Statistical Motion Mask and Sliding Registration

Björn Eiben[1], Elena H. Tran[1], Martin J. Menten[2], Uwe Oelfke[2],
David J. Hawkes[1], and Jamie R. McClelland[1]

[1] Centre for Medical Image Computing, University College London, UK,
`bjoern.eiben.10@ucl.ac.uk`
[2] Joint Department of Physics, The Institute of Cancer Research and The Royal
Marsden NHS Foundation Trust, London, UK

Abstract. Accurate registration of images depicting respiratory motion, e.g. 4DCT or 4DMR, can be challenging due to sliding motion that occurs between the chest wall and organs within the pleural sac (lungs, mediastinum, liver). In this paper we propose a methodology that (1) segments one of the images to be registered (the source or floating/moving image) into two distinct regions by fitting a statistical motion mask, and (2) registers the image with a modified B-spline registration algorithm that can account for sliding motion between the regions. This registration requires the segmentation of the regions in the source image domain as a signed distance map. Two underlying transformations allow the regions to deform independently, while a constraint term based on the transformed distance maps penalises gaps and overlaps between the regions. Although implemented in a B-spline algorithm, the required modifications are not specific to the transformation type and thus can be applied to parametric and non-parametric frameworks alike. The registration accuracy is evaluated using the landmark registration error on the basis of the publicly available DIR-Lab dataset. The overall average landmark error after registration is 1.21mm and the average gap and overlap volumes are 26.4cm^3 and 34.5cm^3 respectively. The fitted statistical motion masks are compared to previously proposed motion masks and the corresponding mean Dice coefficient is 0.96.

Keywords: sliding motion, B-Spline registration, statistical shape model, motion mask

1 Introduction

Registration of images which contain anatomical regions that slide along each other is an ongoing research topic. The major challenge is that registration is an ill-posed problem and thus requires some regularisation which usually constrains the transformation to be smooth. This smoothness assumption however is not true across a sliding interface where discontinuities in the transformation are present. A prominent example where sliding occurs is respiratory motion of the lung along the chest wall, facilitated by the pleural sac that encloses not only the lungs and heart, but also liver and further lower abdominal organs. As a result, registration of images of the thorax depicting respiratory motion is challenging.

Several registration methods for handling sliding motion have been proposed in the literature [1, 5, 7, 10, 13, 14]. Most of these methods require a segmentation of the sliding regions either in the target image or in both target and source image. But for applications such as contour propagation in radiotherapy, it is preferable to segment the source image where the contours have been defined. Furthermore, for combined motion modelling and motion compensated image reconstruction from partial image data [8] it is essential to segment the source image as the target 'images' may be partial image data (e.g. individual slices, projections) which can be very challenging or impossible to accurately segment. Some methods do not require a prior segmentation in either image, but use a regularisation term that permits sliding motion [13]. While such methods are appealing, there is a possibility they may not correctly represent the sliding motion in areas of homogeneous intensities, for instance where the liver meets the chest wall. Furthermore, they do not model a true discontinuous motion, but approximate the sliding as a shear motion. However, it is acknowledged that while the discontinuous motion is more realistic, in practice this may not be important. Therefore, for some applications such as motion modelling from partial image data or radiotherapy dose accumulation, it may be desirable to explicitly specify where the sliding motion should occur, especially if parts of the sliding interface are not easily identifiable in the images. To the best of our knowledge, none of the published methods that require a segmentation allows the sliding regions to be specified only in the source image domain.

Identification of the sliding regions in CT images was proposed by Vandemeulebroucke et al. [12]. Their motion mask includes the lungs and other inferior organs that slide together and is thus anatomically more plausible than just a lung segmentation. However, this motion mask generation is designed specifically for CT images as it utilises features that can be relatively easily segmented from such images, namely lungs, thorax, and bones. Moreover, application of our implementation of this method proved to be time consuming because, (1) level-set evolutions are computationally expensive and (2) we did not find a single set of parameters that worked across all patients. Some manual corrections of the bony structures were also required where pacemakers or high intensity abdominal regions were present and morphologically connected to the bones.

The purpose of this paper is therefore twofold. In section 2.1 we propose a motion mask generation on the basis of a statistical shape model [4] being fit to the source image. A resulting signed distance map where the zero-crossing identifies the interface of sliding image regions is then utilised by our sliding B-spline registration. The sliding framework is presented in section 2.2. The results are quantitatively evaluated on the publicly available DIR-Lab dataset [2, 3].

2 Material and methods

2.1 Statistical motion masks

The motion mask generation by Vandemeulebroucke et al. [12] uses a sequence of level-set evolutions that are controlled by three CT-based feature images,
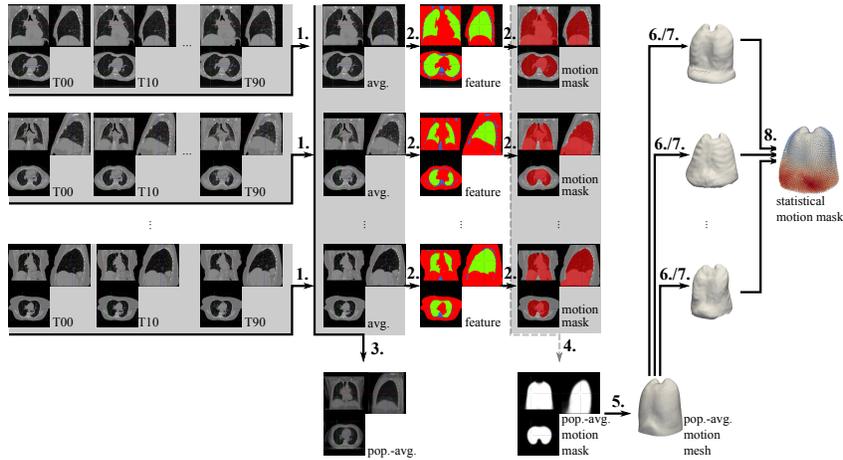
Fig. 1. Flow chart for the generation of the statistical motion mask. (1.) Average 4DCT images (avg.) are generated by intra-patient group-wise registration of the different phases (T00-T90). (2.) From these average images features are extracted and motion masks are calculated for each subject. (3.) Inter-subject group-wise registration of average 4DCT images and (4.) application of the transformations to the mask to create a population average motion mask. (5.) A surface mesh is generated from the average motion mask image. (6./7.) The surface mesh is transformed back to the individual images. (8.) The deformations are captured in a statistical motion mask.

namely lungs, thorax, and bones. This method produces suitable motion masks from CT images, but cannot be used with MR images (which we intend to apply our method to in the future), and can be time consuming. Therefore, we build a statistical shape model of the motion masks from a large number of 4DCT images, that captures the inter-patient variation in the motion masks. This statistical shape model can then be used to segment a new image more rapidly than the original method, and can potentially be used on MR images. This shape model is fitted to directional intensity gradients which are present in both CT and MR images, such as the low-to-high intensity contrast between lung and surrounding tissue or the high-to-low contrast between thorax and air.

The generation of the statistical motion mask comprises of the following steps and is summarised in figure 1:

1. Calculate intra-subject average 4DCT image by group-wise registration,
2. Generate motion mask based on average 4DCT images according to [12],
3. Inter-patient group-wise registration of average 4DCT images,
4. Use registration results to form the population-average motion mask image,
5. Convert population-average motion mask image into mesh representation,
6. Use registration results to transform motion mask mesh to each subject
7. Refine mesh to better match individual motion mask images [6],
8. Statistical motion mask calculation [4].

A prerequisite for building a statistical shape model is the point-to-point correspondence between individual motion mask meshes (7.) [4]. This is why one population-average motion mask mesh is generated (5.) and then transformed back to each subject (6./7.), rather than simply generating the meshes from each individual motion mask image. As the transformed mesh does not perfectly match the individual motion mask images due to registration errors, the mesh is refined by a surface alignment algorithm [6] where the transformed average mesh is iteratively deformed towards a target surface, in this case the motion masks calculated in (2.). The final statistical shape model is formed by applying PCA on the refined mesh coordinates, and comprises the average shape and eigenvectors/principal components that define the modes of variation of the shape.

Once the statistical motion mask exists, an appearance representation in the form of intensity gradients is calculated along the mesh normals determining for each node if positive, negative or mixed intensity gradients dominate. Such gradients are measured in the direction of the outward facing mesh normals on the average 4DCT images. Only those nodes that have a clear positive or negative intensity gradient across all subjects will be used during mask fitting.

To fit the statistical motion mask to an image, i.e. finding the weights for each mode of variation and a global rigid transformation, we follow the iterative multi-scale process by Cootes et al. [4], but make the shape model lock onto either positive or negative image gradients as identified above.

Due to the shape of the motion mask, the positioning and extent in the inferior part of the mask is not well constrained. To improve the SI-scaling of the mask, we also include the lung-diaphragm boundary as two additional surfaces to the statistical motion mask. These surfaces also lock on the directional image gradients in the diaphragm region. This improves the scaling of the fitted mask in the SI-direction. For our application we also want the mask to extend below the image boundaries. Hence during the mask fitting the most inferior nodes of the statistical model are displaced inferiorly until they are below the image boundary (c.f. figure 4(b)).

2.2 Sliding registration framework

The proposed sliding framework uses two separate transformations for the two regions involved, region A and region B. This facilitates independent motion and thus the desired deformation discontinuity at the regions' interface. The corresponding transformations are denoted as $\mathbf{T}_A(\mathbf{x})$ and $\mathbf{T}_B(\mathbf{x})$. The method also requires a signed distance map $D(\mathbf{x})$ which defines the two regions, and can be pre-calculated based on the fitted statistical motion mask. Let region A be identified by negative values $D < 0$ and region B by positive values $D \geq 0$.

In order to calculate the deformed image it is necessary to determine which of the two transformations to use for each voxel in the target image. This is done by deforming the distance map with both transformations, $D_A(\mathbf{x}) = D(\mathbf{T}_A(\mathbf{x}))$ and $D_B(\mathbf{x}) = D(\mathbf{T}_B(\mathbf{x}))$, and summing the resulting deformed distance maps,
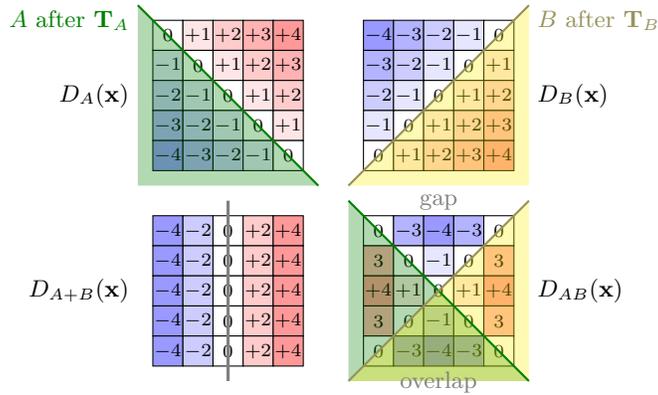
Fig. 2. Schematic of transformed distance maps D_A and D_B, their sum D_{A+B} and product D_{AB} as they are used in the sliding registration approach. The green overlay depicts region A after transformation with $\mathbf{T}_A$ and the yellow overlay shows region B after $\mathbf{T}_B$. The grey line (bottom left) indicates the boundary where the transformation switches from $\mathbf{T}_A$ to $\mathbf{T}_B$ and negative values of D_{AB} indicate gaps and overlaps.

$D_{A+B} = D_A + D_B$. Then

$$\mathbf{T}(\mathbf{x}) = \begin{cases} \mathbf{T}_A(\mathbf{x}) & \text{for } D_{A+B}(\mathbf{x}) < 0 \\ \mathbf{T}_B(\mathbf{x}) & \text{otherwise} \end{cases}. \tag{1}$$

The independent motion of the two regions can result in gaps between or overlap of the two regions. The above approach for determining which transformation should be used at each voxel means that the transition between using $\mathbf{T}_A$ and $\mathbf{T}_B$ will occur in the middle of any gaps and overlaps (see figure 2). However, even with this approach the gaps and overlaps will still lead to a violation of a one-to-one mapping between the images. Therefore, a gap-overlap constraint (GOCT) term is introduced, which is based on the product of the deformed distance maps $D_{AB} = D_A D_B$, and penalises any gaps and overlaps that occur (see figure 2). This term is non-zero only at positions where gaps and overlaps occur.

$$\mathcal{C}_{\text{GOCT}}(\mathbf{x}) = \begin{cases} -D_{AB}(\mathbf{x}) & \text{for } D_{AB}(\mathbf{x}) < 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

The total value of the constraint term is then calculated as the sum of $\mathcal{C}_{\text{GOCT}}$ over all voxels normalised by the number of voxels.

The transformations are both optimised simultaneously by calculating the gradient of the cost function (i.e. sum of similarity metric(s), $\mathcal{C}_{\text{GOCT}}$, and other constraint terms) with respect to each of the individual transformations. When calculating the gradient of the similarity metric(s) with respect to $\mathbf{T}_A$, only the voxels that are transformed by $\mathbf{T}_A$ (i.e. $D_{A+B} < 0$) contribute towards the

gradient calculation. And likewise, only voxels transformed by $\mathbf{T}_B$ contribute towards the gradient with respect to $\mathbf{T}_B$. The gradient of $\mathcal{C}_{\mathrm{GOCT}}$ with respect to $\mathbf{T}_A$ is given by

$$\frac{\partial \mathcal{C}_{\mathrm{GOCT}}(\mathbf{x})}{\partial \mathbf{T}_A(\mathbf{x})} = \begin{cases} -D_B(\mathbf{x})\frac{\partial D_A(\mathbf{x})}{\partial T_A(\mathbf{x})} & \text{for } D_A(\mathbf{x})D_B(\mathbf{x}) < 0 \\ \mathbf{0} & \text{otherwise} \end{cases}, \qquad (3)$$

where $\partial D_A(\mathbf{x})/\partial T_A(\mathbf{x})$ is the spatial gradient of the distance map warped by the deformation field for region A. The gradients of any other constraint terms are calculated for each transformation separately, in exactly the same way as they would be for a standard (non-sliding) registration.

The sliding framework was used to modify NiftyReg, an open source B-spline registration software [9].

2.3   Image and landmark data

To build the statistical motion masks, 4DCT datasets from 32 subjects were used, each comprising of 10 respiratory phases. The first ten image sets were taken from the publicly available *DIR-Lab* dataset [2, 3]. The remaining datasets were acquired as part of standard clinical practice from patients with either early-stage (14 cases) or locally-advanced (8 cases) non-small cell lung cancer.

The DIR-Lab dataset also contains 300 manually selected corresponding landmarks in the end-exhale and end-inhale phases. These were used to quantify the registration accuracy in terms of a landmark registration error (LME). Note that most of the LME values given on the DIR-Lab website have been calculated using the 'snap-to-voxel' approach [3], so for comparability reasons we also follow the same approach here.

3   Results

3.1   Statistical motion mask fitting accuracy

The fitting accuracy of the statistical motion mask with respect to the number of modes included in the model was evaluated against the level-set based motion mask in terms of the Dice coefficient and mean contour distance by using a leave-one-out evaluation strategy. To evaluate the dependency of the fitting accuracy on the number of modes of variation (i.e. number of eigenvectors of the statistical shape model), the fitting was performed in a multi-scale fashion with three resolution levels corresponding to intensity sampling distances along the mesh normal of 4mm, 2mm, and 1mm and 200 iterations per level. The results are shown as box-plots in figure 3. The highest mean Dice coefficient of 0.96 ± 0.02 was achieved with 20 modes, whereas the lowest mean contour distance of 3.67 ± 2.00mm was achieved with 25 modes.

Furthermore the variation covered by the shape model as a function of the number of modes is given in figure 4(a). For a coverage of 95% of the variation
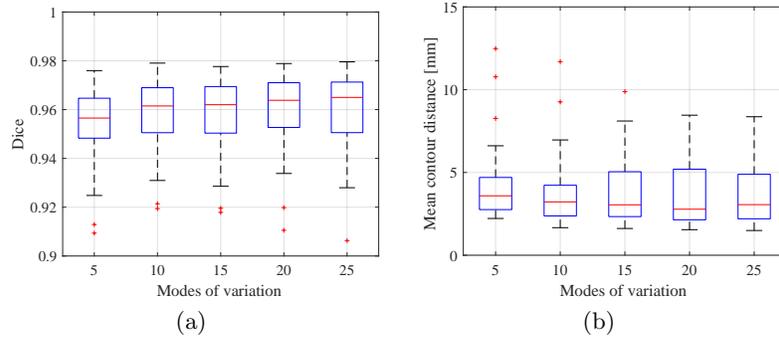
Fig. 3. Motion mask fitting accuracy in terms of the Dice coefficient (a) and mean contour distance (b). Red line represents the median value, the box extends from the first and the third quartile, and the whiskers extend the box by 1.5 times the interquartile range.
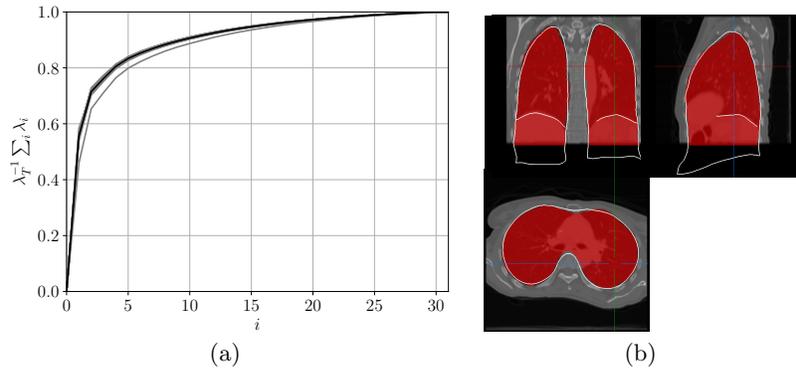
Fig. 4. (a) Cumulative, relative variation described by statistical shape model as a function of the number of modes of variation, i, used in the 32 leave-one-out models. For each model a grey curve is plotted and the average over all models is shown in black. (b) Level-set based motion mask (red) and statistical motion mask (white line) fitted to case 1 of the DIR-Lab dataset.

(a) non-sliding B-spline registration (b) sliding B-spline registration
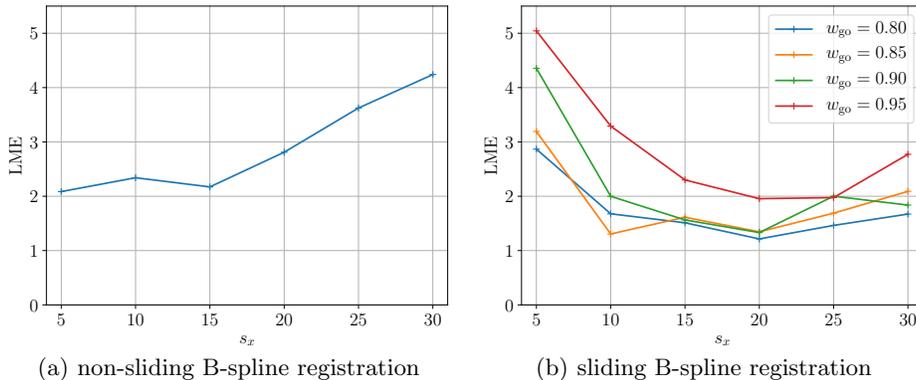
Fig. 5. Mean landmark errors (LME) over all 10 DIR-lab datasets and all landmarks after registration with respect to the control point spacing s_x. For the sliding registration results, (b), errors are given for several gap-overlap-constraint weights, w_{go}. All values are given in mm.

observed in the dataset 16 modes are required, and for 98% coverage 23 modes. After considering the Dice coefficient results, the mean contour distance results, and the percentage of variation covered, it was decided to use 23 modes for all the following experiments. An example result for this fitting configuration for the first case of the DIR-Lab dataset is shown in figure 4(b) along with the level-set based result.

3.2 Registration accuracy

The registration accuracy was quantified using the DIR-Lab datasets as follows: The source image was selected to be the end-exhale image and the statistical motion mask (built leaving out the current dataset) was fitted in the same fashion as described in section 3.1 using 23 modes of variation. A signed distance map was calculated from that motion mask and used as input into the sliding registration. In order to keep the number of parameters for the registration to a minimum, no other constraint terms were used. Different B-spline control point spacings ($s_x = [5, 10, \ldots, 30]$mm) and $\mathcal{C}_{\mathrm{GOCT}}$ weights ($w_{\mathrm{go}} = [0.8, 0.85, 0.9, 0.95]$) were investigated. For comparison, standard, i.e. non-sliding, B-spline registrations with the same control-point spacings and no additional constraint terms were performed and evaluated. For all registrations Local Normalised Cross Correlation (LNCC) was used as the similarity measure.

The landmark errors after alignment with the standard B-spline registration are shown in figure 5(a). Here the lowest average LME is 2.09mm and results from a control-pont spacing of $s_x = 5$mm. The LME increases for coarser control-point grids.

The results for the sliding registration are shown in figure 5(b). The lowest average LME of 1.21mm is achieved for a gap/overlap constraint weight $w_{\mathrm{go}} =$
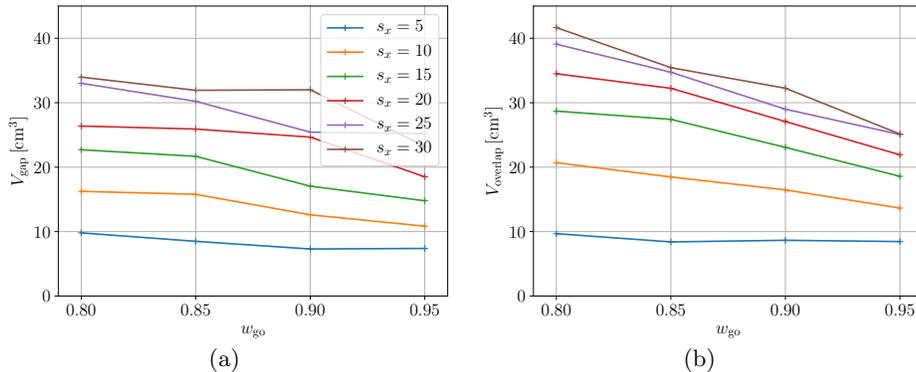
Fig. 6. Gap (a) and overlap (b) volumes in cubic centimetres with respect to the corresponding constraint weight w_{go} for several control point spacings s_x.

Table 1. Gap (V_{gap}) and overlap volumes (V_{overlap}) of segmentation-based methods as reported in the literature [1, 5, 7, 14]. All values are given in cm^3.

	Wu 2008,	Delmon 2013,	Berendsen 2014,	Hua 2017,	our method
V_{gap}	88.2	80.0	76.5	94.0	26.4
V_{overlap}	52.5	55.1	37.4	18.8	34.5
$V_{\mathrm{gap}} \cup V_{\mathrm{overlap}}$	130.7	135.1	113.9	112.8	60.9

0.8 and a control-point spacing of $s_x = 20$mm. However, results with this spacing and $w_{\mathrm{go}} = [0.8, 0.85, 0.9]$ are very close together and below 1.34mm. Compared with other methods, our resulting average landmark registration error is lower than those reported for instance by Papiez et al. [10] (1.95mm), Wu et al. [14] (1.47mm), Delmon et al. [5] (1.66mm), and Berendsen et al. [1] (1.36mm); but slightly larger the one reported by Hua et al. [7] (1.17mm) or Vishnevskiy et al. [13] (0.95mm). A comprehensive list of landmark errors achieved by various algorithms is published on the DIR-Lab website [11].

Figure 6 shows the average gap and overlap volumes given in cubic centimetres as a function of the gap/overlap constraint weight for all tested control-point spacings. Gap and overlap volumes decrease towards higher constraint weights and smaller control point spacings. For the configuration that produces the lowest LME, the average gap and overlap volumes are 26.4cm^3 and 34.5cm^3 respectively. For comparison, gap, overlap and their combined volumes as reported by [1, 5, 7, 14] for segmentation-based methods on the same DIR-Lab dataset are shown in table 1.

4 Discussion and conclusion

In this paper we present a method to register images which contain regions that slide along each other, and demonstrating this on 4DCTs of the thorax. To achieve this source images are segmented into sliding regions by fitting a statistical motion mask which encompasses the lungs, heart and further inferior organs. This segmentation is used as an additional input for a modified B-spline registration. The modification implements a sliding framework that allows for using different transformation types and/or other constraint terms for each region. Defining the sliding regions in the source image domain can be beneficial for applications such as contour propagation for radiotherapy, and is required for motion modelling, and motion compensated image reconstruction from partial image data [8].

The statistical motion mask fitting is an extension of the work of Vandemeulebroucke et al. [12] who calculate a motion mask on the basis of image features that are extracted from CT images and multiple level-set evolution steps. This procedure is computationally expensive and is only applicable to CT images. The statistical motion mask can be fitted relatively fast and the method is designed to be applicable to MR as well as CT images. However, a quantitative evaluation on MR images will be subject of future work.

The sliding registration requires the segmentation of the two regions in the form of a signed distance map that can be pre-calculated. Each region is associated with a separate transformation which allows the regions to move independently from each other. A new constraint term has been used to help prevent gaps and overlaps occurring between the two regions, and hence maintain a one-to-one mapping between the images.

The proposed registration method achieved an average registration accuracy of 1.21mm. These errors are based on manually identified correspondences and thus are subject to inter- and intra-observer errors which range between 0.70mm and 1.13mm [2]. This poses a lower limit on the measurable accuracy. While the DIR-Lab dataset provides an excellent tool for algorithm comparison, tests of statistical significance of algorithm performances however are not possible without gaining access to the full landmark error distributions for all algorithms being compared. Furthermore our method outperformed other segmentation based methods in terms of combined gap and overlap volumes. The reduction in gap and overlap volumes could be important for applications such as radiotherapy dose accumulation, but further work is required to demonstrate this.

One limitation of our method is that currently only two regions can be handled. Nevertheless, from our experiments this is sufficient for registrations of the respiratory motion in the thorax.

References

1. Berendsen, F.F., Kotte, A.N.T.J., Viergever, M.A., Pluim, J.P.W.: Registration of organs with sliding interfaces and changing topologies. Proc.SPIE Medical Imaging 9034, 9034–9034– 7 (2014)
2. Castillo, E., Castillo, R., Martinez, J., Shenoy, M., Guerrero, T.: Four-dimensional deformable image registration using trajectory modeling. Physics in Medicine and Biology 55(1), 305–327 (2010)
3. Castillo, R., Castillo, E., Guerra, R., Johnson, V.E., McPhail, T., Garg, A.K., Guerrero, T.: A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. Physics in Medicine and Biology 54(7), 1849–1870 (2009)
4. Cootes, T.F., Taylor, C.J., et al.: Statistical models of appearance for computer vision (2004), technical report, University of Manchester
5. Delmon, V., Rit, S., Pinho, R., Sarrut, D.: Registration of sliding objects using direction dependent B-splines decomposition. Physics in Medicine and Biology 58(5), 1303–1314 (2013)
6. Eiben, B., Vavourakis, V., Hipwell, J.H., Kabus, S., Lorenz, C., Buelow, T., Williams, N.R., Keshtgar, M., Hawkes, D.J.: Surface driven biomechanical breast image registration. Proc.SPIE Medical Imaging 9786, 9786–9786–10 (2016)
7. Hua, R., Pozo, J.M., Taylor, Z.A., Frangi, A.F.: Multiresolution eXtended Free-Form Deformations (XFFD) for non-rigid registration with discontinuous transforms. Medical Image Analysis 36, 113–122 (2017)
8. McClelland, J.R., Modat, M., Arridge, S., Grimes, H., DSouza, D., Thomas, D., Connell, D.O., Low, D.A., Kaza, E., Collins, D.J., Leach, M.O., Hawkes, D.J.: A generalized framework unifying image registration and respiratory motion models and incorporating image reconstruction, for partial image data or full images. Physics in Medicine and Biology 62(11), 4273–4292 (2017)
9. Modat, M., Ridgway, G.R., Taylor, Z.A., Lehmann, M., Barnes, J., Hawkes, D.J., Fox, N.C., Ourselin, S.: Fast free-form deformation using graphics processing units. Computer Methods and Programs in Biomedicine 98(3), 278–284 (2010)
10. Papiez, B.W., Heinrich, M.P., Fehrenbach, J., Risser, L., Schnabel, J.A.: An implicit sliding-motion preserving regularisation via bilateral filtering for deformable image registration. Medical Image Analysis 18(8, SI), 1299–1311 (2014)
11. The Deformable Image Registration Laboratory: DIR Spatial Accuracy Results. https://www.dir-lab.com/Results.html (2018), [Online; accessed 20-March-2018]
12. Vandemeulebroucke, J., Bernard, O., Rit, S., Kybic, J., Clarysse, P., Sarrut, D.: Automated segmentation of a motion mask to preserve sliding motion in deformable registration of thoracic CT. Medical Physics 39, 1006–1015 (2012)
13. Vishnevskiy, V., Gass, T., Szekely, G., Tanner, C., Goksel, O.: Isotropic total variation regularization of displacements in parametric image registration. IEEE Transactions on Medical Imaging 36(2), 385–395 (2017)
14. Wu, Z., Rietzel, E., Boldea, V., Sarrut, D., Sharp, G.C.: Evaluation of deformable registration of patient lung 4DCTs with subanatomical region segmentations. Medical Physics 35(2), 775–781 (2008)