Non-rigid image registration through efficient discrete optimization

Mattias P. Heinrich[1]
mattias.heinrich@eng.ox.ac.uk
Mark Jenkinson[2]

Sir J. Michael Brady[3]

Julia A. Schnabel[1]

[1] Institute of Biomedical Engineering
Department of Engineering Science
University of Oxford, UK

[2] Oxford Centre of FMRI of the Brain
University of Oxford, UK

[3] Department of Radiation Oncology
and Biology, University of Oxford, UK

Abstract

Discrete optimization offers attractive capabilities for non-rigid medical image registration. Based on previous work on stereo vision with large displacement spaces, we propose a novel non-rigid registration method. The main idea to minimize the computational burden to constant time complexity is to reduce the search space hierarchically from coarse to fine levels. The large search space helps us to avoid local minima in the optimization scheme. We apply our method to a number of challenging registration tasks. Our approach compares very favourably against state-of-the-art registration methods.

1 Introduction

Non-rigid image registration is a key technique in medical image analysis. It is used widely for motion tracking, atlas-based segmentation and multimodal fusion, during radiation treatment for image guided radiotherapy (IGRT), and for improving scan quality by compensating for motion. Generally, the non-rigid alignment of two images has several million degrees of freedom and leads to a non-convex optimization problem. In the past, the great majority of techniques has focused on formulations based on continuous optimization. Generally, a cost function consisting of a dissimilarity term and a regularization penalty is minimized using gradient descent. However, continuous optimization only converges to a global optimum for a convex problem, so that in practice local minima often deteriorate the accuracy of the alignment. Discrete optimization, in contrast, enables search over a larger space of possible solutions and therefore minima, closer to the global optimum, can be found. The increased complexity of the optimization due to the extended search space has previously discouraged wider use in image analysis. However, recent advances in the mathematical foundation and programming techniques for more efficient optimization have spurred the application of discrete optimization for many computer vision problems.

2 Method

The main concept of discrete optimization is to formulate a problem on a Markov random field (MRF), which is an undirected graphical model. The corresponding graph consists of a set of nodes (which correspond to a set of pixels $p \in P$ in an image) and edges N (which connect neighbouring pixels). Each random variable corresponds to a node and takes values from a set of labels $f_p \in \mathscr{L}$. Labels correspond to tissue classes in segmentation, intensities in restoration, disparities in stereo vision or displacements in registration. The optimization process determines a probability for every label assignment. We are interested in finding the label with maximum posterior probability (MAP). This problem can be converted to an equivalent problem of energy minimization, given by the function:

$$E(f) = \sum_{p \in P} S_p(f_p) + \sum_{(p,q) \in N} R(f_p, f_q) \tag{1}$$

In image registration, $S_p(f_p)$ is the cost of the similarity term for assigning label f_p to pixel p. $R(f_p, f_q)$ measures the pairwise cost of assigning labels f_p and f_q to two neighbouring pixels, and is equivalent to a regularization term. To perform the inference on the MRF, we adopt an efficient implementation of the max-product belief propagation (BP) algorithm. A comparative study of other message passing schemes can be found in [5]. The BP algorithm updates and passes messages iteratively for all nodes in parallel. Initially, for each node a message vector $m^0_{p \to q}$ with a length given by the number of labels is set to zero. The update at time t is found by calculating:

$$m^t_{p \to q}(f_q) = min_{f_p}(S_p(f_p) + R(f_p, f_q) + \sum_{s \in N(p) \setminus q} m^{t-1}_{s \to p}(f_p)) \tag{2}$$

The label, which minimizes the final belief vector $b_q(f_q) = S_q(f_q) + \sum_{p \in N(q)} m^T_{p \to q}(f_q)$ after T iterations, is selected for each node. In [2] a hierarchical message passing scheme is introduced: to achieve faster convergence, messages are only calculated and passed on for groups of neighbouring nodes, and the size of the neighbourhood is reduced in subsequent steps. This reduces the computational complexity to $\mathscr{O}(P\mathscr{L}^2)$. However, the memory required is also proportional to $P \cdot \mathscr{L}$ and thus for a medical application impractical.

We adapt a recent approach called *constant space belief propagation* (CSBP), which drastically reduces complexity by also reducing the search space hierarchically [6]. For the coarsest level (largest neighbourhood of grouped nodes) the similarity cost is calculated for each pixel and label, aggregated within the neighbourhood and the k labels with the smallest corresponding cost are selected. This search space reduction is performed at the beginning of each level. Using this reduction, the number of possible labels $\mathscr{L}$ is divided by $2^{d/2}$ (image dimension d). Since the number of individual nodes increases by a factor of 2^d from coarse to a fine levels, the overall complexity remains constant at $\mathscr{O}(P)$ for all levels. [1]

In [3] a different non-rigid registration approach called *drop* is presented. It is based on discrete MRF labeling and applied to 3D medical image registration. The large search space is reduced using a multiresolution and warping scheme. This is different from our approach, where all similarity cost calculations are performed on the original image resolution and small features are not lost, even for large deformations (this is a particular problem in pulmonary image registration). Another main advantage of discrete optimization is the fact that no derivatives of similarity term and regularizer are needed. Therefore a much larger

[1] An implementation of our registration framework can be found at http://users.ox.ac.uk/~shil3388

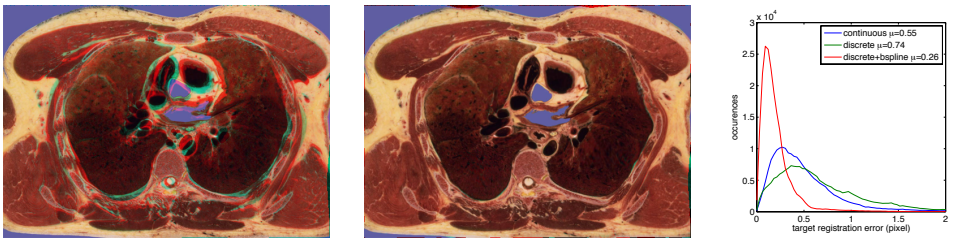

Figure 1: Example of synthetic 2D registration experiment. Left: Non-aligned colour channels of cryosection. Middle: Registered using our approach. Right: The histogram of registration errors demonstrates how the quantization error of the discrete registration (mode ≈ 0.5) can be efficiently reduced using a smoothing spline.

	before reg.		continuous			drop			csbp			csbp+bspline		
	AAE	TRE	AAE	TRE	time	AAE	TRE	time	AAE	TRE	time	AAE	TRE	time
singlemodal A														
SSD	61.3 (23.1)	3.14 (2.13)	2.78 (5.40)	0.18 (0.22)	11	**3.14 (6.52)**	**0.14 (0.21)**	8	12.3 (13.9)	0.67 (0.49)	20.7	3.52(6.07)	0.15 (0.17)	22.4
singlemodal B														
SSD	71.0 (21.3)	6.44 (4.51)	6.17 (12.4)	1.08 (2.05)	11	**2.34 (6.52)**	**0.15 (0.25)**	8.3	8.34 (12.0)	0.70 (0.50)	20.7	2.57 (5.47)	0.17 (0.18)	22.4
multimodal A														
NCC	59.7 (24.6)	3.18 (2.37)	11.6 (16.1)	0.62 (0.79)	25.4	9.20 (15.4)	0.45 (0.72)	9	10.5 (11.2)	0.58 (0.42)	45.8	**4.48 (6.15)**	**0.22 (0.20)**	47.5
GO	59.7 (24.6)	3.18 (2.37)	18.9 (24.1)	1.18 (1.64)	13.2	12.3 (23.4)	0.79 (1.53)	9	13.6 (14.6)	0.75 (0.57)	15.9	6.89 (10.0)	0.35 (0.36)	17.6
multimodal B														
NCC	71.0 (21.3)	6.44 (4.51)	8.08 (12.57)	0.92 (1.19)	25.4	12.7 (21.5)	1.42 (2.28)	9.9	10.03(15.9)	1.05 (1.41)	45.8	**5.04 (11.6)**	**0.52 (1.14)**	47.5
GO	71.0 (21.3)	6.44 (4.51)	12.0 (18.3)	1.59 (2.03)	13.2	-	-		12.6 (19.1)	1.44 (1.93)	15.9	9.26 (17.4)	1.03 (1.79)	17.6

Figure 2: Comparative overview of registration accuracy, measured in average angular error (AAE) in degrees and target registration error (TRE) in pixels. Computation times are given in seconds. Best results per category are typed in bold letters. (Note: We could not find a setting for drop to obtain a good solution for target B using GO.)

variety of metrics can be used without approximations (note that the widely used L1 norm is not differentiable). Additionally, truncated norms can be used without further modification, which allows to deal with outliers (such as noise or missing data) or discontinuities in the deformation field. A disadvantage of discrete optimization frameworks is the lack of subpixel accuracy. We overcome this by fitting a smoothing spline to the obtained discrete vector field, using the `csaps` function of the Matlab Curve Fitting Toolbox, with a weighting parameter $p = 0.001$. This leads to a reduction of the mode of the target registration error (TRE) (Fig. 1 c). Alternatively, a continuous optimization step could be followed up.

3 Experiments and results

We tested our new registration algorithm on both synthetic datasets with known deformations and on expert validated clinical 3D data. We compare our constant space belief propagation *csbp* implementation against the popular *drop*[2] algorithm and a continuous Gauss-Newton optimization framework [4]. For the synthetic registration experiments we deform the target image of an aligned pair of a cryosection slices of the chest using random control point displacements on a uniform B-spline grid. We use a maximum displacement of 24 and 48 pixels for two target images A and B respectively. For multimodal registration tasks the

[2]publicly available for download at www.mrf-registration.net

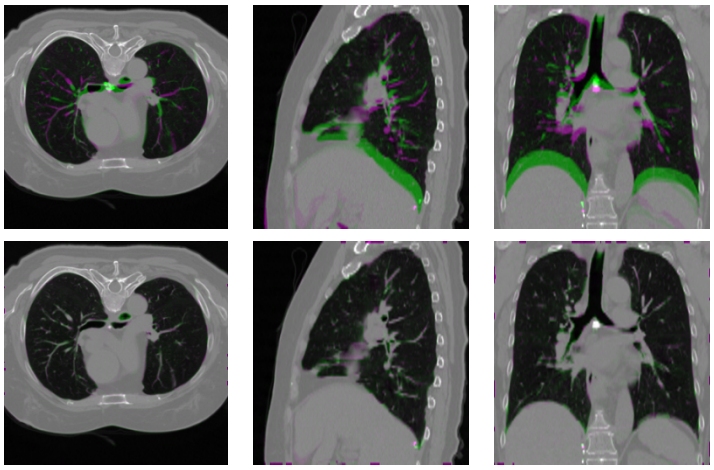

Figure 3: Registration result for Case 5 of 4DCT dataset. Left: axial, middle: sagittal and right: coronal plane. Top row before and bottom row after registration, using the proposed method. Target image displayed in magenta and source in green (complementary colour).

TRE (std) in mm	before reg.	continuous	drop	4DLTMd	csbp
Case 1	4.01 (2.91)	1.020 (0.499)	1.000 (0.519)	0.97 (1.02)	**0.829 (0.944)**
Case 2	4.65 (4.09)	1.177 (0.828)	1.048 (0.603)	0.86 (1.08)	**0.842 (0.953)**
Case 3	6.73 (4.21)	1.950 (1.794)	1.378 (1.018)	1.01 (1.17)	**0.995 (1.058)**
Case 4	9.42 (4.81)	1.895 (1.884)	1.578 (1.389)	1.40 (1.57)	**1.271 (1.241)**
Case 5	7.10 (5.15)	2.504 (2.491)	1.858 (2.180)	1.67 (1.79)	**1.256 (1.520)**

Figure 4: Comparative overview of landmark registration accuracy (in mm) for the extreme phases of clinical 4DCT datasets. Our approach achieves the highest accuracy for all 5 cases.

target and source images are taken from different colour channels, an example image pair before and after registration is shown in Fig. 1. The registration accuracy is measured as both angular error (AAE) (between ground truth and estimated displacements) and TRE.

We use a diffusion regularization function which penalizes the square of the gradient of the deformation field in both our proposed *csbp* method and the continuous framework. *drop* additionally uses B-splines as a transformation model. For single modal registrations we employ sum of squared differences (SSD) as a similarity term and either normalized cross correlation (NCC) or gradient orientation (GO) as similarity terms, respectively, for multi-modal tests. Fig. 2 gives an overview of the obtained registration accuracies. While both discrete methods perform better than the continuous approach, our approach (with spline-fitting for subpixel accuracy) shows its advantages over *drop* in the multimodal cases. The lower TRE demonstrates the effectiveness of the large initial search space using the original resolution images. To evaluate our findings for real clinical data we perform registrations on manually labeled 4D CT datasets [1][3]. The images are taken throughout one breathing cycle with average landmark distances of up to 10 mm for the two extreme phases at inhale and exhale. Manual landmarks are given for 300 anatomical locations. Particular challenges for these registration tasks is changing contrast between tissue and air, because the gas density changes due to compression, discontinuous sliding motion between lung lobes and the lung rib cage interface, and large deformations of small features (lung vessels, airways). The reg-

[3] publicly available for download at www.dir-lab.com

istration result for our approach is displayed in Fig. 3 for the most difficult case 5. One 3D registration with our proposed method for images with dimensions of 256x256x100, using NCC as similarity metric, and an initial label space of $\mathscr{L} = 4725$ takes about 60 minutes (compared to ≈30 minutes for drop), where two thirds of the time are spend on the initial data cost computation. The resulting TRE in mm are given in Fig. 3. 4DLTMd [1] is a specific method for breathing cycle CT registration, which is so far the best ranking algorithm out of 13 for these datasets. Our approach consistently outperforms all other methods.

4 Conclusion

We have presented a novel deformable registration method based on discrete optimization. Our constant space belief propagation enables fast, memory efficient optimization of deformable registration on high resolution medical 3D volumes without the need for resampling or warping schemes. We demonstrate its superior accuracy over continuous optimization methods and previously proposed discrete methods on a range of medical images. Our approach gives very good results particularly for large deformations and multimodal registrations, and it effectively avoids local minima. In the future, we would like to address an alternative solution for the initial data cost computation to improve speed compared to the naïve brute force approach and apply it to clinical multimodal scans.

Acknowledgments

We would like to thank EPSRC and Cancer Research UK for funding this work within the Oxford Cancer Imaging Centre. J.A.S. acknowledges funding from EPSRC EP/H050892/1.

References

[1] R. Castillo, E. Castillo, R. Guerra, V.E. Johnson, T. McPhail, A.K. Garg, and T. Guerrero. A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets. *Physics in Medicine and Biology*, 54(7):1849, 2009.

[2] P. Felzenszwalb and D. Huttenlocher. Efficient belief propagation for early vision. *International Journal of Computer Vision*, 70:41–54, 2006. ISSN 0920-5691.

[3] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, and N. Paragios. Dense image registration through MRFs and efficient linear programming. *Medical Image Analysis*, 12 (6):731 – 741, 2008.

[4] M.P. Heinrich, M. Jenkinson, M. Bhushan, T. Matin, F. Gleeson, J.M. Brady, and Schnabel J.A. Non-local shape descriptor: A new similarity metric for deformable multimodal registration. In *Proc. MICCAI, accepted*, 2011.

[5] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother. A comparative study of energy minimization methods for markov random fields. In *Proc. ECCV*, volume 3952, pages 16–29, 2006.

[6] Q. Yang, L. Wang, and N. Ahuja. A constant-space belief propagation algorithm for stereo matching. In *Proc. CVPR*, pages 1458–1465, 2010.