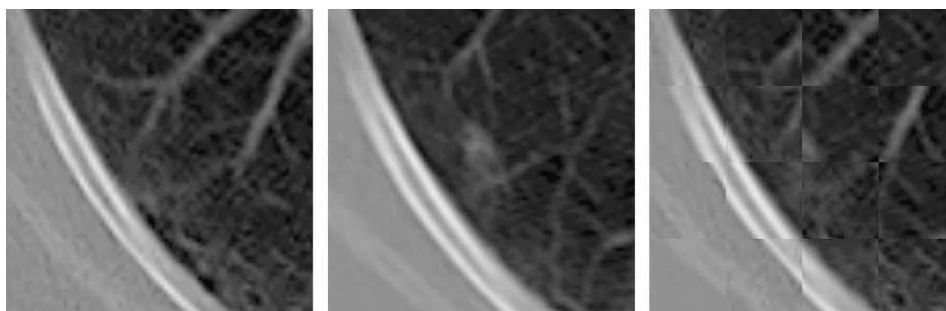

Fig. 20. An example of an LRR result with a large error distance between the mapped location and the annotated location. Very subtle vessels in the two regions and not enough structure resulted in poor feature extraction. This caused a visible misalignment that can be seen from the checkerboard image (right) composed from the mapped moving slice (left) and from the fixed slice (middle).

One potential cause of the LRR failure is an incorrect initialization. It may happen that none of the keypoint pairs in the rank ordering of 20 matches has keypoints from the corresponding image regions. Or, the initializations provided by those matches are too far off from the correct alignment which would cause the estimation to fail. In both cases, the decision algorithm would reject the alignments and no final transformation estimate would be generated. Another failure, although even less common, might occur when the decision step rejects an alignment which would be considered correct based on a manual inspection (false negative) and no further alignments initialized from the rank ordering are verified as correct.

4.5. Affine approximation of the local deformation

The next experiment tests how well the affine transform obtained from the LRR algorithm agrees with the local deformation field (see Eq. (2)). We obtain a set of point locations by regularly sampling in the neighborhood $\mathcal{M}(\mathbf{x}_k)$ around each nodule location $\mathbf{x}_k$ (using nodules from both time points as before). For each location, the average distance between the sampled points mapped with LRR final transform and with the deformation field is computed. Statistics using 134 nodule neighborhoods show that the median error of the average distance between the mapped points is 1.01 mm. Values of the 25th and 75th percentile are 0.84 mm and 1.34 mm, respectively. This indicates a good approximation of the local deformation field by the affine transform, especially when considering that the deformation might be non-linear.

4.6. Nodule registration

We also compared the LRR algorithm to a global nodule registration technique with local refinement. As in Betke et al. (2003), Kawata et al. (2001), Shi et al. (2007), the algorithm starts by global registration of the two scans. First, a multiresolution hierarchy with three resolutions is created by subsampling both volumes. Coarse-to-fine registration algorithm initialized with an identity transformation then runs at each resolution level in succession. The estimated affine transform at the coarse level is used to initialize the registration at the finer level. The error at each iteration is computed as the sum of square intensity differences (SSD) between fixed and mapped moving volumes. Similarly to Shi et al. (2007), Wiemker et al. (2008), the alignment of a volume of interest ($100 \times 100 \times 100$ mm) centered at the nodule location is refined by a local refinement step. The final affine transformation of the region around each given nodule location is verified manually. Since LRR starts from locations specified in only one volume, the corresponding nodule locations in the other volume are also not used during alignment in this experiment. The nodule correspondences are only used for the final quantitative evaluations. We used our own implementation based on the Insight Toolkit (Ibáñez et al., 2003), since publicly accessible code is not available.

The median initial distance between the corresponding nodules computed for all nodule pairs is 38.98 mm. After global coarse-to-fine registration, this distance drops to 6.16 mm. The final mapping error of nodule locations after local refinement is 3.57 mm, which is higher than the result of the Demons algorithm (2.14 mm, Table 2, 2nd column) and higher than the result of the LRR algorithm (1.70 mm, Table 2, 3rd column). The results are summarized in Table 3.

Examples of nodule alignments compared to LRR are in Fig. 21. In many cases, the alignment results of the two algorithms are similar (1st and 2nd column). However, the nodule registration technique produces visible misalignments in several regions, especially neighborhoods with differences due to large breathing

Table 3

Quantitative analysis of mapping errors for the global nodule registration with local refinement. The initial distances are computed between nodules in one volume and corresponding nodules in another volume to obtain the overall statistics (2nd column). The distances after global coarse-to-fine registration are in the 3rd column and after local refinement of the global mapping in the 4th column. Compared to the LRR algorithm (Table 2, 3rd column), median, 25th, and 75th percentile mapping errors computed for all nodules are higher for the nodule registration experiment.

Error (mm)/alignment	Initial	Global	Local
25th percentile	33.52	4.17	2.03
Median	38.98	6.16	3.57
75th percentile	48.49	8.34	5.29

motion and close to the lung wall (3rd column, 1st row) and due to independent motion of the lung and nearby bone structures (3rd column, 3rd row). LRR correctly aligns these regions (4th column).

The global registration techniques with local refinement in nodule regions are attractive because they are relatively easy to implement. One difficulty is the initialization by global alignment of the two volumes. Such initialization might fail when there are large differences between the two scans (Shi et al., 2007). The differences are usually caused by the lung not having been scanned entirely in one of the volumes (Shi et al., 2007), or by large breathing motion. The LRR initialization step computes descriptors from local image regions and is therefore not affected by such global differences.

4.7. Tumor progression

During treatment planning, tumor progression or response to therapy is studied by comparing tumor shapes and sizes across time. To make this comparison possible, LRR can be used to find the corresponding tumor locations in subsequent scans and to locally align the tumor neighborhoods. Our experiments are on RIDER1 and RIDER2 datasets (Section 4.1). The parameter settings remained the same as in the previous experiments.

As in Section 4.4, the LRR algorithm is initialized at nodule locations and the results quantitatively analyzed. Since in the RIDER1 dataset the corresponding nodules in another volume are not known, the quantitative evaluation is computed with respect to the Demons algorithm. The 25th percentile, median, and 75th percentile errors were 1.59 mm, 3.32 mm, and 5.77 mm, respectively. The larger differences when compared to Table 2 are caused by the changes in the large-sized tumors as shown in Fig. 22. This results in a non-linear motion which is estimated by the deformable registration algorithm and causes deviations from the affine mapping. Example alignments are in Fig. 23.

The tumor locations in the RIDER2 dataset are specified in both volumes, but the locations are only approximate (Meyer et al., 2006; Zhao et al., 2009) because of large tumor sizes. The median distance between locations of LRR-mapped tumors and corresponding tumors in another volume is therefore higher than in previous experiments, but still below 5 mm for all 64 alignments (from initialization at both time points).

Figs. 22 and 23 shows that given a tumor location in one scan, LRR can automatically find the corresponding tumor in another scan and locally align the two tumors even when the size or shape of the tumor changes. The result of LRR can be used to initialize the non-rigid registration of lung tumors (Matsopoulos et al., 2005), to initialize the local registration by Diffeomorphic Demons (Vercauteren et al., 2007), or for growth analysis after the tumors have been segmented and separated from the lung wall and vessels (Reeves et al., 2006). The LRR failed in two cases with significant changes in large tumors because the keypoint descriptors computed near
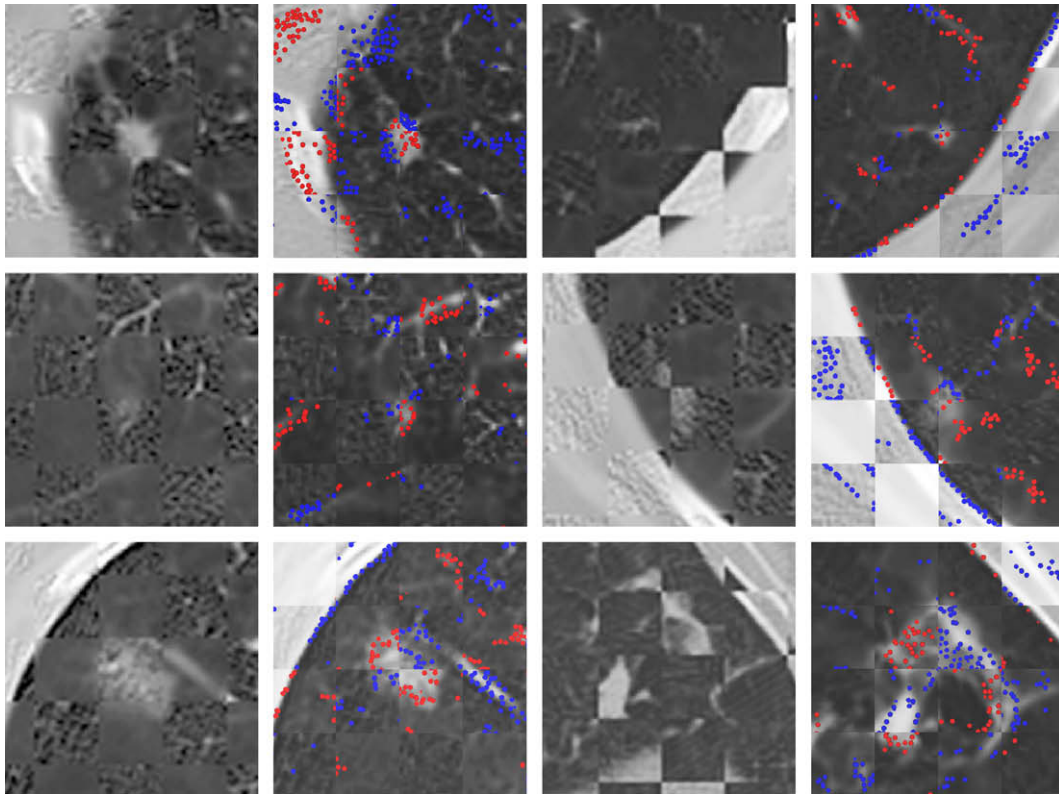
Fig. 21. Examples of nodule registration alignments (1st and 3rd column) and LRR alignments (2nd and 4th column) shown as a checkerboard image alternating fixed and mapped moving axial slices. The two techniques produce comparable results for many nodules (first two columns). The nodule registration fails in the presence of large breathing motion (3rd column, 1st row) and when the motions of the lung region and nearby bone structures are different (3rd column, 3rd row).
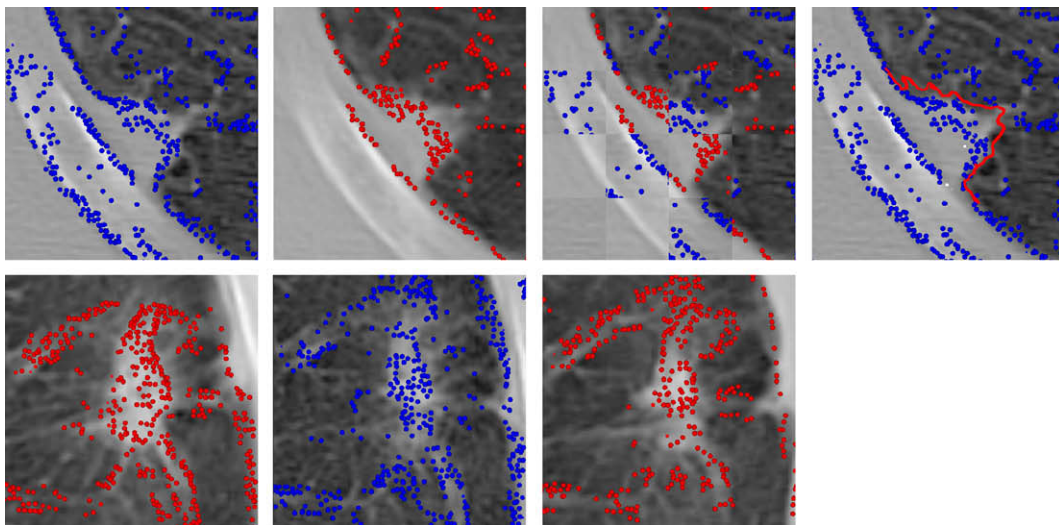

Fig. 22. The tumor shapes and sizes change due to cancer progression or as a response to the therapy. The tumor in the 1st row was scanned again after three months and its size increased as illustrated by a contour of the tumor from the second acquisition (2nd column) overlaying the tumor neighborhood from the first acquisition (1st column). The tumor in the 2nd row was scanned after six months (2nd column) and then again after 2 months (3rd column). The corresponding tumors were automatically found and their neighborhoods locally aligned by LRR.

the tumor centers were so different that the keypoint indexing failed to provide a correct match.

4.8. Timing results

All steps take on average 6 s per location on a laptop PC with Pentium 4, 2.33 GHz processor, and 2 GB of RAM, which includes speedup by discarding initializations with high rotations (10% initializations of the first 20 have rotation more than 45°). Note that further speedups are possible through code profiling and resulting optimizations, running the algorithm on a multi-core or multi-processor machine parallelizing each component, or trying multiple initializations concurrently. The feature detection, keypoint extraction, and descriptor computation takes 5 min and 10 s per volume (multi-threaded execution on an 8-core processor).
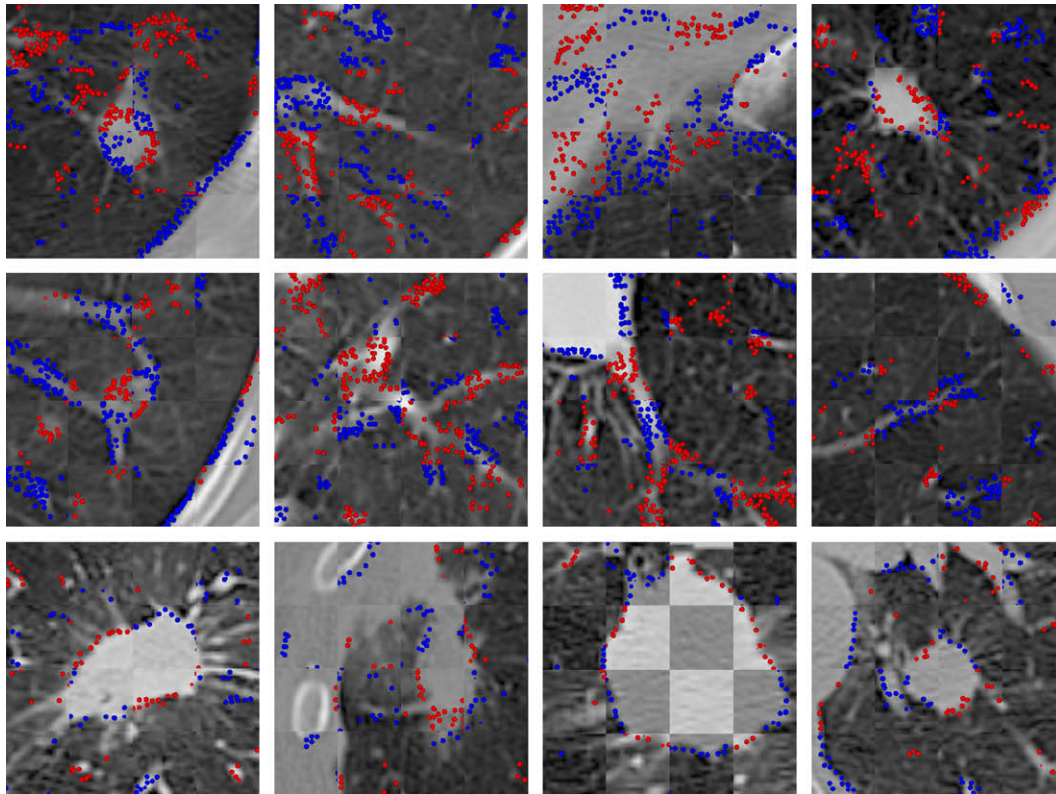
Fig. 23. Examples of tumor alignments shown as a checkerboard image alternating fixed and mapped moving axial slices of the RIDER1 dataset (1st and 2nd row) and the RIDER2 dataset (3rd row). Images have superimposed fixed features (blue) and mapped moving features (red). The tumors of different shapes sizes were correctly recognized in the other volume and accurately aligned by LRR.

There are two possible workflows when using the LRR algorithm. In an *interactive* application, the preprocessing is done for each volume offline. Then, applying LRR at each location takes 6 s on average. In a *batch* processing of multiple locations, the multi-resolution indexing is used (Section 3.2) to save most of the preprocessing time. The preprocessing is done on the coarsest resolution only, which takes 38 s per volume. Then, applying LRR at each location takes 42 s per location, which includes computing features, keypoints, and descriptors at the finer resolution in a region obtained from indexing at the coarser level.

In contrast, a multi-threaded implementation of the Diffeomorphic Demons can take as much as one hour on a computer with 1.86 GHz Quad Core processor to align. The memory requirements for the Demons algorithm are substantial and a 64 bit machine with 8 GB of RAM is recommended (Urschler et al., 2007).

4.9. Results summary

In our experimental evaluation, we have shown the robustness of each of the LRR components. In 95% of the cases, the indexing of keypoint descriptors provides at least one good match within the first 20 sorted based on the inter-descriptor distance. Out of 574 initializations tested, the estimation stage successfully refined most good and several bad initial transforms. The estimation failed on only 13 good initializations. Using the same testing set, the decision stage produced only 2 false positives and 29 false negatives when classifying alignments as successful or unsuccessful. The strongest classifier components were the average of angles between sheet structures and the average alignment error. Overall success rate of the algorithm was 97.3% out of all initial random locations.

Initialization of LRR at 134 nodule locations produced alignments for all nodules. The median mapping error was 1.70 mm

and in only one case, this error was large (12.3 mm). The average distance between regularly sampled locations surrounding the nodules mapped with the Diffeomorphic Demons and with LRR was 1.01 mm. This shows a good agreement of the local affine mapping and the deformation field. In comparison, the median mapping error produced by a global registration technique followed by local refinement was 3.57 mm. We have shown on a database of 116 tumors, that LRR initialized at tumor locations provides accurate alignments which are required for evaluating response to therapy and analyzing nodule growth. Furthermore, the LRR algorithm runs in 6 s per initial location on average.

We computed paired t-tests on the nodule mapping error distances. For each nodule, the distance is computed between the mapped location of a nodule from one volume and the location of a corresponding nodule in another volume. Statistics on the distances (mean and variance) are calculated for each technique and the final critical value t is compared against the tabulated distribution. Doing so, we found out that LRR algorithm provides significantly more accurate alignments than the Diffeomorphic Demons ($p < 0.04$) and the nodule registration technique ($p < 0.0001$).

5. Conclusion

We have presented a novel algorithm for Location Registration and Recognition (LRR) without solving the deformable registration first or simultaneously. We have shown the applicability of Shape-Contexts to indexing and how they can be used to obtain an initial transform. Overall experiments on the longitudinal CT scans of the lung show that the algorithm is able to handle changes between the local regions, is more robust and accurate than a global-to-local registration technique, is at least as accurate as the deformable registration result, and runs at interactive speeds. LRR includes a novel decision compo-

nent which decides whether a region in one scan has been correctly recognized in another scan and aligned with the corresponding region.

The algorithm is motivated by the need for computer systems for diagnosis and treatment monitoring of lung cancer. We have shown the effectiveness of the LRR algorithm in aligning regions surrounding lung nodules. Given a nodule location in one of the volumes, the algorithm correctly finds the corresponding nodule in another volume and successfully and accurately aligns the regions surrounding the nodules. The LRR algorithm is accurate (median error of 1.70 mm, see Table 2), reliable (tested at 250 nodule locations, see Sections 4.4 and 4.7), and fast (6 seconds per initial location after preprocessing of 5 min and 10 s per volume). Due to its robustness, the algorithm ignores changes caused by cancer progression (Section 4.7) and differences in breathing states, scanning procedures, and patient positioning. This makes LRR an effective and efficient tool for aligning regions surrounding lung nodules in the follow-up CT scans of the lung (Fig. 1).

Exploring other applications for the LRR algorithm is imminent since none of the algorithm components makes any assumptions on any particular anatomy. In current colonoscopy practice, the patient is scanned in prone and supine positions so that polyps can be reliably identified by comparing the two scans (Nain et al., 2002). In thoracic CT scans, correspondences between equivalent parts of the airway (Tschirren et al., 2005) or vessel trees (Charnoz et al., 2004) can be used for building atlases (Cool et al., 2003), for atlas-based segmentation and labeling of structures (Chillet et al., 2003), and for registration (Aylward et al., 2003). Aligning pre-operative and intraoperative data can be used for the navigation during surgical procedures (Lange et al., 2004). Many of these alignment methods could potentially benefit from the solution of the LRR problem.

Regions lacking structure to drive the registration may pose difficulty for the algorithm. To correct this, the neighborhood size might need to be chosen adaptively. Faster computation and increased robustness could be achieved by processing multiple neighborhoods concurrently. This is possible because nearby regions, each mapped by an affine transformation, will be close to each other after the mapping. When the neighborhoods are handled sequentially, the earlier results may be exploited in later processing. LRR results from multiple locations might serve as an initialization for a deformable registration algorithm. Alternatively, a piecewise affine deformable model may be defined by interpolating between the LRR affine transformations. This would be particularly effective in areas where anatomical structures move independently from one another, each being mapped individually by LRR. These areas are the main directions in our future work.

Acknowledgements

References

Agam, G., Armato, S.G.I., Wu, C., 2005. Vessel tree reconstruction in thoracic CT scans with application to nodule detection. IEEE Trans. Med. Imag. 24 (4), 486–499.

Armato, S., Meyer, C., Mcnitt-Gray, M., McLennan, G., Reeves, A., Croft, B., Clarke, L., 2008. The reference image database to evaluate response to therapy in lung cancer (RIDER) project: a resource for the development of change-analysis software. Clin. Pharmacol. Ther. 84 (4), 448–456.

Aylward, S., Jomier, J., Weeks, S., Bullitt, E., 2003. Registration of vascular images. Int. J. Comp. Vis. 55 (2–3), 123–138.

Azar, A., Xu, C., Pennec, X., Ayache, N., 2006. An interactive hybrid non-rigid registration framework for 3D medical images. In: IEEE Int. Symp. Biomed. Imaging, pp. 824–827.

Belongie, S., Malik, J., Puzicha, J., 2002. Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. 24 (4), 509–522.

Besl, P., McKay, N., 1992. A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell. 14 (2), 239–256.

Betke, M., Hong, H., Thomas, D., Prince, C., Ko, J.P., 2003. Landmark detection in the chest and registration of lung surfaces with an application to nodule registration. Med. Image Anal. 7, 265–281.

Boldea, V., Sarrut, D., Clippe, S., 2003. Lung deformation estimation with non-rigid registration for radiotherapy treatment. In: Proc. 6th MICCAI. pp. 770–777.

Brown, M., Lowe, D., 2007. Automatic panoramic image stitching using invariant features. Int. J. Comp. Vis. 74 (1), 59–73.

Cahill, N.D., Noble, J.A., Hawkes, D.J., 2009. A demons algorithm for image registration with locally adaptive regularization. In: Proc. 12th MICCAI, vol. 1, London, UK, pp. 574–581.

Camara, O., Delso, G., Colliot, O., Moreno-Ingelmo, A., Bloch, I., 2007. Explicit incorporation of prior anatomical information into a nonrigid registration of thoracic and abdominal CT and 18-FDG whole-body emission PET images. IEEE Trans. Med. Imag. 26 (2), 164–178.

Čech, J., Matas, J., Perďoch, M., 2008. Efficient sequential correspondence selection by cosegmentation. In: Proc. CVPR, Anchorage, AK.

Chang, C.-C., Lin, C.-J., 2001. LIBSVM: a library for support vector machines. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.

Charnoz, A., Agnus, V., Soler, L., 2004. Portal vein registration for the follow-up of hepatic tumours. In: Proc. 7th MICCAI, Saint-Malo, France, pp. 878–886.

Chen, Y., Medioni, G., 1992. Object modeling by registration of multiple range images. IVC 10 (3), 145–155.

Chillet, D., Jomier, J., Cool, D., Aylward, S., 2003. Vascular atlas formation using a vessel-to-image affine registration method. In: Proc. 6th MICCAI, pp. 335–342.

Chui, H., Rangarajan, A., Zhang, J., Leonard, C.M., 2004. Unsupervised learning of an atlas from unlabeled point-sets. IEEE Trans. Pattern Anal. Mach. Intell. 26 (2), 160–172.

Cool, D., Chillet, D., Kim, J., Guyon, J.-P., Foskey, M., Aylward, S., 2003. Tissue-based affine registration of brain images to form a vascular density atlas. In: Proc. 6th MICCAI, pp. 9–15.

Cootes, T., Marsland, S., Twining, C., Smith, K., Taylor, C., 2004. Groupwise diffeomorphic non-rigid registration for automatic model building. In: Proc. Eigth ECCV, pp. 316–327.

Duda, R.O., Hart, P.E., Stork, D.G., 2001. Pattern Classification. John Wiley and Sons.

El-Baz, A., Gimel'farb, G., Falk, R., El-Ghar, M.A., Rainey, S., Heredia, D., Shaffer, T., 2009. Toward early diagnosis of lung cancer. In: Proc. 12th MICCAI, vol. 2, London, UK, pp. 682–689.

Fan, Y., Shen, D., Davatzikos, C., 2005. Classification of structural images via high-dimensional image warping, robust feature extraction, and SVM. In: Proc. 8th MICCAI, Palm Springs, California, USA, pp. 1–8.

Ferrari, V., Tuytelaars, T., Gool, L.V., 2004. Simultaneous object recognition and segmentation by image exploration. In: Proc. Eigth ECCV, pp. 40–54.

Frome, A., Huber, D., Kolurri, R., Buelow, T., Malik, J., 2004. Recognizing objects in range data using regional point descriptors. In: Proc. Eigth ECCV.

Garcia, M., Jemal, A., Ward, E., Center, M., Hao, Y., Siegel, R., Thun, M., 2007. Global Cancer Facts and Figures, 2007. American Cancer Society, Atlanta, GA.

Ginneken, B.V., Romeny, B.M.T.H., Viergever, M.A., 2001. Computer-aided diagnosis in chest radiography: a survey. IEEE Trans. Med. Imag. 20 (12), 1228–1241.

Gorbunova, V., Lo, P., Ashraf, H., Dirksen, A., Nielsen, M., de Bruijne, M., 2008. Weight Preserving Image Registration for Monitoring Disease Progression in Lung CT. New York, NY, pp. 863–870.

Hartkens, T., Rohr, K., Stiehl, H.S., 2002. Evaluation of 3D operators for the detection of anatomical point landmarks in MR and CT images. Comput. Vis. Image Und. 86 (2), 118–136.

Hartley, R., Zisserman, A., 2000. Multiple View Geometry. Cambridge University Press.

Ibáñez, L., Schroeder, W., Ng, L., Cates, J., 2003. The ITK Software Guide: The Insight Segmentation and Registration Toolkit (version 1.4). Kitware Inc.

Jähne, B., 1993. Spatio-Temporal Image Processing: Theory and Scientific Applications. Springer-Verlag New York, Inc.

Kaus, M.R., Netsch, T., Kabus, S., Pekar, V., McNutt, T., Fischer, B., 2004. Estimation of organ motion from 4D CT for 4D radiation therapy planning of lung cancer. In: Proc. 7th MICCAI, Saint-Malo, France, pp. 1017–1024.

Kawata, Y., Niki, N., Ohmatsu, H., Kusumoto, M., Kakinuma, R., Mori, K., Nishiyama, H., Eguchi, K., Kaneko, M., Moriyama, N., 2001. Analysis of pulmonary nodule evolutions using a sequence of three-dimensional thoracic CT images. In: Proc. 4th MICCAI, pp. 103–110. Doi. 10.1007/3-540-45468-3_13.

Kelman, A., Sofka, M., Stewart, C.V., 2007. Keypoint descriptors for matching across multiple image modalities and non-linear intensity variations. In: Proc. IEEE CVPR Workshop on Image Registr. and Fusion, Minneapolis, MN.

Kostis, W.J., Reeves, A.P., Yankelevitz, D.F., Henschke, C.I., 2003. Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images. IEEE Trans. Med. Imag. 22 (10), 1259–1274.

Lai, Z., Hua, J., 2008. 3d surface matching and registration through shape images. In: Proc. 11th MICCAI, New York, NY, pp. 44–51.

Lange, T., Eulenstein, S., Hünerbein, M., Lamecker, H., Schlag, P.-M., 2004. Augmenting intraoperative 3d ultrasound with preoperative models for navigation in liver surgery. In: Proc. 7th MICCAI, Saint-Malo, France, pp. 534–541.

Lhuillier, M., Quan, L., 2005. A quasi-dense approach to surface reconstruction from uncalibrated images. IEEE Trans. Pattern Anal. Mach. Intell. 27 (3), 418–433.

Li, B., Christensen, G.E., McLennan, G., Hoffman, E.A., Reinhardt, J.M., 2003. Establishing a normative atlas of the human lung: inter-subject warping and registration of volumetric CT. Acad. Radiol. 10 (3), 255–265.

Liu, D., Chen, T., 2004. Soft shape context for iterative closest point registration. In: Proc. IEEE Int. Conf. Image Proc., vol. 2, Singapore, pp. 1081–1084.

Liu, C., Yuen, J., Torralba, A., 2009. Nonparametric scene parsing: Label transfer via dense scene alignment. In: Proc. CVPR, Miami, FL, pp. 1972–1979.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. Int. J. Comp. Vis. 60 (2), 91–110.

MacMahon, H., Austin, J.H.M., Gamsu, G., Herold, C.J., Jett, J.R., Naidich, D.P., Edward, F., Patz, J., Swensen, S.J., 2005. Guidelines for management of small pulmonary nodules detected on CT scans: a statement from the Fleischner society. Radiology 237 (2), 395–400.

Matsopoulos, G.K., Mouravliansky, N.A., Asvestas, P.A., Delibasis, K.K., Kouloulias, V., 2005. Thoracic non-rigid registration combining self-organizing maps and radial basis functions. Med. Image Anal. 9 (3), 237–254.

Meer, P., 2004. Robust techniques for computer vision. In: Medioni, G., Kang, S.B. (Eds.), Emerging Topics in Computer Vision. Prentice Hall.

Meyer, C.R. et al., 2006. Evaluation of lung MDCT nodule annotation across radiologists and methods 13 (10), 1254–1265.

Mikolajczyk, K., Schmid, C., 2004. Scale and affine invariant interest point detectors. Int. J. Comp. Vis. 60 (1), 63–86.

Miller, J.V., Stewart, C.V., 18–20 Jun. 1996. MUSE: Robust surface fitting using unbiased scale estimates. In: Proc. CVPR, pp. 300–306.

Mori, G., Belongie, S., Malik, J., 2005. Efficient shape matching using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. 27 (11), 1832–1837.

Murphy, K., van Ginneken, B., Pluim, J., Klein, S., Staring, M., 2008. Semi-automatic reference standard construction for quantitative evaluation of lung CT registration. In: Proc. 11th MICCAI, New York, NY, pp. 1006–1013.

Nain, D., Haker, S., Grimson, W.E.L., Jr, E.C., Wells, W.W., Ji, H., Kikinis, R., Westin, C.-F., 2002. Intra-patient prone to supine colon registration for synchronized virtual colonoscopy. In: Proc. 5th MICCAI, pp. 573–580.

Okada, K., Huang, X., 2007. Robust click-point linking: Matching visually dissimilar local regions. In: Proc. IEEE Int. Workshop on Beyond Multiview Geometry: Robust Estimation and Organization of Shapes from Multiple Cue.

Pock, T., Urschler, M., Zach, C., Beichel, R., Bischof, H., 2007. A duality based algorithm for tv-l1-optical-flow image registration. In: Proc. 10th MICCAI. Brisbane, Australia, pp. 511–518. Doi 10.1007/978-3-540-75759-7_62.

Prokop, M., Galanski, M., Molen, A.V.D., Schaefer-prokop, C., 2000. Spiral and Multislice Computed Tomography of the Body. Thieme.

Reeves, A.P., Chan, A.B., Yankelevitz, D.F., Henschke, C.I., Kressler, B., Kostis, W.J., 2006. On measuring the change in size of pulmonary nodules. IEEE Trans. Med. Imag. 25 (4), 435–450.

Rohr, K., 1997. On 3d differential operators for detecting point landmarks. IVC 15 (3), 219–233.

Rusinkiewicz, S., Levoy, M., 2001. Efficient variants of the ICP algorithm. In: Proc. 3rd Int. Conf. on 3DIM, pp. 224–231.

Seshamani, S., Rajan, P., Kumar, R., Girgis, H., Dassopoulos, T., Mullin, G., Hager, G., 2009. A meta registration framework for lesion matching. In: Proc. 12th MICCAI, vol. 1, London, UK, pp. 582–589.

Shen, D., Davatzikos, C., 2002. HAMMER: hierarchical attribute matching mechanism for elastic registration. IEEE Trans. Med. Imag. 21 (11), 1421–1439.

Shen, H., Fan, L., Qian, J., Odry, B., Novak, C., , Naidich, D., 2002. Real-time correspondence between lung nodules in follow-up multi-slice high resolution CT studies. In: RSNA. Chicago, IL.

Shi, J., Sahiner, B., Chan, H.-P., Hadjiiski, L., Zhou, C., Cascade, P.N., Bogot, N., Kazerooni, E.A., Wu, Y.-T., We, J., 2007. Pulmonary nodule registration in serial CT scans based on rib anatomy and nodule template matching. Med. Phys. 34 (4), 1336–1347.

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

Sofka, M., Stewart, C.V., 2008. Location registration and recognition (LRR) for longitudinal evaluation of corresponding regions in CT volumes. In: Proc. 11th MICCAI, vol. 2, New York, NY, pp. 989–997.

Sofka, M., Yang, G., Stewart, C.V., 2007. Simultaneous covariance driven correspondence (CDC) and transformation estimation in the expectation maximization. In: Proc. CVPR. Minneapolis, MN.

Stewart, C.V., 1999. Robust parameter estimation in computer vision. SIAM Rev. 41 (3), 513–537.

Thirion, J.-P., 1998. Image matching as a diffusion process: an analogy to Maxwell's demons. Med. Image Anal. 2 (3).

Tschirren, J., McLennan, G., Palagyi, K., Hoffman, E.A., Sonka, M., 2005. Matching and anatomical labeling of human airway tree. IEEE Trans. Med. Imag. 24 (12), 1540–1547.

Urschler, M., Bauer, J., Ditt, H., Bischof, H., May 2006a. SIFT and shape context for feature-based nonlinear registration of thoracic CT images. In: Proc. European Conference Computer Vision Workshop on Computer Vision Approaches to Medical Image Analysis, Graz, Austria, pp. 73–84.

Urschler, M., Zach, C., Ditt, H., Bischof, H., 2006b. Automatic point landmark matching for regularizing nonlinear intensity registration: Application to thoracic CT images. In: Proc. 9th MICCAI, Copenhagen, Denmark, pp. 710–717.

Urschler, M., Kluckner, S., Bischof, H., 2007. A framework for comparison and evaluation of nonlinear intra-subject image registration algorithms. In: ISC/NA-MIC Workshop on Open Science at MICCAI 2007.

Vapnik, V.N., 1998. Statistical Learning Theory. Wiley.

Vercauteren, T., Pennec, X., Perchant, A., Ayache, N., 2007. Non-parametric diffeomorphic image registration with the demons algorithm. In: Proc. 10th MICCAI, Brisbane, Australia, pp. 319–326.

Vik, T., Kabus, S., von Berg, J., Ens, K., Dries, S., Klinder, T., Lorenz, C., 2008. Validation and comparison of registration methods for freebreathing 4D lungCT. In: Sahiner, B., Manning, D.J. (Eds.), Proc. SPIE, vol. 6917, Medical Imaging.

Walter, J.Scott, John Howington, S.F.B.M., Pisters, K., 2007. Treatment of non-small cell lung cancer stage I and stage II. Chest 132 (3), 234–242.

Wiemker, R., de Hoop, B., Kabus, S., Gietema, H., Opfer, R., Dharaiya, E., 2008. Performance study of a globally elastic locally rigid matching algorithm for follow-up chest CT. In: Sahiner, B., Manning, D.J. (Eds.), Proc. SPIE, vol. 69, Medical Imaging.

Wörz, S., Rohr, K., 2006. Localization of anatomical point landmarks in 3D medical images by fitting 3D parametric intensity models. Med. Image Anal. 10 (1), 41–58.

Xue, Z., Shen, D., Davatzikos, C., 2004. Determining correspondence in 3-D MR brain images using attribute vectors as morphological signatures of voxels. IEEE Trans. Med. Imag. 23 (10), 1276–1291.

Yang, G., Stewart, C.V., Sofka, M., Tsai, C.-L., 2007. Registration of challenging image pairs: initialization, estimation, and decision. IEEE Trans. Pattern Anal. Mach. Intell. 23 (11), 1973–1989.

Yin, Y., Hoffman, E.A., Lin, C.-L., 2009. Mass preserving nonrigid registration of CT lung images using cubic B-spline. Med. Phys. 36 (9), 4213–4222.

Zhao, B., James, L.P., Moskowitz, C.S., Guo, P., Ginsberg, M.S., Lefkowitz, R.A., Qin, Y., Riely, G.J., Kris, M.G., Schwartz, L.H., 2009. Evaluating variability in tumor measurements from same-day repeat CT scans of patients with nonsmall cell lung cancer. Radiology 252 (1), 263–272.