Computational Anatomy and Machine Learning

John Ashburner

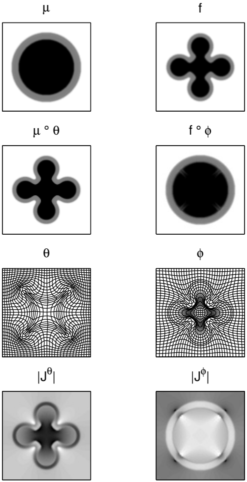
Wellcome Trust Centre for Neuroimaging
UCL Institute of Neurlogy
London, UK.

Nonlinearity of Shapes

According to David Mumford (Fields Medal, 1974):

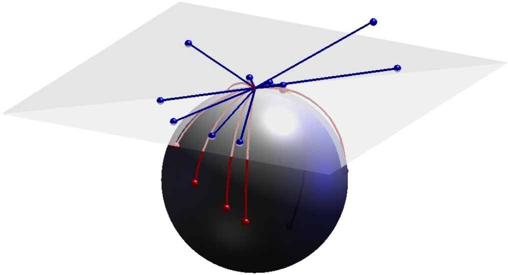
> *"Shapes are the ultimate non-linear sort of thing"*

Relative shapes can not be added and subtracted (ie, they are nonlinear). Instead, deformations should combined by composing them together.
Deformations that are smooth and invertible are known as *diffeomorphisms*, and form a mathematical *group*.



Linear Methods for Data on Manifolds

Dealing with non-Euclidean geometry. Conceptualise curved spaces as manifolds embedded in higher dimensions.



Beg, MF & Khan, A. *Computing an average anatomical atlas using LDDMM and geodesic shooting*. 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006. pp 1116–1119 (2006).

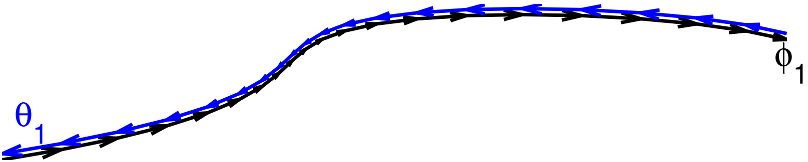
Large Deformations

We can consider a large deformation as the composition of a series of small deformations:

$$\varphi_1 = \left(\mathrm{Id} + \mathbf{v}_{t_{N-1}}\right) \circ \left(\mathrm{Id} + \mathbf{v}_{t_{N-2}}\right) \circ \ldots \circ \left(\mathrm{Id} + \mathbf{v}_{t_1}\right) \circ \left(\mathrm{Id} + \mathbf{v}_0\right)$$

The inverse of the deformation can be computed from:

$$\vartheta_1 = \left(\mathrm{Id} - \mathbf{v}_0\right) \circ \left(\mathrm{Id} - \mathbf{v}_{t_1}\right) \circ \ldots \circ \left(\mathrm{Id} - \mathbf{v}_{t_{N-2}}\right) \circ \left(\mathrm{Id} - \mathbf{v}_{t_{N-1}}\right)$$



Large Deformations

By modelling the trajectories as piecewise linear, distances can be computed by adding the distances from the small deformations:

$$d = \frac{1}{N} \sum_{n=0}^{N-1} ||\mathbf{Lv}_{t_n}||$$

If N approaches infinity (and we use small deformations of $Id + \frac{1}{N}\mathbf{v}_t$), the evolution of a deformation may be conceptualised as integrating the following equation:

$$\frac{d\varphi}{dt} = \mathbf{v}_t(\varphi)$$

Geodesic distances (from zero) are then measured by:

$$d = \int_{t=0}^{1} ||\mathbf{Lv}_t|| dt$$

LDDMM

Large Deformation Diffeomorphic Metric Mapping is an image registration algorithm that minimises the following:

$$\mathcal{E} = \frac{1}{2}\int_{t=0}^{1} ||\mathbf{L}\mathbf{v}_t||^2 dt + \frac{1}{2\sigma^2}||f - \mu\left(\varphi_1^{-1}\right)||^2$$

$$\text{where } \varphi_0 = \mathrm{Id},\ \frac{d\varphi}{dt} = \mathbf{v}_t\left(\varphi_t\right)$$

The first term minimises the squared distance measure of the deformations, whereas the second term simply minimises the difference between the warped template and the individual scan. The objective is to estimate a series of velocity fields ($\mathbf{v}_t$).

Beg, MF, Miller, MI, Trouvé, A & Younes, L. *Computing large deformation metric mappings via geodesic flows of diffeomorphisms*. International Journal of Computer Vision 61(2):139–157 (2005).

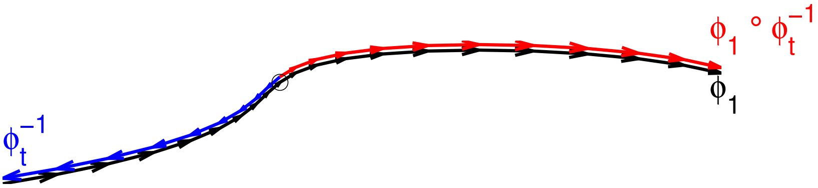
Change of Variables

When we warp images, we should usually account for expansion/contraction via a change of variables.

$$\int_{\mathbf{x}\in\varphi(\Omega)} f(\mathbf{x})d\mathbf{x} = \int_{\mathbf{x}\in\Omega} f(\varphi(\mathbf{x}))\det|(\mathbf{D}\varphi)(\mathbf{x})|d\mathbf{x}$$

where $(\mathbf{D}\varphi)(\mathbf{x})$ means the Jacobian of φ at $\mathbf{x}$.

LDDMM



The matching term of the objective function is:

$$\frac{1}{2\sigma^2}\int_{x\in\Omega}(f\circ\mathbf{x}-\mu\circ\varphi_1^{-1}\circ\mathbf{x})^2 d\mathbf{x}$$

This may be re-written (including a change of variables) as:

$$\frac{1}{2\sigma^2}\int_{x\in\Omega}\det|\mathbf{D}(\varphi_1\circ\varphi_t^{-1})\circ\mathbf{x}|(f\circ\varphi_1\circ\varphi_t^{-1}\circ\mathbf{x}-\mu\circ\varphi_t^{-1}\circ\mathbf{x})^2 d\mathbf{x}$$

This allows the derivatives of the matching term to be computed at any time point.

Euler-Lagrange Equations

At the solution, the derivatives of the objective function are zero, which means that the velocity at any time point is given by:

$$\mathbf{L}^{\dagger}\mathbf{L}\mathbf{v}_t = \frac{1}{\sigma^2}\det|\mathbf{D}(\varphi_1 \circ \varphi_t^{-1})|(\nabla(\mu \circ \varphi_t^{-1}))(\mu \circ \varphi_t^{-1} - f \circ \varphi_1 \circ \varphi_t^{-1})$$

If we introduce something that we'll call initial momentum:

$$\mathbf{u}_0 = \mathbf{L}^{\dagger}\mathbf{L}\mathbf{v}_0 = \frac{1}{\sigma^2}\det|\mathbf{D}\varphi_1|(\nabla\mu)(\mu - f \circ \varphi_1)$$

we see that the velocity at any time point is determined by the initial momentum (or velocity), according to:

$$\mathbf{v}_t = \left(\mathbf{L}^{\dagger}\mathbf{L}\right)^{-1}\left(\det|\mathbf{D}\varphi_t^{-1}|(\mathbf{D}\varphi_t^{-1})^T(\mathbf{u}_0 \circ \varphi_t^{-1})\right)$$

Geodesic Shooting

This all means that we do not need to estimate a series of velocity fields. We just need to estimate an initial velocity ($\mathbf{v}_0$), from which we compute the initial momentum by $\mathbf{u}_0 = \mathbf{L}^\dagger \mathbf{L} \mathbf{v}_0$.

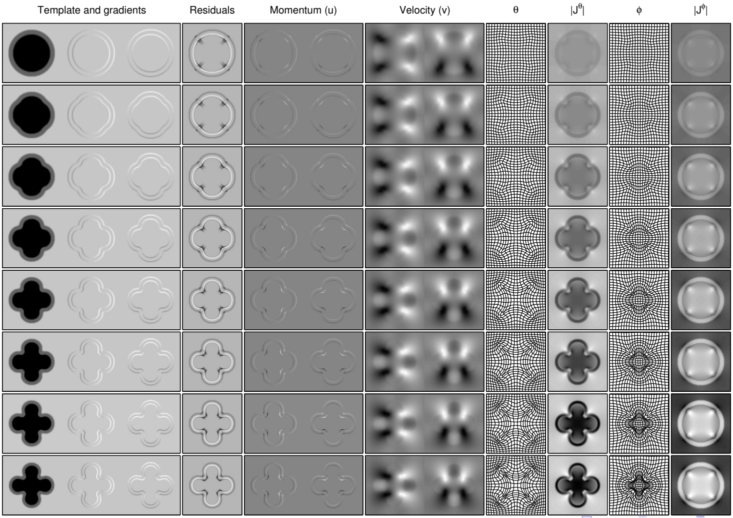
We set the deformation at time 0 to an identity transform ($\varphi_0 = Id$), and then evolve the following dynamical system for unit time:

$$\frac{d\varphi}{dt} = \mathbf{v}_t(\varphi_t)$$

$$\mathbf{v}_t = \left(\mathbf{L}^\dagger \mathbf{L}\right)^{-1} \left(\det |\mathbf{D}\varphi_t^{-1}| (\mathbf{D}\varphi_t^{-1})^T (\mathbf{u}_0 \circ \varphi_t^{-1})\right)$$

Younes, L, Arrate, F & Miller, MI. *Evolutions equations in computational anatomy*. Neuroimage 45(1S1):40–50 (2009).

Evolution


Template and gradients
Residuals
Momentum (u)
Velocity (v)
θ
|J^θ|
ϕ
|J^ϕ|

“Scalar Momentum”

Remember that

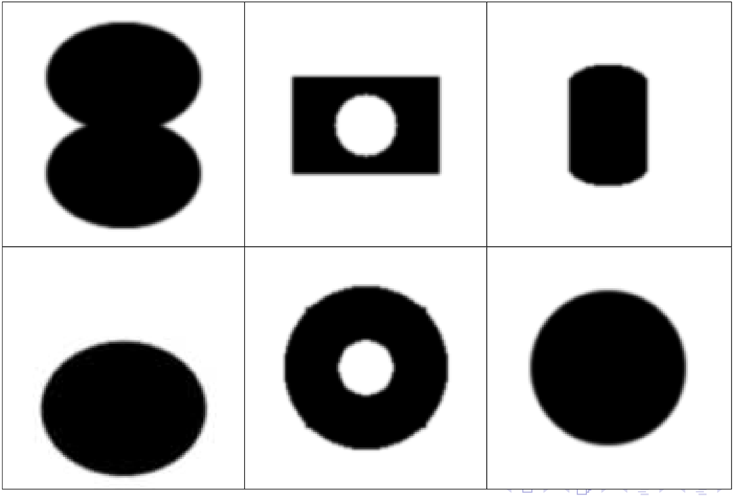
$$\mathbf{u}_0 = \frac{1}{\sigma^2} \det |\mathbf{D}\boldsymbol{\varphi}_1| (\nabla \mu)(\mu - f \circ \boldsymbol{\varphi}_1)$$

If a population of subjects are all aligned with the same template image, $\frac{1}{\sigma^2}(\nabla \mu)$ will be the same for all subjects. Deviations from the template are encoded by the residuals, $\det |\mathbf{D}\boldsymbol{\varphi}_1|(\mu - f \circ \boldsymbol{\varphi}_1)$. This is a scalar field, and in principle is all that is needed (along with the template) to reconstruct the original images.

Singh, Fletcher, Preston, Ha, King, Marron, Wiener & Joshi (2010). *Multivariate Statistical Analysis of Deformation Momenta Relating Anatomical Shape to Neuropsychological Measures*. T. Jiang et al. (Eds.): MICCAI 2010, Part III, LNCS 6363, pp. 529–537, 2010.

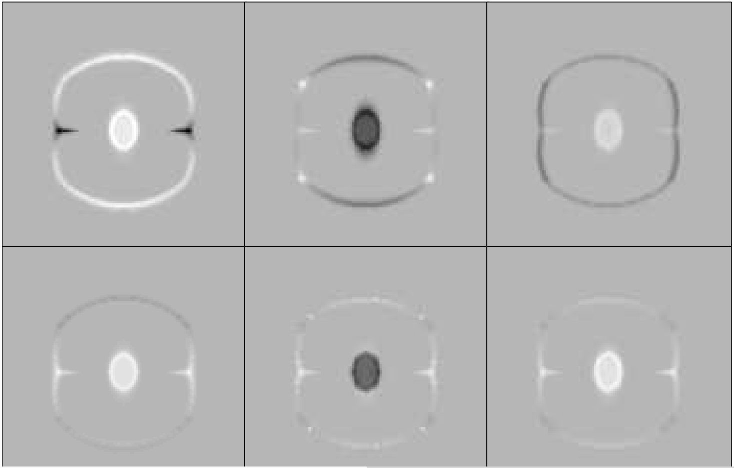
Example Images

Some example images.



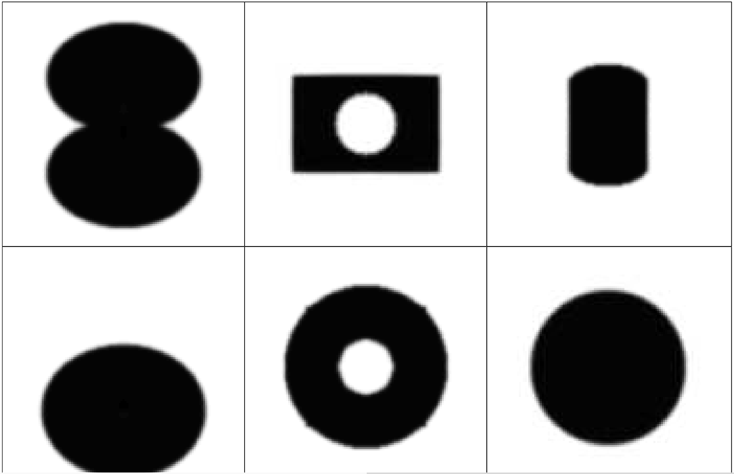
Scalar Momentum

Scalar momentum after aligning the example images to a common template.



Reconstructed Images

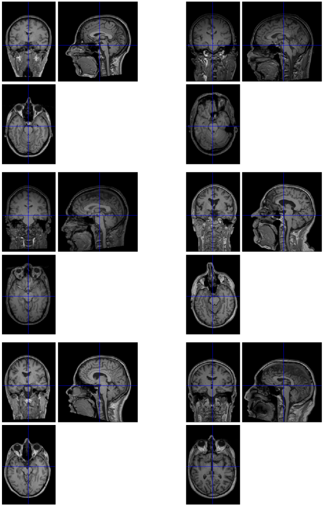
Images reconstructed using just the template and scalar momentum.



IXI Data

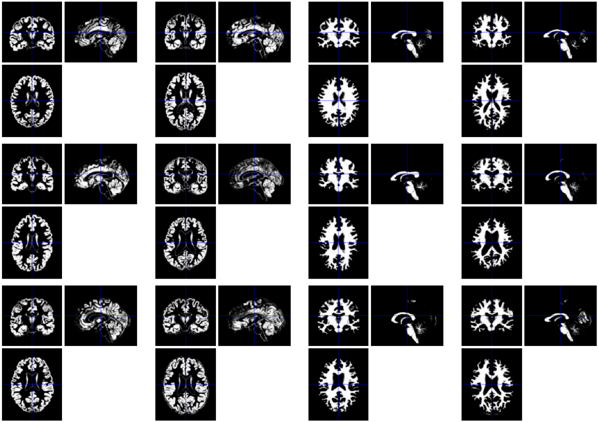
Used 550 T1w brain MRI from IXI (**I**nformation e**X**traction from **I**mages) dataset.
http://www.brain-development.org/
Data from three different hospitals in London:

- Hammersmith Hospital using a Philips 3T system
- Guy's Hospital using a Philips 1.5T system
- Institute of Psychiatry using a GE 1.5T system



Grey and White Matter

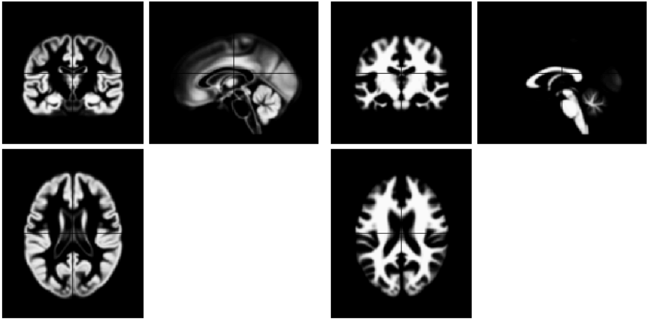
Segmented into GM and WM. Approximately aligned via rigid-body.



Ashburner, J & Friston, KJ. *Unified segmentation*. NeuroImage 26(3):839–851 (2005).

Diffeomorphic Alignment

All GM and WM were diffeomorphically aligned to their common average-shaped template.

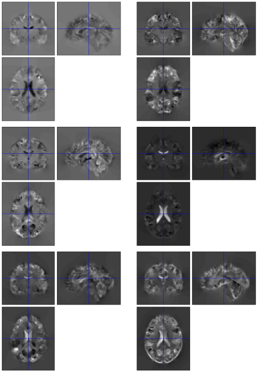


Ashburner, J & Friston, KJ. *Diffeomorphic registration using geodesic shooting and Gauss-Newton optimisation*. NeuroImage 55(3):954–967 (2011).
Ashburner, J & Friston, KJ. *Computing average shaped tissue probability templates*. NeuroImage 45(2):333–341 (2009).

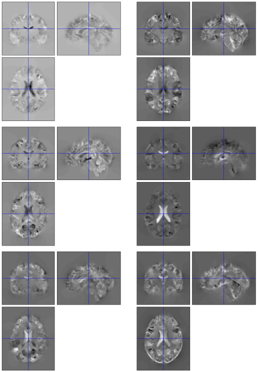
Volumetric Features

A number of features were used for pattern recognition.
Firstly, two features relating to relative volumes.
Initial velocity divergence is similar to logarithms of Jacobian determinants.

Jacobian Determinants

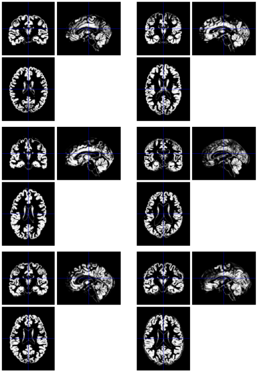


Initial Velocity Divergence

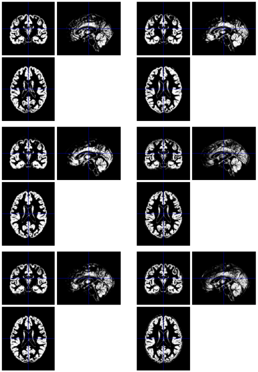


Grey Matter Features

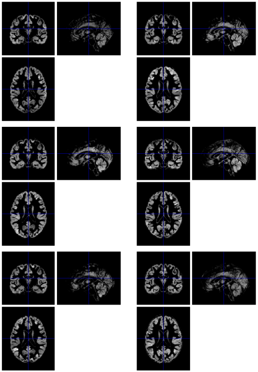
Rigidly Registered GM



Nonlinearly Registered GM



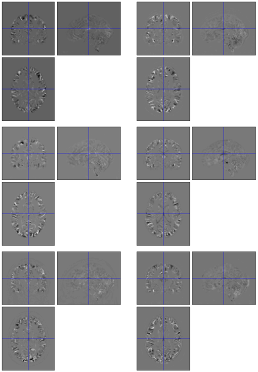
Registered and Jacobian Scaled GM



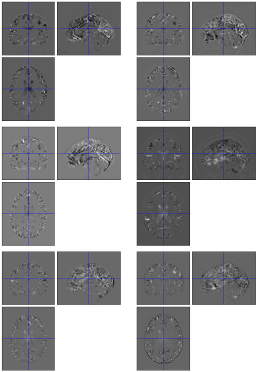
"Scalar Momentum" Features

"Scalar momentum" actually has two components because GM was matched with GM and WM was matched with WM.

First Momentum Component

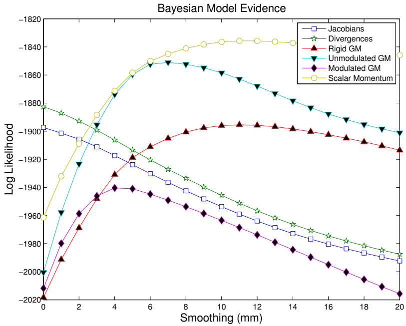


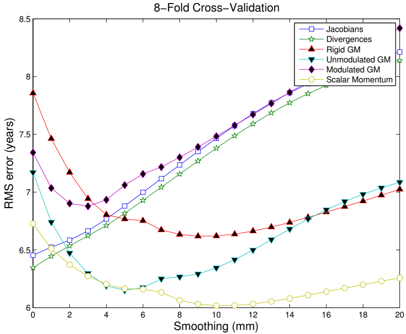
Second Momentum Component



Age Regression

Linear Gaussian Process Regression to predict subject ages.

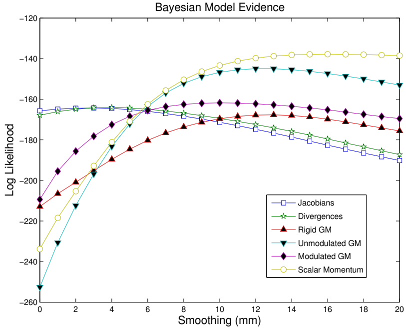


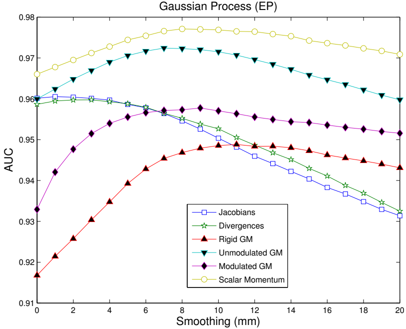


Rasmussen, CE & Williams, CKI. *Gaussian processes for machine learning*. Springer (2006).

Sex Classification

Linear Gaussian Process Classification (EP) to predict sexes.

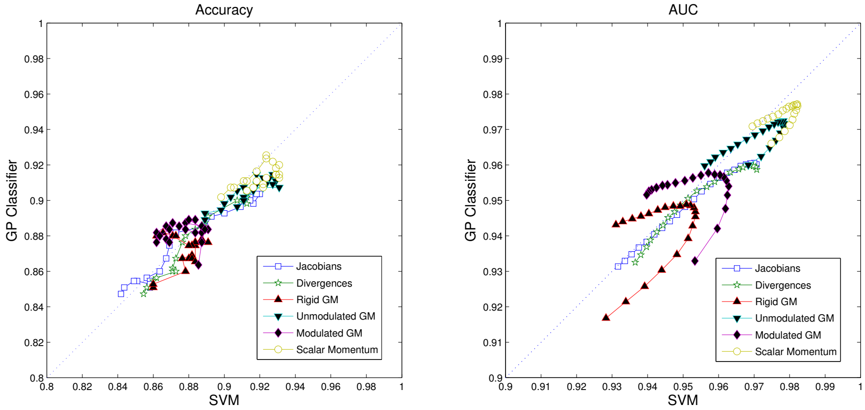




Rasmussen, CE & Williams, CKI. *Gaussian processes for machine learning*. Springer (2006).

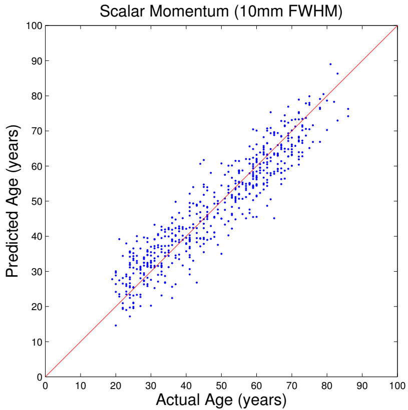
Sex Classification

Linear SVM versus Gaussian Process Classification (EP).

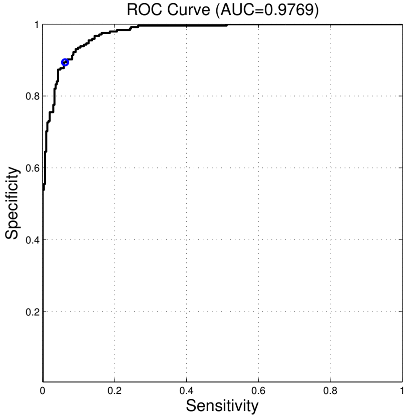


Predictive Accuracies

Age



Sex



Conclusions

- Scalar momentum (with about 10mm smoothing) appears to be a useful feature set.
- Jacobian-scaled warped GM is surprisingly poor.
- Amount of spatial smoothing makes a big difference.
- Further dependencies on the details of the registration still need exploring.