Automatic motion mask extraction for deformable registration of the lungs

Jef Vandemeulebroucke, Olivier BernardJan Kybic, Patrick Clarysse, David Sarrut

Léon Bérard Cancer Center, University of Lyon in France
Creatis Laboratory, University of Lyon in France
Center for Machine Perception, Czech Technical University in Prague, Czech Republic

Abstract

Deformable registration algorithms generally rely on the assumption that the sought transformation is spatially smooth. This can contradict the physiology of the motion, which is the case for the sliding motion of the lung during respiration. We propose an automatic method based for dividing the thorax into regions with homologous breathing motion. We use the level set framework to track the evolution of an interface, constrained by previously extracted anatomical features and regularized by strong geometric priors. We show that the obtained motion masks can facilitate deformable registration. Inner and outer thoracic structures are registered separately and compared to the result when considering the entire thorax simultaneously. For all patients, the mean registration error was reduced, and the improvement was statistically significant for five out of six patients. By preserving the sliding motion, the complexity of the spatial transformation can be reduced considerably while maintaining matching accuracy.

Keywords

Deformable registration, breathing motion, sliding motion, CT imaging, lung cancer

Introduction

There is an increasing demand for accurate deformable registration in radiation therapy. Registration is ill-posed making a direct approach impossible. Explicit parametric restrictions with respect to the sought spatial transformation and suitable regularization of the objective function should encode the physical understanding of the desired deformation properties and drive the optimization to solutions with such characteristics. In particular, the assumption of spatial smoothness of the transformation is widely utilized to estimate motion induced deformations. While required for solving the inverse problem, these mechanisms can contradict the physiology of the organ motion. A example is the sliding motion of the lung with respect to the chest wall during breathing. Homogeneous smoothing of the transformation will in this case blur the estimated transformation across the sliding interface, resulting in locally reduced registration accuracy [1].

The issue of sliding motion in deformable image registration has been addressed in a number of ways, including: anisotropic filtering, shear preserving regularization and finite element modeling. Rietzel *et al.* [2] described an approach in which the thorax is manually segmented, into moving (lungs, mediastinum and abdomen) and less-moving (the remainder) subregions, and each region is registered separately. Wu *et al.* [1] extended the method, introducing a boundary matching criterion that reduces gaps in the deformation field between the separately registered subregions.

We propose a practical method for automatically dividing the thorax into regions with homologous, respiratory-induced motion. Our main objective is to obtain an accurate interface where strong sliding motion occurs, and facilitate subsequent deformable registration. We verify the suitability of the obtained masks by applying them to deformable registration of the lungs. Registration using the motion mask is compared to conventional registration of the end-exhalation and end-inhalation frames of 4D CT images of six lung cancer patients, by looking at the registration error of a well-distributed set of 100 landmarks per image pair.

Material and methods

Taking into account anatomical considerations, the following applies to the sought motion mask. To preserve sliding, an interface should separate the lung from the chest wall. At the medial lung interface, as there is a continuous and smooth transition of motion, the mediastinum can be considered as part of the same motion region as the lung.

The method is based on the level set framework [3]. We exploit the intrinsic interface smoothing which allows to include geometric priors to guide the front propagation in regions with lack of information. The evolution of the interface is fully described by a partial

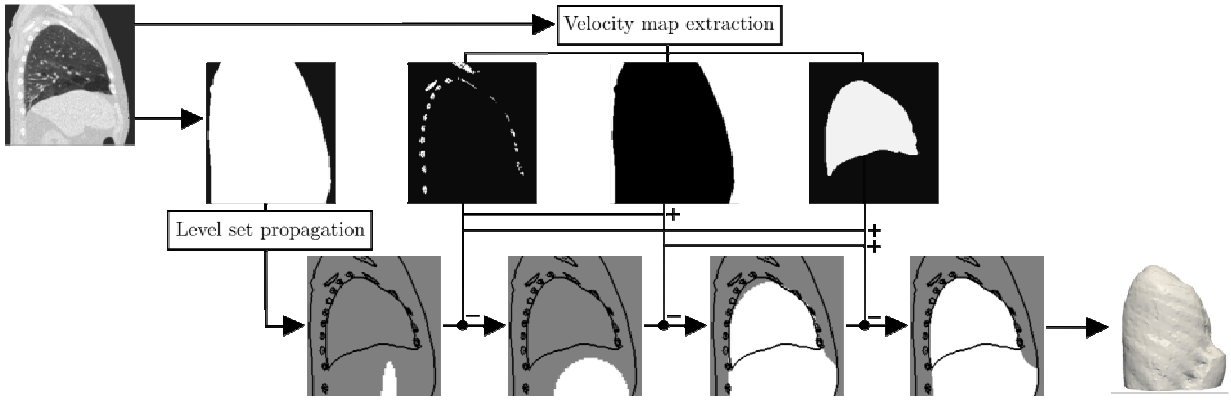

Figure 1: Overview of the method. Velocity maps are computed by reducing the initial CT to binary feature images. The obtained maps are combined (+) and used to constrain (-) the evolving interface in the levelset propagation steps.

differential equation in which two terms appear. The first corresponds to a propagation force, in our case based on a *velocity map* representing information on the boundaries of the object to segment. Depending on the sign of the propagation force, the contour will tend to expand or contract. The second term corresponds to a local interface smoothing force. We can thus divide the method for obtaining the motion mask into two parts (figure 1). First, velocity maps are computed by reducing the original CT image to binary *label* images containing only the relevant features. Next, the velocity maps are combined to guide the evolving interface in consecutive level set propagation steps.

Velocity Map Extraction

The velocity maps are obtained from the input image by extracting features through consecutive thresholding, region growing and mathematical morphology. The retained features are the outer patient body contour, the bony anatomy and the lungs (figure 1, top row).

To extract the outer patient contour, the image is binarized by thresholding at -300 HU, and the regions with lower intensity values are labelled using 3D connected component labelling, using a 26-voxel connectivity. The air surrounding the patient is the principal component and is removed. The remaining labels are binarized and connected component labelling will now yield the patient body as a principal label.

Next, we wish to extract the bony anatomy. Edge preserving smoothing is performed using anisotropic diffusion. The result is binarized with an lower threshold of 100 HU, and principal component corresponds to the connected bony structure (i.e. column, vertebrae, ribs and sternum).

For the lung segmentation method, we refer to the conventional segmentation method described by Rikxoort *et al.* [4], from which the adopted procedure was largely inspired.

Level Set Propagation

Level sets [3] were initially introduced to model the front propagation of an interface. Afterwards, they were applied to medical image segmentation to automatically detect the boundaries of structures of interest. 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 binary velocity maps, defining two types of regions. One where the interface evolves with isotropic speed, and one where the level set is confined to its current state. For each different level set propagation step (figure 1, bottom row), a stopping criteria directly linked to anatomical structures is defined.

The algorithm is initialized with the signed distance function of a ellipsoide, centered with respect to the patient. The goal of this step is to include the abdominal region and reach the anterior patient-to-air interface. We apply a positive propagation force to ensure an expanding interface and guide the evolution with a unit velocity, except at the bony anatomy where velocity is set to zero. The evolution is stopped when a detection point placed 10 mm anterior to the most inferior patient-to-air interface and centered in front of the initial ellipsoid is reached.

Next, we wish to cover the entire thoracic cavity including the lungs and mediastinum. The previous result is propagated further but the underlying velocity map is altered so that in addition to the bony anatomy, velocity is zero outside the patient body. 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 inside the thoracic cavity while the coverage of the extracted lungs is monitored every 50 iterations. When the contour covers at least 95% of the lungs, the algorithm is terminated and the resulting mask is padded to include the full lung region.

Finally, the previous solution is refined to obtain a smooth contour that conforms to the outer shape of the

lungs, but includes the mediastinum. Only the curvature term is retained during this step. Note that in practice this will lead to contraction as curvature is integrated along contour length. The velocity map employed is a unit field everywhere besides outside the body, at the bony anatomy and in the lungs. The algorithm is run for 500 iterations, which was empirically found to be sufficient to smoothen the mask. (figure1, fourth image, bottom row).

Deformable Registration

The suitability of the obtained masks is verified by applying them to deformable registration of the lungs. The results are compared to those obtained using the same registration parameters but without making use of the motion mask. We use B-spline free-form deformation algorithm. A multi-resolution approach with three levels is applied for the images as well as the B-spline control point grid, the final level having a 2 mm and 32 mm spacing, respectively. Similarity is measured through the sum of squared differences and optimized by L-BFGS-B algorithm starting from an initial zero deformation vector field.
As in [1], the intensities in the image pair outside the considered region in the corresponding mask are set to a unique value, here -1200 HU. Similarity is calculated on a region slightly larger than the reference mask in order to include a thin border of voxels of about 10 mm with modified intensity (figure 2). The contribution of this subregion with modified intensity penalizes a potential mismatch between the region borders in reference and target image, but does not constrain the sliding motion. Note however, that masks on both reference and target image are now required. As a consequence their construction should be consistent with respect to anatomical structures.

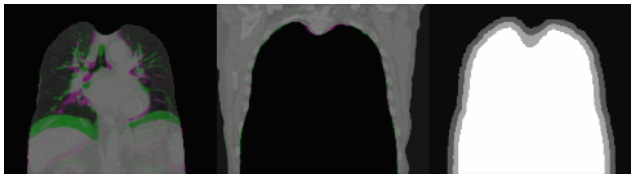

Figure 2: The modified images used for registration. Inhale and exhale are shown in overlay. On the right, the region considered for registration. For the inner (outer) thoracic structures, everything but the black (white) region is considered. The actual mask runs between the gray strips.

The registration performance for the lungs is evaluated by assessing the matching accuracy of a 100 landmarks per image pair. The landmarks were identified using a a semi-automatic software tool [5]. For each end-exhalation frame, a well-distributed set of 100 landmarks in the lungs was detected fully automatically. Using a custom-designed interface, trained observers identified the corresponding anatomical locations in the end-inhalation scans. The interobserver variability over all 600 points, was 0.5 mm (0.9 mm).
Registration was performed between the exhale and inhale frames of 4D CT images of the thorax for six lung cancer patients, acquired as part of the treatment planning protocol on a Brilliance Big Bore 16-slice CT scanner (Philips Medical Systems, Cleveland, OH). Retrospective respiratory-correlated reconstruction was made possible by recording of a respiratory trace using the Pneumo Chest bellows (Lafayette Instrument, Lafayette, IN).

Results and discussion

The calculation of the velocity maps required less then a minute per image. The level set steps processing time, even though different for all patients, remained under 5 minutes (on a single 2.6 GHz CPU).
Table 1 contains the quantitative evaluation of the registration results. For all patients, the use of the motion mask improved the registration accuracy in terms of mean target registration error. For all patients except for Patient 4, the improvement was statistically significant (p-value< 0.05 for a paired t-test).

(mm)	Before	No Mask	Mask
Patient 1	9.4 ± 7.4	2.4 ± 2.2	1.8 ± 1.8
Patient 2	7.3 ± 4.9	2.8 ± 3.9	2.6 ± 4.4
Patient 3	7.1 ± 5.1	1.8 ± 1.8	1.6 ± 1.4
Patient 4	6.7 ± 3.7	1.6 ± 1.5	1.5 ± 1.2
Patient 5	14.0 ± 7.2	2.8 ± 2.9	1.8 ± 1.3
Patient 6	6.8 ± 3.2	2.1 ± 1.5	1.9 ± 1.2

Table 1: Landmark distance before registration, after conventional registration and after registration using the mask. Given is the mean distance and standard deviation.

The mean landmark distance after registration when using the mask ranged from 1.5 to 2.6 mm. This is in good correspondence with previous work addressing sliding motion. In [1], [2] mean registration errors of 2.1 to 3.5 mm and 1.6 to 3.0 mm, respectively, are reported for the lungs when using a B-spline based registration.
Both used manual segmentations similar to the proposed motion mask.
An important parameter for the used B-spline free-form deformation transformation is the spacing of the control point grid. As a finer grid is employed, the deformation space becomes larger and representation of discontinuities such as sliding improves. However, the complexity of the optimization increases rapidly, along with the computation time. In addition, allowing more degrees of freedom increases sensitivity to noise and artifacts since the parametrization of the 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 figure 3, the mean target registration errors obtained

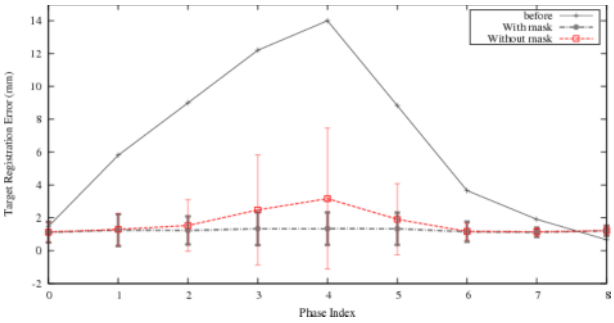

Figure 4: The landmark distance in the consecutive frames of a 4DCT image. Shown are the mean distance before registration, after conventional demons registration and after demons registration using the motion mask.

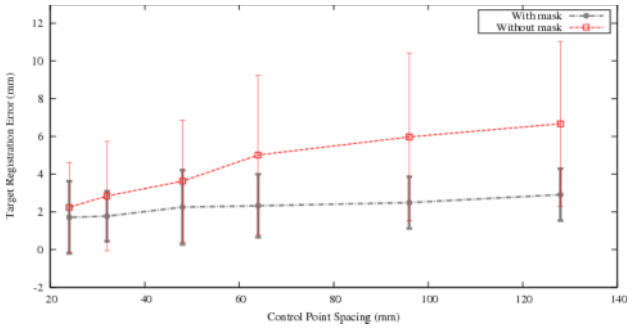

Figure 3: The mean target registration error (1 SD) obtained with and without motion mask as a function of the control point spacing, for Patient 5.

with and without motion mask are shown in function of the control point spacing for patient 5, characterized by large motion amplitudes. We can note that the result obtained with mask using a control point spacing of 128 mm, (2.9 ± 1.4 mm) , is comparable to that obtained without mask using a control point spacing of 32 mm (2.8 ± 2.9 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. 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.

The use of the motion mask is not limited to the presented B-spline registration method. Figure 5 shows the results of the demons algorithms, with and without masking the input images using the motion mask. Shown is the error when registering the exhale phase to the remaining phases of the full 4D CT image. For each registration the performance is assessed by looking at 100 landmarks. The original motion of the landmarks is depicted in the same figure. The graph shows that both methods perform identical for breathing phases near the reference. Near end-inhalation however, the registration error of the conventional approach increases. Registration performed using the motion mask manages to maintain its accuracy throughout the cycle.

Particular attention was paid to making the automatic motion mask extraction robust and reliable, in order to limit the required user interaction in a clinical setting. Even though the described procedure produced good results for all image pairs tested, due to the large anatomical and pathological variability, it might still fail for some patients. In particular, erroneous or incomplete extraction of the lungs and ribs could result in unsatisfactory results. Often, limited manual intervention will suffice to adapt the procedure to the situation. Alternatively, more elaborate (but costly) approaches designed to deal with the pathological lung [5], or specifically devised to label the complete ribcage might reduce the user interaction even further.

Conclusion

We proposed a practical method for automatically dividing the upper thorax into homologously moving regions and facilitate subsequent deformable registration. The suitability of the extracted motion masks was validated by applying them to deformable registration of inhale and exhale frames of 4DCT images. Inner and outer thoracic structures were registered separately using a boundary matching criterion and compared to conventional registration considering the whole thorax simultaneously. For all six patients, the mean target registration error showed improvement with respect to conventional registration. For five patients the improvement was found to be statistically significant. In the particular case of a B-spline based free form deformation, the use of the motion mask can considerably reduce the complexity of the spatial transformation, while maintaining registration accuracy.

References

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

[2] Rietzel, E. & Chen, G. T. Y. (2006), Deformable registration of 4D computed tomography data. *Med Phys* **33** (11), 4423–4430.

[3] Osher, S. & Sethian, J. (1988), Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations., *Journal of Computational Physics* **79** , 12–49.

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

[5] Murphy, K., van Ginneken, B., Pluim, J. P. W., Klein, S. & Staring, M. (2008), Semi-automatic reference standard construction for quantitative evaluation of lung CT registration. *Med Image Comput Comput Assist Interv*, **11**, 1006–1013.