AUTHOR QUERY FORM

American Institute of Physics	Journal: Med. Phys. Article Number: 045202MPH	Please provide your responses and any corrections by annotating this PDF and uploading it to AIP's eProof website as detailed in the Welcome email.

Dear Author,

Below are the queries associated with your article; please answer all of these queries before sending the proof back to AIP. Author please indicate the correct color processing option from the list below:

1. Author, please confirm Figure number(s) that should appear as color in print. Please know that any associated mandatory fees will apply for figures printed in color.
2. Author, please confirm Figure number(s) that should appear as color online only, there will be no fees applied.
3. Author, your paper currently does not include any color figures for online or print. If color is needed please indicate which figures it should be applied to and whether it is color in print or online.

Location in article	Query / Remark: click on the Q link to navigate to the appropriate spot in the proof. There, insert your comments as a PDF annotation.
AQ1	Please check whether all affiliation are correct as given.
AQ2	Please check and confirm whether 'Section II A' refers to the section in Ref. 22.
AQ3	Please note that websites are not allowed in text. So, the website has been moved to the References section and the remaining references renumbered. Please check and confirm.
AQ4	Please provide the volume number for Ref. 10.
AQ5	Please provide the location of the publisher for Refs. 18 and 19.
AQ6	Please provide the initials for the authors in Ref. 31.
AQ7	Please provide the complete page range for Ref. 37 (originally Ref. 36), if possible.
AQ8	Please reword Fig. 4 caption without color words, so that the figure will be understood by readers of the black-and-white print version.

Thank you for your assistance.

Automated segmentation of a motion mask to preserve sliding motion in deformable registration of thoracic CT

Jef Vandemeulebroucke[a)]
University of Lyon, CREATIS; Inserm U1044; CNRS UMR5220; INSA-Lyon, France; Léon Bérard Cancer Center, University of Lyon, F-69373, Lyon, France; and Center for Machine Perception, Czech Technical University in Prague, Czech Republic

Olivier Bernard
University of Lyon, CREATIS; Inserm U1044; CNRS UMR5220; and INSA-Lyon, France

Simon Rit
University of Lyon, CREATIS; Inserm U1044; CNRS UMR5220; INSA-Lyon, France; and Léon Bérard Cancer Center, University of Lyon, F-69373, Lyon, France

Jan Kybic
Center for Machine Perception, Czech Technical University in Prague, Czech Republic

Patrick Clarysse
University of Lyon, CREATIS; Inserm U1044; CNRS UMR5220; and INSA-Lyon, France

David Sarrut
University of Lyon, CREATIS; Inserm U1044; CNRS UMR5220; INSA-Lyon, France; and University of Lyon, Léon Bérard Cancer Center, F-69373, Lyon, France

(Received 30 June 2011; revised 31 October 2011; accepted for publication 2 January 2012; published xx xx xxxx)

Purpose: Deformable registration generally relies on the assumption that the sought spatial transformation is smooth. Yet, breathing motion involves sliding of the lung with respect to the chest wall, causing a discontinuity in the motion field, and the smoothness assumption can lead to poor matching accuracy. In response, alternative registration methods have been proposed, several of which rely on prior segmentations. We propose an original method for automatically extracting a particular segmentation, called a motion mask, from a CT image of the thorax.
Methods: The motion mask separates moving from less-moving regions, conveniently allowing simultaneous estimation of their motion, while providing an interface where sliding occurs. The sought segmentation is subanatomical and based on physiological considerations, rather than organ boundaries. We therefore first extract clear anatomical features from the image, with respect to which the mask is defined. Level sets are then used to obtain smooth surfaces interpolating these features. The resulting procedure comes down to a monitored level set segmentation of binary label images. The method was applied to sixteen inhale-exhale image pairs. To illustrate the suitability of the motion masks, they were used during deformable registration of the thorax.
Results: For all patients, the obtained motion masks complied with the physiological requirements and were consistent with respect to patient anatomy between inhale and exhale. Registration using the motion mask resulted in higher matching accuracy for all patients, and the improvement was statistically significant. Registration performance was comparable to that obtained using lung masks when considering the entire lung region, but the use of motion masks led to significantly better matching near the diaphragm and mediastinum, for the bony anatomy and for the trachea. The use of the masks was shown to facilitate the registration, allowing to reduce the complexity of the spatial transformation considerably, while maintaining matching accuracy.
Conclusions: We proposed an automated segmentation method for obtaining motion masks, capable of facilitating deformable registration of the thorax. The use of motion masks during registration leads to matching accuracies comparable to the use of lung masks for the lung region but motion masks are more suitable when registering the entire thorax. © *2012 American Association of Physicists in Medicine*. [DOI: 10.1118/1.3679009]

Key words: deformable registration, respiratory motion, lung cancer

I. INTRODUCTION

In radiation therapy, deformable image registration of computed tomography (CT) images of the thorax has been extensively used for a variety of tasks[1,2] and is a key enabling tool for 4D radiotherapy.[3] Image registration aims at finding a suitable spatial transformation such that a transformed target image becomes similar to a reference image. The underlying numerical problem is ill-posed, and explicit restrictions should encode the physical understanding of the sought deformation and drive the algorithm to solutions with plausible and desirable properties. In particular, the assumption of spatial smoothness of the transformation is widely used to estimate motion induced deformations. Depending on the registration method, this can be accomplished by expressing the transformation using smooth basis functions, through smoothness constraints and by including regularization penalties in the optimization framework favoring smooth solutions.

Respiratory motion involves sliding of the lung, diaphragm, and liver against the pleural wall. In this case, adjacent structures are moving independently with respect to each other, and the motion field is discontinuous. Homogeneous smoothing of the transformation will result in locally reduced registration accuracy in these regions,[4,5] as it contradicts the physiology of the motion.

The issue of sliding motion in deformable image registration has been addressed in a number of ways. Recently, preliminary results have been reported for specifically designed regularization schemes. Wolthaus *et al.*[6] used tissue-dependent filtering for the deformation field, using the density measure from the CT image to differentiate between regions. Motion estimation improved for the lung region but was still prone to error near the diaphragm and upper abdomen where density is similar to that of the thoracic wall. Ruan *et al.*[7] described a class of discontinuity preserving regularization schemes. Unfortunately, these may preserve other undesirable flow singularities. In response, a robust energy functional was proposed[8] to discriminantly preserve large shear. Similarly, Chun *et al.*[9] modified an invertibility penalty of the craniocaudal deformation component to encourage the preservation of large sliding motion. Visual assessment showed an improved representation of the sliding motion along that direction.

Alternatively, sliding was addressed by using a spatial prior about the sliding interface. Most authors have opted for manual or automatic segmentations of the lungs, which can be extracted easily. Kabus *et al.*[10] and McClelland *et al.*[11] masked the background to the segmented lungs in both images during registration. This procedure removes the influence of the neighboring tissue, but the obtained motion estimate is only valid for the studied object. Siebenthal *et al.*[12] manually segmented the liver in the reference image, while Xie *et al.*[13] used manual segmentations of the thoracic and abdominal cavity in the reference image to account for the sliding motion.

Several authors proposed to use a manual, subanatomical segmentation of the inner thoracic structures in both images.[5,14,15] Rather than segmenting individual organs, this approach divides the thorax into moving (lungs, mediastinum, and abdomen) and less-moving (the remainder) regions, which is why we will refer to this segmentation as a *motion mask*. Each region is registered separately, and the solution is composed for the entire image. To avoid gaps in the composed deformation field, Wu *et al.*[5] introduced a boundary matching criterion, penalizing a potential mismatch between the respective borders. As a consequence, segmentations must be consistent with respect to the patient anatomy to avoid inducing errors during subsequent registration. A similar criterion was used by McClelland *et al.*,[11] who included a mechanism to limit the effect of small segmentation errors of the lungs.

An alternative registration approach was proposed by Schmidt-Richberg *et al.*,[4] who locally modified a diffusive regularization with respect to a given lung segmentation. Their method allows for discontinuities in the motion field in the direction tangential to the interface, required to preserve sliding motion. The continuity in the normal direction is however maintained, eliminating gaps in the motion field. Similar results were obtained by Delmon *et al.*,[16] who decomposed a B-spline deformation grid to represent sliding at a given interface. In both cases, only one segmentation is required, and the entire image can be processed simultaneously.

In this work, we focus on the motion mask, as it has several advantages over other segmentations. With respect to lung masks, motion masks provide a physiologically more complete description of the regions where sliding motion occurs. By continuing below the diaphragm and encompassing the mediastinum, two clearly separated regions are obtained. This facilitates motion estimation for the entire thorax by simplifying the composition of the registration results obtained for each region, or the application of a specific regularization along the interface. Organs with similar motion are grouped, conveniently allowing to jointly estimate the motion for similarly moving tissue; conversely, to adapt the registration to each of the motion regions.

We propose a fully automated method for extracting a motion mask from a CT image of the thorax. To our knowledge, this is the first automated segmentation method to be proposed for this purpose. The obtained masks can be used in combination with any of the previously mentioned registration methods requiring an *a priori* segmentation. The method is applied to inhale and exhale images originating from 4D CT sets of 16 thoracic cancer patients. The suitability of the automatically obtained masks is illustrated by using them during deformable registration following the method proposed by Wu *et al.*,[5] and compared to conventional registration, and registration using lung masks.

II. METHOD

II.A. Motion mask definition

We briefly review the mechanics of respiration with the objective of establishing a physiological prior for the motion mask segmentation. Anatomically (Fig. 1), each lung is located within a pleural sac, which is made up of two membranes called the *pleurae*. The outer parietal pleura is adherent to the internal surface of the thoracic cavity, the diaphragm, and the mediastinum. The inner visceral pleura covers the lung and is adherent to its surface. Both inner and outer pleura join at the root of the lung, which is the point of entry of bronchi, vessels, and nerves into the lung. The space enclosed between the pleurae is called the pleural cavity, which is filled with liquid.[17] Below the diaphragm, the parietal pleura continues to the costodiaphragmatic recess, a potential space not occupied by the lung during normal tidal breathing, where the costal and diaphragmatic part of the parietal pleura meet.[18]

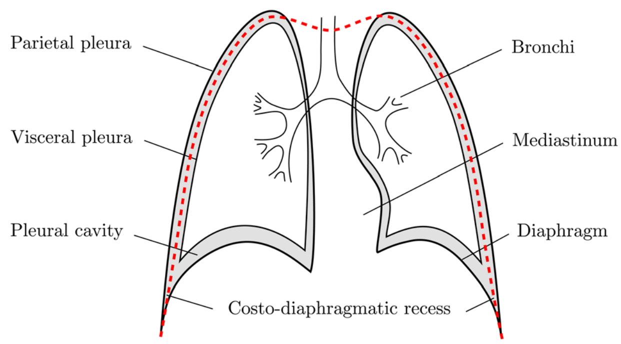

FIG. 1. Overview of the relevant anatomy of the thorax. The dashed line defines the contour of the motion mask.

During respiration, the breathing muscles—mainly the diaphragm and intercostal muscles—contract, which causes the thoracic cage and subsequently the lungs to expand. As the lungs inflate inside the thorax, sliding can occur between the membranes. At the lung-to-mediastinum interface, sliding is limited due to the entry of the vessels, bronchi, and nerves. These, along with the heart and other structures in the mediastinum, tend to move with the lung, though usually with reduced amplitude. At the interface of the lungs with the chest wall, the pleurae are free to slide with respect to each other. The inferior, posterior part of the lungs near the diaphragm tends to exhibit the largest sliding. At the anterior side of the lung-to-chest interface, sliding motion is small as the diaphragm is attached to the sternum, limiting the extent of motion (see Gray,[19] chapter IV.6.c). Below the diaphragm, the presence of the parietal pleura allows sliding of the liver and upper abdomen against the chest wall. Figure 4 shows sagittal views of the exhale and inhale images in color overlay, allowing to identify the regions where strong sliding motion occurs.

We define the sought motion mask as follows (Fig. 1). To preserve sliding motion, the segmentation should provide a separation between the lungs and the chest wall. At the medial lung interface, there is a continuous and smooth transition of motion, making it more convenient to consider the mediastinum together with the lungs. Below the diaphragm, the segmentation should continue downwards. Though the extent of the costodiaphragmatic recess is usually not visible on CT images, it should at least reach below the diaphragm position of the inhale frame, thereby including the liver and upper abdomen. The strong correlation of the motion of this region with the diaphragm and lower lungs justifies this choice. Further below the diaphragm, the motion mask is not defined and should therefore not be used for registering the entire abdomen.

II.B. Motion mask extraction

The core of the method is based on the level set framework,[20] from which we exploit the intrinsic smoothing property, which allows to include geometric priors in the definition of the motion mask. The conventional level set segmentation problem is simplified by applying it to binary images. The available image information is strongly reduced prior to processing, only retaining clear anatomical structures with respect to which we define the location of the segmentation.

We can thus divide the method for obtaining the motion mask into two parts (Fig. 2). First, a preprocessing step is performed during which the CT images are reduced to binary *label* images containing only the relevant anatomical features.

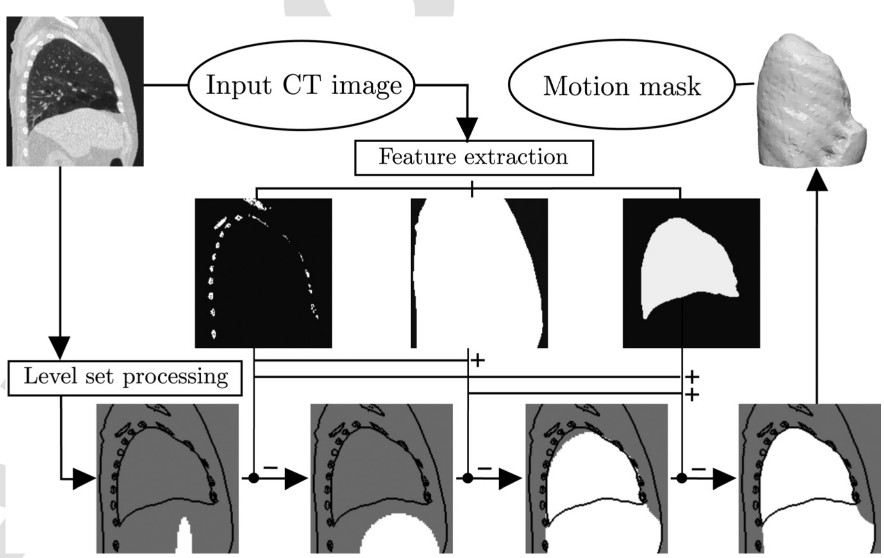

FIG. 2. Overview of the proposed method for extracting the motion mask. The figure shows sagittal views through the right lung of the results obtained for patient 1. The top row shows the input CT image, and a 3D surface rendering of the corresponding motion mask obtained using the method. The second row shows the label images, obtained by extracting anatomical features from the CT image. This yields the bony anatomy, the patient body and the lungs. The label images are then combined (+) and used to constrain (−) the evolving interface during consecutive level set processing steps, the results of which are shown in the bottom row. The current mask (*white*) is shown in overlay with the edges of the extracted features (*black*). From left to right are shown: the centered ellipsoid used to initialize the level set, the contour after reaching the detection point just in front of the anterior patient-to-air interface (the detection point—located in the center sagittal plane—projected onto this plane, would be located in the lower right corner of the image), the contour after having covered 95% of the lungs, and the final motion mask.

Next, these images are combined and used to control the evolving interface in consecutive level set processing steps.

II.B.1. Feature extraction

The method requires binary label images of the outer patient body contour, the bony anatomy, and the lungs. In radiotherapy, some or all of these segmentations are often readily available due to the treatment planning process. Alternatively, methods described in literature, allowing to automatically detect these features can be used. For completeness, we will give a brief overview of the procedure followed here and refer to the relevant work for a detailed description. The basic operations are thresholding, mathematical morphology, and region growing. In particular, three-dimensional (3D) connected component labeling using a 26-voxel connectivity is frequently applied, which amounts to labeling each distinct object in a binary image. By sorting the labeled objects with respect to the number of voxels, a selection can be made based on the object size. In the following text, we will assume that patients were imaged from above the lungs to approximately 15 cm below the diaphragm in the exhale position.

The patient body: The image is first binarized by thresholding at -300 Hounsfield units (HU), and the largest connected component of the low intensity regions is the air surrounding the patient. Connected component labeling of all remaining regions yields the patient body as the principal label.

The bony anatomy: Depending on the image quality, extracting the complete bony anatomy can be challenging. However, only an approximate segmentation of the rib cage is required, since this label is only used for constraining the evolving level set (see Sec. II B 1). We first perform edge preserving smoothing, using anisotropic diffusion.[21] The largest connected component, after binarizing with a lower threshold of 100 HU, corresponds to the main connected bony structure, i.e., column, vertebrae, ribs, and sternum. For images obtained using a contrast agent, this step might have to be modified.

The lungs: The quality of the lung label image directly influences the aspect of the final motion mask (see Sec. II B 2). The procedure adopted is largely based on a segmentation method described by van Rikxoort *et al.* (Ref. 22, Section II A). First, thresholding at -300 HU is applied and only the second largest label is retained, corresponding to the lungs, bronchi, and trachea. The trachea is detected in the top axial slices, and trachea and bronchi are identified using explosion-controlled region growing.[23] The lungs and airways are then segmented using Otsu thresholding[24] from which the previously detected trachea and bronchi are removed. Morphological closing using a 4 mm kernel radius is applied on the result.

II.B.2. Level set processing

Level sets, originally proposed by Osher and Sethian,[20] correspond to a numerical method for tracking the evolution of an interface. Let Ω be a bounded open subset of $\mathbb{R}^d$. In the level set formalism, the evolving interface $\Gamma \subset \mathbb{R}^d$ at time τ is embedded as the zero level of a Lipschitz-continuous *level set function* $\varphi : \mathbb{R}^{d+1} \mapsto \mathbb{R}$, that satisfies

$$\begin{cases} \varphi(\boldsymbol{x},\tau) < 0 & \text{for} \quad \boldsymbol{x} \in \Omega_{\text{in}} \\ \varphi(\boldsymbol{x},\tau) > 0 & \text{for} \quad \boldsymbol{x} \notin \Omega_{\text{in}} \\ \varphi(\boldsymbol{x},\tau) = 0 & \text{for} \quad \boldsymbol{x} \in \Gamma, \end{cases} \tag{1}$$

where Ω_{in} is a region in Ω bounded by $\Gamma = \partial\Omega_{\text{in.}}$ In this work, the evolution of the level set is governed by the following expression:[25]

$$\frac{\partial \varphi(\boldsymbol{x},\tau)}{\partial \tau} = (\alpha v(\boldsymbol{x}) + \beta v(\boldsymbol{x})\kappa)\|\nabla\varphi(\boldsymbol{x},\tau)\|. \tag{2}$$

In Eq. (2), κ is the curvature calculated on the zero-level, v is a scalar velocity map derived from the image, and α and β are scalar constants introduced to balance the relative influence of each of the terms. The first term provides a propagation force, favoring an expansion or contraction of the contour depending on the sign of α. The second term will penalize high curvature and serves as a spatial regularization limiting the complexity of the shape of the interface.

Note that initially, level set methods were introduced to model the front propagation of an interface.[20] Afterwards, they were applied to medical image segmentation to automatically detect the boundaries of structures of interest[26,27]—in this case obtained as the steady state solution $\frac{\partial\varphi}{\partial t} = 0$. We propose to take benefit of both uses. The level set framework is used as a high-level tool to propagate a 3D interface with a global regularization of its shape. The propagation of the interface is controlled through velocity maps v, obtained as a combination of the extracted binary label images. These velocity maps will define two types of regions in Ω: one where the interface evolves with isotropic speed ($v(\boldsymbol{x}) = 1$), and one where the level set is confined to its current state ($v(\boldsymbol{x}) = 0$). For each of the level set processing steps, we define a stopping criteria directly linked to the extracted anatomical structures.

Following the considerations made in Sec. II A, the sought region Ω_{in} extends beyond the field of view at the inferior end of the image, making Ω unbounded with respect to the original image size. To remedy this, all velocity maps are mirrored with respect to the inferior axial plane (Fig. 3). From the resulting mask, only the part covering the original region is retained. A strong geometric prior was established during the entire procedure by means of a high curvature scaling: $\beta = 30$ while $|\alpha| \leq 1$.

Initializing the level set in the abdomen: The main goal of this step is to provide a stable initialization, to ensure a proper inclusion of the abdominal region and to reach the anterior patient-to-air interface. The initial contour is taken from a small ellipsoid centered at the patient body, and moved forward until there is no more overlap with the bony anatomy. The level set is initialized with the signed distance map of the ellipsoid and let to evolve with $v(\boldsymbol{x}) = 1$, except at the bony anatomy where $v(\boldsymbol{x}) = 0$. A positive propagation ($\alpha = 1$) ensures a growing interface. The evolution of the contour is monitored and stopped when a detection point is reached, placed 10 mm outside of the patient contour, in front of the most inferior patient-to-air interface, and centered with respect to the patient.

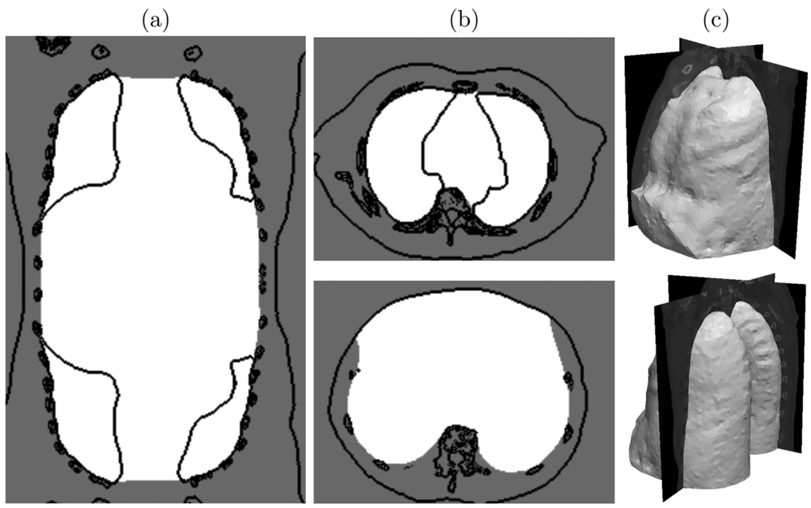

FIG. 3. The final motion mask: (a) a coronal view of the final mask shown on the edges of the used *mirrored* version of the binary label images; (b) two axial views of the mask: the top one taken halfway through the lungs and the bottom one taken from the most inferior plane of the image; (c) 3D surface renderings of the anterior (*top*) and posterior (*bottom*) view of the motion mask.

Filling the thoracic cavity: Next, we wish to fill the entire thoracic cavity including the lungs and mediastinum. To this end, the previous result is propagated further ($\alpha = 1$), but the underlying velocity field is altered so that in addition to the bony anatomy, everything outside the patient body yields $v(\boldsymbol{x}) = 0$. The part of the interface which has evolved outside the patient body is now confined to its current position. The remainder of the level set is let to propagate with unit velocity inside the thoracic cavity while the coverage of the extracted lungs is monitored. When the contour covers at least 95% of the lungs, the algorithm is terminated, and the full lung region is included by calculating the union of the resulting mask with the lung mask. The execution is terminated at 95% rather than 100% for reasons of efficiency, the upper part of the lungs requiring a lot of iterations while only marginally modifying the aspect of the contour.

Smoothing to the lungs: In this final step, we refine the previous solution to obtain a smooth contour that adopts the outer shape of the lungs but includes the mediastinum and upper abdomen. The velocity map employed is a unit field everywhere except outside the body, at the bony anatomy and in the lungs. Only the curvature term is retained during this phase ($\beta = 30$), no propagation force was included ($\alpha = 0$). In practice, this will lead to contraction as curvature is integrated along contour length. As the driving force is now only determined by the local curvature, regions with strong curvature initially evolve quickly while others remain virtually unchanged. The evolution conveniently slows down as the contour becomes smoother, making the total number of iterations not very critical. This step is run for 500 iterations, which was empirically found to be sufficient to smoothen the mask. Note that, by propagating the contour outside the patient body in the first level set processing step, and applying $v(\boldsymbol{x}) = 0$ to the air around the body afterwards, we can avoid that the contour recedes and no longer includes the entire abdomen. The motion mask is limited to the patient body using the corresponding mask.

III. EXPERIMENTS

III.A. Motion mask extraction

The method was applied to the exhale and inhale frames of 4D CT images of the thorax of 16 patients with thoracic malignancies (esophagus or lung cancer). All images were part of a radiotherapy planning protocol.

The first six data sets were acquired on a Brilliance Big Bore 16-slice CT scanner (Philips Medical Systems, Cleveland, OH) at our institute, the Léon Bérard Cancer Center in Lyon, France. Retrospective respiratory-correlated reconstruction into ten 3D CT images was made possible by simultaneous recording of a respiratory trace using the Pneumo Chest Bellows (Lafayette Instrument, Lafayette, IN).

The remaining 10 data sets are part of a publicly available deformable image registration reference database.[28] They AQ3 were acquired on a Discovery ST PET/CT scanner (GE Medical Systems, Waukesha, WI), at the University of Texas M. D. Anderson Cancer Center in Houston, Texas. The respiratory signal was obtained from the Real-Time Position Management Respiratory Gating System (Varian Medical Systems, Palo Alto, CA).

The original resolution was approximately $1 \times 1 \times 2$ mm for the first sets, and $1 \times 1 \times 2.5$ mm for the remaining sets. Features were extracted using the original images to ensure optimal image quality. Prior to level set processing and registration, all images were resampled to a 2 mm isotropic voxel size.

III.B. Deformable registration per region

III.B.1. *Registration method*

The suitability of the obtained masks was verified by applying them to deformable registration of the lungs. We used the publicly available ITK (Ref. 29) implementation of free-form deformations based on cubic B-splines.[30] A multi-resolution approach with three levels was used for the images as well as the B-spline control point grid, the final level having a 2 mm and 32 mm spacing, respectively. Image intensities were interpolated using cubic B-splines. Similarity was measured through the sum of squared differences and optimized by the limited memory BFGS algorithm[31] starting from an initial zero deformation vector field.

The motion masks were incorporated in the registration framework using the method proposed by Wu *et al.*[5] In this procedure, a mask is calculated for both the reference and target image and used to mask the background of the considered region. Each region is then registered separately, and the full deformation field is obtained by composing the results. During registration, a thin border of voxels of 10 mm outside the considered region is included in the calculation of the similarity measure. This contribution penalizes a potential mismatch between the region borders, reducing the presence of gaps in the composed deformation field but without constraining the sliding motion. The results were compared to conventional registration without the use of a mask, and to registration using the lung masks extracted in Sec. II B 1. The registration method for the latter was the same as for the motion masks.

III.B.2. Evaluation method

The registration performance for the lungs was evaluated by assessing the matching accuracy of anatomical landmarks in the lungs. We calculated the target registration error (TRE), which is defined as the distance between the manual annotation in the target image, and the corresponding point in the reference image after being displaced by the registration results.

The landmarks for the first six patients were identified using a semiautomatic software tool proposed by Murphy *et al.*[32] The landmark identification procedure followed is described in Ref. 33 and resulted in approximately 100 landmarks per image pair. The landmarks for the remaining images were identified following the methodology described in Refs. 34, 35 and consisted of 300 landmarks per image pair. In total, 3620 landmarks were available inside the lungs. Two subsets of the landmarks were created in order to allow local evaluation of the registration accuracy in the lungs. For the first, all landmarks within 10 mm of the chest wall were selected, leading to a total of 757 landmarks. The second subset was based on all points within 10 mm of the remaining lung borders, i.e., the diaphragm and mediastinum and consisted of 636 landmarks.

In addition, we assessed the registration performance outside the lungs by evaluating the overlap between anatomical features extracted from the image pair. We used the bony anatomy, and the trachea along with the bronchi, detected as described in Sec. II B 1. The features in the target image were deformed using the registration results and we calculated the Dice similarity coefficient[36] (DSC) with the features in the reference image. The DSC of two label images is defined as the ratio of the number of voxels in the intersection to the mean label volume.

IV. RESULTS

IV.A. Motion mask extraction

The procedure gave satisfying results for all tested images. All masks complied with the general requirements given in Sec. II A, i.e., the segmentations encompassed the mediastinum and upper abdomen and conformed to the lungs near the chest wall. The calculation of the label images required less than a minute per image. The level set processing time required 6 min 44 s on average while the longest execution time recorded was 9 min 04 s (on a single 2.4 GHz CPU).

Sagittal views of the final and intermediate results of the feature extraction and the level set processing steps, obtained for patient 1, are shown in Fig. 2. In Fig. 3(a), the final segmentation obtained for patient 1 is depicted in overlay with the edges of the full, mirrored label image used during the final level set propagation step. In addition, two axial views [Fig. 3(b)] and an anterior and posterior view of a 3D surface rendering of the motion mask [Fig. 3(c)] are shown.

Figure 4 shows color overlays of the exhale and inhale images for three patients, along with the contours of the motion masks extracted for each of the images. The masks corresponding to exhale and inhale are overall very similar as a consequence of the predominantly diaphragmatic respiration and consistent mask extraction. Only in Fig. 4(c), stronger differences are noticeable due to larger chest and rib motion.

Note that, as the motion mask is a subanatomical segmentation relying on geometric and physiological priors, it is difficult to directly evaluate the accuracy of the obtained segmentations. As an alternative, their usefulness will be quantified in the following section.

IV.B. Deformable registration per region

The masks were constructed for both the end-exhalation frame and the end-inhalation frame of each 4D CT. They were then used to modify the images as described in Wu *et al.*,[5] and the inner and outer thoracic regions were registered separately.

In Fig. 5, a qualitative comparison of the registration results obtained for patient 1 is given. Difference images between the reference and the warped target image are shown in a coronal plane [Fig. 5(a)]. Enlarged views of the diaphragm are shown

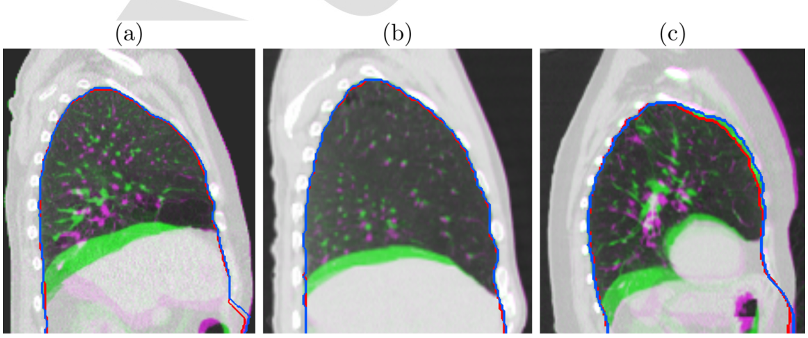

FIG. 4. Sagittal views of green-purple color overlays in of the exhale and inhale image pairs for three patients, corresponding to patients 1, 8, and 14. The contours of the motion mask extracted for each of the images is also shown for the exhale (*red*) and the inhale image (*blue*).

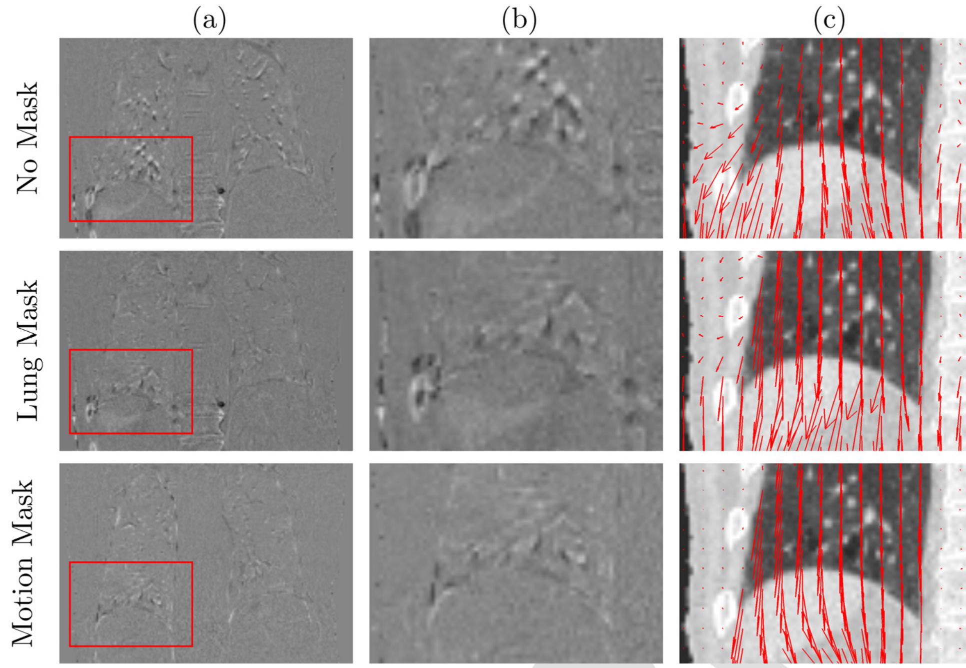

FIG. 5. Comparison of the results obtained for patient 1 for conventional registration, registration using a lung mask, and using the motion mask. Column (a) shows the difference of the reference and the target images when compensated with the obtained motion estimate, while (b) shows an enlarged view of the highlighted region, and (c) is the deformation vector field for that same region.

in Fig. 5(b), along with the respective deformation vector fields [Fig. 5(c)].

Conventional registration shows large motion estimates for the lower ribs and column, leading to poor matching accuracy in these regions. Using a lung mask, the matching of the lungs clearly improves, but errors are still present near the diaphragm. Using a motion mask, a discontinuity in the deformation field can be represented below the diaphragm position of the exhale frame, resulting in an improved matching of the surrounding tissue and bony anatomy.

Table I contains the quantitative evaluation of the registration results for the lungs using all landmarks. We compare the distance between the landmarks before registration, the mean TRE after conventional registration, after registration using the lung mask, and using the motion mask. Table II contains the mean TRE over all patients, the mean TRE based on subsets of the landmarks and the overlap measures. The registration results are further compared by computing a paired t-test over all patients and considered statistically different for p-values lower than 5×10^{-2}.

The use of the motion mask improved the registration accuracy for all patients, with respect to registration without using a mask. The mean TRE over all patients and all landmarks (Table III, A) improved from 2.76 mm $\pm$ 3.14 mm to 1.75 mm $\pm$ 1.52 mm, and the improvement was statistically significant (p-value $= 6.0 \times 10^{-3}$). The difference in performance was greater when only considering the landmarks near the chest wall (Table II, B). We also evaluated the signed average along all components (not shown in Table I), which revealed that the largest errors were found along the craniocaudal direction, which is also where the largest displacements take place. The relatively large bias of 1.75 mm for the conventional registration is reduced to below 0.3 mm when using the motion mask.

Registration using a lung mask performs comparable to the motion mask when looking at the entire lung region. Slightly better results were obtained when using the motion mask, but the difference was not significant (p-value $= 3.2 \times 10^{-1}$). Results become virtually identical near the chest wall, but when limiting the evaluation to the remaining lung borders (i.e., the diaphragm and mediastinum, Table II, C), results were found to be significantly better for the motion mask (p-value $= 1.2 \times 10^{-2}$). The TRE based on the complementary subset of landmarks (i.e., all landmarks within the lungs, not within a 10 mm distance of the diaphragm and mediastinum) was still lower for the motion mask, but the difference was not significant.

Similar observations can be made regarding the performance outside the lung region. The overlap of the bony anatomy (Table II, D)—subject to little motion as can be seen from the large DSC before registration—is significantly better for the motion mask (p-value $< 10^{-3}$). In fact, the initial overlap is not improved by conventional registration nor registration using a lung mask. The overlap of the highly mobile trachea and bronchi (Table II, E) was improved by all registration methods. The best match was obtained using the motion mask, and the difference with the lung mask was significant (p-value $< 10^{-3}$).

V. DISCUSSION

Particular attention was paid to making the automatic motion mask extraction reliable and robust, in order to limit the required user interaction in a clinical setting. To this end,

TABLE I. The distance between all landmarks in the lungs before registration (*BR*), the TRE after conventional registration without using a mask (*NM*), after registration using a lung mask (*LM*) and using the motion mask (*MM*). Given are the mean values (μ) and standard deviation (σ), and the maximum value over all points (Max).

Method	Patient	TRE (mm) $\mu \pm \sigma$	TRE (mm) Max	Patient	TRE (mm) $\mu \pm \sigma$	TRE (mm) Max
BR	1	14.00 ± 7.17	32.4	9	6.94 ± 4.05	16.6
NM		2.91 ± 3.45	21.7		2.36 ± 2.02	10.9
LM		1.43 ± 1.06	8.0		1.93 ± 1.44	11.8
MM		1.42 ± 1.13	6.5		1.69 ± 1.12	7.1
BR	2	6.34 ± 2.94	16.0	10	9.83 ± 4.85	20.3
NM		1.05 ± 0.75	4.5		2.64 ± 2.03	11.9
LM		1.02 ± 0.62	3.3		2.00 ± 1.35	11.5
MM		1.00 ± 0.55	3.0		1.82 ± 1.14	10.7
BR	3	7.67 ± 5.03	24.5	11	7.48 ± 5.50	24.8
NM		1.84 ± 1.91	9.9		3.30 ± 3.13	16.2
LM		1.27 ± 1.19	11.2		2.60 ± 2.47	18.3
MM		1.44 ± 1.56	13.5		2.75 ± 2.45	16.9
BR	4	6.68 ± 3.67	14.2	12	10.90 ± 6.96	27.6
NM		1.49 ± 1.46	8.8		3.63 ± 3.40	19.0
LM		1.31 ± 1.16	9.1		2.06 ± 1.37	13.0
MM		1.29 ± 0.95	6.7		2.01 ± 1.16	6.7
BR	5	7.09 ± 5.08	19.8	13	11.00 ± 7.42	30.6
NM		1.67 ± 1.77	12.1		4.63 ± 4.46	24.2
LM		1.44 ± 1.46	11.1		2.23 ± 1.68	10.2
MM		1.49 ± 1.46	11.0		2.15 ± 1.59	14.0
BR	6	7.33 ± 4.86	24.1	14	15.00 ± 9.00	30.6
NM		2.36 ± 3.12	21.2		7.13 ± 7.71	30.3
LM		2.01 ± 3.42	25.8		1.96 ± 1.72	19.4
MM		1.88 ± 2.86	20.8		2.11 ± 1.79	18.0
BR	7	3.89 ± 2.78	10.9	15	7.92 ± 3.97	15.8
NM		1.53 ± 1.11	6.4		3.03 ± 2.20	10.3
LM		1.62 ± 1.11	8.7		2.13 ± 1.26	7.1
MM		1.52 ± 0.92	6.1		2.05 ± 1.20	8.0
BR	8	4.34 ± 3.90	17.7	16	7.30 ± 6.34	27.8
NM		1.61 ± 1.66	11.2		2.91 ± 2.93	18.3
LM		1.31 ± 1.07	11.4		2.19 ± 2.08	18.7
MM		1.30 ± 1.03	8.9		2.12 ± 1.66	12.0

TABLE II. The top part of the table shows the group mean of the distance between landmarks before registration (*BR*), of the TRE after conventional registration without using a mask (*NM*), after registration using a lung mask (*LM*) and using the motion mask (*MM*). The TRE is calculated for all points in the lung based on 3620 measurements (*A*), for all points within 10 mm of the chest wall based on 757 measurements (*B*), and for all points within 10 mm of the diaphragm and mediastinum using 636 landmarks (*C*). The bottom part of the table lists the DSC for the extracted bony anatomy (*D*), and the trachea and bronchi (*E*). Given are the mean values (μ) and standard deviation (σ).

Measure	*BR*	*NM*	*LM*	*MM*
TRE (mm)	$\mu \pm \sigma$	$\mu \pm \sigma$	$\mu \pm \sigma$	$\mu \pm \sigma$
A	8.36 ± 5.49	2.76 ± 3.14	1.78 ± 1.66	1.75 ± 1.52
B	8.92 ± 5.71	4.82 ± 3.97	2.59 ± 2.71	2.59 ± 2.45
C	8.38 ± 5.17	2.03 ± 2.01	1.89 ± 1.68	1.71 ± 1.60
DSC (%)	$\mu \pm \sigma$	$\mu \pm \sigma$	$\mu \pm \sigma$	$\mu \pm \sigma$
D	91.3 ± 4.8	91.0 ± 2.5	91.0 ± 2.5	92.3 ± 2.3
E	57.0 ± 9.1	80.5 ± 4.5	79.2 ± 4.7	81.1 ± 4.2

the level set procedure was applied to label images identifying clear anatomical features. The evolution of the level set was monitored by defining stopping criteria directly related to these structures, thus eliminating additional convergence parameters. The proposed procedure comes down to a controlled level set segmentation of binary images. By design, the obtained segmentation is confined between the ribs and the lungs, includes the upper abdomen, and continues smoothly between the lungs and below the diaphragm.

The procedure requires previously extracted feature images, and the result can be affected by incorrect detection of these features. Within our group, the segmentation procedure has been applied to other images, outside this study. These include all 60 frames of the first six 4D CT data sets,[33] and the 60 images used in the Empire lung registration challenge.[37] This allowed us to identify issues when confronted to input images with varying quality and characteristics, and evaluate the sensitivity of the method to erroneous feature detection.

Few problems were encountered when segmenting the patient body. Depending on the patient set-up, the procedure described in Sec. II B 1 might include the scanner couch, which is not a problem. Images cropped to contain only the lungs should be padded prior to processing, to include a −1000 HU border on all sides.

The aspect of the lung mask directly influences the final aspect of the motion mask. While the lungs are usually easily segmented, malignancies in the lungs may be excluded when they are located at the pleural wall. More elaborate approaches, designed to deal with the pathological lung,[22] can handle such configurations and include the entire lung region. For subsequent registration, however, it will depend on whether the tumor moves together with the lung, or is adherent to the chest wall, which of the previous is the most favorable solution. We currently do not have an automated mechanism to deal with these cases. Motion-induced artifacts, frequently present in 4D CT images, sometimes affected the lung segmentations. The corresponding motion masks were however not influenced, as the impact on the outer shape of the lungs was small.

The method is less sensitive to incomplete detection of the bony anatomy, as this feature is only used to constrain the evolving interface. When entire ribs are missing from the label image, however, the impact will become noticeable. Depending on the image quality, detecting the complete rib cage can be challenging and problems have been encountered for images characterized by low resolution, low dose, and artifacts. Manual adaptation of the threshold for the bony anatomy extraction described in Sec. II B 1 resolved these issues. Similar interventions were required for contrast-enhanced CT. Alternatively, methods specifically devised to label the complete rib cage[38] or atlas-based approaches could reduce this influence to image quality.

An important parameter for the used B-spline free-form deformation transformation is the spacing of the control point

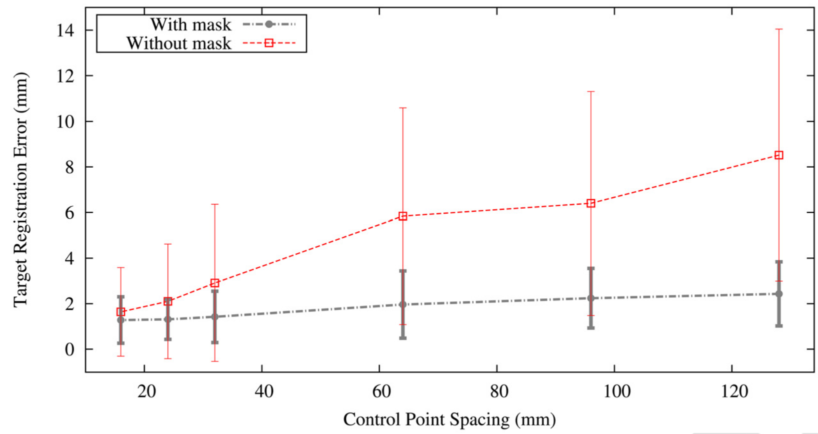

FIG. 6. The mean target registration error with respect to the landmarks is plotted in function of the B-spline control point spacing. The error bars correspond to one standard deviation.

grid. The issue of sliding motion has brought several authors to lower this spacing. As a finer grid is employed, the representation of discontinuities such as sliding improves. However, the complexity of the optimization increases rapidly with the number of parameters, along with the computation time. In addition, allowing more degrees of freedom increases sensitivity to noise and artifacts since the parameterization of the spatial transformation becomes less restrictive. The choice of the control point spacing is thus a trade-off between matching accuracy on one hand, and robustness and efficiency on the other.

In Fig. 6, the mean TRE (and standard deviation) obtained with and without motion mask are shown in function of the control point spacing for patient 1, characterized by large motion. We note that the result obtained with mask using a control point spacing of 128 mm (2.43 ± 1.41 mm), is better than the result obtained without mask using a control point spacing of 32 mm (2.91 ± 3.45 mm). This indicates that, despite the large motion, the lung deformation is inherently smooth and the improved registration accuracy—obtained by increasing the number of control points—is mainly due to a better representation of the sliding motion. Considering this, the role of the motion mask can thus be viewed as facilitating the registration by lowering the complexity for the spatial transform, while maintaining accuracy.

While this work specifically focused on the sliding motion of the lungs with respect to the chest wall, some principles may be generalized. Other anatomical sites present organs that deform and move independently with respect to the neighboring tissue, such as the bladder and prostate or the esophagus. In these cases, performing separate registrations with adapted parameters and using physiologically compatible subanatomical segmentations may improve registration results. Ding *et al.*[39] measured sliding between lung lobes using breath-hold exhale and inhale images. Registration accuracy was shown to improve when registering the segmented lobes separately. In 4D CT images, acquired during normal tidal breathing, we assumed lobar sliding was small and did not explicitly take it into account.

Evaluating the overlap of the bony anatomy revealed that for several patients, the initial overlap did not improve after registration using a motion mask. Visual inspection of the registration results showed that small reproducibility errors in the motion mask extraction, very near to the bony anatomy, were causing local mismatches. While registration using the motion mask still gave better results than conventional registration and registration using lung masks, this issue brings forward a drawback of the registration method used in this work. The fact masks are needed for both images, in combination with a boundary matching penalty, raises the requirements for the segmentations. Registration methods relying on one segmentation, as proposed by Schmidt-Richberg *et al.*[4] and Delmon *et al.*,[16] not only require less segmentations to be performed, but are expected to be less prone to errors induced by that segmentation.

VI. CONCLUSION

We proposed a method for automatically dividing the upper thorax into similarly moving regions, capable of facilitating deformable registration of the thorax in combination with any registration method relying on a prior segmentation. Compared to using lung masks, motion masks were shown to be more suited when registering the entire thorax.

ACKNOWLEDGMENT

This work was supported by the Région Rhône-Alpes (France) via the Simed project of the ISLE research cluster. Jef Vandemeulebroucke was funded by the EC Marie Curie grant WARTHE. Jan Kybic was sponsored by the Czech Ministry of Education Project MSM6840770012.

[a]Electronic mail: jefvdmb@gmail.com.

[1]M. L. Kessler, "Image registration and data fusion in radiation therapy," Br. J. Radiol. **79**, S99–S108 (2006).
[2]D. Sarrut, "Deformable registration for image-guided radiation therapy," Z. Med. Phys. **16**, 285–297 (2006).
[3]P. Keall, "4-dimensional computed tomography imaging and treatment planning," Semin. Radiat. Oncol. **14**(1), 81–90 (2004).
[4]A. Schmidt-Richberg, J. Ehrhardt, R. Werner, and H. Handels, "Slipping objects in image registration: improved motion field estimation with direction-dependent regularization," Med. Image Comput. Comput. Assist. Interven. **12**(Pt 1), 755–762 (2009).

[5]Z. Wu, E. Rietzel, V. Boldea, D. Sarrut, and G. C. Sharp, "Evaluation of deformable registration of patient lung 4D CT with subanatomical region segmentations," Med. Phys. **35**(2), 775–781 (2008).

[6]J. W. H. Wolthaus, J. J. Sonke, M. van Herk, and E. M. F. Damen, "Reconstruction of a time-averaged midposition CT scan for radiotherapy planning of lung cancer patients using deformable registration," Med. Phys. **35**(9), 3998–4011 (2008).

[7]D. Ruan, J. A. Fessler, and S. Esedo, "Discontinuity preserving regularization for modeling sliding effects in medical image registration," in *Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference*, 2008, pp. 5304–8.

[8]D. Ruan, S. Esedoglu, and J. A. Fessler, "Discriminative sliding preserving regularization in medical image registration," in *Proceedings of the IEEE International Symposium on Biomedical Imaging*, 2009, pp. 430–3.

[9]S. Y. Chun, J. A. Fessler, and M. L. Kessler, "A simple penalty that encourages local invertibility and considers sliding effects for respiratory motion," in *Proceedings of SPIE Medical Imaging Processing*, 2009, p. 72592U.

[10]S. Kabus, T. Klinder, K. Murphy, B. van Ginneken, C. Lorenz, and J. P. W. Pluim, "Evaluation -CT lung registration," Med. Image Comput.
AQ4 Comput. Assist. Interv. 7–754 (2009).

[11]J. R. McClelland, S. Hughes, M. Modat, A. Qureshi, S. Ahmad, D. B. Landau, S. Ourselin, and D. J. Hawkes, "Inter-fraction variations in respiratory motion models," Phys. Med. Biol. **56**(1), 251–272 (2011).

[12]M. von Siebenthal, G. Székely, U. Gamper, P. Boesiger, A. Lomax, and P. Cattin, "4D MR imaging of respiratory organ motion and its variability," Phys. Med. Biol. **52**(6), 1547–1564 (2007).

[13]Y. Xie, M. Chao, and G. Xiong, "Deformable image registration of liver with consideration of lung sliding motion," Med. Phys. **38**(10), 5351 (2011).

[14]S. Flampouri, S. B. Jiang, G. C. Sharp, J. Wolfgang, A. A. Patel, and N. C. Choi, "Estimation of the delivered patient dose in lung IMRT treatment based on deformable registration of 4D-CT data and Monte Carlo simulations," Phys. Med. Biol. **51**, 2763–2779 (2006).

[15]Eike Rietzel and George T. Y. Chen, "Deformable registration of 4D computed tomography data," Med. Phys. **33**(11), 4423–4430 (2006).

[16]V. Delmon, S. Rit, R. Pinho, D. Sarrut, and V. Delmon, "Direction dependent b-splines decomposition for the registration of sliding objects," in *Proceedings of the Fourth International Workshop on Pulmonary Image Analysis*, Toronto, Canada, 2011, pp. 45–55.

[17]R. Werner, J. Ehrhardt, R. Schmidt, and H. Handels, "Patient-specific finite element modeling of respiratory lung motion using 4D CT image data," Med. Phys. **36**(5), 1500–1511 (2009).

[18]M. Schuenke, E. Schulte, and U. Schumacher, *Neck and Intern gans*
AQ5 *(THIEME Atlas of Anatomy)* (Thieme Medical Publishers Inc., 0).

[19]H. Gray. *Anatomy of the Human Body*, 20th ed. (Lea & Febiger, 918).

[20]S. Osher and J. Sethian, "Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations," J. Computat. Phys. **79**, 12–49 (1988).

[21]P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," IEEE Trans. Pattern Anal. Mach. Intell. **12**(7), 629–639 (1990).

[22]E. M. van Rikxoort, B. deHoop, M. A. Viergever, M. Prokop, and B. van Ginneken, "Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection," Med. Phys. **36**(7), 2934–2947 (2009).

[23]K. Mori, J. Hasegawa, J. Toriwaki, H. Anno, and K. Katada, "Recognition of bronchus in three-dimensional X-ray CT images with application to virtualized bronchoscopy system," Proc. Int. Conf. Pattern Recognit. **3**, 528–532 (1996).

[24]N. Otsu, "A threshold selection method from gray-level histograms," IEEE Trans. Syst. Man Cybern. **9**(1), 62–66 (1979).

[25]S. J. Osher and R. P. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces* (Springer, New York, 2002).

[26]V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," Int. J. Comput. Vis. **22**, 61–79 (1997).

[27]R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: a level set approach," IEEE Trans. Pattern Anal. Mach. Intell. **17**(2), 158–175 (1995).

[28]http://www.DIR-LAB.com.

[29]L. Ibanez, W. Schr , L. Ng, and J. Cates, *The ITK Software Guide*, 2nd ed. (Kitware, Inc., 05). http://www.itk.org/ItkSoftwareGuide.pdf.

[30]D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast MR images," IEEE Trans. Med. Imaging **18**(8), 712–721 (1999).

[31]Byrd, Lu, and Nocedal, "A limited memory algorithm for bound-constrained optimization," SIAM J. Sci. Comput. **16**(5), 1190–1208 (1995). AQ6

[32]K. Murphy, B. van Ginneken, S. Klein, M. Staring, B. deHoop, M. Viergever, and J. Pluim, "Semi-automatic construction of reference standards for evaluation of image registration," Med. Image Anal. **15**, 71–84 (2011).

[33]J. Vandemeulebroucke, S. Rit, J. Kybic, P. Clarysse, and D. Sarrut, "Spatio-temporal motion estimation for respiratory-correlated imaging of the lungs," Med. Phys. **38**(1), 166–178 (2011).

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

[35]R. Castillo, E. Castillo, R. Guerra, V. E. Johnson, T. McPhail, A. K. Garg, and T. Guerrero, "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets," Phys. Med. Biol. **54**(7), 1849–1870 (2009).

[36]L. R. Dice, "Measures of the amount of ecologic association between species," Ecology **26**(3), 297–302 (1945).

[37]J. Vandemeulebroucke, S. Rit, J. Schaerer, and D. Sarrut, "Deformable image registration with automated motion-mask extraction," in *Medical Image Analysis for the Clinic: A Grand Challenge 2010 of the MICCAI Conference*, September 2010, AQ7

[38]J. Staal, B. van Ginneken, and M ergever, "Automatic rib segmentation and labeling in computed tomography scans using a general framework for detection, recognition and segmentation of objects in volumetric data," Med. Image Anal. **11**(1), 35–46 (2007).

[39]K. Ding, Y. Yin, K. Cao, G. E. Christensen, C.-L. Lin, E. A. Hoffman, and J. M. Reinhardt, "Evaluation of lobar biomechanics during respiration using image registration," Med. Image Comput. Comput. Assist. Interv. **12**(Pt 1), 739–746 (2009).