Bayesian Estimation of Regularization and Atlas Building in Diffeomorphic Image Registration

Miaomiao Zhang, Nikhil Singh, and P. Thomas Fletcher

Scientific Computing and Imaging Institute, University of Utah, Salt Lake City, UT

Abstract. This paper presents a generative Bayesian model for diffeomorphic image registration and atlas building. We develop an atlas estimation procedure that simultaneously estimates the parameters controlling the smoothness of the diffeomorphic transformations. To achieve this, we introduce a Monte Carlo Expectation Maximization algorithm, where the expectation step is approximated via Hamiltonian Monte Carlo sampling on the manifold of diffeomorphisms. An added benefit of this stochastic approach is that it can successfully solve difficult registration problems involving large deformations, where direct geodesic optimization fails. Using synthetic data generated from the forward model with known parameters, we demonstrate the ability of our model to successfully recover the atlas and regularization parameters. We also demonstrate the effectiveness of the proposed method in the atlas estimation problem for 3D brain images.

1 Introduction

Deformable image registration is often formulated as a maximum a posteriori (MAP) optimization problem, in which an image match likelihood term is regularized by a prior that encourages smooth deformations. In the diffeomorphic image registration setting, the log prior is in the form of the geodesic energy arising from a metric on an infinite-dimensional manifold of diffeomorphisms. In this framework, the level of smoothness is typically controlled by parameters describing the metric on the tangent space of the diffeomorphism group, as well as the noise variance in the image match term. However, despite the probabilistic motivation for the diffeomorphic registration problem, these model parameters are not estimated in current practice, but rather specified in an ad hoc manner. Part of the reason for this is that the estimation problem is inherently difficult, due to the fact that the log posterior of the metric parameters does not have a closed form and is computationally problematic to solve using direct optimization.

Further issues arise in the MAP formulation of diffeomorphic atlas building, where a template image, or atlas, is estimated along with the diffeomorphic registrations between the template and each input image. Current approaches [9, 18] optimize over both the template image and the diffeomorphic transformations. However, in the MAP formulation the diffeomorphisms should be treated as hidden random variables and not parameters to be estimated. The current

practice of optimizing over the diffeomorphisms is a mode approximation of the posterior distribution. Allassonnière et al. [1] shows that the common mode approximation scheme performs poorly under image noise, even for a simple 1D template estimation problem where the transformations are discrete shifts. As we show in this paper, the mode approximation in the diffeomorphism setting has similar difficulties when atlas estimation is combined with estimation of the metric and noise variance.

In this paper we propose a truly probabilistic formulation of the diffeomorphic atlas building problem. We develop an algorithm that can for the first time estimate the parameters controlling the smoothness of the diffeomorphisms and the image noise variance. This estimation procedure is a Monte Carlo Expectation Maximization (MCEM) algorithm, where the expectation step integrates over the posterior distribution of the diffeomorphic image transformations. We sample from this distribution using a novel Hamiltonian Monte Carlo (HMC) method on the space of the diffeomorphisms. Because we have a generative Bayesian model, we generate a synthetic data set from known parameters by sampling from the forward model, and then we show that our MCEM estimation procedure is able to recover those parameters. We also demonstrate that, unlike our MCEM algorithm, the mode approximation algorithm is unable to jointly estimate the atlas and the correct parameters. Finally, we show an example of an atlas and smoothness parameters estimated from real 3D brain images.

2 Related Work

Several works have proposed probabilistic motivations of the "groupwise" image registration problem, both in the small deformation [5, 21] and diffeomorphic [9, 17, 18] setting. In these approaches a set of input images are registered to a template, which is simultaneously estimated in an alternating optimization strategy. Allassonniére et al. [1] were the first to point out that atlas estimation via this alternating optimization scheme is not completely faithful to the probabilistic interpretation. They go on to propose a fully generative probability model for an image atlas and population. Later, Allassonniére et al. [2] developed a stochastic approximative expectation maximization (SAEM) algorithm to estimate the atlas and registration parameters. This estimation was done by appropriately marginalizing over the posterior distribution for the image deformations using a Monte Carlo sampling procedure.

Another related area of research involves Bayesian models of the segmentation problem. Van Leemput [10] developed a Bayesian model of the image segmentation problem that includes an atlas image and a generative deformation and image intensity model. He introduced a sampling procedure for image deformations also based on HMC, although his registration is based on a small deformation model and ours is in the diffeomorphic setting. Iglesias et al. [8] later extended this work to include uncertainty in the registration parameters, by introducing hyperpriors on the parameters and integrating over their posterior. Risholm et al. [13, 14] also formulated a Bayesian model for elastic image

registration and provided an MCMC method for sampling deformations, with the goal of quantifying uncertainty in the image registrations. Simpson et al. [15] furthermore inferred the level of regularization in non-rigid registration by a hierarchical Bayesian model.

Our work is the first in the diffeomorphic setting to bring MCMC sampling and correct parameter estimation via marginalization of the image transformations. Ma et al. [11] introduced a Bayesian formulation of the diffeomorphic image atlas problem, but also estimated the atlas using a mode approximation to alternate between atlas and registration optimizations. They do not estimate the registration parameters. There has been some work on stochastic flows of diffeomorphisms [6], which are Brownian motions, i.e., small perturbations integrated along a time-dependent flow. This differs from the prior distribution in our work, which is on the tangent space of initial velocity fields, rather than on the entire time-dependent flow. Our formulation leads to random geodesics in the space of diffeomorphisms, and makes possible an efficient sampling procedure for MCMC sampling.

3 A Bayesian Model for Diffeomorphic Atlas Building

We define a generative probabilistic model for atlas building in the setting of large deformation diffeomorphic metric mappings (LDDMM) [4], which we begin by reviewing. In this framework, the registration between two images, $I_0, I_1 \in L^2(\Omega, \mathbb{R})$, is the minimizer of the energy,

$$E(v, I_0, I_1) = \int_0^1 (Lv_t, v_t)\, dt + \frac{1}{2\sigma^2} \|I_0 \circ \phi_1^{-1} - I_1\|^2. \tag{1}$$

Here $v \in L^2([0,1], V)$ is a time-varying velocity field in a reproducing kernel Hilbert space, V, equipped with a metric, $L : V \to V^*$, a positive-definite, self-adjoint, differential operator, mapping to the dual space, V^*. The notation (m, v) denotes the pairing of a momentum vector $m \in V^*$ with a tangent vector $v \in V$. The deformation ϕ is defined as the integral flow of v, that is, $(d/dt)\phi(t, x) = v(t, \phi(t, x))$. We use subscripts for the time variable, i.e., $v_t(x) = v(t, x)$, and $\phi_t(x) = \phi(t, x)$. When the energy above is minimized over all initial velocities, it yields a squared distance metric between the two input images, i.e.,

$$d(I_0, I_1)^2 = \min_{v \in V} E(v, I_0, I_1).$$

Using this distance metric between images, the atlas estimation problem can be formulated as a least-squares estimation problem, or in other words, a Fréchet mean. Given input images $I_1, \ldots, I_N$, the diffeomorphic atlas building problem is to find a template image $\hat{I}$ and initial velocities v^k that minimize the sum-of-squared distances function, i.e.,

$$\hat{I} = \arg\min_I \frac{1}{N} \sum_{k=1}^{N} d(I, I_k)^2. \tag{2}$$

Because the distance function between images is itself a minimization problem, the atlas estimation is typically done by alternating between the minimization in (1) to find the optimal v^k and the minimization in (2) to update the atlas $\hat{I}$. However, in a probabilistic interpretation of the energy (1) as a negative log posterior, the initial velocities v^k should be regarded as latent random variables. The maximization step is only a mode approximation to this posterior.

For a continuous domain $\Omega \subset \mathbb{R}^n$, direct interpretation of (1) as a negative log posterior is problematic, as the image match term would be akin to isotropic Gaussian noise in the infinite-dimensional Hilbert space $L^2(\Omega, \mathbb{R})$. This is not a well-defined probability distribution as it has infinite measure. More appropriately, we can instead consider our input images, I_k, and our atlas image, I, to be measured on a discretized grid, $\Omega \subset \mathbb{Z}^n$. That is, images are elements of the finite-dimensional Euclidean space $l^2(\Omega, \mathbb{R})$. We will also consider velocity fields v^k and the resulting diffeomorphisms ϕ^k to be defined on the discrete grid, Ω. Now our noise model is i.i.d. Gaussian noise at each image voxel, with likelihood given by

$$p(I_k \,|\, v^k, I) = \frac{1}{(2\pi)^{M/2}\sigma^M} \exp\left(-\frac{\|I \circ (\phi^k)^{-1} - I_k\|^2}{2\sigma^2}\right), \tag{3}$$

where M is the number of voxels, σ^2 is the noise variance, and the norm inside the exponent is the Euclidean norm of $l^2(\Omega, \mathbb{R})$.

The negative log prior on the v^k is a discretized version of the squared Hilbert space norm above. Now consider L to be a discrete, self-adjoint, positive-definite differential operator on the domain Ω. The prior on each v^k is given by a multivariate Gaussian,

$$p(v^k) = \frac{1}{(2\pi)^{\frac{M}{2}}|L^{-1}|^{\frac{1}{2}}} \exp\left(-\frac{(Lv^k, v^k)}{2}\right), \tag{4}$$

where d is the dimension of v^k, and $|L|$ is the determinant of L. In this work, we use a metric of the form $L = -\alpha\Delta + \beta$, where Δ is the discrete Laplacian, and α and β are positive numbers. In the sequel, we consider $\theta = (\alpha, \sigma, I)$ to be parameters that we wish to estimate. We fix β to a small number to ensure that the L operator is nonsingular. Putting together the likelihood (3) and prior (4), we arrive at the log joint posterior for the diffeomorphisms, via initial velocities, v^k, as

$$\begin{aligned} \log \prod_{k=1}^{N} p\left(v^k \,|\, I_k; \theta\right) \;\propto\; & \frac{N}{2}\log|L| - \frac{1}{2}\sum_{k=1}^{N}(Lv^k, v^k) \\ & - \frac{MN}{2}\log\sigma - \frac{1}{2\sigma^2}\sum_{k=1}^{N}\|I \circ (\phi^k)^{-1} - I_k\|^2. \end{aligned} \tag{5}$$

4 Estimation of Model Parameters

We now present an algorithm for estimating the parameters, θ, of the probabilistic image atlas model specified in the previous section. These parameters include

the image atlas, I, the smoothness level, or metric parameter, α, and the standard deviation of the image noise, σ. We treat the v^k, i.e., the initial velocities of the image diffeomorphisms, as latent random variables with log posterior given by (5). This requires integration over the latent variables, which is intractable in closed form. We thus develop a Hamiltonian Monte Carlo procedure for sampling v^k from the posterior and use this in a Monte Carlo Expectation Maximization algorithm to estimate θ. It consists of two main steps:

1. E-step We draw a sample of size S from the posterior distribution (5) using HMC with the current estimate of the parameters, $\theta^{(i)}$. Let v^{kj}, $j = 1, \ldots, S$, denote the jth point in this sample for the kth velocity field. The sample mean is taken to approximate the Q function,

$$\begin{aligned} Q(\theta \,|\, \theta^{(i)}) &= E_{v^k | I_k; \theta^{(i)}} \left[\sum_{k=1}^{N} \log p \left(v^k \,|\, I_k; \theta \right) \right] \\ &\approx \frac{1}{S} \sum_{j=1}^{S} \sum_{k=1}^{N} \log p \left(v^{kj} \,|\, I_k; \theta \right). \end{aligned} \tag{6}$$

2. M-step Update the parameters by maximizing $Q(\theta \,|\, \theta^{(i)})$. The maximization is closed form in I and σ, and a one-dimensional gradient ascent in α.

4.1 Background on Geodesic Shooting of Diffeomorphisms

Before presenting our MCEM estimation algorithm, we provide a brief background on the computations we will use for geodesic shooting and gradients for diffeomorphic image matching. Details of these methods are found in [19, 20, 16].

Deformation momenta: The tangent space at identity, $V = T_{\mathrm{Id}}\mathrm{Diff}(\Omega)$, consists of all vector fields with finite Sobolev norm. Let $V^* = T^*_{\mathrm{Id}}\mathrm{Diff}(\Omega)$ denote its dual space. The velocity, $v \in V$, maps to its dual deformation momenta, $m \in V^*$, via the operator L such that $m = Lv$. The operator $K : V^* \to V$ denotes the inverse of L, so that $v = Km$. Note that constraining ϕ to be a geodesic with initial momentum $m_0 = m(0)$ implies that ϕ, m and I all evolve in a way entirely determined by the metric L, and that the deformation is determined entirely by the initial momenta, m_0.

EPDiff for geodesic evolution: Given the initial velocity, $v_0 \in V$, or equivalently, the initial momentum, $m_0 \in V^*$, the geodesic path $\phi(t)$ is constructed via integration of the following EPDiff equation [3, 12]:

$$\frac{\partial m}{\partial t} = -\mathrm{ad}^*_v m = -(Dv)^T m - Dm\, v - m \,\mathrm{div}(v), \tag{7}$$

where D denotes the Jacobian matrix, and the operator ad^* is the dual of the negative Lie bracket of vector fields,

$$\mathrm{ad}_v w = -[v, w] = Dvw - Dwv.$$

The deformed image $I(t) = I_0 \circ \phi^{-1}(t)$ evolves via the equation

$$\frac{\partial I}{\partial t} = -v \cdot \nabla I.$$

Image matching gradient: In our HMC sampling procedure, we will need to compute gradients, with respect to initial momenta, of the diffeomorphic image matching problem in (1), for matching the atlas I to an input image I_k.

Following the optimal control theory approach in [19], we add Lagrange multipliers to constrain the diffeomorphism $\phi^k(t)$ to be a geodesic path. This is done by introducing time-dependent adjoint variables, $\hat{m}, \hat{I}$ and $\hat{v}$, and writing the augmented energy,

$$\begin{aligned}\tilde{E}(m_0) =& E(Km_0, I, I_k) + \\ & \int_0^1 \langle \hat{m}, \dot{m} + \mathrm{ad}_v^* m \rangle dt + \int_0^1 \langle \hat{I}, \dot{I} + \nabla I \cdot v \rangle dt + \int_0^1 \langle \hat{v}, m - Lv \rangle dt,\end{aligned}$$

where E is the diffeomorphic image matching energy from (1), and the other terms correspond to Lagrange multipliers enforcing: a) the geodesic constraint, which comes from the EPDiff equation (7), b) the image transport equation, $\dot{I} = -\nabla I \cdot v$, and c) the constraint that $m = Lv$, respectively.

The optimality conditions for m, I, v are given by the following time-dependent system of ODEs, termed the *adjoint equations*:

$$-\dot{\hat{m}} + \mathrm{ad}_v \hat{m} + \hat{v} = 0, \qquad -\dot{\hat{I}} - \nabla \cdot (\hat{I} v) = 0, \qquad -\mathrm{ad}_{\hat{m}}^* m + \hat{I} \nabla I - L\hat{v} = 0,$$

subject to initial conditions

$$\hat{m}(1) = 0, \qquad \hat{I}(1) = \frac{1}{\sigma^2}(I(1) - I_k).$$

Finally, after integrating these adjoint equations backwards in time to $t = 0$, the gradient of $\tilde{E}$ with respect to the initial momenta is

$$\nabla_{m_0} \tilde{E} = Km_0 - \hat{m}_0. \tag{8}$$

4.2 Hamiltonian Monte Carlo (HMC) Sampling

Hamiltonian Monte Carlo [7] is a powerful MCMC sampling methodology that is applicable to a wide array of continuous probability distributions. It utilizes Hamiltonian dynamics as a Markov transition probability and efficiently explores the space of a target distribution. The integration through state space results in more efficient, global moves, while it also uses gradient information of the log probability density to sample from higher probability regions. In this section, we derive a HMC sampling method to draw a random sample from the posterior distribution of our latent variables, v^k, the initial velocities defining the diffeomorphic image transformations from the atlas to the data.

To sample from a pdf $f(x)$ using HMC, one first sets up a Hamiltonian $H(x,\mu) = U(x) + V(\mu)$, consisting of a "potential energy", $U(x) = -\log f(x)$, and a "kinetic energy", $V(\mu) = -\log g(\mu)$. Here $g(\mu)$ is some proposal distribution (typically isotropic Gaussian) on an auxiliary momentum variable, μ. An initial random momentum μ is drawn from the density $g(\mu)$. Starting from the current point x and initial random momentum μ, the Hamiltonian system is integrated forward in time to produce a candidate point, $\tilde{x}$, along with the corresponding forward-integrated momentum, $\tilde{\mu}$. The candidate point $\tilde{x}$ is accepted as a new point in the sample with probability

$$P(\text{accept}) = \min(1, \exp(-U(\tilde{x}) - V(\tilde{\mu}) + U(x) + V(\mu)).$$

This acceptance-rejection method is guaranteed to converge to the desired density $f(x)$ under fairly general regularity assumptions on f and g.

In our model, to sample v^k from the posterior in (5), we equivalently sample m^k from the dual momenta, using $v^k = Km^k$, so we define our potential energy as $U(m^k) = -\log p(m^k|I_k;\theta)$. We use the prior distribution on the dual momenta as our proposal density, in other words, we use $p(K\mu)$ defined as in (4), taking care to include the appropriate change-of-variables. This gives the kinetic energy, $V(\mu) = (\mu, K\mu)$. This gives us the following Hamiltonian system to integrate in the HMC:

$$\begin{aligned} \frac{dm^k}{dt} &= \frac{\partial H}{\partial \mu} = K\mu, \\ \frac{d\mu}{dt} &= -\frac{\partial H}{\partial m^k} = -\nabla_{m^k}\tilde{E}, \end{aligned}$$

where the last term comes from the gradient defined in (8). As is standard practice in HMC, we use a "leap-frog" integration scheme, which better conserves the Hamiltonian and results in high acceptance rates.

4.3 The Maximization Step

We now derive the M-step for updating the parameters $\theta = (\alpha, \sigma, I)$ by maximizing the HMC approximation of the Q function, which is given in (6). This turns out to be a closed-form update for the noise variance σ^2 and the atlas I, and a simple one-dimensional gradient ascent for α.

From (5) and (6), it is easy to derive the closed-form update for σ as

$$\sigma^2 = \frac{1}{MNS}\sum_{j=1}^{S}\sum_{k=1}^{N}\|I_0 \circ (\phi^{kj})^{-1} - I^k\|^2. \tag{9}$$

For updating the atlas image I, we set the derivative of the Q function approximation which with respect to I to zero. The solution for I gives a closed-form update,

$$I = \frac{\sum_{j=1}^{S}\sum_{k=1}^{N} I^k \circ \phi^{kj}|D\phi^{kj}|}{\sum_{j=1}^{S}\sum_{k=1}^{N}|D\phi^{kj}|}.$$

The gradient ascent over α requires that we take the derivative of the metric $L = -\alpha\Delta + \beta I$, with respect to α. We do this in the Fourier domain, where the discrete Laplacian is a diagonal operator. For a 3D grid, the coefficients A_{xyz} of the discrete Laplacian at coordinate (x, y, z) in the Fourier domain is

$$A_{xyz} = -2\left(\cos\frac{2\pi x}{W-1} + \cos\frac{2\pi y}{H-1} + \cos\frac{2\pi z}{D-1}\right) + 6,$$

where W, H, D are the dimension of each direction. Hence, the determinant of the operator L is

$$|L| = \prod_{x,y,z} A_{xyz}\alpha + \beta.$$

The gradient of the HMC approximated Q function, with respect to α, is

$$\nabla_\alpha Q(\theta \,|\, \theta^{(i)}) \approx \frac{1}{2}\sum_{j=1}^{S}\sum_{k=1}^{N}\left[\sum_{x,y,z}\frac{A_{xyz}}{A_{xyz}\alpha+\beta} - \left\langle -\Delta v^{kj}, v^{kj}\right\rangle\right].$$

5 Results

We demonstrate the effectiveness of our proposed model and MCEM estimation routine using both 2D synthetic data and real 3D MRI brain data. Because we have a generative model, we can forward simulate a random sample of images from a distribution with known parameters $\theta = (\alpha, \sigma, I)$. Then, in the next subsection, we test if we can recover those parameters using our MCEM algorithm. Figure 1 illustrates this process. We simulated a 2D synthetic dataset starting from a atlas image, I, of a binary circle with resolution 100×100. We then generated 20 smooth initial velocity fields from the prior distribution, $p(v^k)$, defined in (4), setting $\alpha = 0.025$ and $\beta = 0.001$. Deformed circle images were constructed by shooting the initial velocities by the EPDiff equations and transforming the atlas by the resulting diffeomorphisms, ϕ^k. Finally, we added i.i.d. Gaussian noise according to our likelihood model (3). We used a standard deviation of $\sigma = 0.05$, which corresponds to an SNR of 20 (which is more noise than typical structural MRI).

Parameter estimation on synthetic data In our estimation procedure, we initialized α with 0.002 for noise free, and 0.01 for noise corrupted images. The step size of 0.005 for leap-frog integration is used in HMC with 10 units of time discretization in integration of EPDiff equations.

Figure 2 compares the true atlas and estimated atlases in the clean and noisy case. Figure 3 shows the convergence graph for α and σ estimation. It shows that our method recovers the model parameters fairly well. However, the iterative mode approximation algorithm does not recover the α parameter as nicely as our method. In the noisy case, the mode approximation algorithm estimates α as 0.0152, which is far from the ground truth value of 0.025. This is compared with our estimation of 0.026. In addition, in the noise free example, the mode approximation algorithm blows up due to the σ dropping close to 0, thus making the image match term numerically too high and the geodesic shooting unstable.

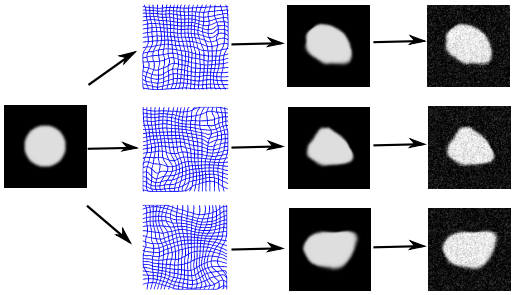

Fig. 1. Simulating synthetic 2D data from the generative diffeomorphism model. From left to right: the ground truth template image, random diffeomorphisms from the prior model, deformed images, and final noise corrupted images.

Fig. 2. Atlas estimation results. Left: ground-truth template. Center: estimated template from noise free dataset. Right: estimated template from noise corrupted dataset.

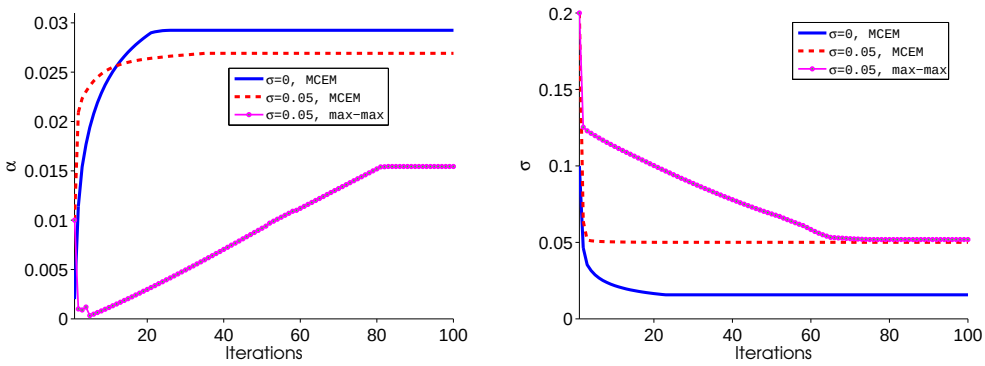

Fig. 3. Estimation of α, σ. Left: α estimation. Right: σ estimation. In our MCEM method, final estimated α and σ for noise free data are 0.028, 0.01, and for noise data are 0.026, 0.0501. Compared with max-max method, for the noise data, estimated α and σ are 0.0152, 0.052.

Atlas building on 3D brain images To demonstrate the effectiveness of our method on the real data, we apply our MCEM atlas estimation algorithm to

a set of brain MRI from ten healthy subjects. The MRI have resolution $108 \times 128 \times 108$ and are skull-stripped, intensity normalized, and co-registered with rigid transforms. We set the initial $\alpha = 0.01$, $\beta = 0.001$ with 15 time-steps.

The left side of Figure 4 shows coronal and axial slices from the 3D MRI used as input. The right side shows the initialization (greyscale average of the input images), followed by the final atlas estimated by our method. The final atlas estimate correctly aligns the anatomy of the input images, producing a sharper average image. The algorithm also jointly estimated the smoothness parameter to be $\alpha = 0.028$ and the image noise standard deviation to be $\sigma = 0.031$.

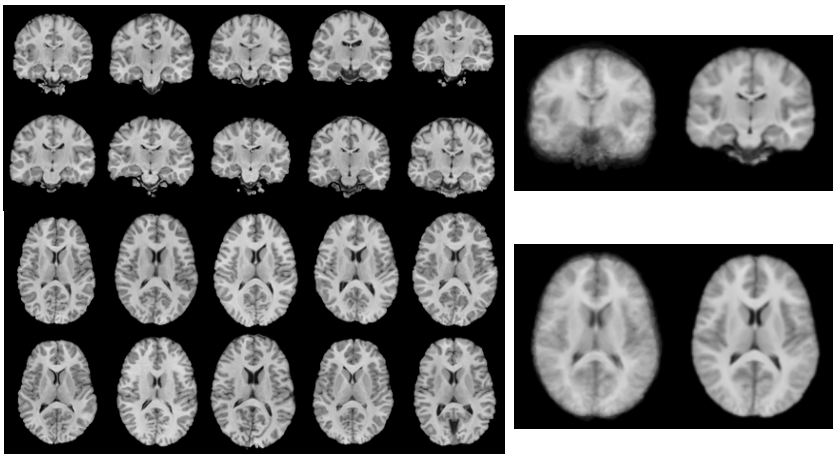

Fig. 4. Left: coronal and axial slices from the input 3D MRIs. Middle: initial greyscale average of the input images. Right: final atlas estimated by our MCEM estimation procedure.

Image matching accuracy Finally, we demonstrate that another benefit of our HMC sampling methodology is improved performance in the standard image registration problem under large deformation shooting. Rather than use a direct gradient descent to solve the image registration problem, we instead can find the posterior mean of the model (5), where for image matching we fix the "atlas", I, as the source image and have just one target image, I_1. The stochastic behavior in the sampling helps to get out of local minima, where the direct gradient descent can get stuck. We compared our proposed method with direct gradient descent image registration by geodesic shooting from [19]. We used the authors' uTIlzReg package for geodesic shooting, which is available freely online. For the comparison, we registered the image pair shown in the first two panels of Figure 5, which requires a large deformation. The source and target images are 50×50. We used $\alpha = 0.02, \beta = 0.001$ for smoothing kernel, and $h = 40$ time-

steps between $t = 0$ and $t = 1$. Note that we only want to compare the image matching here, so we fix the α and σ parameters.

Fig. 5. The first two images from left to right are the source and target image respectively. Third is the matched image obtained by geodesic shooting method using [19]. Last image is the matched image from our MCEM method.

Figure 5 demonstrates the results of the direct geodesic shooting registration with our HMC posterior mean. It shows that the geodesic shooting method gets stuck in a local minima and cannot make it to the target image even with a large number of time-steps ($h = 60$) in the time discretization (we tried several time discretizations up to 60, and none worked). Though our method did not match perfectly in the tip of the "C", it still recovers the full shape while retaining a diffeomorphic transformation.

6 Conclusion

We presented a novel generative model of the diffeomorphic atlas estimation problem. Our method is the first to jointly estimate the regularity parameter, noise variance, and image atlas. It faithfully treats the diffeomorphic transformations from the atlas to the input images as unobserved random variables. We introduced a MCMC sampling scheme to integrate over these transformations. While we chose a particular parameterized form for the metric operator L, other metrics are also possible in our framework. This work opens up the possibility of extensions for rigorous probabilistic modeling of shape variability through diffeomorphisms.

Acknowledgements. This work was supported by NIH Grant R01 MH084795, NIH Grant 5R01EB007688, and NSF CAREER Grant 1054057.

References

1. S. Allassonnière, Y. Amit, and A. Trouvé. Toward a coherent statistical framework for dense deformable template estimation. *Journal of the Royal Statistical Society, Series B*, 69:3–29, 2007.

2. S. Allassonnière and E. Kuhn. Stochastic algorithm for parameter estimation for dense deformable template mixture model. *ESAIM-PS*, 14:382–408, 2010.
3. V. I. Arnol'd. Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Ann. Inst. Fourier*, 16:319–361, 1966.
4. M.F. Beg, M.I. Miller, A. Trouvé, and L. Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision*, 61(2):139–157, 2005.
5. K. Bhatia, J. Hajnal, B. Puri, A. Edwards, and D. Rueckert. Consistent groupwise non-rigid registration for atlas construction. In *ISBI*, 2004.
6. A. Budhiraja, P. Dupuis, and V. Maroulas. Large deviations for stochastic flows of diffeomorphisms. *Bernoulli*, 16:234–257, 2010.
7. S. Duane, A. Kennedy, B. Pendleton, and D. Roweth. Hybrid Monte Carlo. *Physics Letters B*, pages 216–222, 1987.
8. J. E. Iglesias, M. R. Sabuncu, K. Van Leemput, and ADNI. Incorporating parameter uncertainty in Bayesian segmentation models: application to hippocampal subfield volumetry. In *MICCAI*, 2012.
9. S. Joshi, B. Davis, M. Jomier, and G. Gerig. Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage*, 23, Supplement1:151–160, 2004.
10. Koen Van Leemput. Encoding probabilistic brain atlases using Bayesian inference. *IEEE Transactions on Medical Imaging*, 28:822–837, 2009.
11. J. Ma, M. I. Miller, A. Trouvé, and L. Younes. Bayesian template estimation in computational anatomy. *NeuroImage*, 42:252–261, 2008.
12. M. I. Miller, A. Trouvé, and L. Younes. Geodesic shooting for computational anatomy. *Journal of Mathematical Imaging and Vision*, 24(2):209–228, 2006.
13. P. Risholm, S. Pieper, E. Samset, and W. M. Wells. Summarizing and visualizing uncertainty in non-rigid registration. In *MICCAI*, 2010.
14. P. Risholm, E. Samset, and W. M. Wells. Bayesian estimation of deformation and elastic parameters in non-rigid registration. In *WBIR*, 2010.
15. Ivor J. A. Simpson, Julia A. Schnabel, Adrian R. Groves, Jesper L. R. Andersson, and Mark W. Woolrich. Probabilistic inference of regularisation in non-rigid registration. *NeuroImage*, 59:2438–2451, 2012.
16. Nikhil Singh, Jacob Hinkle, Sarang Joshi, and P. Thomas Fletcher. A vector momenta formulation of diffeomorphisms for improved geodesic regression and atlas construction. In *International Symposium on Biomedial Imaging (ISBI)*, April 2013.
17. C. Twining, T. Cootes, S. Marsland, V. Petrovic, R. Schestowitz, and C. Taylor. A unified information-theoretic approach to groupwise non-rigid registration and model building. In *IPMI*, pages 1–14, 2005.
18. F.-X. Vialard, L. Risser, D. Holm, and D. Rueckert. Diffeomorphic atlas estimation using Kärcher mean and geodesic shooting on volumetric images. In *MIUA*, 2011.
19. F.-X. Vialard, L. Risser, D. Rueckert, and C. J.Cotter. Diffeomorphic 3d image registration via geodesic shooting using an efficient adjoint calculation. In *International Journal of Computer Vision*, pages 229–241, 2012.
20. L. Younes, F. Arrate, and M.I. Miller. Evolutions equations in computational anatomy. *NeuroImage*, 45(1S1):40–50, 2009.
21. L. Zöllei, M. Jenksinon, S. Timoner, and W.M. Wells III. A marginalized map approach and em optimization for pair-wise registration. In *IPMI*, pages 662–667, 2007.