Estimating Large Lung Motion in COPD Patients by Symmetric Regularised Correspondence Fields

Mattias P. Heinrich[1], Heinz Handels[1], and Ivor J.A. Simpson[2]

[1] Institute of Medical Informatics, University of Lübeck, Germany
[2] Centre for Medical Image Computing, University College London, UK
heinrich@imi.uni-luebeck.de
www.mpheinrich.de

Abstract. This paper presents a new and highly efficient approach for finding correspondences across volumes with large motion. Most existing registration approaches are set in the continuous optimisation domain, which has severe limitations for estimating larger deformations. Feature-based approaches that rely on finding corresponding keypoints have been proposed, but they are prone to erroneous matching due to repetitive features and low contrast areas. This can be overcome by using a discrete optimisation approach. However, finding a constrained search space and regularisation strategy is still an open problem. Our method calculates a dissimilarity distribution over a densely sampled space of displacements for a small number of distinctive keypoints (found in only one volume). A parts-based model is used to infer smooth motion of connected keypoints and regularise the correspondence field. This effective and highly accurate approach is further improved by enforcing the symmetry of uncertainty estimates of displacements. Our method ranks first on one of the most challenging medical registration benchmarks for breath-hold CT scan-pairs of COPD patients, where accurate motion estimation is important for diagnosis.

Keywords: Deformable registration, Parts-based model, Discrete optimisation.

1 Introduction

Finding correspondences between scans is a fundamental process in medical image analysis. In particular, in the diagnosis of COPD (Chronic Obstructive Pulmonary Disease), the detection of air trapped in localised malfunctioning regions of the lung can be assisted by registering an inhale and exhale CT scan pair. Motion of small amplitude can be robustly estimated with widely studied intensity-based continuous optimisation approaches, which often use coarse-to-fine schemes. However, such approaches are severely limited when the expected motion vectors are of larger scale than the image features (>40 mm for respiration). The difficulty of registering inhale and exhale images is highlighted in a comparison of many state-of-the-art methods for deformable registration of

© Springer International Publishing Switzerland 2015
N. Navab et al. (Eds.): MICCAI 2015, Part II, LNCS 9350, pp. 338–345, 2015.
DOI: 10.1007/978-3-319-24571-3_41

lung motion [19]. Discrete optimisation registration strategies have been demonstrated to be helpful in the estimation of larger scale deformation in computer vision [16] and recently also for medical image registration [15,11,13]. However, choosing an appropriate search space of deformations and regularisation strategy are still open problems for efficient, accurate and plausible registration.

In feature-based matching, the search space is restricted and correspondences across images are only seeked for sparse keypoints. These discrete matches can act as constraints in a continuous motion estimation [3]. However, it may be difficult in many applications to find enough corresponding features that can be reliably detected directly in both images. Feature matching has also been used in [20], where additionally the mutual saliency of interest points was optimised. In [14], 3D volumes were represented by supervoxels and used together with a graphical model to match corresponding feature regions across images. Block-matching is an alternative solution to find correspondence fields for medical scans [4,8]. It has, however, two main disadvantages. First, no spatial regularity between matches of neighbouring patches is enforced, so unrealistic deformations from these point matches have to be compensated for. Second, usually only the best match is retained (with the exception of e.g. [23] who use a small set of possible correspondences), and a following smoothing step of the motion field usually ignores any probabilistic information from the initial discrete search.

In this work, we present a novel framework that addresses these limitations using the following three steps. First, we extract features from one image using a keypoint operator. Second, a similarity measure is calculated over a densely quantised space of displacements, which translate to potentially corresponding locations in the other image (in contrast to previous work using only sparse samples [3,14,23]). Third, we exploit contextual information by using a parts-based model in order to regularise the differences of neighbouring displacements vectors using a Markov Random Field (MRF). Inference of regularisation for general MRFs is NP-hard, however, tree approximations have been shown to produce accurate and fast solutions [24]. Parts-based models [10], which employ interactions between different parts of an object to stabilise their localisation, have so far only been employed for landmark localisation in medical images [22]. We evaluate our method in Sec. 3 on the challenging COPD dataset.

2 Method

We aim to find a dense correspondence field between a fixed volume F and a moving volume M. Our method consists of three main steps, which are detailed below. First, a set of distinctive keypoints $\mathbf{k}_F \in K$ with spatial coordinates (k_x, k_y, k_z) are extracted from F, the inhale scan (total lung capacity) and a graph is built that uses edges $e_{ij} \in \mathcal{E}$ to connect nodes i and j with similar coordinates and intensities. Second, a similarity metric is calculated for voxels within a small patch around the keypoint $\mathbf{k}_F$ in the fixed volume and a corresponding patch displaced by a 3D motion vector $\mathbf{d} = (d_x, d_y, d_z)$. This is repeated for every displacement in the set $\mathcal{L}$ yielding a 3D distribution of similarity energies. Third, spatial regularity is inferred over the full displacement distributions of all

keypoints together using belief propagation with a squared penalty for deviations of displacements, which yields an exact solution in linear time (c.f.[10]).

Keypoint Detection: Following previous work in lung registration [21], we employ the Foerstner operator to find keypoints in the 3D intensity volumes. After smoothing the volume with a Gaussian kernel G_σ the spatial gradients ∇F are calculated. A distinctiveness $D(\mathbf{x}) = 1/\operatorname{trace}\left((G_\sigma * (\nabla F \nabla F^T))^{-1}\right)$ can then be determined for each voxel $\mathbf{x}$, where a high response indicates good candidates (see Fig. 1). In order to obtain a good dispersion of points over the whole volume of interest, we apply a grey-value dilation over a 3D cubic region G_M with side-length c_D to the Foerstner response and obtain $D^* = \max_{\mathbf{y}\in G_M} D(\mathbf{y})$. Only points with equal response in D and D^* are added to K.

Similarity Evaluation over Search Space: The aim of image registration is to find a correspondence (or displacement) field, which assigns a motion vector to every control point: $\mathbf{u}(\mathbf{k}_F) \leftarrow \mathbf{d}$. The new location $\mathbf{l} = (l_x, l_y, l_z) = (k_x + d_x, k_y + d_y, k_z + d_z)$ is chosen to maximise the similarity between images. We define the space of potential displacements for every keypoint $\mathbf{k}_F$ to be the densely sampled 3D set $\mathbf{d} \in \mathcal{L} = \{0, \pm q, \pm 2q, \ldots, \pm l_{\max} q\}^3$, where q is a very small quantisation step and $l_{\max} q$ large enough to cover all potential motion between both images. Note, that in contrast to previous work on medical image registration [11] the motion is not decoupled into 1D displacements, as this may lead to inferior matching. Since, medical volumes with large motion are often also degraded by local intensity variations and influenced by strong noise, we use the self-similarity descriptors (SSC) proposed in [12] for matching. The dissimilarity metric $\mathcal{D}$, the L_1 norm between 64 bit binary descriptor representations SSC_F and SSC_M at locations $\mathbf{k}$ and $\mathbf{l}$, can be efficiently calculated in the Hamming space: $\mathcal{D}(\mathbf{k}, \mathbf{l}) = 1/|\mathcal{P}| \sum_{p\in\mathcal{P}} \Xi\{SSC_F(\mathbf{k}+p) \oplus SSC_M(\mathbf{l}+p)\}$, where $\oplus$ defines an exclusive OR, Ξ a bit-count, and $\mathcal{P}$ a local patch.

Parts-Based Model for Inference of Regularisation: In traditional block-matching approaches [4,8] or hybrid approaches [3], the motion vector corresponding to the minimum dissimilarity is selected and a heuristic or variational regularisation is then applied to the correspondence field. It is, however, beneficial

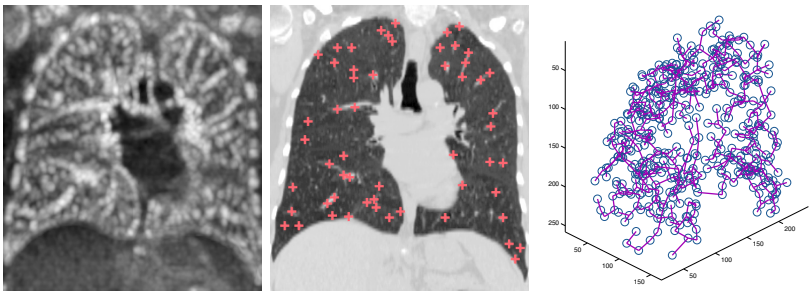

Fig. 1. Left: Response of Foerstner operator with log scale. Keypoints found in the respective coronal plane of the CT volume (and 3 neighbouring slices each) are marked with red crosses. Right: Minimum-spanning-tree, which connects keypoints.

to evaluate the spatial regularity of the correspondence field over the whole space of displacements as done in MRF-based medical image registration [11,13]. While the previous approaches use a regular grid of equally spaced control points, we employ only a very sparse set of keypoints. Building a graph on this irregular set of points is not straightforward. While a general k-nearest neighbour graph would be possible, it is not necessarily spanning. Following parts-based models [10], we opt to use a minimum-spanning-tree (MST) to define the set of edges $\mathcal{E}$ in our model. MSTs have been used for inference in many applications including stereo estimation [24] and 3D registration [13] on regular grids. The regularisation cost assigns a penalty $\mathcal{R}$ for every pair of keypoints, which is connected in the graph by an edge e_{ij} and has unequal displacements:

$$\mathcal{R}(\mathbf{d}_i, \mathbf{d}_j) = \frac{||\mathbf{d}_i - \mathbf{d}_j||^2}{\sqrt{||\mathbf{x}_i - \mathbf{x}_j||^2 + |I(\mathbf{x}_i) - I(\mathbf{x}_j)|/\sigma_I}} \tag{1}$$

This term encourages smoothly varying displacements in the estimated correspondence field and makes our method robust against missing correspondences or image artefacts. The denominator accounts for different lengths of edges, which includes both their spatial Euclidean distance and the absolute difference of their local intensity means. Combining the regularisation penalty with the self-similarity based dissimilarity (weighted by α) yields the total energy of a certain correspondence field $\mathbf{u}$: $E(\mathbf{u}) = \alpha \sum_{\mathbf{k} \in K} \mathcal{D}(\mathbf{k}, \mathbf{d_k}) + \sum_{e_{ij} \in \mathcal{E}} \mathcal{R}(\mathbf{d}_i, \mathbf{d}_j)$.

Belief propagation can be used to obtain exact marginal distributions after two passes of messages. Starting from the leaf nodes, messages $\mathbf{m}$ are passed along the edges e_{ij} of the tree and updated with the following computation for a node i and its parent j [10]:

$$\mathbf{m}_i(\mathbf{d}_j) = \min_{\mathbf{d}_i} \left(\alpha \mathcal{D}(\mathbf{d}_i) + \mathcal{R}(\mathbf{d}_i, \mathbf{d}_j) + \sum_c \mathbf{m}_c(\mathbf{d}_i) \right), \tag{2}$$

where c are the children of i. The evaluation of Eq. 2 is particularly efficient (linear in $|\mathcal{L}|$), when using distance transforms.

Symmetry of Marginal Distribution: After globally optimising the regularised cost function, there may still be spurious errors. To reduce their influence we make use of the estimated marginal energies. An additional estimation of the marginal distribution is performed in opposite direction (from moving to fixed volume) using the translated keypoints of the first optimisation as control points. Ideally, the forward estimate was correct and so the energies of the backwards displacement space are symmetrical to the forward ones and averaging re-enforces them. If the correspondence from fixed to moving image pointed to a close neighbour of the correct displacement and the motion in a small neighbourhood is smooth the backward search can improve the match. Forward and backward marginal energies are averaged as follows. Let $M^f_{\mathbf{k}_F}$ by a vector of marginal energies of length $|\mathcal{L}|$ using the forward search for a certain point $\mathbf{k}_F$ and $M^b_{\mathbf{k}_M}$ the marginals from the backward search of the translated control point

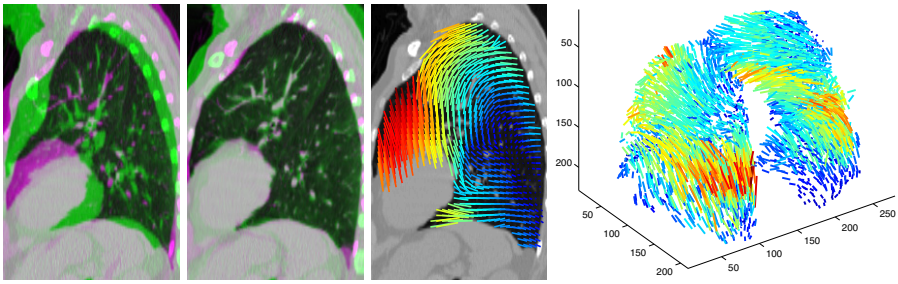

Fig. 2. Visual results of dense correspondence fields using our proposed method for case #4. Colour overlay of inhale (green) and exhale (magenta) sagittal slices. 3D motion vectors (magnitude from blue to red) are shown for the most challenging case #2.

in the moving volume $\mathbf{k}_M$. The symmetric marginal energies $M^s_{\mathbf{k}}$ are then defined as $M^s_{\mathbf{k}}(i) = \frac{1}{2}(M^f_{\mathbf{k}_F}(i) + M^b_{\mathbf{k}_M}(|\mathcal{L}| - i))$.

Dense Field from Sparse Matches: In the previous steps, we have described how an accurate, sparse correspondence field can be established by minimising the registration cost function. We estimate the marginal distributions for every displacements and node, so a first refinement of the integer motion vectors to sub-voxel precision can be obtained by a parabolic fit around the minimum. We employ thin-plate splines (TPS) [2] to interpolate motion vectors between known keypoints. This approach yields a smooth dense displacement field.

Refinement of Correspondences: A further improvement in accuracy can be obtained by refinement of the correspondence field. Using the dense displacement field, the moving volume is warped towards the fixed volume. A new displacement space $\mathcal{L}_{\text{refine}}$ with a smaller search range and quantisation q is then defined. We increase the number of keypoints, since this second optimisation will be of much lower complexity due to the reduced label space. In order to correctly regularise the combined motion, the concept of offsets in the message computation [13] is employed. As before, first forward and then backward correspondence fields are estimated for improved accuracy and robustness.

3 Experiments and Results

To evaluate our method for very large 3D motion, we use the recently published dataset of [6]. It consists of ten inhale and exhale scan pairs of patients from the COPDgene study, with severe breathing disorders. One difficulty is the poor SNR of the exhale scan due to an ultra-low dose CT protocol. The lung volume change between full exhalation and inhalation can be more than 100 % and the motion magnitude of individual features within the lungs is much larger than in other medical image registration benchmarks (4DCT motion [5] or inter-subject brain mapping), where the performance of new algorithms has nearly converged. The challenges of the COPD dataset have been highlighted by a comparison of three of the most widely used medical registration tools in [18]. The quantitative

Table 1. Results for 3D registration of 10 inhale and exhale CT scan pairs of COPD data set. The average TRE in mm is evaluated for 300 expert selected landmark pairs per case as defined by [6]. The p-values show the statistical significance of the difference in landmark error from the proposed method (calculated with Wilcoxon rank-sum test).

case #	initial	gsyn [1]	LMP [21]	MILO [6]	SGM [15]	without inference	Laplacian [7]	uniform points	asym-metric	**proposed**
1	26.3	1.21	1.26	**0.93**	1.22	5.86	2.90	1.13	1.10	1.00±0.93
2	21.8	3.01	2.02	1.77	2.48	8.39	6.70	2.27	2.10	**1.62**±1.78
3	12.6	1.24	1.14	**0.99**	1.01	4.18	1.41	1.22	1.08	1.00±1.06
4	29.6	1.38	1.62	1.14	2.42	7.85	2.24	1.39	1.20	**1.08**±1.05
5	30.1	1.31	1.47	1.02	1.93	6.36	2.61	1.21	1.15	**0.96**±1.13
6	28.5	1.49	1.39	**0.99**	1.45	4.49	2.30	1.03	1.20	1.01±1.25
7	21.6	1.24	1.22	**1.03**	1.05	5.13	2.11	1.09	1.07	1.05±1.07
8	26.5	2.09	1.63	1.31	1.16	5.29	1.74	1.35	1.15	**1.08**±1.24
9	14.9	1.18	1.12	0.86	0.81	5.29	1.37	1.16	0.84	**0.79**±0.80
10	21.8	1.63	1.45	1.23	1.28	7.98	4.40	1.47	1.47	**1.18**±1.31
avg	23.4	1.58	1.43	1.13	1.48	6.08	2.78	1.33	1.24	**1.08**
std	10.1	1.93	1.45	1.29	2.19	7.00	4.57	1.67	1.60	**1.21**
time		>30 m	-	>10 m	1.5 m	1.5 m	3.5 m	3 m	1.8 m	3 m
p-**val**		$6{\cdot}10^{-4}$	$1{\cdot}10^{-3}$	0.45	0.03	$9{\cdot}10^{-5}$	$2{\cdot}10^{-4}$	$6{\cdot}10^{-3}$	0.03	-

evaluation revealed errors of 1.79±2.1 mm, 2.19±2.0 mm and 4.68±4.1 mm for the ANTS gsyn [1], NiftyReg [17] and DROP [11] methods, respectively (results are after applying the boosting method of [18]).

Implementation Details: The original size of a voxel is between 0.586 and 0.742 mm in-plane and 2.5 mm in z-direction. The fixed volume was manually cropped to include the full lung and resampled to an isotropic voxel-size of 1 mm. The moving volume is also cropped and then resampled to match the dimensions of the fixed volume. A lung segmentation of the fixed volume is automatically obtained using thresholding and morphological filters. This mask is only used to restrict keypoint detection to the inner lung volume. The following parameters were used for the initial correspondence search: $G_\sigma = 1.4$, $c_D = 6$, σ_I=150, $l_{\max}q$=32, $q = 2$, which yielded ≈3000 keypoints and $|\mathcal{L}| = 35'937$ displacement labels. The regularisation weight was manually tuned for one case (#4) to $\alpha = 1$ and left unaltered for all other cases. For the refinement stage we reduce $c_D = 3$, $l_{\max}$=8, $q = 1$. Computation times for the dissimilarity calculation are 60 sec. and 32 sec. for inference (per direction using one CPU core). Our source code is made publicly available at `http://mpheinrich.de/software.html`.

Discussion of Results: Table 1 (right) lists the obtained target registration error (TRE) of 1.08 mm for our proposed method over 300 manual selected landmark pairs per case [6]. Visualisations of the outcome for two cases are shown in Fig. 2. The obtained deformations have a small amount, 0.14% on average, of voxels with negative Jacobians, and the average deformation Jacobian std. dev. is 0.26. We also evaluated the TRE for the following four variants of our algorithm to show the usefulness of our algorithmic choices. No inference or regularisation (classical block-matching) leads to 6.08 mm; random-walk regularisation

with the graph Laplacian as used in [7] to 2.78 mm; belief propagation without symmetric constraint to 1.24 mm; and uniform keypoint sampling without Foerstner operator to a TRE of 1.33 mm.

The task of aligning COPD breath-hold scans has been addressed by the following methods. In [21] a block-matching approach is used for keypoints, but without MRF regularisation, and a TPS transform to initialise a variational optimisation method (LMP). The TPS has an accuracy of 3.20 mm, the refinement reduces this to 1.43 mm. Hermann [15] used a discrete optimisation (semi-global matching, SGM), with a coarse-to-fine multi resolution scheme, and Census cost function and reached a TRE of 1.48 mm. Finally, Castillo et al. [4] used a very large number of control points (every 2nd voxel in each dimension) for block-matching followed by an optimisation based filtering of outliers (MILO). They obtained a TRE of 1.13 mm, which is slightly inferior to our results of 1.08 mm, and they report relatively high computation times of 10 to 20 minutes per case. For the more commonly used 4DCT data of [5], we achieve an average TRE of 0.97 mm, which compares well with [15]: 0.95mm, [13]: 1.43mm and [21]: 1.10mm.

4 Conclusion

We have presented a new MRF-based approach to estimate accurate correspondence fields for images with large motion. Our main contribution is a sparse, but informative, sampling of keypoints locations for which motion vectors are determined from a large set of displacements. The regularity of the displacements is enforced by inference over all pair-wise combinations of close locations using a parts-based model. Devoting much effort only to interesting positions in the image, yields both accurate and fast results. Our method currently achieves the lowest registration error of 1.08 mm on the public COPD dataset of [6]. Future work, will include the application to whole-body registration using organ-surface keypoints and comparisons to graph matching [9].

References

1. Avants, B., Epstein, C., Grossman, M., Gee, J.: Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. Med. Imag. Anal. 12(1), 26–41 (2008)
2. Bookstein, F.L.: Principal warps: Thin-plate splines and the decomposition of deformations. IEEE Trans. Pat. Anal. Mach. Intel. 11(6), 567–585 (1989)
3. Brox, T., Malik, J.: Large displacement optical flow: descriptor matching in variational motion estimation. IEEE Pattern Anal. Mach. Intell. 33(3), 500–513 (2011)
4. Castillo, E., Castillo, R., Fuentes, D., Guerrero, T.: Computing global minimizers to a constrained B-spline image registration problem from optimal L1 perturbations to block match data. Med. Phys. 41(4), 041904 (2014)
5. Castillo, R., Castillo, E., Guerra, R., Johnson, V., McPhail, T., Garg, A., Guerrero, T.: A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. Phys. Med. Biol. 54(7), 1849 (2009)
6. Castillo, R., Castillo, E., Fuentes, D., Ahmad, M., Wood, A.M., Ludwig, M.S., Guerrero, T.: A reference dataset for deformable image registration spatial accuracy evaluation using the COPDgene study archive. Phys. Med. Biol. 58(9) (2013)

7. Cobzas, D., Sen, A.: Random walks for deformable image registration. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part II. LNCS, vol. 6892, pp. 557–565. Springer, Heidelberg (2011)
8. Commowick, O., Arsigny, V., Isambert, A., Costa, J., Dhermain, F., Bidault, F., Bondiau, P.Y., Ayache, N., Malandain, G.: An efficient locally affine framework for the smooth registration of anatomical structures. Med. Imag. Anal. 12(4) (2008)
9. Duchenne, O., Bach, F., Kweon, I.S., Ponce, J.: A tensor-based algorithm for high-order graph matching. IEEE Pattern Anal. Mach. Intell. 33(12), 2383–2395 (2011)
10. Felzenszwalb, P.F., Huttenlocher, D.P.: Pictorial structures for object recognition. Int. J. Comp. Vis. 61(1), 55–79 (2005)
11. Glocker, B., Komodakis, N., Tziritas, G., Navab, N., Paragios, N.: Dense image registration through MRFs and efficient linear programming. Med. Imag. Anal. 12(6), 731–741 (2008)
12. Heinrich, M.P., Jenkinson, M., Papież, B.W., Brady, S.M., Schnabel, J.A.: Towards realtime multimodal fusion for image-guided interventions using self-similarities. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013, Part I. LNCS, vol. 8149, pp. 187–194. Springer, Heidelberg (2013)
13. Heinrich, M., Jenkinson, M., Brady, M., Schnabel, J.: MRF-based deformable registration and ventilation estimation of lung CT. IEEE Trans. Med. Imag. 32(7), 1239–1248 (2013)
14. Heinrich, M.P., Jenkinson, M., Papież, B.W., Glesson, F.V., Brady, S.M., Schnabel, J.A.: Edge- and detail-preserving sparse image representations for deformable registration of chest MRI and CT volumes. In: Gee, J.C., Joshi, S., Pohl, K.M., Wells, W.M., Zöllei, L. (eds.) IPMI 2013. LNCS, vol. 7917, pp. 463–474. Springer, Heidelberg (2013)
15. Hermann, S.: Evaluation of scan-line optimization for 3D medical image registration. In: CVPR 2014, pp. 1–8 (2014)
16. Liu, C., Yuen, J., Torralba, A.: SIFT flow: Dense correspondence across scenes and its applications. IEEE Pattern Anal. Mach. Intell. 33(5), 978–994 (2011)
17. 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. Comp. Meth. Prog. Bio. 98(3), 278–284 (2010)
18. Muenzing, S.E., van Ginneken, B., Viergever, M.A., Pluim, J.P.: DIRBoost–an algorithm for boosting deformable image registration: Application to lung CT intra-subject registration. Med. Imag. Anal. 18(3), 449–459 (2014)
19. Murphy, K., et al.: Evaluation of registration methods on thoracic CT: The EMPIRE10 challenge. IEEE Trans. Med. Imag. 30(11), 1901–1920 (2011)
20. Ou, Y., Sotiras, A., Paragios, N., Davatzikos, C.: DRAMMS: deformable registration via attribute matching and mutual-saliency weighting. Med. Imag. Anal. 15(4), 622–639 (2011)
21. Polzin, T., Rühaak, J., Werner, R., Strehlow, J., Heldmann, S., Handels, H., Modersitzki, J.: Combining automatic landmark detection and variational methods for lung CT registration. In: MICCAI Workshop: Pulmonary Image Analysis (2013)
22. Potesil, V., Kadir, T., Platsch, G., Brady, M.: Personalization of pictorial structures for anatomical landmark localization. In: Székely, G., Hahn, H.K. (eds.) IPMI 2011. LNCS, vol. 6801, pp. 333–345. Springer, Heidelberg (2011)
23. Taquet, M., Macq, B., Warfield, S.K.: Spatially adaptive log-euclidean polyaffine registration based on sparse matches. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011, Part II. LNCS, vol. 6892, pp. 590–597. Springer, Heidelberg (2011)
24. Veksler, O.: Stereo correspondence by dynamic programming on a tree. In: CVPR 2005, vol. 2, pp. 384–390 (2005)