High-dimensional Multi-modal Image Registration

P. Lorenzen and S. Joshi

UNC Departments of Computer Science and Radiation Oncology

Abstract. This paper presents a Bayesian framework for generating inter-subject high-dimensional transformations between two multi-modal image sets of the brain. In this framework, the estimated transformations are generated by using the maximal information about the underlying neuroanatomy present in each of the different modalities. This modality independent registration framework is achieved by using the Bayesian paradigm and jointly estimating the posterior densities associated with the multi-modal image sets and the high dimensional registration transformation mapping the two subjects. The methods presented do not assume that the same modalities were used to image the two subjects. To maximally use the information present in all the modalities, relative entropy (or Kullback Leibler divergence) between the estimated posteriors is minimized to estimate the registration. The high-dimensional registration is constrained to be diffeomorphic by using the large deformation fluid formulation.

We also show that the general framework optimally unifies the processes of segmentation and registration of human brain anatomy. Given a set of multi-modal intensity images of an individual subject and a geometric atlas prior, we jointly estimate the posterior that represents the structure of that subject's neuroanatomy and the transformation that maps the space of the atlas prior to the space of the estimated posterior, capturing high-dimensional local variability.

Key Words: Multi-modal image registration, image segmentation, information theory, medical image analysis, computational anatomy.

1 Introduction

Modern imaging techniques provide an array of imaging modalities which enable the acquisition of complementary information representing underlying neuroanatomy. Certain imaging sensors, such as MRI, produce inherently registered multi-modal images. In the case were the images are derived from disparate sensors with differing geometry (e.g. CT and MRI), rigid registration is required to bring the modalities into correspondence. Multi-modal rigid image registration has received much attention and a thorough review and comparison of methods is presented in [18].

Inter-subject high-dimensional image registration has proven to be a powerful tool in understanding brain anatomy [5,3,7]. The book *Brain Warping* [15] discusses the application of high-dimensional image registration to the understanding of neuroanatomy.

Although inter-subject high-dimensional image registration has received much attention [13,10,6,14], no algorithm exists that uses multi-modal image sets of subjects to estimate the registration transformation. Most image registration techniques use a single imaging modality for the generation of high-dimensional transformations between images of two subjects. An algorithm for cross modality inter-subject high-dimensional image registration that map image intensities from one modality to another has been proposed [8]. This algorithm only works with a single image from each subject.

In this paper, we present a framework based on the Bayesian paradigm that enables us to generate inter-subject high-dimensional transformations given two acquired sets of multi-modal images. We assume that each subject has associated with it a set of imaging modalities (e.g. PET, MRI, CT, etc.). We develop an algorithm in which the maximal amount of information present in the different modalities is used to generate the transformation. Throughout this paper we assume that, for each subject, the multi-modal images of that subject are registered to one another.

The framework is based on the assumption that human brain anatomy consists of finitely enumerable structures such as grey matter (GM), white matter (WM), and cerebrospinal fluid (CSF). These structures present with varying radiometric intensity values across disparate image modalities. Given multi-modal image sets representing two studies, we jointly estimate, for each subject, the posterior distributions associated with each of the structures along with the diffeomorphic high-dimensional registration map that relates the coordinate spaces of the two subjects. The Kullback-Leibler divergence is used as a metric for the posterior densities to estimate the transformation. The use of the posterior probability densities provides an image intensity independent approach to image registration.

In this paper, we also show that the framework described above optimally unifies the processes of segmentation and registration of human brain anatomy. Given a set of multi-modal intensity images representing an individual subject, and a geometric atlas prior, the same framework can be used to jointly estimate the posterior that represents the structure of that subject's neuroanatomy and the transformation that maps the space of the atlas prior to the space of the estimated posterior.

The remainder of the paper is organized in the following manner. Section 2 presents the high-dimensional multi-modal image registration framework describing the posterior and transformation estimation. In Section 3, this framework is applied to the problem of atlas-based image registration and segmentation. In Section 4, we evaluate the performance of the segmentation algorithm.

2 High-dimension Multi-modal Image Registration

We first consider the problem of registering two acquired sets of multi-modal images. We assume that the underlying neuroanatomy, represented in these images, consists of N separate structures (or classes), c_i, $i = 1, \cdots, N$. Let subject $\mathcal{A}$ be characterized by m multi-modal images so that $\bar{I}_{\mathcal{A}}(x) \in I\!R^m$ and subject $\mathcal{B}$ be characterized by n multi-modal images so that $\bar{I}_{\mathcal{B}}(x) \in I\!R^n$, where $c_i(x)$ is the class associated spatial position $x = [x_1, x_2, x_3]^T \in \Omega_{A,B}$, respectively. From the multi-modal images, $\bar{I}_{\mathcal{A}}(x)$ and $\bar{I}_{\mathcal{B}}(x)$, for each class c_i we jointly estimate the posterior distributions $p_{\mathcal{A}}(c_i(x))|\bar{I}_{\mathcal{A}})$ and $p_{\mathcal{B}}(c_i(x)|\bar{I}_{\mathcal{B}})$ along with the registration map, $h(x)$, that maps the space of subject $\mathcal{A}$, $\Omega_{\mathcal{A}} \subset I\!R^3$, into the space of subject $\mathcal{B}$, $\Omega_{\mathcal{B}} \subset I\!R^3$. This method is independent of the choice of m and n. Optimal inter-subject multi-modal image registration is estimated by an alternating iterative algorithm which is motivated by an expectation maximization method used in [17,12]. Our algorithm interleaves the estimation of the posteriors associated with studies $\mathcal{A}$ and $\mathcal{B}$ and the estimation of the registration map that maps $\Omega_{\mathcal{A}}$ to $\Omega_{\mathcal{B}}$.

Following [17], for each class c_i the associated data likelihood is modeled as a normal distribution with mean, μ_i, and covariance, Σ_i. That is, the probability that a voxel at spatial location x having m-dimensional intensity values $\bar{I}(x)$ belongs to class c_i, is given by

$$p(\bar{I}(x)|c_i(x), \mu_i, \Sigma_i) = \frac{1}{\sqrt{(2\pi)^N |\Sigma_i|}} e^{-\frac{1}{2}(\bar{I}(x)-\mu_i)^T \Sigma_i^{-1} (\bar{I}(x)-\mu_i)},$$

where $c_i(x)$ is the class associated with the voxel at position $x = [x_1, x_2, x_3]^T \in I\!R^3$.

Given a transformation, $h(x)$, mapping the coordinate systems $\Omega_{\mathcal{A}}$ and $\Omega_{\mathcal{B}}$ and current estimates μ_i, Σ_i, the posterior densities of the two studies can be associated with each other by using the posterior of one as the prior for the other. Using Bayes's Rule, the posterior of subject $\mathcal{A}$ becomes,

$$p_{\mathcal{A}}(c_i(x)|\bar{I}_{\mathcal{A}}) = \frac{p_{\mathcal{A}}(\bar{I}_{\mathcal{A}}(x)|c_i(x), \mu_i^{\mathcal{A}}, \Sigma_i^{\mathcal{A}}) p_{\mathcal{B}}(c_i(h(x))|\bar{I}_{\mathcal{B}})}{\sum_{k=1}^{N} p_{\mathcal{A}}(\bar{I}_{\mathcal{A}}(x)|c_k(x), \mu_k^{\mathcal{A}}, \Sigma_k^{\mathcal{A}}) p_{\mathcal{B}}(c_k(h(x))|\bar{I}_{\mathcal{B}})},$$

where $p_{\mathcal{B}}(c_i(h(x))|\bar{I}_{\mathcal{B}})$ is the transformed posterior of $\mathcal{B}$ viewed as the prior for $\mathcal{A}$. Similarly, the posterior, under the transformation $h^{-1}(x)$, of subject $\mathcal{A}$ is viewed as the prior for subject $\mathcal{B}$ is given by,

$$p_{\mathcal{B}}(c_i(x)|\bar{I}_{\mathcal{B}}) = \frac{p_{\mathcal{B}}(\bar{I}_{\mathcal{B}}(x)|c_i(x), \mu_i^{\mathcal{B}}, \Sigma_i^{\mathcal{B}}) p_{\mathcal{A}}(c_i(h^{-1}(x))|\bar{I}_{\mathcal{A}})}{\sum_{k=1}^{N} p_{\mathcal{B}}(\bar{I}_{\mathcal{B}}(x)|c_k(x), \mu_k^{\mathcal{B}}, \Sigma_k^{\mathcal{B}}) p_{\mathcal{A}}(c_k(h^{-1}(x))|\bar{I}_{\mathcal{A}})}.$$

Having defined the posteriors, the parameters $\mu_i^{\mathcal{A}}, \mu_i^{\mathcal{B}}, \Sigma_i^{\mathcal{A}}$, and $\Sigma_i^{\mathcal{B}}$ are updated via their expected values:

$$\mu_i^{\mathcal{A}} = \frac{\int_{\Omega_{\mathcal{A}}} p_{\mathcal{A}}(c_i(x)|\bar{I}_{\mathcal{A}}) \bar{I}_{\mathcal{A}}(x) dx}{\int_{\Omega_{\mathcal{A}}} p_{\mathcal{A}}(c_i(x)|\bar{I}_{\mathcal{A}}) dx}, \quad \mu_i^{\mathcal{B}} = \frac{\int_{\Omega_{\mathcal{B}}} p_{\mathcal{B}}(c_i(x)|\bar{I}_{\mathcal{B}}) \bar{I}_{\mathcal{B}}(x) dx}{\int_{\Omega_{\mathcal{B}}} p_{\mathcal{B}}(c_i(x)|\bar{I}_{\mathcal{B}}) dx},$$

$$\Sigma_i^{\mathcal{A}} = \frac{\int_{\Omega_{\mathcal{A}}} p_{\mathcal{A}}(c_i(x)|\bar{I}_{\mathcal{A}}) \left[(\bar{I}_{\mathcal{A}}(x) - \mu_i)(\bar{I}_{\mathcal{A}}(x) - \mu_i)^T\right] dx}{\int_{\Omega_{\mathcal{A}}} p_{\mathcal{A}}(c_i(x)|\bar{I}_{\mathcal{A}}) dx},$$

and

$$\Sigma_i^{\mathcal{B}} = \frac{\int_{\Omega_{\mathcal{B}}} p_{\mathcal{B}}(c_i(x)|\bar{I}_{\mathcal{B}}) \left[(\bar{I}_{\mathcal{B}}(x) - \mu_i)(\bar{I}_{\mathcal{B}}(x) - \mu_i)^T\right] dx}{\int_{\Omega_{\mathcal{B}}} p_{\mathcal{B}}(c_i(x)|\bar{I}_{\mathcal{B}}) dx}.$$

Using the Kullback-Leibler divergence, $D_{KL}(\cdot,\cdot)$, the transformation is estimated via the minimization

$$\hat{h}(x) = \arg\min_{h(x)} \frac{1}{|\Omega_{\mathcal{A}}|} \int D_{KL}(p_{\mathcal{A}}(\bar{c}(x)|\bar{I}_{\mathcal{A}}), p_{\mathcal{B}}(\bar{c}(h(x))|\bar{I}_{\mathcal{B}}))dx.$$

The transformation $h(x)$ is constrained to be a diffeomorphism following the fluid formulation of [2]. The whole process is repeated until convergence.

2.1 Kullback-Leibler Divergence

As a distance measure between the distributions $p_{\mathcal{A}}(\bar{c}(x)|\bar{I}_{\mathcal{A}})$ and $p_{\mathcal{B}}(\bar{c}(h(x))|\bar{I}_{\mathcal{B}})$, Kullback-Leibler divergence (relative entropy), is defined as

$$\begin{aligned} &D_{KL}(p_{\mathcal{A}}(\bar{c}(x)|\bar{I}_{\mathcal{A}}), p_{\mathcal{B}}(\bar{c}(h(x))|\bar{I}_{\mathcal{B}})) = \\ &\sum_{i=1}^{N} p_{\mathcal{A}}(\bar{c}(x)|\bar{I}_{\mathcal{A}}) \log \frac{p_{\mathcal{A}}(\bar{c}(x)|\bar{I}_{\mathcal{A}})}{p_{\mathcal{B}}(\bar{c}(h(x))|\bar{I}_{\mathcal{B}})} = \\ &\qquad \mathcal{E}\left[\log \frac{p_{\mathcal{A}}(\bar{c}(x)|\bar{I}_{\mathcal{A}})}{p_{\mathcal{B}}(\bar{c}(h(x))|\bar{I}_{\mathcal{B}})}\right]. \end{aligned}$$

From an information theoretic viewpoint [4], this measure can be interpreted as the inefficiency of assuming that $p_{\mathcal{B}}(\bar{c}(h(x))|\bar{I}_{\mathcal{B}})$ is true when $p_{\mathcal{A}}(\bar{c}(x)|\bar{I}_{\mathcal{A}})$ is true. That is, if we have a model expressed as a probability density $p_{\mathcal{B}}(\bar{c}(h(x))|\bar{I}_{\mathcal{B}})$, we can then measure how far an observation, also expressed by a probability density, $p_{\mathcal{A}}(\bar{c}(x)|\bar{I}_{\mathcal{A}})$, deviates from $p_{\mathcal{B}}(\bar{c}(h(x))|\bar{I}_{\mathcal{B}})$ using Kullback-Leibler divergence. The second equality above simply highlights the notion that Kullback-Leibler divergence is the expected logarithm of the likelihood ratio of $p_{\mathcal{A}}(\bar{c}(x)|\bar{I}_{\mathcal{A}})$ to $p_{\mathcal{B}}(\bar{c}(h(x))|\bar{I}_{\mathcal{B}})$.

Kullback-Leibler divergence is non-negative, and is zero if and only if the probabilities are equal. However, Kullback-Leibler divergence is not a true distance in the strict mathematical sense since it is not symmetric. We are also investigating the use of a symmetric generalized form of the Kullback-Leibler divergence motivated by the work of Renyi [16]. Such a symmetric form would eliminate the asymmetry in the divergence computation.

2.2 Registration

In this section, we discuss the process of registering the posterior of subject $\mathcal{B}$ to the posterior of subject $\mathcal{A}$. We estimate the transformation $h(x)$ that maps the space of the prior, $\Omega_{\mathcal{A}}$, into that of the posterior, $\Omega_{\mathcal{B}}$, via the following average Kullback-Leibler minimization:

$$\hat{h}(x) = \arg\min_{h(x)} \frac{1}{|\Omega_A|} \int_{\Omega_{\mathcal{A}}} D_{KL}(p_{\mathcal{A}}(\bar{c}(x)|\bar{I}_{\mathcal{A}}), p_{\mathcal{B}}(\bar{c}(h(x))|\bar{I}_{\mathcal{B}}))dx$$

where $h(x) \in I\!R^3$ is the high-dimensional vector field describing local variation between the prior spatial coordinate system and the posterior spatial coordinate system and $D_{KL}(\cdot,\cdot)$ is the Kullback-Leibler divergence.

We apply the theory of large deformation fluid diffeomorphisms of [9,11] to require that the deformation, $h(x)$, be the solution of the Lagrangian o.d.e.,

$$\frac{d}{dt}h(x,t) = v(h(x,t),t).$$

The optimal transformation is found by estimating the velocity field following the minimization:

$$\hat{v}(\cdot,\cdot) = \arg\min_{v(\cdot)} \frac{1}{|\Omega_{\mathcal{A}}|} \int_{\Omega_{\mathcal{A}}} D_{KL}(p_{\mathcal{A}}(\bar{c}(x)|\bar{I}_{\mathcal{A}}), p_{\mathcal{B}}(\bar{c}(h(x))|\bar{I}_{\mathcal{B}}))dx + \int_0^T \int_{\Omega_{\mathcal{A}}} ||Lv(x,t)||^2 dxdt$$

where the second term is the fluid formulated regularization. Following Christensen's greedy algorithm for propagating templates [1], we compute the variation of the average Kullback-Leibler divergence term,

$$\frac{\partial}{\partial h}\frac{1}{|\Omega_{\mathcal{A}}|} \int_{\Omega_{\mathcal{A}}} D_{KL}(p_{\mathcal{A}}(\bar{c}(x)|\bar{I}_{\mathcal{A}}), p_{\mathcal{B}}(\bar{c}(h(x))|\bar{I}_{\mathcal{B}}))dx.$$

To compute this variation, we approximate $\log x$ near $x = 1$ using the Maclaurin series expansion. The point $x = 1$ is chosen for the expansion as it is the desired ratio of $\frac{p_{\mathcal{A}}(c_i(x)|\bar{I}_{\mathcal{A}})}{p_{\mathcal{B}}(c_i(h(x))|\bar{I}_{\mathcal{B}})}$ for ideal registration. The logarithm can be approximated up to the second order by

$$\log x \approx \frac{1}{\ln 2}(-\frac{1}{2}x^2 + 2x - \frac{3}{2}).$$

Thus, the second order approximation to Kullback-Leibler divergence becomes

$$D_{KL}(p_{\mathcal{A}}(\bar{c}(x)|\bar{I}_{\mathcal{A}}), p_{\mathcal{B}}(\bar{c}(h(x))|\bar{I}_{\mathcal{B}})) \approx -\frac{1}{\ln 2}\sum_{i=1}^{N} p_{\mathcal{A}}(c_i(x)|\bar{I}_{\mathcal{A}})$$

$$\cdot \left[-\frac{1}{2} \left(\frac{p_{\mathcal{B}}(c_i(h(x))|\bar{I}_{\mathcal{B}})}{p_{\mathcal{A}}(c_i(x)|\bar{I}_{\mathcal{A}})} \right)^2 + 2 \left(\frac{p_{\mathcal{B}}(c_i(h(x))|\bar{I}_{\mathcal{B}})}{p_{\mathcal{A}}(c_i(x)|\bar{I}_{\mathcal{A}})} \right) - \frac{3}{2} \right] =$$

$$-\frac{1}{\ln 2} \sum_{i=1}^{N} \left[-\frac{1}{2} \frac{p_{\mathcal{B}}(c_i(h(x))|\bar{I}_{\mathcal{B}})^2}{p_{\mathcal{A}}(c_i(x)|\bar{I}_{\mathcal{A}})} + 2p_{\mathcal{B}}(c_i(h(x))|\bar{I}_{\mathcal{B}}) - \frac{3}{2} \right].$$

Therefore, we compute the variation of the average Kullback-Leibler divergence term as follows,

$$\frac{\partial}{\partial h} \frac{1}{|\Omega_{\mathcal{A}}|} \int_{\Omega_{\mathcal{A}}} D_{KL}(p_{\mathcal{A}}(\bar{c}(x)|\bar{I}_{\mathcal{A}}), p_{\mathcal{B}}(\bar{c}(h(x))|\bar{I}_{\mathcal{B}}))dx \approx$$

$$-\frac{1}{\ln 2} \frac{1}{|\Omega_{\mathcal{A}}|} \frac{\partial}{\partial h} \int_{\Omega_{\mathcal{A}}} \sum_{i=1}^{N} \left[-\frac{1}{2} \frac{p_{\mathcal{B}}(c_i(h(x))|\bar{I}_{\mathcal{B}})^2}{p_{\mathcal{A}}(c_i(x)|\bar{I}_{\mathcal{A}})} + 2p_{\mathcal{B}}(c_i(h(x))|\bar{I}_{\mathcal{B}}) - \frac{3}{2} \right] =$$

$$\frac{1}{\ln 2} \frac{1}{|\Omega_{\mathcal{A}}|} \int_{\Omega_{\mathcal{A}}} \sum_{i=1}^{N} \left[\frac{p_{\mathcal{B}}(c_i(h(x))|\bar{I}_{\mathcal{B}})}{p_{\mathcal{A}}(c_i(x)|\bar{I}_{\mathcal{A}})} - 2 \right] \nabla p_{\mathcal{B}} \Big|_{c_i(h(x))}^{T}.$$

The velocity field at each iteration is updated by solving the p.d.e.,

$$Lv(x,t) = \frac{\partial}{\partial h} \int_{\Omega_{\mathcal{A}}} D_{KL}(p_{\mathcal{A}}(\bar{c}(x)|\bar{I}_{\mathcal{A}}), p_{\mathcal{B}}(\bar{c}(h(x))|\bar{I}_{\mathcal{B}}))dx,$$

where $L = a\nabla^2 + b\nabla \cdot \nabla + cI$ is the Navier-Stokes operator.

3 Atlas-based Image Segmentation

The above framework for multi-modal image registration can be applied to optimally unify atlas-based image segmentation and registration. The problem of high-dimensional multi-modal image registration reduces to the problem of atlas-based image segmentation where subject $\mathcal{B}$, in Section 2, is replaced by an atlas prior with known probability densities.

Given a multi-modal image data set, $\bar{I}(x)$, and a geometric prior $p(c_i(x))$, image segmentation lends itself to the Bayesian paradigm in which inference about a class, c_i, of each of the voxels is based on the posterior distribution $p(c_i(x)|\bar{I})$ given by Bayes's Rule. To apply Bayes's Rule, the prior and likelihood need to be in the same coordinate system, and hence, a registration map is required. If the prior and likelihood are not in the same coordinate system then the estimation of the posterior is subject to overlap error due to spatial misalignment. Working within the probabilistic framework defined in Section 2 unifies the approach of estimating both the posterior and the registration map that matches the coordinate spaces of the prior and posterior.

As described in Section 2, the data likelihood associated with each of the N classes is modeled as a normal distribution with mean, μ_i, and covariance, Σ_i. That is, the probability that a voxel at spatial location x having m-dimensional intensity values $\bar{I}(x)$ belongs to class c_i, is given by

$$p(\bar{I}(x)|c_i(x),\mu_i,\Sigma_i) = \frac{1}{\sqrt{(2\pi)^N|\Sigma_i|}} e^{-\frac{1}{2}(\bar{I}(x)-\mu_i)^T\Sigma_i^{-1}(\bar{I}(x)-\mu_i)}.$$

Given a registration transformation, $h(x)$, mapping the coordinate system of the atlas, Ω_{atlas}, into that of the study, Ω_{study}, the posterior distribution, $p_{study}(c_i(x)|\bar{I}_{study})$, is given by Bayes's Rule:

$$p_{study}(c_i(x)|\bar{I}_{study}) = \frac{p_{study}(\bar{I}_{study}(x)|c_i(x),\mu_i^{study},\Sigma_i^{study})p_{atlas}(c_i(h(x)))}{\sum_{k=1}^{N} p_{study}(\bar{I}_{study}(x)|c_k(x),\mu_k^{study},\Sigma_k^{study})p_{atlas}(c_k(h(x)))}.$$

The atlas-based image segmentation algorithm interleaves the estimation of the posterior, $p_{study}(c_i(x)|\bar{I}_{study})$ by using a fixed atlas with the estimation of the registration map $h(x)$. In this algorithm the parameter estimation is considered only for the study. Given $h(x)$, the parameters of the likelihood are estimated via the empirical expectations:

$$\mu_i^{study} = \frac{\int_{\Omega_{study}} p_{study}(c_i(x)|\bar{I}_{study})\bar{I}_{study}(x)dx}{\int_{\Omega_{study}} p_{study}(c_i(x)|\bar{I}_{study})dx}$$

and

$$\Sigma_i^{study} = \frac{\int_{\Omega_{study}} p_{study}(c_i(x)|\bar{I}_{study})\left[(\bar{I}_{study}(x)-\mu_i^{study})(\bar{I}_{study}(x)-\mu_i^{study})^T\right]dx}{\int_{\Omega_{study}} p_{study}(c_i(x)|\bar{I}_{study})dx}$$

Given these parameter estimates, the registration map is estimated to minimize the Kullback-Leibler divergence between the posterior and prior,

$$\hat{h}(x) = \arg\min_{h(x)} \frac{1}{|\Omega_{study}|}\int_{\Omega_{study}} D_{KL}(p_{study}(\bar{c}(x)|\bar{I}_{study}), p_{atlas}(\bar{c}(h(x))))dx.$$

As with the inter-subject registration, the transformation $h(x)$ is constrained to be a diffeomorphism using the fluid formulation. Again, the whole process is repeated until convergence.

4 Results

To evaluate the performance of the algorithm we define a geometric prior, a known transformation, and two synthetic images whose radiometric characteristics are statistically similar to actual T1- and T2-weighted MR images. A four class atlas prior comprising of concentric ellipses was generated using Matlab. A composite label image is given by the superposition of the individual classes. Both the atlas and composite label image are shown in Figure 1. Additionally, a transformation, $h(x)$, was constructed using sinusoidal displacements which was then applied to the composite label image. The deformed label image is shown in

the left panel of Figure 2. The two synthetic images were simulated by sampling from a multi-variate Gaussian distribution with different means and covariances for each of the classes in the deformed label image. These means and covariances are given as follows,

$$\mu_{c_1} = \begin{bmatrix} 182.28 \\ 150.12 \end{bmatrix}, \mu_{c_2} = \begin{bmatrix} 114.05 \\ 67.62 \end{bmatrix}, \Sigma_{c1} = \begin{bmatrix} 341.75 & 43.20 \\ 43.20 & 58.05 \end{bmatrix}, \Sigma_{c2} = \begin{bmatrix} 900.25 & -78.51 \\ -78.51 & 641.67 \end{bmatrix},$$

$$\mu_{c_3} = \begin{bmatrix} 30.11 \\ 117.52 \end{bmatrix}, \mu_{c_3} = \begin{bmatrix} 30.11 \\ 117.52 \end{bmatrix}, \Sigma_{c3} = \begin{bmatrix} 450.8 & 1375 \\ 1375 & 7286.3 \end{bmatrix}, \text{and } \Sigma_{c4} = \begin{bmatrix} 447.3 & 255.9 \\ 255.9 & 352.1 \end{bmatrix}.$$

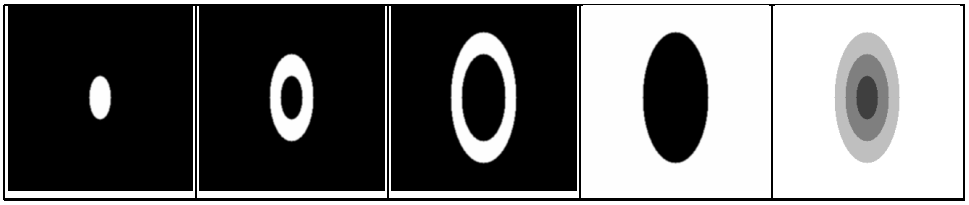

Fig. 1. Geometric Atlas Prior
Manually generated geometric four class atlas prior (first four images) and the corresponding composite labeled image.

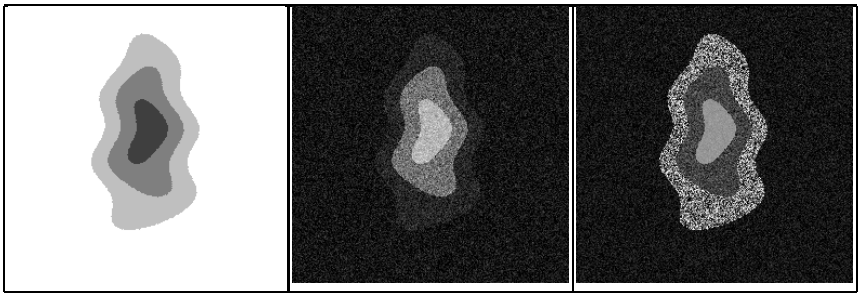

Fig. 2. Synthetic Images
Deformed label image (left), synthetic image derived from T1 samples (middle), and synthetic image derived from T2 samples (right).

The algorithm was was run for fifty iterations with ten steps of the high-dimensional registration per iteration. The final segmentation and deformation estimates shown in Figure 3.

In order to evaluated the success of the high-dimensional registration the algorithm is run, again with fifty iterations, holding the transformation fixed to the identity map. The final estimated segmentation is then compared to the one

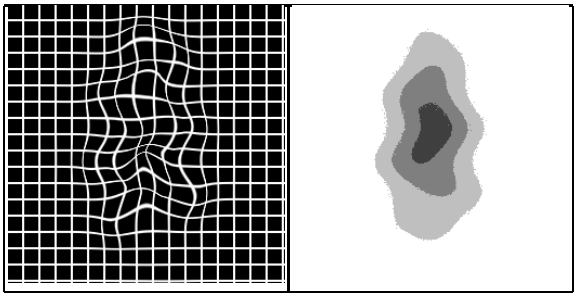

Fig. 3. Estimated deformation (left) and segmentation (right) following fifty iterations of the algorithm.

based on registration, with results shown in Figure 4. By examining the regions where the two segmentations differ from the ground truth label image it is clear the registration has improved the segmentation.

In both invocations of the algorithm, the class means and covariances were collected and compared. Table 1 shows the final relative norms for the estimated and actual means and covariances at the final iteration. For all classes the registration has improved estimates for both the means and covariances. The convergence of the means and covariance estimates using registration is shown in Figure 5. From this figure we see that the estimation of the means and covariances have converged quickly when the transformation is fixed to the identity map. When registration is added the estimates of the means and covariances continue to improve as estimation of the transformation between atlas and the subject converges. This exemplifies the effectiveness of the alternating nature of the algorithm.

Class	$\frac{\|\hat{\mu}_{c_i} - \mu_{c_i}\|}{\|\mu_{c_i}\|}$ w/ reg.	$\frac{\|\hat{\mu}_{c_i} - \mu_{c_i}\|}{\|\mu_{c_i}\|}$ w/o reg.	$\frac{\|\hat{\Sigma}_{c_i} - \Sigma_{c_i}\|}{\|\Sigma_{c_i}\|}$ w/ reg.	$\frac{\|\hat{\Sigma}_{c_i} - \Sigma_{c_i}\|}{\|\Sigma_{c_i}\|}$ w/o reg.
c_1	0.0047	0.1216	0.0450	7.0588
c_2	0.0152	0.1168	0.2104	1.8164
c_3	0.0960	0.0939	0.0206	0.0811
c_4	0.0046	0.0046	0.0081	0.0266

Table 1. Relative Norm Statistics

The first two columns of numbers are the relative means using registration (left) and fixed identity map (right) as the final iteration. The last two columns show the same for the relative covariances.

From these results we see that the registration improves the segmentation by accommodating local variability. The multi-modal inter-subject image registration results are currently being developed and will be available by the time of the conference.

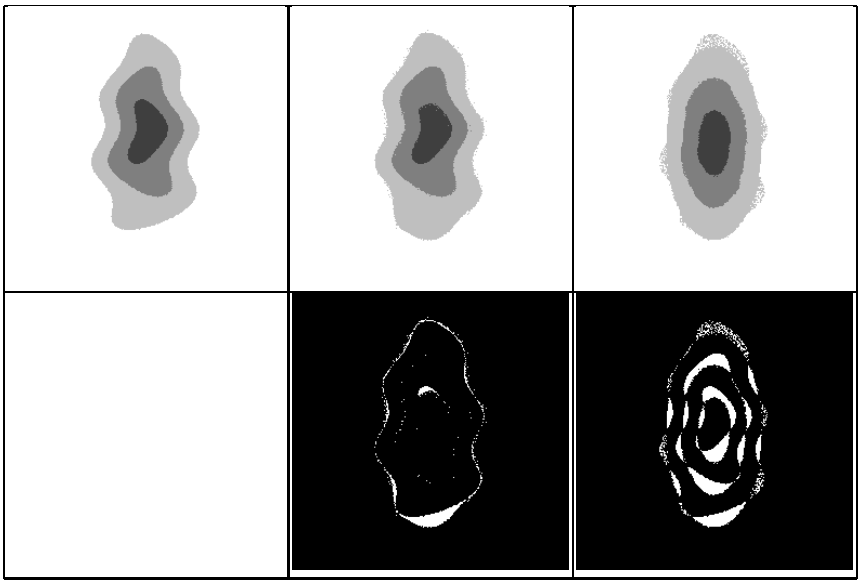

Fig. 4. Final Segmentation
The image in the upper left is the ground truth label image. The second column shows the final segmentation estimation using registration and the regions where this segmentation differs from the ground truth. The right column shows the same where the transformation is fixed to the identity map.

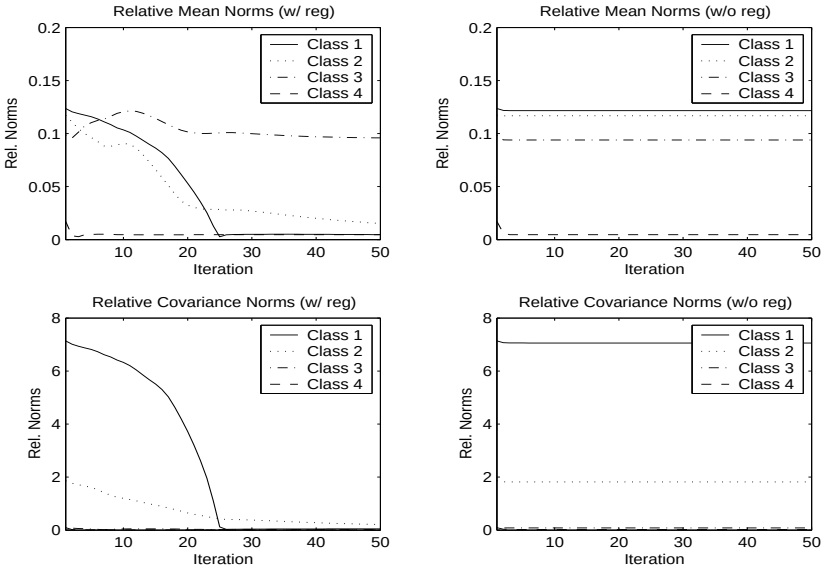

Fig. 5. Convergence
The left column shows the convergence of means and covariances using registration. The right column shows the same using the fixed identity transformation.

5 Acknowledgements

The authors would like to thank Dr. Guido Gerig, of the department of computer science, Dr. Elizabeth Bullitt, of the department of surgery, and Marcel Prastawa, of the department of computer science for their insightful suggestions and assistance during the development of this manuscript. This work was supported by NIBIB-NIH grant R01 EB000219.

References

1. G. E. Christensen, R. D. Rabbitt, and M. I. Miller. Deformable templates using large deformation kinematics. *IEEE Transactions on Image Processing*, 5(10):1435–1447, October 1996.
2. Gary Christensen, R Rabbit, and Michael Miller. Deformable templates using large deformation kinematics. *Transactions on Image Processing*, 5(10):1435–1447, 1996.
3. G.E. Christensen, S.C. Joshi, and M.I. Miller. Deformable templates using large deformation kniematics. *IEEE Transactions on Medical Imaging*, 16:864–877, 1997.
4. Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, 1991.
5. JG Csernansky, S Joshi, LE Wang, J Haller, JP Gado, M andM̃iller, U Grenander, and MI Miller. Hippocampal morphometry in schizophrenia via high dimensional brain map ping. *Proc. Natl. Acad. Sci. USA*, 95:11406–11411, September 1998.
6. Tom Gaens, Frederik Maes, Dirk Vandermeulen, and Paul Suetens. Non-rigid multimodal image registration using mutual information. *Proceedings of Medical Image Computing and Computer-Assisted Intervention*, pages 1099–1106, 1998.
7. J. C. Gee and R. K. Bajcsy. Elastic matching: Continuum mechanical and probabilistic analysis. In A. W. Toga, editor, *Brain Warping*. Academic Press, 1999.
8. A. Guimond, A. Roche, N. Ayache, and J. Meunier. Multimodal Brain Warping Using the Demons Algorithm and Adaptative Intensity Corrections. *IEEE Transaction on Medical Imaging*, 20(1):58–69, 2001.
9. S. Joshi and M. I. Miller. Landmark matching via large deformation diffeomorphisms. *IEEE Transactions on Image Processing*, 9(8):1357–1370, August 2000.
10. Frederik Maes, Andre Collignon, and at al. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging (TMI)*, 16:187–198, April 1997.
11. Michael Miller, Sarang Joshi, and Gary Christensen. Large deformation fluid diffeomorphisms for landmark and image matching. In *Brain Warping*, pages 115–131. Wiley-Interscience, 1999.
12. Nathan Moon, E Bullitt, and Guido Gerig. Automatic brain and tumor segmentation. *Medical Image Computing and Computer–Assisted Intervention (MICCAI)*, LNCS–0000:0–0, 00–00 September 2002.
13. D. Rueckert, C. Hayes, C. Studholme, P. Summers, M. Leach, and D. J. Hawkes. Non-rigid registration of breast mr images using mutual information. *Proceedings of Medical Image Computing and Computer-Assisted Intervention*, pages 1144–1152, 1998.
14. C. Studholme, D. L. G. Hill, and D. J. Hawkes. An overlap invariant entropy measure of 3d medical image alignment. *Pattern Recognition*, pages 71–86, December 1998.
15. Arthur Toga. *Brain Warping*. Wiley-Interscience, New York, 1999.

16. P. Turan. On measures of entropy and information. In *Selected Papers of Alfred Renyi*, pages 565–579. Akademiai Kiado, Budapest, 1976.
17. K. v. Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based tissue classification of mr images of the brain. *IEEE Transactions on Medical Imaging (TMI)*, 18:897–908, 1999.
18. J. West, J. Fitzpatrick, M. Wang, B. Dawant, C. Maurer, R. Kessler, and R. Maciunas. Comparison and evaluation of retrospective intermodality image registration techniques, 1996.