Multi-Atlas Segmentation Propagation with Uncertainty Estimates from Belief Propagation

Mattias P. Heinrich[1], Ivor J.A. Simpson[2],
Sir Michael Brady[3], and Julia A. Schnabel[1]

[1] Institute of Biomedical Engineering,
Department of Engineering, University of Oxford, UK
[2] Centre for Medical Image Computing, University College London, UK
[3] Department of Oncology, University of Oxford, UK

Abstract. Accurate automatic segmentation of medical scans is of great value to study developmental-, or pathological differences of the volume and/or shape of anatomical structures across subjects. Manual delineations are often too time consuming for larger datasets. Registration-based segmentation propagation is an important approach for medical image segmentation. The difficulty and reliability of registering a given annotated atlas to a subject will vary due to morphological variability. However, few registration algorithms explicitly deal with the spatial uncertainty of an obtained transformation. We apply our recent method to incorporate uncertainty estimates for improved segmentation accuracy to two of the SATA challenge datasets. Based on an efficient discrete optimisation framework, marginal distributions for a large number of local displacements are calculated and converted into probabilities. We employ the self-similarity based metric MIND to achieve robustness to various noise distributions and imaging sequences. Our method achieves good performance with very low computational cost.

Keywords: marginal distributions, discrete optimisation, SATA

1 Introduction

The MICCAI Challenge Workshop on Segmentation: Algorithms, Theory and Applications (**SATA**) aims at studying general purpose methods for atlas-based segmentation. The focus lies on approaches, which can reliably segment images acquired with a diversity of imaging sequences and of different anatomical locations for clinical practice. We employ our recent method [6] that incorporates uncertainty estimates, which are obtained using an efficient discrete optimisation framework, to improve segmentation accuracy. Our approach combines registration and atlas fusion in a single framework, therefore none of the registered labels provided by the challenge using the ANTs package [1] are used. A review of related work that includes the quantification of uncertainty distribution and a more detailed explanation of our approach is given in [6]. In the next section, we present an overview of our method and a description of the modifications made

to adapt this approach to more general segmentation problems. The choice of parameters for the experimental setup and the results are discussed thereafter.

2 Registration Model, Optimisation and Results

Similarity metric: In contrast to our technical workshop paper [6], we use a more robust similarity metric, the modality independent neighbourhood descriptor (MIND) [3]. A quantised descriptor is calculated for each voxel in both scans, which describes its self-similarity context [4]. For a patch centred at $\mathbf{x}$, the self-similarity descriptor is given by: $\mathcal{S}(I, \mathbf{x}, \mathbf{y}) = \exp(-\frac{SSD(\mathbf{x},\mathbf{y})}{\sigma^2})$, for $\mathbf{x}, \mathbf{y} \in \mathcal{N}$. Where σ^2 is a local noise estimate and $\mathbf{y}$ defines the centre location of a patch within a certain neighbourhood $\mathcal{N}$. This descriptor is nearly independent of locally varying contrast (such as MRI bias fields) and robust to noise. Distance evaluations can be efficiently performed using Hamming weights.

Linear registration: In order to initialise the deformable registration, we align each atlas image to the target scan with an affine transformation. Here, a block-matching approach, similar to [8], is adapted based on the MIND similarity metric using block sizes of 7^3 and 5^3 voxels.

Deformable registration: We use a parametric deformable registration model with a first order B-spline transformation grid. The space of displacements $\mathbf{u}_p$ is discretised with a quantisation step of d yielding a three dimensional displacement space of $\mathcal{L} = \{0, \pm d, \ldots, \pm \max(|\mathbf{u}|)\}^3$. For more details see [5]. Smooth transformations can be ensured by avoiding large differences in displacement vectors for neighbouring nodes with a penalty $\mathcal{R}$, which is weighted with the parameter λ. Here, absolute differences (TV regularisation) are used.

Inference and optimisation: We use belief propagation on a relaxed graph structure of a spanning tree [2] to find approximately globally optimal marginals for each displacement for each node. A message vector $\mathbf{m}_p$ containing the cost for the best displacement $\mathbf{u}_p^*$, given the displacement $\mathbf{u}_q$ of its parent node q and the messages of its children c, can be found by evaluating: $\mathbf{m}_p(\mathbf{u}_q) = \min_{\mathbf{u}_p} \left(\mathcal{S}(\mathbf{x}_p, \mathbf{u}_p) + \lambda\mathcal{R}(\mathbf{u}_p, \mathbf{u}_q) + \sum_c \mathbf{m}_c(\mathbf{u}_p)\right)$. Only two passes of belief propagation are required to obtain the marginal distributions. Multiple trees are chosen, the inference is performed independently for each tree, and the resulting marginal distributions are averaged for each node.

Uncertainty estimates: The estimation of local uncertainties allows us to overcome some of the difficulties of registration-based segmentation, e.g.: missing one-to-one correspondences, mismatch of structures and image acquisition noise and artefacts. The marginal distributions enable the use of many different probable mappings, which can be weighted by their probability. The probability $p(\mathbf{x}_i, \mathbf{u}_i)$ for each voxel $i \in \Omega$ in the image and each displacement $\mathbf{u}_i \in \mathcal{L}$ can be directly obtained from the min-marginal energies $E(\mathbf{x}, \mathbf{u})$[7]: $p(\mathbf{x}_i, \mathbf{u}_i) = \exp(-\frac{\beta \cdot E(\mathbf{x}_i, \mathbf{u}_i)}{\mathrm{std}(E(\mathbf{x}_j, \mathbf{u}_j))})$. The parameter β varies the spread of the probability estimates: low values result in smoother distributions, and high values in narrower peaks. Finding the best segmentation label is possible by summing the

Table 1. Trained re-weighting of labels to avoid shrinkage of small structures. Background is weighted with 1. Label # is given in increments rather than actual values.

label #	1	2	3	4	5	6	7	8	9	10	11	12	13	14
diencephalon	1.31	1.24	1.30	1.28	1.01	1.01	1.24	1.23	1.21	1.18	1.11	1.11	1.09	1.06
canine legs	2.00	1.00	1.00	0.95	1.05	1.20	1.00							

probabilities for each possible segmentation label (from the atlas) and choosing the arg max of all labels.

In our approach incorporating multiple atlases is straightforward once the marginals are estimated between the target scan and each labelled moving scan. The votes from each atlas are aggregated for each voxel and the segmentation label with highest probability is chosen thereafter.

Experimental details: The range of displacements ($\max(|\mathbf{u}|)$) was defined to be 5 voxels (with unit spacing) in each direction and dimension. A control-point grid with spacing of 3 voxels was used for the diencephalon (mid-brain) data and 5 voxels for the canine leg scans, yielding roughly 1.5×10^7 degrees of freedom per registration. The weighting of the regularisation was empirically chosen to be $\lambda = 4$. The parameter $\beta = 5$ had been tuned on a different task in [6] and was left unaltered. The number of spanning trees was chosen to be 5.

It has been previously found that imperfections in label voting approaches can result in shrinkage of smaller segmentation labels. To reduce this adverse effect, we follow a similar approach to [10] and introduce label specific weights for the voxel-wise aggregation of probabilities. We used a simple tuning process, based on a subset of the training data, to find weights that result in an approximate preservation of the average volumes of all labels, see Table 1.

Implementation: Our presented method has a comparatively very low computational complexity, due to the efficient calculation of both similarity metric and inference of regularisation. For each registration on a single CPU core less than 30 seconds are required for the diencephalon scans and less than a minute for the canine legs data. It is also straightforward to make use of parallel architectures, both within each registration and when using multiple atlases.

Results: Segmentation accuracy is measured with the DICE similarity coefficient (DSC). Using the proposed probabilistic label weighting scheme, the segmentation accuracy on the challenge testing data was on average 84.02 % for the diencephalon data and 68.71 % for the canine scans. This results are overall good, but slightly short of the leading competitor using the registrations of [1].

3 Discussion and Conclusion

We applied our recently proposed framework, which incorporates uncertainty estimates of displacement parameters, on the **SATA** challenge data. Compared to [6], we have added an affine pre-registration and replaced the similarity metric with a more robust variant. The same settings were used for both datasets (except for the spacing of control points, which depends on the image dimensions),

and only the weighting of the regularisation term has been specifically tuned on the training data. This indicates the general applicability of our approach to a variety of clinical data. We left the evaluation of our method on the cardiac atlas data for future work, because it is four-dimensional and therefore not straightforward to use within our framework. Further improvements could be achieved for the canine data by including both MRI sequences (we only used T2w).

In particular for the more challenging canine dataset, we found that the reliability of segmentation propagations of some of the training atlases is very low. Therefore, the uniform weighting of training data used in our experiments may not be appropriate. A simple choice or weighting could be included based on the DSC of each individual segmentation propagation with the initial label fusion result, in order to further improve accuracy. More sophisticated label fusion approaches, e.g. [9], will also be considered in the future.

Acknowledgements: We would like to thank EPSRC and Cancer Research UK for funding this work within the Oxford Cancer Imaging Centre.

References

1. Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C.: Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med. Imag. Anal., 12(1) pp. 26–41 (2008)
2. Felzenszwalb, P., Huttenlocher, D.: Pictorial structures for object recognition International Journal of Computer Vision 61(1), pp. 55–79 (2005)
3. Heinrich, M.P., Jenkinson, M., Bhushan, M., Matin, T., Gleeson, F., Brady, M., Schnabel, J.A.: MIND: Modality independent neighbourhood descriptor for multimodal deformable registration. Med. Imag. Anal. 16(7) pp. 1423-1435 (2012)
4. Heinrich, M.P., Jenkinson, M., Papiez, B.W., Brady, M., Schnabel, J.A.: Towards Realtime Multimodal Fusion for Image- Guided Interventions using Self-Similarities In: Mori, K., Sakuma, I., MICCAI 2013, LNCS, in press, (2013)
5. Heinrich, M.P., Jenkinson, M. Brady, M., Schnabel, J.A.: MRF-based Deformable Registration and Ventilation Estimation of Lung CT. IEEE Trans. Med. Imag., 32(7) pp. 1239–48, (2013)
6. Heinrich, M.P., Simpson, I.J.A., Jenkinson, M., Brady, M., Schnabel, J.A.: Uncertainty Estimates for Improved Accuracy of Registration-Based Segmentation Propagation using Discrete Optimisation. In: MICCAI Challenge Workshop on Segmentation: Algorithms, Theory and Applications, (2013)
7. Kohli, P., Torr, P.H.S.: Measuring uncertainty in graph cut solutions. Computer Vision and Image Understanding 112(1), pp. 30–38 (2008)
8. Ourselin, S., Roche, A., Prima, S., Ayache, N.: Block matching: A general framework to improve robustness of rigid registration of medical images. In: Delp, S.L., DiGoia, A.M, Jaramaz, B., MICCAI 2000, LNCS, pp. 557–566, (2000)
9. Sabuncu, M.R., Yeo, B.T.T., van Leemput, K., Fischl, B., Golland, P.: A generative model for image segmentation based on label fusion. IEEE Trans. Med. Imag. 29(10) 1714–1729 (2010)
10. Smith, B.M., Yang, L., Brandt, J., Lin, Z., Yang, J.: Exemplar-Based Face Parsing IEEE CVPR, pp. 1–8, (2013)