

Deformation Estimation with Automatic Sliding Boundary Computation

Joseph Samuel Preston^{3(✉)}, Sarang Joshi^{1,3}, and Ross Whitaker^{2,3}

¹ Department of Bioengineering, University of Utah, Salt Lake City, USA

² School of Computing, University of Utah, Salt Lake City, USA

³ Scientific Computing and Imaging (SCI) Institute, University of Utah, Salt Lake City, USA

{jsam,sjoshi,whitaker}@sci.utah.edu

Abstract. We present a novel method for image registration via a piecewise diffeomorphic deformation which accommodates sliding motion, such as that encountered at organ boundaries. Our method jointly computes the deformation as well as a coherent sliding boundary, represented by a segmentation of the domain into regions of smooth motion. Discontinuities are allowed only at the boundaries of these regions, while invertibility of the total deformation is enforced by disallowing separation or overlap between regions. Optimization alternates between discrete segmentation estimation and continuous deformation estimation. We demonstrate our method on chest 4DCT data showing sliding motion of the lungs against the thoracic cage during breathing.

Keywords: Image registration · Sliding motion · Motion segmentation

1 Introduction

Dense image registration driven by a pointwise dissimilarity term is an ill-posed problem, requiring regularization to produce meaningful results. For medical images, smoothness and invertibility of the resulting deformations have become the standard requirements used to define regularization penalties. This represents assumptions regarding the physical system being modeled – that objects should not appear or disappear, and that nearby points should remain nearby. However, certain anatomical motions violate these assumptions. In particular, many organs are not directly attached to adjacent anatomical structures, but are able to slide freely along a boundary. Most notable in CT imaging is the sliding of the lower lungs against the thoracic cage.

We propose a method for jointly computing a segmentation representing regions of smooth motion, and a set of deformations modeling the motion in each region. Driven by a novel formulation of the sliding constraint as a discrete pairwise penalty, optimization alternates between a graph-cut based estimation of the motion segmentation and a continuous optimization of the constituent deformations. The resulting composite deformation is globally invertible and

piecewise-smooth. The motion segmentation guarantees that discontinuities due to sliding occur only along coherent boundaries in the image, and the joint estimation of mask and deformations allows anatomical structures to be automatically grouped based on the similarity of their motion.

Previous work has attempted to address the errors caused by using globally smooth deformations to represent lung motion. Wu *et al.* [11] segment regions and separately register them, while adding a penalty to avoid ‘gaps’ (or overlaps) in the resulting deformation. Recent work in this area has either required a precomputed segmentation of the sliding boundary (*e.g.* [2, 6]) or used a discontinuity preserving regularization on the deformation (such as Total Variation [10] or Bilateral Filtering [5]), which may introduce spurious discontinuities and does not guarantee the invertibility of the deformation. Vandemeulebroucke *et al.* [9] give an automated method for creating a ‘motion mask’ of the anatomy defining the sliding boundaries, but base this on anatomical features, and not the observed motion. Schmidt-Richberg *et al.* [7] propose a diffusion-based regularization which allows sliding along organ boundaries. Boundaries are either precomputed or estimated based on an ad-hoc measure of deformation discontinuity at precomputed image edge locations. Work from computer vision has also considered motion-based segmentation (*e.g.* [8]), but the modalities dictate that deformations model occlusion instead of invertibility/sliding. The most similar to our work is [2], which also formulates the sliding conditions in terms of a region segmentation, but assumes a predefined segmentation on one of the images.

2 Methods

We first formulate our registration problem and constraints for a piecewise diffeomorphic deformation $\bar{\phi}$. Given two D -dimensional scalar-valued images $f_0 : \Omega_0 \rightarrow \mathbb{R}$ and $f_t : \Omega_t \rightarrow \mathbb{R}$ (assigned artificial time indices 0 and t) where $\Omega_0, \Omega_t \subset \mathbb{R}^D$, we attempt to find a deformation $\bar{\phi} : \Omega_t \rightarrow \Omega_0$ such that $f_0 \circ \bar{\phi} \approx f_t$, where the dissimilarity is measured by a functional of the form

$$E_{\text{data}}(f_0 \circ \bar{\phi}, f_t) := \int_{\Omega_t} D(f_0(\bar{\phi}(\mathbf{x})), f_t(\mathbf{x})) d\mathbf{x}, \quad (1)$$

where $D : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is a pointwise dissimilarity term.

Our method formulates piecewise-diffeomorphic deformation $\bar{\phi}$ via K diffeomorphic deformations, $\phi := \{\phi^k\}_{k=1\dots K}$, $\phi^k : \Omega_t \rightarrow \Omega_0$ (where we assume identical codomains Ω_0 for simplicity), and a segmentation $\ell_0 : \Omega_0 \rightarrow \{1 \dots K\}$ defining the region over which each deformation has effect.

We define indicator functions of each label as

$$\chi_0^k(\mathbf{y}) := \begin{cases} 1 & \text{if } \ell_0(\mathbf{y}) = k \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

and the indicator function deformed by the corresponding deformation as

$$\chi_t^k := \chi_0^k \circ \phi^k.$$

We assume that each ϕ^k is diffeomorphic, and therefore invertible. In order to maintain invertibility of the composite deformation we require that it does not result in ‘tearing’ or ‘overlap’ between regions. This can be expressed as a pointwise constraint on the sum of indicator functions

$$\sum_{k=1}^K \chi_t^k(\mathbf{x}) = 1 \quad \forall \mathbf{x} \in \Omega_t. \quad (3)$$

Intuitively, $\chi_t^k = 1$ asserts that deformation ϕ^k is in effect at point \mathbf{x} . The constraint stated in terms of the labeling is then *at each \mathbf{x} , $\ell_0(\phi^k(\mathbf{x})) = k$ should be true for exactly one k* . If the statement is not true for any label k , then we have tearing in the deformation at time t , and if it is true for two or more labels we have overlap at this point.

We can now define the segmentation at time t

$$\ell_t(\mathbf{x}) := k \text{ s.t. } \chi_t^k(\mathbf{x}) = 1 \quad (4)$$

which, for $\mathbf{x} \in \Omega_t$, is guaranteed to exist and be unique by (3). We can then define a single piecewise-smooth deformation as

$$\bar{\phi}(\mathbf{x}) := \phi^{\ell_t(\mathbf{x})}(\mathbf{x}). \quad (5)$$

It is easy to verify that $\bar{\phi}(\mathbf{x})$ is invertible. It is surjective since each point $\mathbf{y} \in \Omega_0$ can be mapped to a point $\mathbf{x} \in \Omega_t$ via $\mathbf{x} = (\phi^{\ell_0(\mathbf{y})})^{-1}(\mathbf{y})$ via the invertibility of the constituent deformations, and injective because $\bar{\phi}(\mathbf{x})$ s.t. $\mathbf{x} \in \chi_t^k$ is a unique mapping among $\hat{\mathbf{x}} \in \chi_t^k$ via the invertibility of the constituent deformations, and unique among $\hat{\mathbf{x}} \notin \chi_t^k$ as $\hat{\mathbf{x}} \notin \chi_t^k$ implies $\chi_0^k(\bar{\phi}(\hat{\mathbf{x}})) \neq k$, contradicting $\mathbf{x} \in \chi_t^k$.

2.1 Optimization Criteria

We have now given constraints on the labeling ℓ_0 and deformations ϕ guaranteeing a piecewise-diffeomorphic deformation, and we choose among these valid deformations by optimizing over a composite objective function that balance the tradeoff between data matching and the regularity of the deformation and labeling

$$\mathbb{E}(\phi, \ell_0) := \mathbb{E}_{\text{data}}(\phi, \ell_0) + \lambda_{\text{per}} \mathbb{E}_{\text{per}}(\ell_0) + \lambda_{\text{pdf}} \mathbb{E}_{\text{pdf}}(\ell_0) + \lambda_{\text{reg}} \sum_{k=1}^K \mathbb{E}_{\text{reg}}(\phi^k), \quad (6)$$

where we reformulate the data term (1) as a function of ϕ and ℓ_0

$$\mathbb{E}_{\text{data}}(\phi, \ell_0) := \sum_{k=1}^K \int_{\Omega_t} \chi_t^k(\mathbf{x}) D(f_0(\phi^k(\mathbf{x})), f_t(\mathbf{x})) d\mathbf{x}, \quad (7)$$

and the regularization is taken over each constituent deformation (the form of the deformation regularization \mathbb{E}_{reg} is dependent on the registration method used).

The term E_{per} is a ‘perimeter’ penalty on the size of the segmentation boundaries, enforcing the notion that discontinuity should be allowed at a limited set of points. In addition, the penalty E_{pdf} is a region-based intensity segmentation penalty, encouraging the labeling ℓ_0 to segment regions of similar intensity. This helps produce ‘anatomically reasonable’ results (where segmentations follow organ boundaries) in regions such as the top of the lungs, where no sliding is observed and therefore motion segmentation is ambiguous.

The problem we wish to solve can then be stated as

$$\operatorname{argmin}_{\phi, \ell_0} E(\phi, \ell_0) \quad \text{s.t.} \quad \sum_{k=1}^K \chi_t^k(\mathbf{x}) = 1 \quad \forall \mathbf{x} \in \Omega_t. \quad (8)$$

This is clearly a challenging nonconvex optimization problem. In order to achieve an approximate solution, the hard constraint (3) is relaxed to a penalty,

$$E_{\text{sum}}(\phi, \ell_0) := \left\| \sum_{k=1}^K \chi_t^k - \mathbf{1} \right\|_1, \quad (9)$$

We write the relaxed version of (8) as

$$\operatorname{argmin}_{\phi, \ell_0} E(\phi, \ell_0) + \lambda_{\text{sum}} E_{\text{sum}}(\phi, \ell_0), \quad (10)$$

and optimization alternates between finding an optimal labeling ℓ_0 for a fixed set of deformations ϕ , and optimizing over ϕ while keeping ℓ_0 fixed. This alternating optimizations do not increase the objective (10) at each step, guaranteeing convergence.

2.2 Segmentation Estimation

In the segmentation estimation step we hold the deformations ϕ fixed and optimize (10) over ℓ_0

$$\operatorname{argmin}_{\ell_0} E_{\text{data}}(\phi, \ell_0) + \lambda_{\text{per}} E_{\text{per}}(\ell_0) + \lambda_{\text{pdf}} E_{\text{pdf}}(\ell_0) + \lambda_{\text{sum}} E_{\text{sum}}(\phi, \ell_0). \quad (11)$$

Note that for a large enough value of λ_{sum} , this becomes a hard constraint; since a single-label segmentation \mathbf{k} s.t. $\mathbf{k}(\mathbf{x}) = k \forall \mathbf{x}$ is guaranteed to satisfy the constraints and also have zero boundary cost ($E_{\text{per}}(\mathbf{k}) = 0$), setting $\lambda_{\text{sum}} = \min_{\mathbf{k}} (E_{\text{data}}(\phi, \mathbf{k})) + \epsilon$ guarantees that the optimal solution satisfies the constraint.

As the segmentation ℓ_0 takes discrete values, we propose a discrete optimization formulation for its estimation. We discretize our problem by defining f_0 and ℓ_0 only on a discrete set of points arranged on a regular grid, $\{\mathbf{y}_n\}_{n=1}^N \subset \Omega_0$, and similarly define f_t , ℓ_0 , ϕ^k , etc. on $\{\mathbf{x}_n\}_{n=1}^N \subset \Omega_t$. For simplicity we will assume equal numbers of points indexed by n . For interpolation of images linear interpolation is used, and for label images nearest-neighbor interpolation is used. The objective is written as a discrete set of unary and pairwise functions

of node labelings which can be solved exactly for binary labelings ($K=2$), and approximately otherwise, via graph-cut algorithms [3].

The data matching term E_{data} (7) can be written discretely as a unary point-wise penalty on the labeling ℓ_0 , as can E_{pdf} . We write E_{pdf} as a negative log likelihood penalty on ℓ_0

$$E_{\text{pdf}}(\ell_0) := -\log(p(f_t|\ell_0)) = -\sum_{n=1}^N p(f_t(\mathbf{y}_n)|\ell_0(\mathbf{y}_n)), \quad (12)$$

where $p(f_t(\mathbf{y}_n)|\ell_0(\mathbf{y}_n) = k)$ is a kernel density estimate of the distribution of $\{f_t(\mathbf{y}_n) \text{ s.t. } (\ell_0)^-(\mathbf{y}_n) = k\}$, and $(\ell_0)^-$ is the segmentation from the previous iteration. As this penalty is intended only to resolve ambiguous situations, and not lead to an intensity based segmentation, its influence is kept relatively weak.

We now formulate E_{per} and E_{sum} discretely as sums of pairwise terms. For ease of notation when working with label values, we define the functions $q : \{1 \dots K\} \rightarrow \{0, 1\}$ (and similarly $\neg q$) for indicating equality (or inequality) of labelings (q for **equal**s):

$$q(i, j) := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad \neg q(i, j) := \begin{cases} 1 & \text{if } i \neq j \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

Our ‘perimeter’ penalty E_{per} is written as a simple Potts model smoothness term on ℓ_0 ,

$$E_{\text{per}}(\ell_0) := \sum_{n=1}^N \sum_{\mathbf{y}_m \in \mathcal{N}(\mathbf{y}_n)} \neg q(\ell_0(\mathbf{y}_n), \ell_0(\mathbf{y}_m)), \quad (14)$$

where $\mathcal{N}(\mathbf{y}_n)$ is a set containing the neighbors of \mathbf{y}_n , in our case a 4- or 6-connected neighborhood in 2D or 3D, respectively.

We can directly write the discretized version of $E_{\text{sum}}(\phi, \ell_0)$ (9) as the sum of N terms (one associated with each \mathbf{x}_n), term n taking as inputs the labelings $\{\ell_0(\phi^k(\mathbf{x}_n))\}_{k=1}^K$

$$E_{\text{sum}}(\phi, \ell_0) := \sum_{n=1}^N \left| \left(\sum_{k=1}^K q(\ell_0(\phi^k(\mathbf{x}_n)), k) \right) - 1 \right|. \quad (15)$$

Due to nearest-neighbor interpolation of the labeling, each $\ell_0(\phi^k(\mathbf{x}_n))$ is equivalent to $\ell_0(\mathbf{y}_n)$ for some $n \in \{1 \dots N\}$. However, the K -input terms associated with each point are difficult and costly to optimize over. Instead, we introduce auxiliary variables representing the labeling ℓ_t , and optimize over both ℓ_0 and ℓ_t . Under the sum constraint (9), ℓ_t is uniquely determined via ℓ_0 and ϕ , so we have not introduced any additional unknowns. However it allows us to rewrite each K -input term in (15) as the sum of K pairwise terms

$$E_{\text{sum}}(\phi, \ell_0, \ell_t) := \sum_{n=1}^N \sum_{k=1}^K c^k(\ell_0(\phi^k(\mathbf{x}_n)), \ell_t(\mathbf{x}_n)), \quad (16)$$

where $c^k(i, j) := -\mathbf{q}(\mathbf{q}(i, k), \mathbf{q}(j, k))$. The function $c^k(\ell_0(\phi^k(\mathbf{x}_n)), \ell_t(\mathbf{x}_n))$ indicates whether labels $\ell_0(\phi^k(\mathbf{x}_n))$ and $\ell_t(\mathbf{x}_n)$ agree that deformation k is (or is not) in effect at $\phi^k(\mathbf{x}_n)$. It takes value zero if $\ell_0(\phi^k(\mathbf{x}_n)) = \ell_t(\mathbf{x}_n) = k$ (the labelings agree that $\bar{\phi}(\mathbf{x}_n) = \phi^k(\mathbf{x}_n)$) or if $\ell_0(\phi^k(\mathbf{x}_n)) \neq k$ and $\ell_t(\mathbf{x}_n) \neq k$ (the labelings agree that $\bar{\phi}(\mathbf{x}_n) \neq \phi^k(\mathbf{x}_n)$), but takes value 1 if the labelings are in disagreement.

We now have the objective function (11) written as unary (E_{data} and E_{pdf}) and pairwise (E_{per} and E_{sum}) terms on the label values $\ell_0(\mathbf{y}_n)$ and $\ell_t(\mathbf{x}_n)$, which can be optimized by efficient discrete methods [3].

In practice, the effects of discretization mean that enforcing hard constraints give unsatisfactory results, even if the true deformations are known. Instead, λ_{sum} is chosen to allow some constraint violations. In locations of constraint violation the labeling of ℓ_t is ambiguous, but the given optimization chooses among ambiguous labels by minimizing the data energy.

2.3 Deformation Estimation

In the deformations estimation step, we hold the labeling ℓ_0 fixed and optimize (10) over the deformations ϕ

$$\operatorname{argmin}_{\phi} E_{\text{data}}(\phi, \ell_0) + \lambda_{\text{reg}} \sum_{k=1}^K E_{\text{reg}}(\phi^k) + \lambda_{\text{sum}} E_{\text{sum}}(\phi, \ell_0), \quad (17)$$

We represent each ϕ^k via a b-spline control grid, and impose additional regularization via $E_{\text{reg}}(\phi^k) := \|\nabla \phi^k\|^2$. While invertibility is not guaranteed by this regularization, it can be easily verified. In practice the regularization is significant (as allowed by the piecewise formulation) and the deformations are sufficiently well-behaved that we do not observe noninvertibility during optimization. However, the optimization is agnostic towards the deformation model, and in particular a flow-based diffeomorphic model could be substituted without additional changes.

The L^1 norm in E_{sum} (9) is replaced by a differentiable approximation, and optimization is performed via gradient descent.

2.4 Segmentation Initialization

As we are optimizing a highly nonconvex function, a reasonable initialization is needed. As the goal is to segment deformations along regions of motion discontinuity, we automatically create an initial segmentation which separates regions showing different motion. We first generate a deformation on downsampled versions of the input data enforcing only the anisotropic TV penalty

$$\operatorname{argmin}_{\nu} \int_{\Omega_t} D(\mathbf{f}_0(\mathbf{x} + \nu(\mathbf{x})), \mathbf{f}_t) + \sum_{d=1}^D |\nabla_d \nu(\mathbf{x})| \, d\mathbf{x}, \quad (18)$$

where ∇_d is the gradient along dimension d .

This penalty results in a discontinuous motion field containing piecewise-constant regions. The elements of ν are then grouped into K clusters using the k-means algorithm, and the cluster labels are then mapped back to the corresponding spatial locations creating an initial segmentation.

3 Results

We test our method on the publicly available DIR-lab dataset [1], which consists of ten 4DCT datasets as well as landmark correspondences. Resolution (voxel size) is $\approx 1 \times 1 \times 2.5$ mm. Datasets 6–10 contain a large volume of empty space around the patient, and were cropped to a region of interest around the patient body. We use a binary motion segmentation ($K = 2$) to represent the breathing motion. A multiscale optimization was used, both on the image resolution and the b-spline grid resolution (two image resolutions, with a total of three b-spline grid refinements for cases 1–5 and four for cases 6–10). Other parameters for both the initialization and optimization steps were empirically chosen, but fixed across all datasets. In order to better register fine features and overcome intensity differences due to density changes in the lung during breathing, we chose a Normalized Gradient Field (NGF) dissimilarity [4] as the pointwise image dissimilarity measure. Figure 1 shows the computed motion segmentation for each dataset. The segmented region includes the lungs as well as regions of visible motion continuing into the organs of the abdominal cavity, while excluding structures such as bone which show little motion. Table 1 gives the landmark errors on these datasets, showing similar results to methods which rely on a precomputed segmentation, or do not guarantee the invertibility of the deformation (average correspondence error of 1.07 mm across all ten datasets, compared to 1.00 mm for method NGF(b) of [4] and 1.01 mm for method pTV of [10]). Figure 2 compares the effect of the piecewise-smooth deformation to that of a single smooth deformation run with the same parameters.

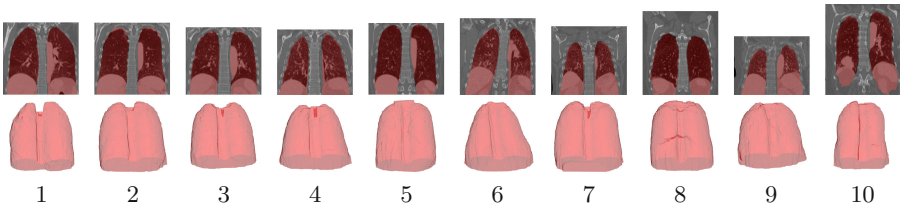
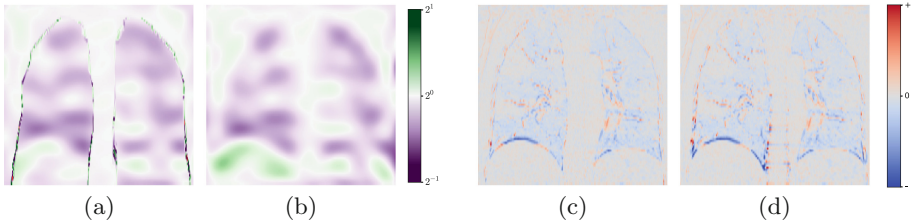


Fig. 1. Motion segmentation results for all datasets. Top row shows a coronal slice near the spine with motion mask region highlighted in red. Bottom row shows transparent 3D rendering of this region. Bottom number indicates the dataset. (Color figure online)

Table 1. landmark errors (in millimeters) for the $n = 300$ landmarks defined at the extreme breathing phases for each dataset

dataset	1	2	3	4	5	6	7	8	9	10	ttl
err mean (mm)	0.79	0.79	0.87	1.35	1.25	1.05	1.12	1.22	1.12	1.12	1.07
err std (mm)	0.90	0.95	1.04	1.28	1.53	1.04	1.18	1.26	1.03	1.38	1.17

**Fig. 2.** Comparison of our piecewise-smooth deformation ((a) and (c)), and a single globally smooth deformation ((b) and (d)), on the DIR-lab dataset 1. The figures on the left show the jacobian determinant of the deformation on a log-scale colormap. Note the clear sliding boundary (dark/red colors where the jacobian is undefined along the tearing boundary) in (a), and the nonphysical deformation of the spine region in (b). On the right are difference images between exhale and deformed inhale for the two methods. Again notice the motion of the spine with a single deformation. (Color figure online)

Acknowledgements. This work was supported in part by NIH R01 CA169102-01A13 and a grant from GE Medical Systems.

References

- 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. *Phys. Med. Biol.* **54**(7), 1849 (2009)
- Heldmann, S., Polzin, T., Derksen, A., Berkels, B.: An image registration framework for sliding motion with piecewise smooth deformations. In: Aujol, J.-F., Nikolova, M., Papadakis, N. (eds.) *SSVM 2015*. LNCS, vol. 9087, pp. 335–347. Springer, Heidelberg (2015). doi:[10.1007/978-3-319-18461-6_27](https://doi.org/10.1007/978-3-319-18461-6_27)
- Kolmogorov, V., Zabini, R.: What energy functions can be minimized via graph cuts? *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(2), 147–159 (2004)
- Konig, L., Ruhaak, J.: A fast and accurate parallel algorithm for non-linear image registration using normalized gradient fields. In: *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, pp. 580–583. IEEE (2014)
- Papiež, 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. *Med. Image Anal.* **18**(8), 1299–1311 (2014)
- Risser, L., Vialard, F.X., Baluwala, H.Y., Schnabel, J.A.: Piecewise-diffeomorphic image registration: Application to the motion estimation between 3d ct lung images with sliding conditions. *Med. Image Anal.* **17**(2), 182–193 (2013)

7. Schmidt-Richberg, A., Werner, R., Handels, H., Ehrhardt, J.: Estimation of slipping organ motion by registration with direction-dependent regularization. *Med. Image Anal.* **16**(1), 150–159 (2012)
8. Sun, D., Sudderth, E.B., Black, M.J.: Layered segmentation and optical flow estimation over time. In: *IEEE CVPR*, pp. 1768–1775. IEEE (2012)
9. 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. *Med. Phys.* **39**(2), 1006–1015 (2012)
10. Vishnevskiy, V., Gass, T., Székely, G., Goksel, O.: Total variation regularization of displacements in parametric image registration. In: Yoshida, H., Näppi, J.J., Saini, S. (eds.) *Abdominal Imaging: Computational and Clinical Applications*. LNCS, vol. 8676, pp. 211–220. Springer, Heidelberg (2014)
11. Wu, Z., Rietzel, E., Boldea, V., Sarrut, D., Sharp, G.C.: Evaluation of deformable registration of patient lung 4DCT with subanatomical region segmentations. *Med. Phys.* **35**(2), 775–781 (2008)