A PROBABILISTIC NON-RIGID REGISTRATION FRAMEWORK USING LOCAL NOISE ESTIMATES

Ivor J.A. Simpson[1,2], *Mark W. Woolrich*[2], *Jesper L.R Andersson*[2], *Adrian R. Groves*[2], *Julia A. Schnabel*[1]

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

ABSTRACT

Accurate inter-subject registration of magnetic resonance (MR) images of the human brain is required to allow meaningful comparisons across groups of subjects. Some anatomical structures can be very difficult to match and this can result in intensity based registration approaches inferring complex and implausible mappings in some regions. In this work, we propose a generic probabilistic framework for non-rigid registration with a spatially varying trade-off between image information and regularisation. This trade-off is based on local estimates of misalignment "noise", which effectively increases regularisation in regions which are difficult to register. We demonstrate that the proposed method infers smoother, more plausible and slightly more accurate mappings for inter-subject registration of MR images of the human brain.

Index Terms— Image registration, regularisation, probabilistic modelling, brain MRI

1. INTRODUCTION

Non-rigid registration is widely used in the analysis of MR images of the human brain, most commonly for mapping subjects to a common anatomical space (spatial normalisation). Intensity based registration requires a cost function to measure image similarity, which drives the optimisation. A common problem is the presence of non-matching image structures, particularly in inter-subject brain registration, where there may be regions where an accurate and plausible mapping using a chosen transformation model does not exist.

The relative weighting of image information and regularisation can cause the inference of complex, non-smooth and implausible mappings in some regions. Conversely, although increasing the global level of regularisation can avoid such phenomena, it leads to a more restricted registration across the whole image. This may therefore infer an inaccurate, and overly smooth mapping between images.

A solution to this problem is to use a spatially varying weight between image information and regularisation. Several approaches have been previously proposed addressing this. They either require image segmentations e.g. [1][2], which is likely to be impractical, or they use a data driven approach. One such data driven method proposed a local image reliability measure based on image structure and estimated noise levels, which was then used to derive spatially localised regularisation [3]. A limitation of that approach is that only the information in the individual images are used to define the local regularisation weighting, which may still lead to problems dealing with non-matching structures. Methods of assessing and appropriately weighting image data based on mutually salient features has been demonstrated to improve registration [4], although this approach is computationally highly expensive.

In this work we propose the use of a local Gaussian noise model to provide a more flexible and robust approach to image alignment. This approach infers localised estimates of the noise in model fit across the image, yielding a spatially varying trade-off between image information and regularisation. The modelled "noise" is based on the registration misalignment, thus is inferred from both images. This means that regions with non-matching structures will be more strongly affected by the regularisation, and thus the mapping will be smoother and more constrained in this region. This will restrain the inference of unnecessarily complex warps, which may not be biologically plausible.

This work extends the generic probabilistic registration framework introduced in [5]. This framework has intrinsic benefits in being able to infer a global level of spatial regularisation, as well as estimating the uncertainty of the registration. This framework is similar to that of Risholm et al. [6], although they choose to infer the model parameters using a computationally intensive sampling approach. Conversely, we use a variational Bayesian inference scheme to allow a more efficient high resolution 3D anatomical alignment.

In the following section we describe our proposed probabilistic registration model and inference scheme. In Section 3 we demonstrate that the proposed registration model infers smoother, more plausible mappings for inter-subject registrations, with a slight improvement in structural overlap compared to using a global noise model. We conclude with a discussion of benefits and future work.

IJAS acknowledges funding from the EPSRC via the Life Science Interface Doctoral Training Centre, University of Oxford.

2. METHODS

Our approach extends the probabilistic registration framework described in [5] by allowing the inference of local Gaussian noise estimates which model image misalignment. The generic generative model for image registration is written as:

$$\mathbf{y} = \mathbf{t}(\mathbf{x}, \mathbf{w}) + \mathbf{e} \tag{1}$$

where **y** and **x** are the target and source images, respectively, and **t** is a non-linear transformation function parametrised by **w**. This approach is valid for any choice of transformation model, but in this work we choose to use a free form deformation (FFD) transformation model for **t**, where **w** describes the displacement of the b-spline knots. **e** is the noise in model fit which is modelled as Gaussian noise with spatially varying precision (inverse variance) and is defined as:

$$\mathbf{e} = \mathcal{N}(0, \mathrm{diag}(\Phi^T \mathbf{b})^{-1}) \tag{2}$$

where $\Phi = \{\alpha_1\phi_1, \alpha_2\phi_2, ..., \alpha_L\phi_L\}$, is a $L\times 1$ vector of noise precisions where L is the number of noise components and **b** is a $L \times N_v$ matrix representing the basis set which assigns the weighting of each ϕ across the N_v voxels in the image. **b** must represent a non-negative basis set that has a degree of spatial smoothness to allow estimates of the image gradients. We choose to use a set of equally spaced normalised Gaussian kernels as our basis set. α_l is a locally estimated virtual decimation factor which compensates for spatial smoothness in the noise [5]. α is an estimate of the proportion of independent voxels in the residual. We now have a probabilistic model with spatially smooth noise which we can infer upon. The log likelihood for this model is given as:

$$\log P(\mathbf{y}|\mathbf{x}, \mathbf{w}, \Phi) = \sum_l^L \left(\frac{\alpha_l N_{v,l}}{2} \log \frac{\alpha_l \phi_l}{2\pi} - \frac{1}{2} \mathbf{m}_l^T \alpha_l \phi_l \mathbf{m}_l \right) \tag{3}$$

where $\mathbf{m}_l = (\mathbf{b}_{l,:})^{\frac{1}{2}} \circ (\mathbf{y} - \mathbf{t}(\mathbf{x}, \mathbf{w}))$, $\circ$ refers to the elementwise matrix product (Hadamard product). $N_{v,l}$ refers to the number of partial voxels weighted by ϕ_l, $N_{v,l} = \mathrm{rowsum}(\mathbf{b}_{(l,:)})$.

We have a prior distribution on ϕ_l which is described using a Gamma distribution:

$$P(\phi_l) = Ga(a_0, b_0) \tag{4}$$

where a_0 is the scale, and b_0 the shape parameter of the distribution. The prior distribution on noise precision is initially set to be uninformative with $a_0 = 10^{10}$,$b_0 = 10^{-10}$. Once we have an estimate of Φ, the prior becomes informative, and models the current distribution of Φ.

The prior on **w** is normally distributed:

$$P(\mathbf{w}) = \mathcal{N}(0, (\lambda\Lambda)^{-1}) \tag{5}$$

where Λ is a matrix encoding a bending energy model and λ is the global level of regularisation. λ has a fixed Gamma prior $P(\lambda) = Ga(s_0, c_0)$. A graphical description of the probabilistic model is given in Fig. 1.

Variational Bayes [7] is used to find analytic iterative updates for approximate posterior distributions of the model parameters. The posterior distribution for the noise precisions is approximated using Gamma distributions
$P(\Phi|\mathbf{y}) \sim \prod_l^L q(\phi_l)$, where $q(\phi_l) = Ga(a_l, b_l)$. The updates for the hyper-parameters are given below:

$$b_l = b_0 + \alpha_l N_{v,l} \tag{6}$$

$$\frac{1}{a_l} = \frac{1}{a_0} + \frac{\alpha_l}{2}(\mathbf{r}_l^T \mathbf{r}_l + \mathrm{Tr}(\boldsymbol{\Upsilon}^{-1} \mathbf{J}_l^T \mathbf{J}_l) \tag{7}$$

where $\mathbf{J}_l$ is the $N_c \times N_v$ Jacobian matrix of partial derivatives of the transformation parameters, calculated on the source image which has been weighted by the basis function $\mathbf{b}_{(l,:)}$. $\mathbf{r}_l = ((\mathbf{b}_{(l,:)})^{\frac{1}{2}} \circ \mathbf{k})$ where **k** is a $1 \times N_v$ vector containing the difference image $\mathbf{y} - \mathbf{t}(\mathbf{x}, \boldsymbol{\mu})$.

The approximate posterior distribution on **w** is normally distributed $P(\mathbf{w}|\mathbf{y}) \sim q(\mathbf{w}) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Upsilon}^{-1})$. The updates for these parameters are given as:

$$\boldsymbol{\Upsilon} = \sum_l^L \alpha_l \bar{\phi}_l \mathbf{J}_l^\mathsf{T} \mathbf{J}_l + \boldsymbol{\Lambda}\bar{\lambda} \tag{8}$$

$$\boldsymbol{\Upsilon}\boldsymbol{\mu}_{new} = \sum_l^L \alpha_l \bar{\phi}_l \mathbf{J}_l^\mathsf{T} (\mathbf{J}_l \boldsymbol{\mu}_{old} + \mathbf{r}_l) \tag{9}$$

where $\bar{\phi}_l$ is the expectation of the approximate noise distribution $\bar{\phi}_l = \mathrm{E}[\phi_l] = a_l b_l$ and similarly $\bar{\lambda} = \mathrm{E}[\lambda] = sc$. The approximate posterior distribution of λ is Gamma distributed, $q(\lambda) = Ga(\lambda; s, c)$. The hyper-parameter updates are as follows:

$$c = c_0 + \frac{N_c}{2} \tag{10}$$

$$\frac{1}{s} = \frac{1}{s_0} + \frac{1}{2}\left(\mathrm{Tr}(\boldsymbol{\Upsilon}^{-1}\boldsymbol{\Lambda}) + \boldsymbol{\mu}^T \boldsymbol{\Lambda} \boldsymbol{\mu}\right) \tag{11}$$

The trace terms in eq. 11 and eq. 7 involve a multiplication by the full inverse of $\boldsymbol{\Upsilon}$, which can be very large, and is non-sparse after inversion. For the purposes of eq. 11 and eq. 7 we treat $\boldsymbol{\Upsilon}$ as diagonalised with the only covariance between directions to ensure computational tractability. We iteratively apply these updates until convergence in order to fit the model parameters. Since VB is essentially an EM update, it is guaranteed to converge [8]. In practise convergence at each multi-resolution level generally occurs within 20 iterations.

We implemented our probabilistic registration model in FNIRT[1], a multi-resolution FFD registration tool, which has been previously demonstrated to provide good anatomical overlap in an independent study [9]. In our experiments we use 4 variants of FNIRT, standard FNIRT, FNIRT VB which uses a global noise model and FNIRT VB LN and FNIRT

[1] http://www.fmrib.ox.ac.uk/fsl/fnirt

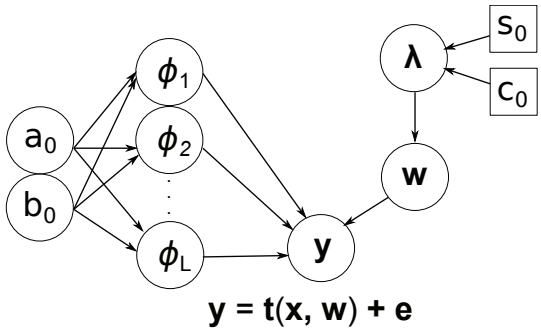


Fig. 1. A graphical description of the probabilistic dependencies of the registration model parameters. The variables in square boxes are constants, and the variables in circles are random variables.

VBLN2 which both use a local noise model. FNIRT VBLN uses 48 smooth Gaussian kernels as the basis set for the final multi-resolution level, and FNIRT VBLN2 uses 125. The basis set kernels have a full-width at half-maximum that is 50% of the size of the spacing between kernel centres. There is an approximately 2 fold increase in runtime for FNIRT VBLN, and a 4 fold increase for FNIRT VBLN2 over the FNIRT VB approach. However, as the majority of the computation is in calculating $\mathbf{J}_l^T\mathbf{J}_l$ which is separable, this could be made significantly faster if parallelised.

3. EXPERIMENTS AND RESULTS

We present an evaluation of inter-subject registration using 18 real 3D MR scans taken from the publicly available IBSR dataset[2]. An example slice and its corresponding segmentation is given in Fig. 2 a) and b). We use 23 of the annotated structures in these experiments, and the overlap of these registered structures can be treated as a gold standard of registration accuracy. We use the Dice score [10] to calculate overlap between structures. In our experiments each image was registered to every other, giving a total of 306 registrations.

All of the VB extensions of FNIRT inferred generally better mappings than the original FNIRT in terms of improved smoothness (Fig. 3 a)), less folding of the deformation field (Fig. 3 b)) and improved structural volume overlap (Table 1).

Furthermore, we find similar advantages when using a local noise model as opposed to a global one. As shown in figure 3, a local noise model infers smoother mappings, with much less image folding. The results for FNIRT VBLN2 are almost always diffeomorphic. We also find a slight, but not statistically significant improvement in structural overlap when using FNIRT VBLN2 over FNIRT VB (0.37%) across all structures. These improvements occur mostly in subcortical regions, e.g. the Amygdala (0.44%), although there is a slight decrease in overlap for cortical regions e.g. the cerebral cortex (-0.45%).

[2]http://www.cma.mgh.hardvard.edu/ibsr

Table 1. Improvement in the accuracy in structural overlap when using the various VB FNIRT approaches over FNIRT.

Method	Average improvement %	Maximum improvement (%)	Number of significant improvements ($p<0.05$)
FNIRT VB	1.83%	6.26% (Amygdala)	7/30
FNIRT VBLN	1.75%	6.62% (Pallidum)	8/30
FNIRT VBLN2	2.20%	6.70% (Amygdala)	11/30

FNIRT VBLN2 has an average 0.45% higher overlap across all structures compared to FNIRT VBLN, as well as a reduction in image folding. This indicates that the use of a more spatially localised noise model may infer more accurate and more appropriately constrained mappings.

An example map of inferred noise variance using FNIRT VBLN2 is given in Figure 2 c). The local noise model tends to lower the weighting in regions of complex image data which do not match well. For a volumetric registration approach, such as we use, this is most commonly in cortical regions. This decrease in image weighting does results in a small decrease in structural overlap in these regions, but it allows a smoother, more plausible mapping to be inferred. As the image data is weighted highly in subcortical regions, which are much easier to align using a volumetric registration method, these tend to get improved overlap.

4. CONCLUSIONS

In this work we have proposed the use of a generic probabilistic registration framework with a spatially varying noise model. We have shown that for the problem of inter-subject brain registration, this method infers mappings that are smoother, more plausible and more accurate.

Although we choose to implement this framework using a FFD, it could be similarly implemented with a variety of transformation models, including diffeomorphic strategies. Using a diffeomorphic transformation model would mean regional under-regularisation would no longer manifest itself as folding, however it will still result in the inference of overly complex and potentially inaccurate mappings

In future work we will investigate the use of different basis sets, with varying levels of smoothness. We will also look at the effects of inferring local noise on the estimated registration uncertainty.

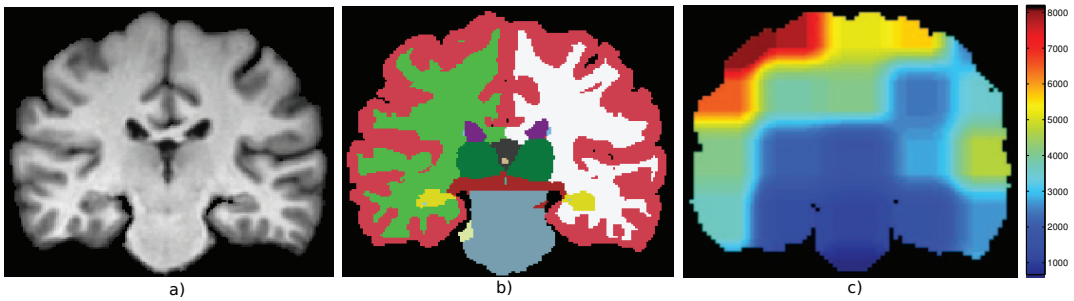


Fig. 2. a) An example image from the IBSR data. b) The associated segmentation map. c) The inferred local noise variance, corrected for spatial smoothness, from an example registration using FNIRT VBLN2 which uses a noise basis set of 125 Gaussian kernels.

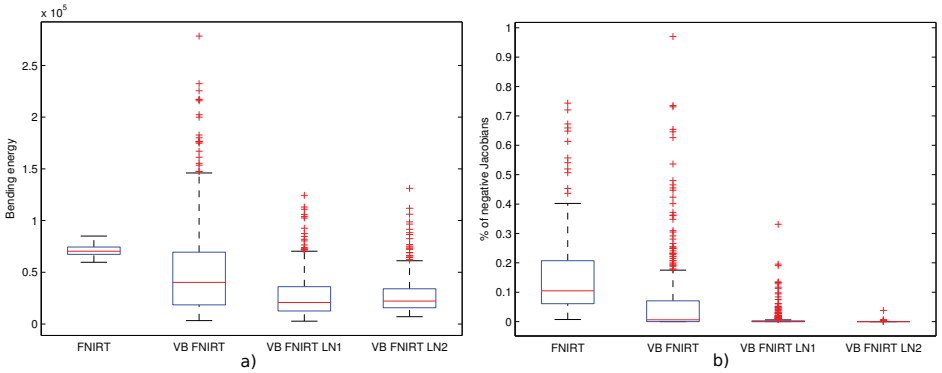


Fig. 3. a) Boxplot of the level of bending energy in the inferred mapping across the 306 registrations. Bending energy describes the smoothness of the mapping (lower values are smoother). b) Boxplot of the % of folded image voxels due to an inappropriate level of regularisation.

5. REFERENCES

[1] H. Lester, S. Arridge, K. Jansons, L. Lemieux, J. Hajnal, and A. Oatridge, "Non-linear registration with the variable viscosity fluid algorithm," in *IPMI*, vol. 1613 of *LNCS*, pp. 238–251. 1999.

[2] C. Davatzikos, "Spatial transformation and registration of brain images using elastically deformable models," *Computer Vision and Image Understanding*, vol. 66, no. 2, pp. 207–222, 1997.

[3] L. Tang, G. Hamarneh, and R Abugharbieh, "Reliability-driven, spatially-adaptive regularization for deformable registration," *WBIR: LNCS*, pp. 173–185, 2010.

[4] Y. Ou, A. Sotiras, N. Paragios, and C. Davatzikos, "Dramms: Deformable registration via attribute matching and mutual-saliency weighting," *Medical Image Analysis*, vol. 15, no. 4, pp. 622 – 639, 2011.

[5] I.J.A. Simpson, J.A. Schnabel, A.R. Groves, J.L.R. Andersson, and M.W. Woolrich, "Probabilistic inference of regularisation in non-rigid registration," *NeuroImage*, vol. 59, no. 3, pp. 2438–2451, 2012.

[6] P. Risholm, E. Samset, and W. Wells, "Bayesian Estimation of Deformation and Elastic Parameters in Non-rigid Registration," *WBIR: LNCS*, pp. 104–115, 2010.

[7] M.I. Jordan, Z. Ghahramani, T.S. Jaakkola, and L.K. Saul, "An introduction to variational methods for graphical models," *Machine learning*, vol. 37, no. 2, pp. 183–233, 1999.

[8] H. Attias, "A variational Bayesian framework for graphical models," *Advances in neural information processing systems*, vol. 12, no. 1-2, pp. 209–215, 2000.

[9] A. Klein, J. Andersson, B.A. Ardekani, J. Ashburner, B. Avants, M.C. Chiang, G.E. Christensen, D.L. Collins, J. Gee, P. Hellier, et al., "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration," *Neuroimage*, vol. 46, no. 3, pp. 786–802, 2009.

[10] L.R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.