

Longitudinal Brain MRI Analysis with Uncertain Registration

Ivor J.A. Simpson^{1,2,*}, Mark W. Woolrich^{2,3}, Adrian R. Groves²,
and Julia A. Schnabel¹

¹ Institute of Biomedical Engineering, University of Oxford, Oxford
`{ivor.simpson,julia.schnabel}@eng.ox.ac.uk`

² Oxford Centre for Functional MRI of the Brain, University of Oxford, Oxford
`adriang@fmrib.ox.ac.uk`

³ Oxford Centre for Human Brain Activity, University of Oxford, Oxford
`mark.woolrich@ohba.ox.ac.uk`

Abstract. In this paper we propose a novel approach for incorporating measures of spatial uncertainty, which are derived from non-rigid registration, into spatially normalised statistics. Current approaches to spatially normalised statistical analysis use point-estimates of the registration parameters. This is limiting as the registration will rarely be completely accurate, and therefore data smoothing is often used to compensate for the uncertainty of the mapping. We derive localised measurements of spatial uncertainty from a probabilistic registration framework, which provides a principled approach to image smoothing. We evaluate our method using longitudinal deformation features from a set of MR brain images acquired from the Alzheimer’s Disease Neuroimaging Initiative. These images are spatially normalised using our probabilistic registration algorithm. The spatially normalised longitudinal features are adaptively smoothed according to the registration uncertainty. The proposed adaptive smoothing shows improved classification results, (84% correct Alzheimer’s Disease vs. controls), over either not smoothing (79.6%), or using a Gaussian filter with $\sigma = 2\text{mm}$ (78.8%).

1 Introduction

Alzheimer’s disease (AD) is a progressive neurological disease, and the most common to be associated with the symptoms of dementia. The Alzheimer’s Disease Neuroimaging Initiative (ADNI) [7] is a large multi-site study whose primary goal is to test whether serial magnetic resonance imaging (MRI) and other biological and imaging markers can be combined to measure the progression of mild cognitive impairment (MCI) and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials. ADNI’s initial goal was to recruit 800

* This work was funded by the EPSRC through the LSI DTC.

adults, ages 55 to 90, to participate in the research, approximately 200 cognitively normal older subjects to be followed for 3 years, 400 people with MCI to be followed for 3 years and 200 people with early AD to be followed for 2 years.

Longitudinal studies of this scale provide a platform to analyse the progression of the anatomical effects of neurodegenerative diseases. The use of imaging tools on longitudinal data allows quantification of the rate of brain atrophy, e.g. SIENA [11]. Feature maps can be derived from the deformation fields required to warp between follow-up images and baseline scans, providing high-resolution information illustrating the anatomical changes taking place over time [9]. The use of deformation derived features is known as tensor-based morphometry (TBM). A popular feature one can choose to use is the determinant of the Jacobian of the deformation fields, which is interpreted as the local expansion/contraction at a particular location and in the case of longitudinal brain imaging this would describe atrophy. Once these longitudinal features have been created, they need to be spatially normalised to enable comparison. This spatial normalisation is required to be sufficiently accurate to ensure robust inference. Accurate spatial normalisation can be provided using non-rigid registration methods [6]. A drawback in current approaches is that a point-estimate of the registration parameters is used, e.g. in [2], which assumes that the mapping is exactly correct. This however, in the case of inter-subject brain registration, is highly unlikely. In a recent comparative study of volumetric brain registration algorithms, it was shown that no algorithm was capable of entirely correctly registering large labelled regions [5]. This situation can be improved upon by using a probabilistic registration method, which includes a level of uncertainty on the inferred transformation. The concept of spatial uncertainty in low-level vision has been previously explored [12]. More recently, probabilistic registration methods have been shown to produce a map of registration uncertainty in intra-subject brain registration following tumour resection [8]. However, although their registration model is capable of estimating the weighting between the similarity and regularisation terms, in this work they prefer to use an ad-hoc weighting of these factors. This is probably because their approach employs Markov chain Monte Carlo (MCMC) to infer the registration model parameters, which is particularly computationally expensive. The significance of this weighting is that it will influence both the registration result, and the distribution of the spatial uncertainty.

In this work we propose the use of a generic and adaptive approach for probabilistic non-rigid registration to provide a more principled estimate of the spatial uncertainty of a transformation. Based on these estimates, local Gaussian smoothing kernels can be automatically estimated and used to smooth image features over the set of probable locations, rather than just the most likely. This approach is demonstrated using longitudinal data from the ADNI dataset, where Jacobian maps are used as image features. These are spatially normalised using a principled probabilistic approach. This transformed data is adaptively smoothed based on the registration derived uncertainty. Our adaptive approach to image smoothing is compared, and performs favourably to an ad-hoc Gaussian smoothing, or no image smoothing for multi-variate disease state classification.

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 \mathbf{w} parametrises the transformation. The specific form of \mathbf{T} used in this implementation is a Free-Form Deformation (FFD) model, where \mathbf{w} is the set of control point displacements in each direction.

As the model will have residual error throughout the registration process, this error estimate needs to be included. The noise is assumed to have zero mean and be independent and identically distributed (i.i.d.) across image voxels. In this case, the noise is assumed to be normally distributed $\mathbf{e} \approx N(0, \phi^{-1}\mathbf{I})$, where \mathbf{I} represents the matrix identity. The noise distribution is assumed to have global precision (inverse variance) across the image, ϕ . The generic generative model for registration is therefore given as $\mathbf{Y} = \mathbf{T}(\mathbf{X}, \mathbf{w}) + \mathbf{e}$. Using a normally distributed noise model, as we use here, is equivalent to using the sum of squared differences (SSD) as the image similarity term.

2.1 Priors

In 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 eq. 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. λ is modelled as an unknown parameter, and determined adaptively from the data resulting in an automated approach to regularisation. Where λ is constant, this method of regularisation is seen in other probabilistic approaches to registration [2]. In a related approach in groupwise registration, the covariance matrix of the deformations used to warp a template to a set of observed images is calculated from the data [1]. This provides an alternative principled approach to inferring the covariance between deformation parameters, however the model inference is computationally infeasible on full 3D volumes.

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

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

$$P(\lambda) = Ga(\lambda; s_0, c_0) \quad (2)$$

$$P(\phi) = Ga(\phi; a_0, b_0) \quad (3)$$

2.2 Model Inference

The model parameters are inferred upon using Variational Bayes [3] which uses an objective function of the Variational Free Energy (VFE). The VFE

measures model fit and complexity. Model fit can be approximately considered to be minimising the sum of squared difference, and model complexity as the level of bending energy of the transformation. Using the VB framework, analytic updates are derived for the approximate posterior distributions of the transformation, regularisation and noise parameters, which maximise the VFE.

VB uses the mean-field approximation, hence the posterior parameter distribution $p(\mathbf{w}, \lambda, \phi | \mathbf{Y})$ is approximated as $q(\mathbf{w})q(\lambda)q(\phi)$. 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 (eqs. 4-9) are derived by integrating out the factorised parameters from the log posterior model distribution.

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

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

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

Here, \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 distribution and $\bar{\phi}$ is the expectation of the estimated noise precision. N_c is the number of active control points in the model and N_v is the number of active voxels, α is the virtual decimation factor which accounts for the correlation in the image noise [4].

2.3 Spatial Uncertainty

The probabilistic registration method intrinsically provides a measurement of uncertainty on the posterior distribution of the model parameters. This is particularly interesting for the transformation parameters, given by Υ , which has units of mm^2 . The uncertainty of \mathbf{w} 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 accurately estimating Υ , and interpolating the variance and cross-directional covariance across the image, a multivariate normal distribution illustrating the spatial uncertainty can be calculated at each voxel. The spatial uncertainty of a particular control point is governed by 5 factors:

- The local image information which is affected by this control point's movement ($\mathbf{J}^T \mathbf{J}$).
- The noise precision: how much the image data is trusted, related to SSD.
- The spatial precision: how similar the transformation is to the spatial prior.

- The form of the spatial prior, e.g. bending energy, membrane energy.
- The uncertainty of neighbouring control points.

This distribution provides both the magnitude and direction of the uncertainty at a given point. As the uncertainty is dependent on the image information, it is lower across an image boundary than along it. This results in an anisotropic measure of spatial uncertainty, which varies across the image. The scale of uncertainty will also vary across individual registrations. A straightforward approach to compensate for the spatial uncertainty in the mapping of any given voxel is to smooth the data according to the local uncertainty distribution, which is an anisotropic Gaussian kernel. Where the uncertainty of the mapping is high, smoothing the data compensates for this, but still retains potentially discriminating information. We propose this novel method to help reduce the inter-subject variability due to mis-registration of spatially normalised image data.

3 Materials

Data was provided from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) [7], a large multi-site longitudinal study containing MR imaging data. For this study 125 random control subjects, and AD patients, of both sexes and a range of ages AD (mean age 76.6 std.dev 7.67, 67M, 58F) and Normal controls (mean age 78.6 std.dev 5.61, 66 M, 59F) were drawn from the database. Pre-processed images which have been corrected for geometric distortions, bias fields and geometric scaling are present on the ADNI website, and were used in this work. Subjects were chosen with at least 2 scans with a minimum interval of 1 year.

4 Experiments

Tools from the publicly available open-source software library, FSL¹ were used to pre-process the images. To correct for difference in size and location, each of the scans was registered to the MNI 152 template using 9 degrees of freedom. Each scan was also re-sampled to have 1mm isotropic voxels and was processed to remove non-brain tissue. An initial atlas was created by averaging 40 healthy individual control scans after initial affine alignment to the MNI152 template. To create a sharper atlas, each of these scans was then non-rigidly registered to the affine atlas using the probabilistic registration tool with a 5mm knot spacing.

The most recent follow-up scan was registered to the baseline scan of each individual using the probabilistic registration tool with a 5mm knot spacing, which yielded individual maps of local atrophy. As the interval between scans varied across subjects, the Jacobian values were linearly scaled to a single year.

The probabilistic non-rigid registration algorithm was used to provide accurate spatial normalisation, which allows measurement of spatial uncertainty. Each baseline scan was registered to the atlas, an example is given in figure 2. Each registration result yields a map of the variance and cross-directional covariance of the transformation parameters. An illustration of the uncertainty in the mapping between an AD subject to the atlas is given in figure 1.

¹ <http://www.fmrib.ox.ac.uk/fsl/>

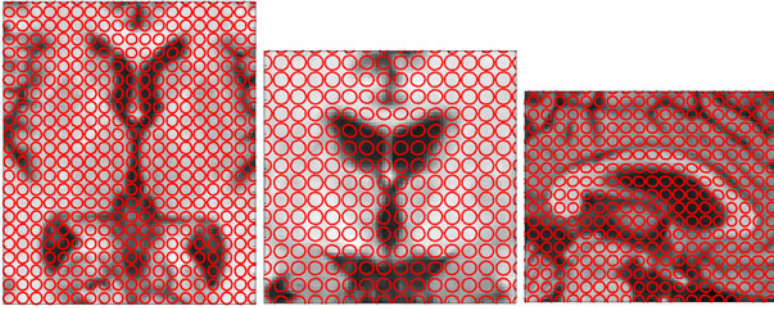


Fig. 1. An example image showing the distribution of spatial uncertainty in the mapping between an AD patient and the atlas. Ellipoids are plotted every 5mm showing the full-width at half-maximum of the uncertainty distribution at each point. The underlying image is the patient image transformed to the atlas.

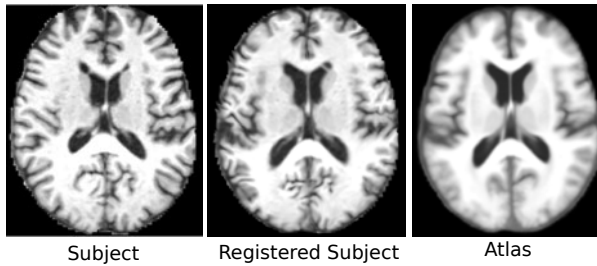


Fig. 2. An example registration from the AD patient shown in figure 1 to the atlas

Each individual subject's longitudinal Jacobian map was transformed to the atlas space. It is then adaptively smoothed using a unique 3-D Gaussian kernel for each spatial location derived from the uncertainty of the registration from the baseline to atlas space. As control experiments, we compare this strategy against not smoothing the images, and Gaussian smoothing with $\sigma = 2mm$.

5 Results

Data decomposition and feature selection was required to provide computationally tractable classification on the 250 spatially normalised and smoothed Jacobian maps, each of which had 2068979 voxels within the anatomical mask. The concatenated data was processed using a principal component analysis (PCA) algorithm to reduce the data dimensionality. The feature maps were masked prior to the PCA using a voxelwise t-test between the two populations with a threshold of $p < 0.05$ uncorrected. The percentage of voxels within the anatomical atlas region included in the mask was 26.82% for the unsmoothed data, 31.91% for the Gaussian smoothed data and 29.74% for the proposed method. Thresholded z-stat maps of the 3 smoothing methods are presented in figure 3. The lowest

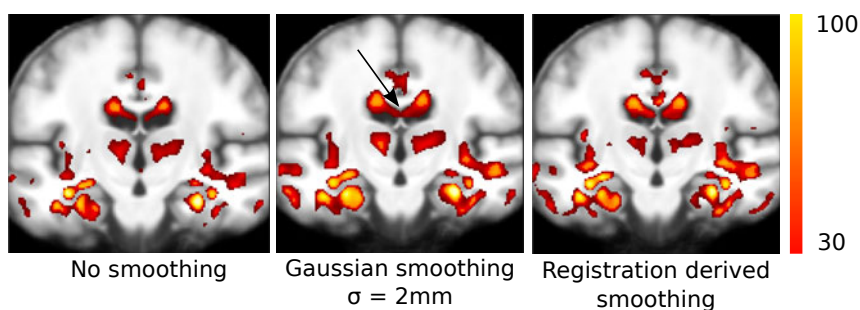


Fig. 3. Z-stat maps showing the results from an uncorrected, two-tail t-test on the spatially normalised Jacobian maps between AD patients and control subjects. The registration derived smoothing does not blur the boundaries between the ventricles, unlike Gaussian smoothing, and it increases the signal to noise over no smoothing.

Table 1. Classification results using an SVM with a radial basis function kernel for the different smoothing methods

Smoothing method	Correct rate	Sensitivity	Specificity
No smoothing	0.796	0.704	0.888
2mm Gaussian Smoothing	0.788	0.904	0.67
Adaptive Smoothing	0.84	0.968	0.71

uncorrected p-value observed for the individual methods is 6.8736×10^{-20} with no smoothing, 1.1289×10^{-20} for Gaussian smoothing and 1.5601×10^{-21} when using the proposed adaptive smoothing. In each case the most significant voxel was located in the left hippocampus which is consistent with previous work [6].

The components which explained 99.0% of the sample variance were used to project the data. The projected data were classified using a support vector machine with a radial basis function kernel using $\sigma_{RBF} = 2$. A leave-two-out methodology was used, where an instance of each class was used for testing at each iteration, while the rest of the data were used for training. The average classification results are given in table 1 and show an improvement in the correct classification rate when using the proposed approach.

6 Conclusions and Future Work

Adaptive smoothing of spatially normalised image feature data, based on registration derived uncertainty, has been demonstrated to provide an increase in the ability to classify between patients with Alzheimer’s Disease and normal controls using longitudinal Jacobian maps as features. Although the uncertainty measurements we derive from the image registration process are only a surrogate indication of the true anatomical uncertainty, we have shown that they still provide useful information for reducing the inter-subject variability due

to mis-registration. As the methods presented here are generic, they could be employed in alternative applications, including probabilistic segmentation propagation [10]. Future work includes a more rigorous experimental design and including MCI subjects. Additionally, we will consider the use of the independent uncertainty distribution for each FFD control point, as opposed to the unfactorised estimates used in this work. This may provide more accurate estimates of the spatial uncertainty for use as a smoothing kernel.

References

1. Allasonnière, S., Amit, Y., Trounev, A.: Toward a coherent statistical framework for dense deformable template estimation. *Journal of the Royal Statistical Society, Series B* 69(2) (2007)
2. Ashburner, J.: A fast diffeomorphic image registration algorithm. *Neuroimage* 38(1), 95–113 (2007)
3. Attias, H.: A variational Bayesian framework for graphical models. In: Leen, T., Dietterich, T., Tresp, V. (eds.) *NIPS 2000*, vol. 12, pp. 209–215. MIT Press, Cambridge (2000)
4. Groves, A.R., Beckmann, C.F., Smith, S.M., Woolrich, M.W.: Linked independent component analysis for multimodal data fusion. *NeuroImage* 54(3), 2198–2217 (2011)
5. Klein, A., Andersson, J., Ardekani, B., Ashburner, J., Avants, B., Chiang, M., Christensen, G., Collins, D., Gee, J., Hellier, P., Song, J., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R., Mann, J., Parsey, R.: Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* 46(3), 786–802 (2009)
6. Leow, A., Yanovsky, I., Parikshak, N., Hua, X., Lee, S., Toga, A., Jack Jr., C., Bernstein, M., Britson, P., Gunter, J., Ward, C., Borowski, B., Shaw, L., Trojanowski, J., Fleisher, A., Harvey, D., Kornak, J., Schuff, N., Alexander, G., Weiner, M., Thompson, P.: Alzheimer's disease neuroimaging initiative: a one-year follow up study using tensor-based morphometry correlating degenerative rates, biomarkers and cognition. *Neuroimage* 45(3), 645–655 (2009)
7. Mueller, S., Weiner, M., Thal, L., Petersen, R., Jack, C., Jagust, W., Trojanowski, J., Toga, A., Beckett, L.: Alzheimer's Disease Neuroimaging Initiative. *Advances in Alzheimer's and Parkinson's Disease*, 183–189 (2008)
8. Risholm, P., Pieper, S., Samset, E., Wells, W.: Summarizing and visualizing uncertainty in non-rigid registration. In: Jiang, T., Navab, N., Pluim, J., Viergever, M. (eds.) *MICCAI 2010. LNCS*, vol. 6362, pp. 554–561. Springer, Heidelberg (2010)
9. Scahill, R., Schott, J., Stevens, J., Rossor, M., Fox, N.: Mapping the evolution of regional atrophy in Alzheimer's disease: unbiased analysis of fluid-registered serial MRI. *Proceedings of the National Academy of Sciences of the United States of America* 99(7), 4703 (2002)
10. Simpson, I., Woolrich, M., Schnabel, J.: Probabilistic segmentation propagation. In: *Medical Image Understanding and Analysis 2011* (2011)
11. Smith, S., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P., Federico, A., De Stefano, N.: Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *Neuroimage* 17(1), 479–489 (2002)
12. Szeliski, R.: Bayesian modeling of uncertainty in low-level vision. *International Journal of Computer Vision* 5(3), 271–301 (1990)