Probabilistic Segmentation Propagation from Uncertainty in Registration

Ivor J.A. Simpson[1]
ivor.simpson@eng.ox.ac.uk
Mark W. Woolrich[2]
mark.woolrich@ohba.ox.ac.uk
Julia A. Schnabel[1]
julia.schnabel@eng.ox.ac.uk

[1] Institute of Biomedical Engineering
Old Road Campus Research Building
Oxford OX3 7DQ

[2] Oxford Centre for Human Brain Activity
Warneford Hospital
Oxford OX3 7JX

Abstract

In this paper we propose a novel approach for incorporating measures of spatial uncertainty which are derived from non-rigid registration, into propagated segmentation labels. In current approaches to segmentation via label propagation, a point-estimate of the registration parameters is used. However, this is limited by the registration accuracy achieved. In this work, we derive local measurements of the uncertainty of a non-rigid mapping from a probabilistic registration framework. This allows us to consider the set of probable locations for a segmentation label to hold. We demonstrate the use of this method on the propagation of accurately delineated cortical labels in inter-subject brain MRI using the NIREP dataset. We find that accounting for the spatial uncertainty of the mapping increases the sensitivity of correctly classifying anatomical labels.

1 Introduction

Automated medical image segmentation can be achieved by propagating manually delineated labels from an annotated source image registered to a particular target. This registration is required to be accurate to ensure a high level of anatomical correspondence, and therefore non-rigid registration methods are often used. A drawback in current approaches is that a point-estimate of the registration parameters is used, e.g. in [6]. This is limited in assuming that the mapping is entirely correct. This however, in the case of inter-subject brain registration, is unlikely and in a recent comparative study of volumetric brain registration algorithms, it was shown that none of the tested algorithms were capable of entirely correctly registering large labelled regions [4]. The situation can be improved upon by using a probabilistic registration method, which models a level of uncertainty on the inferred transformation. Risholm et al.[5] previously showed that probabilistic registration methods can produce a map of registration uncertainty. However, their estimates of uncertainty depended upon the definition of an ad-hoc weighting between the similarity and regularisation terms, the weighting of which will influence the registration result, and the distribution of spatial uncertainty. Their approach was also aimed specifically at coping with resolving brain-shift after tumour resection rather than for generic application. Moreover, the elastic registration method they employed is not well suited for inter-subject registration.

In this work we propose the use of a generic and adaptive approach for probabilistic non-rigid registration in order to provide a more principled estimate of the spatial uncertainty of a registration. Based on these estimated uncertainties, local Gaussian smoothing kernels can be automatically estimated and used to smooth propagated segmentation labels over the set of probable locations, rather than propagating only the most likely ones. We demonstrate this method on inter-subject brain registration, with cortical labels and illustrate the potential of our new approach using an example probability map of a propagated cortical label and showing an increase in the sensitivity of correctly classifying the true segmentation label.

2 Methods

Image registration can be described probabilistically using a generative model, where it is assumed that the target image data $\mathbf{Y}$ can be generated from a source image $\mathbf{X}$, which is deformed by a transformation, where $\mathbf{T}(\mathbf{X},\mathbf{w})$ is the transformed source image, and where $\mathbf{w}$ parametrises the transformation. The specific form of $\mathbf{T}$ used in this implementation is a Free-Form Deformation (FFD) model [6], where $\mathbf{w}$ is the set of control point displacements.

The model has residual error throughout the registration process, which needs to be modelled. The noise is assumed to have zero mean and be independent and identically distributed across image voxels. In this case, the noise is assumed to be normally distributed error $\mathbf{e} \approx N(0,\phi^{-1}\mathbf{I})$, where $\mathbf{I}$ is the matrix identity and ϕ is the global noise precision (inverse variance). The generic generative model for registration is given as $\mathbf{Y} = \mathbf{T}(\mathbf{X},\mathbf{w}) + \mathbf{e}$.

2.1 Priors

Since we use a probabilistic model for registration, regularisation can be incorporated as a prior on the transformation parameters which is modelled using a Multivariate Normal Distribution (MVN). The prior on $\mathbf{w}$ is described in equation 1 where Λ encodes the regularisation as a spatial kernel matrix providing bending energy regularisation. λ is the spatial precision parameter, controlling the level of spatial regularisation, and is inferred from the data resulting in an automated approach to regularisation.

Non-informative priors on the spatial (eq. 3) and noise precision (eq. 2) are specified using a Gamma distribution Ga, where the subscript $_0$ denotes initial parameter estimates. We use wide priors over λ and ϕ $s_0 = 10^{10}, c_0 = 10^{-10}, a_0 = 10^{10}, b_0 = 10^{-10}$.

$$P(\mathbf{w}|\lambda) = MVN(\mathbf{w};0,(\lambda\Lambda)^{-1}) \quad (1)$$

$$P(\phi) = Ga(\phi;a_0,b_0) \quad (2)$$

$$P(\lambda) = Ga(\lambda;s_0,c_0) \quad (3)$$

2.2 Model Inference

The model parameters are inferred using Variational Bayes [1]. VB was chosen as it provides a computationally efficient approach to fully Bayesian inference, unlike numerical methods such as MCMC. VB uses an objective function of the Variational Free Energy (VFE). The VFE measures both the model fit and model complexity. Model fit quantifies how well the target data is explained by the model parameters, and for this model is related to the SSD. Model complexity in this case is approximately equivalent to the bending energy in the transformation. Using the VB framework, analytic updates are derived for the approximate

posterior distributions of the transformation, regularisation and noise parameters. VB uses the mean-field approximation so the posterior distribution $p(\mathbf{w},\lambda,\phi|\mathbf{Y})$ can be approximated as $q(\mathbf{w}|\mathbf{Y})q(\lambda|\mathbf{Y})q(\phi|\mathbf{Y})$. The functional forms of the approximate posterior distributions are constrained to be conjugate to the priors, and are given as:$q(\mathbf{w}) = MVN(\mathbf{w};\mu,\Upsilon)$, $q(\lambda) = Ga(\lambda;s,c)$ and $q(\phi) = Ga(\phi,a,b)$. The hyper-parameter updates (eq. 4-9) are derived by integrating out factorised parameters from the log posterior distribution of the model.

$$\Upsilon = (\alpha\bar{\phi}\mathbf{J}^{\mathsf{T}}\mathbf{J} + \bar{\lambda}\Lambda)^{-1} \quad (4) \qquad \mu_{new} = \Upsilon\left[\alpha\bar{\phi}\mathbf{J}^{\mathsf{T}}(\mathbf{J}\mu_{old} + \mathbf{k})\right] \quad (7)$$

$$c = c_0 + \frac{N_c}{2} \quad (5) \qquad \frac{1}{s} = \frac{1}{s_0} + \frac{1}{2}\left(Tr(\Lambda\Upsilon^{-1}) + \mu^{\mathsf{T}}\Lambda\mu\right) \quad (8)$$

$$b = b_0 + \frac{N_v\alpha}{2} \quad (6) \qquad \frac{1}{a} = \frac{1}{a_0} + \frac{1}{2}\alpha(\mathbf{k}^{\mathsf{T}}\mathbf{k} + Tr(\Upsilon^{-1}\mathbf{J}^{\mathsf{T}}\mathbf{J})) \quad (9)$$

where $\mathbf{J}$ is the matrix of first order partial derivatives of the transformation parameters with respect to the transformed image $\mathbf{T}(\mathbf{X},\mu_{old})$, centred about the previous estimate of the mean μ_{old}. $\mathbf{k}$ is the vector representing the residual image $\mathbf{Y} - \mathbf{T}(\mathbf{X},\mathbf{w})$. μ_{new} describes the current estimated transformation parameters, and is dependent on the old estimated values. The approximate posterior covariance matrix of the set of transformation parameters is given by Υ. $\bar{\lambda}$ is the expectation of the posterior spatial precision and $\bar{\phi}$ is the expectation of the noise precision. N_c is the number of control points in the model and N_v is the number of active voxels, α is the virtual decimation factor modelling correlation of the image noise [3].

2.3 Spatial Uncertainty

The probabilistic registration model intrinsically provides a measurement of uncertainty on the posterior distribution of the model parameters. This is particularly interesting for the transformation parameters whose posterior covariance matrix is given by Υ. The spatial uncertainty of the transformation parameters represents how certain we are that a given point in the source image should be transformed to a particular point in the target image. By calculating Υ, and interpolating the variance and cross-directional co-variance across the image, an MVN distribution illustrating the spatial uncertainty can be calculated at each voxel, indicating the magnitude and direction of the uncertainty at that point.

The spatial uncertainty of a particular control point is governed by 5 factors: 1. the **local image information** which is affected by this control points' movement ($\mathbf{J}^T\mathbf{J}$); 2. the **noise precision**, which indicates how noisy the image data is; 3. the **spatial precision**, the similarity of the transformation to the spatial prior; 4. the form of the **spatial prior**, e.g. Bending Energy and 5. the **uncertainty of neighbouring control points**.

As the uncertainty depends upon the image information, it is anisotropic by nature, displaying less uncertainty across an image boundary than along it. A straightforward approach to compensate for the spatial uncertainty in the mapping of any given voxel is to smooth the label data according to the local uncertainty distribution, which is an anisotropic Gaussian kernel. We propose the use of this novel method to account for the spatial uncertainty of each transformed voxel.

2.4 Experiments

In this work we evaluate our method using the NIREP database [2]. This consists of 16 healthy control subjects, each with 32 accurately delineated cortical labels. Each NIREP

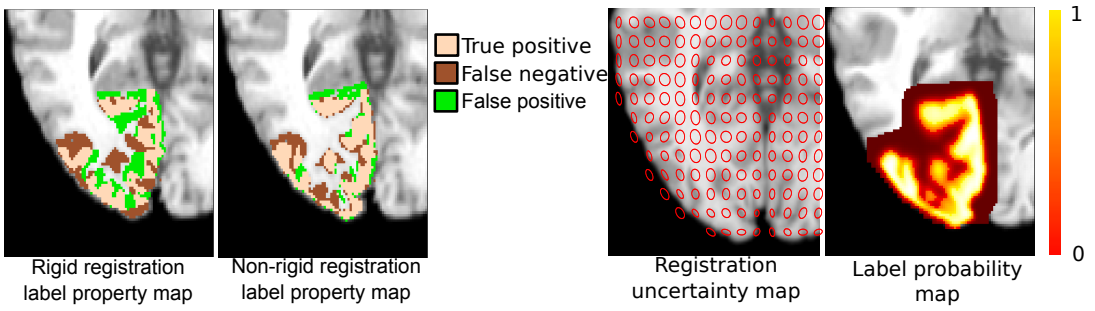

Figure 1: An example registration, illustrating label propagation for the right occipital lobe. The first two panels show the label property for the rigid, and non-rigid registered images. The third image illustrates a map of the local uncertainty of the transformation overlaid on the transformed source image. The final panel shows the derived label probability map.

image was registered affinely with 9 degrees of freedom to the MNI152 atlas. Each image pair was then pairwise registered to one another using an initial rigid registration. This was followed by a non-rigid registration using the probabilistic method outlined previously with a 5mm FFD control point spacing. The 32 cortical labels were transformed from the source image space to the target image using the mean of the transformation parameters. In order to consider the full posterior distribution of the transformation parameters, the transformed labels were adaptively smoothed using a local Gaussian kernel. An example registration with label propagation is shown in figure 1.

We have thus obtained a map showing the probable locations of the label. An informative method of quantifying this information is treating each segmentation map as a binary classifier and drawing a receiver operator characteristic curve. In figure 2 we show the average true positive rate (sensitivity) and false positive rate (1 - specificity) of our method for two labelled cortical regions against the mean of the transformation over 240 subject-to-subject registrations. As can be seen, in both plots the two curves look reasonably similar, with the

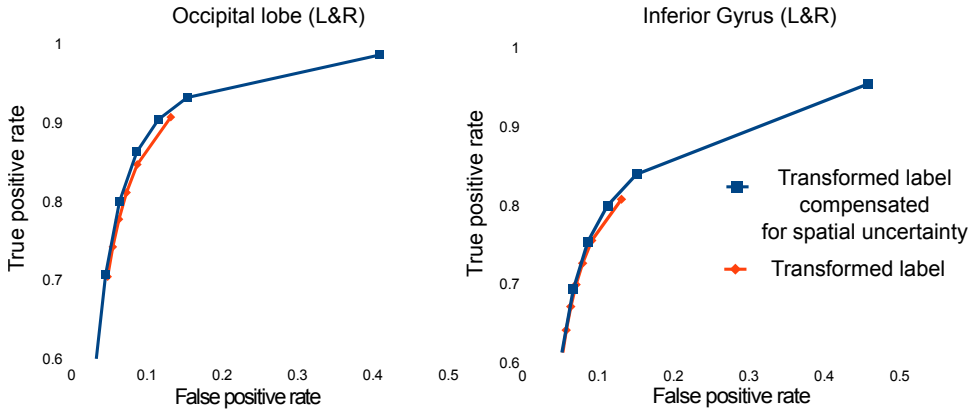

Figure 2: ROC curves for the occipital lobe and inferior gyrus segmentation labels. Below a true positive rate of 0.6, the two lines are practically identical.

exception that by compensating for the uncertainty in the spatial mapping the maximum sensitivity is improved as more correctly labelled voxels are considered probable. Additionally, the optimal trade-off between the true, and false positive rate is slightly better when using the proposed method. This results in a small improvement in label classification accuracy when using the simple method of selecting the most likely segmentation at each voxel. The

proposed approach yields an average increase in Dice overlap of 0.23% over all labels for the 240 pairwise registrations, with a maximum average improvement of 1.7% for the left insula gyrus. 6 of the 32 labels showed statistically significant improvement at the 0.05 significance level, with an average overlap improvement of 1.35%. These 6 regions had larger scale spatial structures, which may be better captured at the resolution of registration .

3 Discussion and Conclusions

In this work we have introduced a generic, principled mechanism for considering the uncertainty in non-rigid registration. This was demonstrated with application to propagated cortical segmentations in inter-subject brain MRI, allowing a probabilistic segmentation representation. This method has also been shown to improve disease classification rates in a longitudinal study of Alzheimer's disease [7]. The incorporation of the spatial uncertainty of a non-rigid mapping in segmentation propagation provides a principled method for allowing high sensitivity to the true segmentation label as more nearby voxels are considered to be probable. In order to demonstrate the full utility of probabilistic segmentation propagation, potential areas of future work are to use this information as a prior for an intensity based segmentation scheme, or to incorporate this additional knowledge within a multiple-atlas classifier fusion framework.

4 Acknowledgements

The authors would like to acknowledge the LSI DTC and the EPSRC for funding this work.

References

[1] H. Attias. A variational Bayesian framework for graphical models. *Advances in neural information processing systems*, 12(1-2):209–215, 2000.

[2] G.E. Christensen, X. Geng, J.G. Kuhl, J. Bruss, T.J. Grabowski, I.A. Pirwani, M.W. Vannier, J.S. Allen, and H. Damasio. Introduction to the non-rigid image registration evaluation project (NIREP). *WBIR2006*, 4057:128–135, 2006.

[3] A.R. Groves, C.F. Beckmann, S.M. Smith, and M.W. Woolrich. Linked independent component analysis for multimodal data fusion. *NeuroImage*, 54(3):2198 – 2217, 2011.

[4] A. Klein, J. Andersson, B.A. Ardekani, J. Ashburner, B. Avants, M.C. Chiang, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage*, 46(3):786–802, 2009.

[5] P. Risholm, S. Pieper, E. Samset, and W. Wells. Summarizing and visualizing uncertainty in non-rigid registration. *MICCAI*, 6362:554–561, 2010.

[6] D. Rueckert, LI Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, and D.J. Hawkes. Nonrigid registration using Free-Form Deformations: application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999.

[7] I.J.A. Simpson, M.W. Woolrich, and J.A. Schnabel. Longitudinal Brain MRI Analysis with Uncertain Registration. *MICCAI (Accepted)*, 2011.