



MRI Whole Heart Segmentation Using Discrete Nonlinear Registration and Fast Non-local Fusion

Mattias P. Heinrich¹(✉) and Julien Oster²

¹ Institute of Medical Informatics, University of Lübeck, Lübeck, Germany
heinrich@imi.uni-luebeck.de

² IADI, U947, Inserm, CHRU de Nancy, Vandoeuvre les Nancy, France
<http://www.mpheinrich.de>

Abstract. We present a robust and accurate method for multi-atlas segmentation of whole heart MRI scans. After preprocessing, which includes resampling to isotropic voxel sizes and cropping or padding to same dimensions, all training scans are registered linearly and nonlinearly to an unseen set of test scans. We employ the efficient discrete registration framework called *deeds* that captures large shape variations across scans, performed best in a recent registration comparison on abdominal scans and requires less than 2 min of computation time per scan. Subsequently, we perform multi-atlas label fusion using a non-local means approach with a normalised SSD metric and a fast implementation using boxfilters. Subsequently, a multi-label random walk is performed on the obtained probability maps for an edge-preserving smoothing. Without performing any domain-specific parameter tuning, we obtained a Dice accuracy of 86.0% (averaged across 7 labels) and 87.0% for the whole heart on the MRI test dataset, which is the first rank of the MICCAI 2017 challenge. The segmentations are also visually very smooth using this fully automatic method.

1 Introduction and Related Work

The automatic segmentation of patient geometry from MRI scans has numerous applications in clinical practice and research of cardiac physiology. Cardiac MRI enables an accurate estimation of the cardiac function [1], by acquiring CINE images throughout the cardiac cycle. Radiologists are then required to manually delineate the left ventricle blood cavity, for at least the end-diastole and end-systole phase and all (≈ 10) slices covering the heart. The left ventricle blood cavity volume can then be computed and the left ventricle ejection fraction can be estimated. This process is thus very cumbersome for the radiologists and automatic delineation and segmentation of the MRI scans is therefore appealing for clinicians.

We are in particular interested in deriving a robust anatomical surface model from tomographic scans to enable electro-physiological modelling and perspective electrocardiographic imaging (ECGI) [2]. ECGI consists in fusing images

(morphological models) with electrophysiological signals (hundreds of ECG signals acquired across the patient torso) in order to estimate a map of the myocardial electrical biopotentials. Applications of ECGI range from ablation procedure planning for both atrial and ventricular arrhythmias, and localisation of premature ventricular contractions. Cardiac MRI images are rarely used for ECGI, but may offer advantages over CT scans, namely dynamic images for myocardial motion estimation [3], but also offer an in-depth tissue characterization, such as ischemic scars [4] and fiber orientation with Diffusion Tensor Images [5].

The goal of the “*MM-WHS 2017: Multi-Modality Whole Heart Segmentation*” challenge held in conjunction with MICCAI 2017 is the delineation of the following seven anatomical structures in medical scans: ■ the myocardium of the left ventricle, ■ the left atrium blood cavity, ■ the left ventricle blood cavity, ■ the right atrium blood cavity, ■ the right ventricle blood cavity, ■ the ascending aorta, and ■ the pulmonary artery. The challenge offers two datasets of each 20 CT and MRI scans for training with corresponding ground truth labels and 40 scans each for testing and benchmarking of different methods. We are mainly interested in MRI scans and do not consider the CT data at all.

Numerous automatic algorithms have been proposed for heart segmentation in the past (see e.g. [6] for an overview of a recent left atrium challenge), which can be roughly divided into three classes: registration-based, model-based approaches (cf. [7]) or deep learning with convolutional neural networks (CNN). The strengths of statistical models and deep learning are their low processing time when applying the model to unseen data and a computational demand in test that is generally independent of the number of atlases. Whereas registration-based approaches require fewer labelled atlases and often achieve the highest accuracies [8]. Since, individual registrations sometimes fail, it is important to find a robust label fusion strategy, e.g. by employing patch-similarity [9]. This may, however, result in more irregular surfaces for difficult to segment subjects (see Fig. 5 of [6] for the registration approach UCL-1C). Statistical shape models or locally adaptive models can more robustly capture physiologically plausible shapes, but sometimes are inaccurate when certain rare anatomies are under-represented in the training data [7]. CNN models have only very recently been successfully addressed to cardiac image segmentation (e.g. in [10]) and excel in directly deriving clinically relevant information from cine MR (as evident from <https://www.kaggle.com/c/second-annual-data-science-bowl>). But, deep models require large training datasets, additional data augmentation strategies and substantial computing resources. Furthermore, they cannot directly encode shape or deformation priors as well as registration- or statistical-model based approaches. Newer work, therefore, focuses on incorporating shape priors [11] to improve accuracy for scans with low quality.

In this work, we aim to address the remaining challenges of MRI cardiac segmentation, which include: varying image quality, (relatively) low numbers of training scans and inconsistent image contrast. Our approach relies on a very effective discrete deformable registration approach that is combined with a specifically adapted non-local fusion strategy and probability map

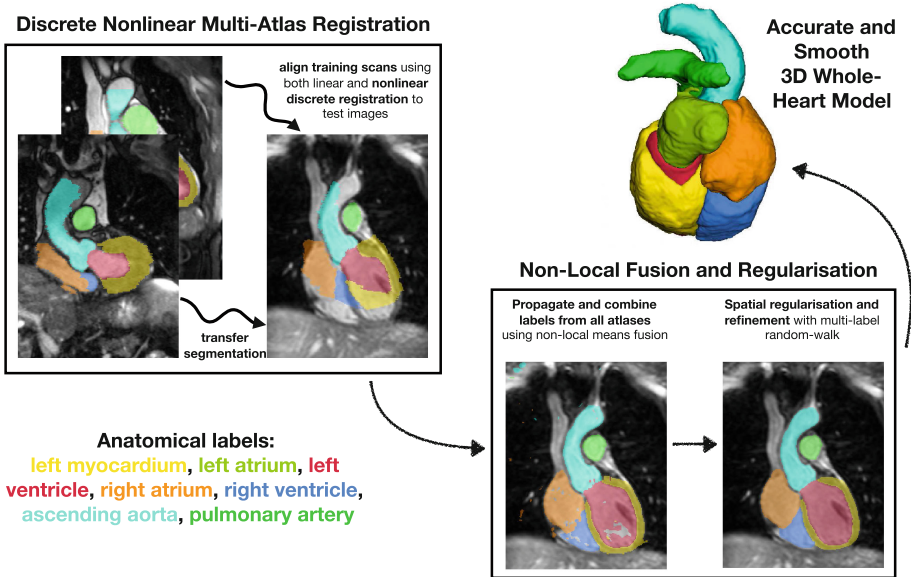


Fig. 1. Flow-chart of our proposed algorithm for accurate segmentation of the whole heart in MRI. First, all 20 training scans are linearly and nonlinearly aligned to the unseen test scan. Second, the transformed segmentations from all atlases are weighted (over a non-local search region) using their patch similarity. Finally, a spatial edge-preserving regularisation is applied to compensate for minor inconsistencies.

regularisation. The discrete registration *deeds* [12] overcomes the limitations of long-processing times for registration-based and performs substantially better than common methods using continuous optimisation e.g. ANTS, IRTK and NiftyReg. A detailed evaluation of these state-of-the-art approaches for non-brain scans was presented in [13]. A graphical overview of our approach is given in Fig. 1. Each part of the approach is then described in detail in Sect. 2 (discrete registration), Sect. 3 (non-local fusion) and Sect. 4 (multi-label random walk regularisation). A preliminary evaluation of the influence of those consecutive parts on the segmentation accuracy and robustness of surface estimation is presented alongside the description of the methods and discussed in Sect. 5.

2 Discrete Registration

We resampled the scans to an isotropic resolution of $\approx 1.0\text{mm}^3$ and padded or cropped them to have same dimensions of $288 \times 288 \times 180$ and enable pairwise registration of all images. We employ the publicly available discrete deformable registration tool *deeds* [12] based on the implementation found at github.com/mattiaspaul/deedsBCV. The choice is motivated by its very high computational efficiency and the encouraging accuracies for inter-subject abdominal registration in [13]. An affine pre-registration is performed using discrete

block-matching on four scale levels as defined per default. Subsequently the non-linear registration is called using with the following (default) parameters that were found to be very suitable for abdominal registration in [13]: number of displacement steps $l_{\max} = [8, 7, 6, 5, 4]$, quantisation steps $q = [5, 4, 3, 2, 1]$, and B-spline grid spacings of $[8, 7, 6, 5, 4]$ voxels. Using a dense displacement sampling covering a large range of potential displacements together with a combinatorial optimisation based on dynamic programming, enables this approach to capture strong shape deformations across patients. The self-similarity context (SSC) [14] was used as a similarity metric with a weighting of $\alpha = 1.6$. Initially developed for multi-modal registration, SSC enables contrast invariance, which is particularly beneficial for MRI scans and focuses the alignment on image edges. The computations are performed in less than two minutes per scan-pair (and many registrations can be run in parallel on a multi-core system). We store both the deformed MRI scans and the transformed segmentations for the following label fusion.

3 Non-local Label Fusion

Label fusion of multiple atlases is a common approach in multi-atlas segmentation [9, 15, 16]. The rationale behind this is to exclude poorly registered atlases and refine the segmentation boundary. Due to the regularisation constraint within deformable registration, the deformation of the anatomies is restricted and certain parts are therefore often not perfectly aligned. One could directly incorporate the displacement uncertainty into a multi-atlas fusion [17], but usually a sequential step that performs an (unregularised) non-local search is simpler to implement.

We follow the approach of [18] and calculate non-local weights, independently for each patch (and atlas scan), based on a similarity metric, the normalised sum of squared differences (NSSD), between the central patch of the target image and a patch within the non-local search region $\mathbf{n} \in \mathcal{N}$ of the atlas scan. Let \mathbf{x} be a 3D location in an image I and $X(\mathbf{x})$ an intensity patch in I with a patch radius r , which defines $\Omega_{\mathbf{x}}$ with a size of $R = (2r + 1)^3$. The corresponding patch at \mathbf{n} within the search region of the atlas scan is defined as $Y(\mathbf{x} + \mathbf{n})$. To reduce the impact of local contrast variations (cf. [15]), we subtract the mean of each patch and divide by their standard deviation to obtain a normalised patch $X'(\mathbf{x})$, $Y'(\mathbf{x})$, using:

$$X'(\mathbf{x}) = \frac{X(\mathbf{x}) - \mu_X}{\sigma_X} \text{ with } \mu_X = \frac{1}{R} \sum_{\mathbf{y} \in \Omega_{\mathbf{x}}} X(\mathbf{y}) \text{ and } \sigma_X = \sqrt{\frac{1}{R} \sum_{\mathbf{y} \in \Omega_{\mathbf{x}}} (X(\mathbf{y}) - \mu_X)^2} \quad (1)$$

In practice, we use a fast implementation to calculate local means and standard deviations using the boxfilter approach of [19]. It can be shown that the computation time is then independent of the patch size R and thus substantially lower than in [15, 18]. To calculate the NSSD between all patches in two scans, only one additional pointwise multiplication between the intensity images followed by

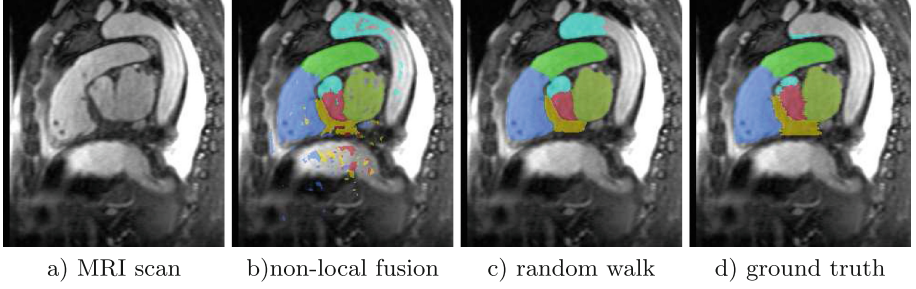


Fig. 2. Visual comparison of segmentation results. ■ myocardium of the left ventricle, ■ left atrium, ■ left ventricle, ■ right atrium, ■ right ventricle, ■ ascending aorta, and ■ pulmonary artery

another boxfilter is required. The value at each location can then be computed using binomial expansion as (the sum being implemented as filter):

$$\text{NSSD}(\mathbf{x}, \mathbf{n}) = 2R - \frac{2(\sum_{\mathbf{y} \in \Omega_{\mathbf{x}}} (X(\mathbf{y}) \cdot Y(\mathbf{y} + \mathbf{n})) - \mu_X \mu_Y R)}{\sigma_X \sigma_Y} \quad (2)$$

We shift the atlas image by each of the $\mathbf{n} \in \mathcal{N}$ displacements within the non-local search region and achieve around 25 boxfilter operations (for scans with 15 million voxels) per second using advanced vector instructions. Finally the weights are computed by the exponential term $w(\mathbf{x}, \mathbf{n}) = \exp(-\text{NSSD}(\mathbf{x}, \mathbf{n}) / (2\sigma_{nlm}^2))$, with $\sigma_{nlm} = 0.25$. The size of the local patch and search region are empirically chosen to be both $7 \times 7 \times 7$ voxels.

We anticipate that considering fusion weights jointly as proposed in [15] would lead to better results. However, this algorithm requires much longer computation times and, due to limited processing resources, we would have been unable to finish these experiments in time. Alternatively, the SIMPLE approach for atlas pre-selection as e.g. employed in [20, 21] could further reduce the computational demand and negative influence of poorly aligned atlases.

4 Multi-label Random Walk Regularisation

The non-local fusion helps to correct minor registration inaccuracies, but its normalised sum of squared difference metric is also more sensitive to image edges. Therefore, the output may yield some isolated misclassifications or holes in solid objects, hence the segmentation is not necessarily spatially consistent. To counter these effects it is therefore beneficial to regularise the probability maps $P^c(\mathbf{x})$ spatially that were obtained using the label fusion in the previous section. We aim to obtain smooth maps $P(\mathbf{x})_{reg}^c$ for every label $c \in C$, while adhering to image edges and employ the multi-label random walk [22] to minimise the objective $E(P(\mathbf{x})_{reg}^c)$:

$$\sum_{\mathbf{x}} \frac{1}{2} (P(\mathbf{x})^c - P(\mathbf{x})_{reg}^c)^2 + \sum_{\mathbf{x}} \frac{\lambda}{2} \|\nabla P(\mathbf{x})_{reg}^c\|^2 \quad (3)$$

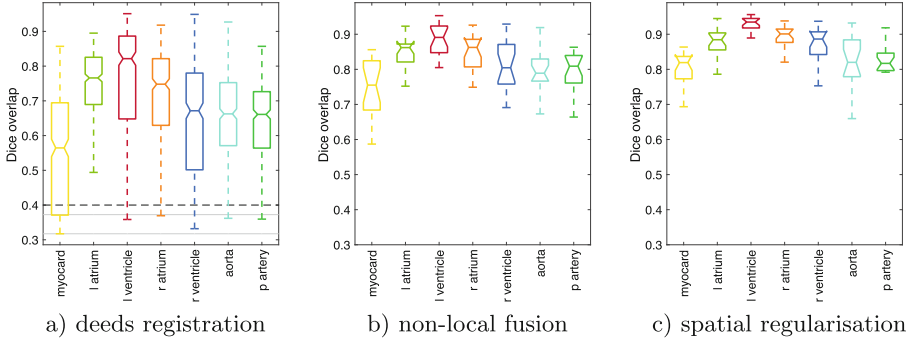


Fig. 3. Numerical results for Dice overlap, of the three different steps in our multi-atlas segmentation framework. Note that the first boxplot (a) consists of the results for single-atlas propagation (380 nonlinear registrations) and has a larger range due to sporadic outliers. It is evident that both non-local multi-atlas fusion and random walk regularisation can substantially improve the registration outcome.

The implementation follows [23] and is publicly available. To preserve edges, the gradient of the probability map is weighted by $w_j = \exp(-(I(\mathbf{x}_i) - I(\mathbf{x}_j))^2 / (2\sigma_w^2))$. We empirically chose $\sigma_w = 50$ and a regularisation weight of $\lambda = 50$ as suggested in [23] for multi-organ segmentation. This weight depends on the differences of image intensities I of \mathbf{x}_i and its neighbouring voxels $\mathbf{x}_j \in \mathcal{N}_i$. Alternatively, optimisation techniques such as fully-connected conditional random fields (CRF) have been proposed in the literature, which are particularly powerful when enforcing long-range spatial clues [24]. However, for the segmentation task at hand, we found the random walk algorithms provided visually very good results (see Fig. 2) with computation times as low as a few seconds per scan.

Numerical results for segmentation accuracy: We use the Dice overlap $D = 2|A \cap M| / (|A| + |M|)$ to measure the accuracy of our automatic segmentations A compared to the provided manual label images M . For 20 training scans of the challenge, we performed a leave-one-out validation. The distribution of

Table 1. Numerical results of the label fusion (NLM) and spatial regularisation in Dice overlap in % and mean surface distance (MSD) in mm, obtained using leave-one-out cross validation on the MRI training dataset and the submitted results for the test set of the challenge. The spatial random walk (RW) regularisation improves all labels. On the test dataset only the complete method including RW is employed.

Label	myocard	l atrium	l ventricle	r atrium	r ventricle	aorta	p artery	Average
NLM train	74.9%	84.3%	86.7%	84.8%	79.5%	78.8%	79.0%	81.1%
RW train	79.6%	87.8%	92.0%	89.3%	85.8%	81.7%	80.3%	85.2%
Dice test	78.1%	88.6%	91.8%	87.3%	87.1%	87.8%	80.4%	86.0%
MSD test	1.46	1.38	1.32	1.80	1.74	1.23	1.64	1.51

scores for the discrete registration (single-atlas propagation), the non-local label fusion and the spatial regularisation are presented in Fig. 3. The average over all 20 training cases is additionally reported in Table 1.

5 Discussion and Conclusion

The obtained segmentations are visually very convincing, in particular for the left ventricle and atrium. By using a relatively strong regularisation, we can obtain smooth surfaces that are important for mesh generation and motion or electrophysiological modelling. Depending on the available number of processing cores, the complete multi-atlas fusion for one test scan could be accomplished in less than 5 min. The Dice overlap for all 7 labels is given in Table 1. Averaged across all structures, we achieve a Dice of 85.2% (training) and 86.0% (test dataset), which is the first rank for the “*MM-WHS 2017: Multi-Modality Whole Heart Segmentation*” challenge. It is, however, slightly below the current state-of-the-art that reached 88.3% on the same 7 structures [9]. We obtained a relatively low mean surface result across all structures of 1.51 mm.

We argue that part of the remaining inaccuracies are due to ambiguities for the definition of the ascending aorta and pulmonary artery (see Fig. 2), which are hard to consistently segment manually (leading to inter-observer errors of up to 24% in [9]). Furthermore, the papillary muscles are often visually hard to distinguish from the myocardium, which leads to further uncertainties.

Prospectively, a more elaborated parameter tuning and adaptation to this specific domain of registration, label fusion and regularisation could lead to further improvements. In addition, the automatic cropping of the MRI scans sometimes led to a field-of-view that did not completely cover the whole heart and this should be addressed in the future. Two areas of particular interest would be a better delineation of the myocardium and the exact separation between left and right atria e.g. by employing a shape prior during label fusion [25] or considering contextual features [26].

Acknowledgements. We would like to thank the organisers of the MM-WHS 2017 for providing this rich new dataset to the public, which enables the evaluation of new algorithms for the problem of detailed 3D heart segmentation.

References

1. Grothues, F., Smith, G.C., Moon, J.C., Bellenger, N.G., Collins, P., Klein, H.U., Pennell, D.J.: Comparison of interstudy reproducibility of cardiovascular magnetic resonance with 2D echocardiography in normal subjects and in patients with heart failure or left ventricular hypertrophy. *Am. J. Cardiol.* **90**(1), 29–34 (2002)
2. Ramanathan, C., Ghanem, R.N., Jia, P., Ryu, K., Rudy, Y.: Noninvasive electrocardiographic imaging for cardiac electrophysiology and arrhythmia. *Nat. Med.* **10**(4), 422–428 (2004)

3. Vuissoz, P.A., Odille, F., Fernandez, B., Lohezic, M., Benhadid, A., Mandry, D., Felblinger, J.: Free-breathing imaging of the heart using 2D cine-GRICS with assessment of ventricular volumes and function. *J. Magn. Reson Imaging* **35**(2), 340–351 (2012)
4. Nazarian, S., Bluemke, D.A., Lardo, A.C., Zviman, M.M., Watkins, S.P., Dickfeld, T.L., Meininger, G.R., Roguin, A., Calkins, H., Tomaselli, G.F., et al.: Magnetic resonance assessment of the substrate for inducible ventricular tachycardia in non-ischemic cardiomyopathy. *Circulation* **112**(18), 2821–2825 (2005)
5. Nielles-Vallespin, S., Mekkaoui, C., Gatehouse, P., Reese, T.G., Keegan, J., Ferreira, P.F., Collins, S., Speier, P., Feiweier, T., Silva, R., et al.: In vivo diffusion tensor MRI of the human heart: reproducibility of breath-hold and navigator-based approaches. *Magn. Reson. Med.* **70**(2), 454–465 (2013)
6. Tobon-Gomez, C., Geers, A.J., Peters, J., Weese, J., Pinto, K., Karim, R., Ammar, M., Daoudi, A., Margeta, J., Sandoval, Z., et al.: Benchmark for algorithms segmenting the left atrium from 3D CT and MRI datasets. *IEEE Trans. Med. Imag.* **34**(7), 1460–1473 (2015)
7. Kutra, D., Saalbach, A., Lehmann, H., Groth, A., Dries, S.P.M., Krueger, M.W., Dössel, O., Weese, J.: Automatic multi-model-based segmentation of the left atrium in cardiac MRI scans. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) *MICCAI 2012. LNCS*, vol. 7511, pp. 1–8. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33418-4_1
8. Zhuang, X., Rhode, K.S., Razavi, R.S., Hawkes, D.J., Ourselin, S.: A registration-based propagation framework for automatic whole heart segmentation of cardiac MRI. *IEEE Trans. Med. Imag.* **29**(9), 1612–1625 (2010)
9. Zhuang, X., Shen, J.: Multi-scale patch and multi-modality atlases for whole heart segmentation of MRI. *Med. Imag. Anal.* **31**, 77–87 (2016)
10. Wolterink, J.M., Leiner, T., Viergever, M.A., Išgum, I.: Dilated convolutional neural networks for cardiovascular MR segmentation in congenital heart disease. In: Zuluaga, M.A., Bhatia, K., Kainz, B., Moghari, M.H., Pace, D.F. (eds.) *RAMBO/HVSMR -2016. LNCS*, vol. 10129, pp. 95–102. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-52280-7_9
11. Oktay, O., Ferrante, E., Kamnitsas, K., Heinrich, M., Bai, W., Caballero, J., Guerrero, R., Cook, S., de Marvao, A., O'Regan, D., et al.: Anatomically constrained neural networks (ACNN): Application to cardiac image enhancement and segmentation. *arXiv preprint arXiv:1705.08302* (2017)
12. Heinrich, M., Jenkinson, M., Brady, J., Schnabel, J.: MRF-based deformable registration and ventilation estimation of lung CT. *IEEE Trans. Med. Imag.* **32**(7), 1239–1248 (2013)
13. Xu, Z., Lee, C., Heinrich, M., Modat, M., Rueckert, D., Ourselin, S., Abramson, R., Landman, B.: Evaluation of six registration methods for the human abdomen on clinically acquired CT. *IEEE Trans. Biomed. Eng.* 1–10 (2016)
14. 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. LNCS*, vol. 8149, pp. 187–194. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40811-3_24
15. Wang, H., Suh, J.W., Das, S.R., Pluta, J.B., Craige, C., Yushkevich, P.A.: Multi-atlas segmentation with joint label fusion. *IEEE Trans. Patt. Anal. Mach. Intell.* **35**(3), 611–623 (2013)
16. Asman, A.J., Landman, B.A.: Non-local statistical label fusion for multi-atlas segmentation. *Med. Imag. Anal.* **17**(2), 194–208 (2013)

17. Heinrich, M.P., Simpson, I., Papież, B., Brady, J., Schnabel, J.: Deformable image registration by combining uncertainty estimates from supervoxel belief propagation. *Med. Imag. Anal.* **27**, 57–71 (2016)
18. Coupé, P., Manjón, J.V., Fonov, V., Pruessner, J., Robles, M., Collins, D.L.: Patch-based segmentation using expert priors: application to hippocampus and ventricle segmentation. *NeuroImage* **54**(2), 940–954 (2011)
19. Heinrich, M.P., Papież, B.W., Schnabel, J.A., Handels, H.: Non-parametric discrete registration with convex optimisation. In: Ourselin, S., Modat, M. (eds.) *WBIR 2014. LNCS*, vol. 8545, pp. 51–61. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-08554-8_6
20. Langerak, T., Van Der Heide, U., Kotte, A., Viergever, M., Van Vulpen, M., Pluim, J.: Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (SIMPLE). *IEEE Trans. Med. Imag.* **29**(12), 2000–2008 (2010)
21. Xu, Z., Asman, A.J., Shanahan, P.L., Abramson, R.G., Landman, B.A.: SIMPLE is a good idea (and better with context learning). In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) *MICCAI 2014. LNCS*, vol. 8673, pp. 364–371. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10404-1_46
22. Grady, L.: Multilabel random walker image segmentation using prior models. In: *CVPR*, pp. 763–770 (2005)
23. Heinrich, M.P., Blendowski, M.: Multi-organ segmentation using vantage point forests and binary context features. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) *MICCAI 2016. LNCS*, vol. 9901, pp. 598–606. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_69
24. Krähenbühl, P., Koltun, V.: Efficient inference in fully connected CRFs with Gaussian edge potentials. In: *Proceedings of NIPS*, pp. 2–9 (2011)
25. Oguz, I., Kashyap, S., Wang, H., Yushkevich, P., Sonka, M.: Globally optimal label fusion with shape priors. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) *MICCAI 2016. LNCS*, vol. 9901, pp. 538–546. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_62
26. Bai, W., Shi, W., Ledig, C., Rueckert, D.: Multi-atlas segmentation with augmented features for cardiac MR images. *Med. Imag. Anal.* **19**(1), 98–109 (2015)