

Bayesian Atlas Building with Hierarchical Priors for Subject-specific Regularization

Jian Wang¹ and Miaomiao Zhang^{1,2}

¹ Computer Science, University of Virginia, USA

² Electrical and Computer Engineering, University of Virginia, USA

Abstract. This paper presents a novel hierarchical Bayesian model for unbiased atlas building with subject-specific regularizations of image registration. We develop an atlas construction process that automatically selects parameters to control the smoothness of diffeomorphic transformation according to individual image data. To achieve this, we introduce a hierarchical prior distribution on regularization parameters that allows multiple penalties on images with various degrees of geometric transformations. We then treat the regularization parameters as latent variables and integrate them out from the model by using the Monte Carlo Expectation Maximization (MCEM) algorithm. Another advantage of our algorithm is that it eliminates the need for manual parameter tuning, which can be tedious and infeasible. We demonstrate the effectiveness of our model on 3D brain MR images. Experimental results show that our model provides a sharper atlas compared to the current atlas building algorithms with single-penalty regularizations. Our code is publicly available at <https://github.com/jw4hv/HierarchicalBayesianAtlasBuild>.

1 Introduction

Deformable atlas building is to create a “mean” or averaged image and register all subjects to a common space. The resulting atlas and group transformations are powerful tools for statistical shape analysis of images [12,18], template-based segmentation [21,22,13], or object tracking [17,16], just to name a few. A good quality of atlas heavily relies on the registration process, which is typically formulated as a regularized optimization to solve [5,7,25,30]. An issue in the current process of registration-based atlas construction is how to regularize model parameters. Having an appropriate regularization is critical to the “sharpness” of the atlas, as well as ensuring a set of desirable properties of transformations, i.e., a smooth and invertible smooth mapping between images, also known as diffeomorphisms, to preserve the topology of original images.

Current atlas building models either exhaustively search for an optimal regularization in the parameter space, or treat it as unknown variables to estimate from Bayesian models. While ad hoc parameter-tuning may yield satisfactory results, it requires expert domain knowledge to guide the tuning process [14,26,18,27]. Inspired by probabilistic models, several works have proposed Bayesian models of atlas building with automatically estimated regularizations

[1,2,31]. These approaches define a posterior distribution that consists of an image matching term between a deformed atlas and each individual as a likelihood, and a regularization as a prior to support the smoothness of transformation fields. The regularization parameter is then jointly estimated with atlas after carefully integrating out the image deformations using Monte Carlo sampling. However, sampling in a high-dimensional transformation space (i.e., on a dense 3D image grid 128^3) is computationally expensive and often leads to a long execution time with high memory consumption. More importantly, the aforementioned methods are limited to regularizations with single-penalty for population studies. This prohibits the model’s ability to adaptively search for the best regularization parameter associated with an individual subject, which is critical to images with various degrees of geometric transformations. The typical “one-fits-all” fails in cases where large geometric variations occur, i.e., brain shape changes of Alzheimer’s disease group. Allowing the subject-specific (data-driven) regularization can substantially affect the sharpness and quality of the atlas [29].

In this paper, we propose a hierarchical Bayesian model of atlas building with subject-specific regularizations in the context of Large Deformation Diffeomorphic Metric Mapping (LDDMM) algorithm [7]. In contrast to previous approaches treating the regularization of individual subjects as a single-penalty function with adhoc parameters, we develop a data-adaptive algorithm to automatically adjust the model parameters accordingly. To achieve this, we introduce a novel hierarchical prior that features (i) prior distributions with multiple regularization parameters on the group transformations in a low-dimensional bandlimited space; and (ii) a hyperprior to model the regularization parameters as latent variables. We then develop a Monte Carlo Expectation Maximization (MCEM) algorithm, where the expectation step integrates over the regularization parameters using Hamiltonian Monte Carlo (HMC) sampling. The joint estimation of model parameters including atlas, registration, and hyperparameters in the maximization step successfully eliminates a massive burden of multi-parameters tuning. We demonstrate the effectiveness of our algorithm on both 2D synthetic images and 3D real brain MRIs.

To the best of our knowledge, we are the first to extend the atlas building to a data-adaptive and parameter-tuning-free framework via hierarchical Bayesian learning. Experimental results show that our model provides an efficient atlas construction of population images, particularly with large variations of geometric transformations. This paves a way for an improved quality of clinical studies where atlas building is required, for example, statistical shape analysis of brain changes for neurodegenerative disease diagnosis [12], or atlas-based segmentation for in-utero placental disease monitoring [16].

2 Background: Atlas Building with Fast LDDMM

We first briefly review an unbiased atlas building algorithm [14] based on Fourier-approximated Lie Algebra for Shooting (FLASH), a fast variant of LDDMM

with geodesic shooting [30]. Given a set of images I_1, \dots, I_N with N being the number of images, the problem of atlas building is to find a template image I and transformations ϕ_1, \dots, ϕ_N that minimize the energy function

$$E(I, \phi_n) = \sum_{n=1}^N \text{Dist}(I \circ \phi_n, I_n) + \text{Reg}(\alpha, \phi_n). \quad (1)$$

The $\text{Dist}(\cdot, \cdot)$ is a distance function that measures the dissimilarity between images, i.e., sum-of-squared differences [7], normalized cross correlation [6], and mutual information [28]. The $\text{Reg}(\cdot)$ is a weighted regularization with parameter α that guarantees the diffeomorphic properties of transformation fields.

Regularization In Tangent Space of Diffeomorphisms. Given an open and bounded d -dimensional domain $\Omega \subset \mathbb{R}^d$, we use $\text{Diff}(\Omega)$ to denote a space of diffeomorphisms and its tangent space $V = T\text{Diff}(\Omega)$. The regularization of LDDMM is defined as an integral of the Sobolev norm of the time-dependent velocity field $v(t) \in V(t \in [0, 1])$ in the tangent space, i.e.,

$$\text{Reg}(\alpha, \phi_n) = \int \langle \mathcal{L}(\alpha)v_n(t), \mathcal{L}(\alpha)v_n(t) \rangle dt, \text{ with } \frac{d\phi_n(t)}{dt} = -D\phi_n(t) \cdot v_n(t). \quad (2)$$

Here \mathcal{L} is a symmetric, positive-definite differential operator, with parameter α controlling the smoothness of transformation fields. In this paper, we use the Laplacian operator $\mathcal{L} = (-\alpha\Delta + \text{Id})^3$, where Id is an identity matrix. The operator D is a Jacobian matrix and \cdot denotes an element-wise matrix multiplication.

According to the geodesic shooting algorithm [25], the minimum of LDDMM is uniquely determined by solving a Euler-Poincaré differential equation (EPDiff) [3,19] with initial conditions. This inspires a recent model FLASH to reparameterize the regularization of Eq. (2) in a low-dimensional bandlimited space of initial velocity fields, which dramatically reduces the computational complexity of transformation models with little to no loss of accuracy [30].

Fourier Computation of Diffeomorphisms. Let $\widetilde{\text{Diff}}(\Omega)$ and \tilde{V} denote the space of Fourier representations of diffeomorphisms and velocity fields respectively. Given time-dependent velocity field $\tilde{v}(t) \in \tilde{V}$, the diffeomorphism $\tilde{\phi}(t) \in \widetilde{\text{Diff}}(\Omega)$ in the finite-dimensional Fourier domain can be computed as

$$\tilde{\phi}(t) = \tilde{\text{Id}} + \tilde{u}(t), \quad \frac{d\tilde{u}(t)}{dt} = -\tilde{v}(t) - \tilde{\mathcal{D}}\tilde{u}(t) * \tilde{v}(t), \quad (3)$$

where $\tilde{\text{Id}}$ is the frequency of an identity element, $\tilde{\mathcal{D}}\tilde{u}(t)$ is a tensor product $\tilde{\mathcal{D}} \otimes \tilde{u}(t)$, representing the Fourier frequencies of a Jacobian matrix $\tilde{\mathcal{D}}$ with central difference approximation, and $*$ is a circular convolution³.

³ To prevent the domain from growing infinity, we truncate the output of the convolution in each dimension to a suitable finite set.

The Fourier representation of the geodesic shooting equation (EPDiff) is

$$\frac{\partial \tilde{v}(t)}{\partial t} = -\tilde{\mathcal{K}} \left[(\tilde{\mathcal{D}}\tilde{v}(t))^T \star \tilde{\mathcal{L}}\tilde{v}(t) + \tilde{\nabla} \cdot (\tilde{\mathcal{L}}\tilde{v}(t) \otimes \tilde{v}(t)) \right], \quad (4)$$

where \star is the truncated matrix-vector field auto-correlation. The operator $\tilde{\nabla} \cdot$ is the discrete divergence of a vector field. Here $\tilde{\mathcal{K}}$ is an inverse operator of $\tilde{\mathcal{L}}$, which is the Fourier transform of a Laplacian operator in this paper.

The regularization in Eq. (2) can be equivalently formulated as

$$\text{Reg}(\alpha, \phi_n) = \langle \tilde{\mathcal{L}}(\alpha)\tilde{v}_n(0), \tilde{\mathcal{L}}(\alpha)\tilde{v}_n(0) \rangle, \quad \text{s.t. Eq. (3)\&Eq. (4)}.$$

We will drop off the time index in remaining sections for notational simplicity, e.g., defining $\tilde{v}_n \triangleq \tilde{v}_n(0)$.

3 Our Model: Bayesian Atlas Building with Hierarchical Priors

This section presents a hierarchical Bayesian model for atlas building that allows subject-specific regularization with no manual effort of parameter-tuning. We introduce a hierarchical prior distribution on the initial velocity fields with adaptive smoothing parameters followed by a likelihood distribution on images.

Likelihood. Assuming an independent and identically distributed (i.i.d.) Gaussian noise on image intensities, we formulate the likelihood of each observed image I_n as

$$p(I_n | I, \tilde{v}_n, \sigma^2) = \frac{1}{(\sqrt{2\pi}\sigma^2)^M} \exp \left(-\frac{1}{2\sigma^2} \|I \circ \phi_n - I_n\|_2^2 \right). \quad (5)$$

Here σ^2 denotes a noise variance, M is the number of image voxels, and ϕ_n is an inverse Fourier transform of $\tilde{\phi}_n$ at time point $t = 1$. It is worth mentioning that other noise models such as spatially varying noises [24] can also be applied.

Prior. To ensure the smoothness of transformation fields, we define a prior on each initial velocity field \tilde{v}_n as a complex multivariate Gaussian distribution

$$p(\tilde{v}_n | \alpha_n) = \frac{1}{(2\pi)^{\frac{M}{2}} |\tilde{\mathcal{L}}_n^{-1}(\alpha_n)|} \exp \left(-\frac{1}{2} \langle \tilde{\mathcal{L}}_n(\alpha_n)\tilde{v}_n, \tilde{\mathcal{L}}_n(\alpha_n)\tilde{v}_n \rangle \right), \quad (6)$$

where $|\cdot|$ is matrix determinant. The Fourier coefficients of a discrete Laplacian operator is $\tilde{\mathcal{L}}_n(\xi_1, \dots, \xi_d) = \left(-2\alpha_n \sum_{j=1}^d (\cos(2\pi\xi_j) - 1) + 1 \right)^3$, with (ξ_1, \dots, ξ_d) being a d-dimensional frequency vector.

Hyperprior. We treat the subject-specific regularization parameter α_n of the prior distribution (6) as a random variable generated from Gamma distribution, which is a commonly used prior to model positive real numbers [23]. Other prior such as inverse Wishart distribution [11] can also be applied. The hyperprior of our model is formulated as

$$p(\alpha_n | k, \beta) = \frac{\alpha_n^{k-1} \exp(-\alpha_n/\beta)}{\Gamma(k)\beta^k}, \quad (7)$$

with k and β being positive numbers for shape and scale parameters respectively. The Gamma function $\Gamma(k) = (k-1)!$ for all positive integers of k . We finally arrive at the log posterior of the diffeomorphic transformation and regularization parameters as

$$\begin{aligned} E(\tilde{v}_n, \alpha_n, I, \sigma, k, \beta) &\triangleq \ln \prod_{n=1}^N p(I_n | I, \tilde{v}_n, \sigma^2) \cdot p(\tilde{v}_n | \alpha_n) \cdot p(\alpha_n | k, \beta) \\ &= \sum_{n=1}^N \frac{1}{2} \ln |\mathcal{L}_n| - M \ln \sigma - \frac{\|I \circ \phi_n - I_n\|_2^2}{2\sigma^2} - \frac{1}{2} (\tilde{\mathcal{L}}\tilde{v}_n, \tilde{\mathcal{L}}\tilde{v}_n) \\ &\quad (k-1) \ln \alpha_n - \frac{\alpha_n}{\beta} - k \ln \beta - \ln \Gamma(k) + \text{const}. \end{aligned} \quad (8)$$

3.1 Model Inference

We develop an MCEM algorithm to infer the model parameter Θ , which includes the image atlas I , the noise variance of image intensities σ^2 , the initial velocities of diffeomorphic transformations \tilde{v}_n , and the hyperparameters k and β . We treat the regularization parameter α_n as latent random variables and integrate them out from the log posterior in Eq. (8). Computations of two main steps (expectation and maximization) are illustrated below.

Expectation: HMC. Since the E-step does not yield a closed-form solution, we employ a powerful Hamiltonian Monte Carlo (HMC) sampling method [9] to approximate the expectation function Q with respect to the latent variables α_n . For each α_n , we draw a number of S samples from the log posterior (8) by using HMC from the current estimated parameters $\hat{\Theta}$. The Monte Carlo approximation of the expectation Q is

$$Q(\Theta | \hat{\Theta}) \approx \frac{1}{S} \sum_{n=1}^N \sum_{j=1}^S \ln p(\alpha_{nj} | I_n; \hat{\Theta}). \quad (9)$$

To produce samples of α_n , we first define the potential energy of the Hamiltonian system $H(\alpha_n, \gamma) = U(\alpha_n) + W(\gamma)$ as $U(\alpha_n) = -\ln p(\alpha_n | I_n; \Theta)$. The kinetic energy $W(\gamma)$ is a typical normal distribution on an auxiliary variable γ . This gives us Hamilton's equations to integrate

$$\frac{d\alpha_n}{dt} = \frac{\partial H}{\partial \gamma} = \gamma, \quad \frac{d\gamma}{dt} = -\frac{\partial H}{\partial \alpha_n} = -\nabla_{\alpha_n} U. \quad (10)$$

Since α_n is a Euclidean variable, we use a standard “leap-frog” numerical integration scheme, which approximately conserves the Hamiltonian and results in high acceptance rates. The gradient of U with respect to α_n is

$$\nabla_{\alpha_n} U = \frac{3}{2S} \sum_{j=1}^S \left[\sum_{i=1}^d \frac{\tilde{\mathcal{A}}_i}{\alpha_{nj} \tilde{\mathcal{A}}_i + 1} - \langle 2(\alpha_{nj} \tilde{\mathcal{A}} + 1)^5 \tilde{\mathcal{A}} \tilde{v}_{nj}, \tilde{v}_{nj} \rangle \right], \quad (11)$$

where $\tilde{\mathcal{A}} = -2 \sum_{i=1}^d (\cos(2\pi\xi_i) - 1)$. Here $\tilde{\mathcal{A}}$ denotes a discrete Fourier Laplacian operator with a d -dimensional frequency vector.

Starting from the current point α_n and initial random auxiliary variable γ , the Hamiltonian system is integrated forward in time by Eq. (10) to produce a candidate point $(\hat{\alpha}_n, \hat{\gamma})$. The candidate point $\hat{\alpha}_n$ is accepted as a new point in the sample with probability $p(\text{accept}) = \min(1, -U(\hat{\alpha}_n) - W(\hat{\gamma}) + U(\alpha_n) + W(\gamma))$.

Maximization: Gradient Ascent. We derive the maximization step to update the parameters $\Theta = \{I, \tilde{v}_n, \sigma^2, k, \beta\}$ by maximizing the HMC approximation of the expectation Q in Eq. (9).

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

$$I = \frac{\sum_{j=1}^S \sum_{n=1}^N (I_n \circ \phi_{nj}^{-1}) \cdot |D\phi_{nj}^{-1}|}{\sum_{j=1}^S \sum_{n=1}^N |D\phi_{nj}^{-1}|}. \quad (12)$$

Similarly, we obtain the closed-form solution for the noise variance σ^2 after setting the gradient of Q w.r.t. σ^2 to zero

$$\sigma^2 = \frac{1}{MNS} \sum_{n=1}^N \sum_{j=1}^S \|I \circ \phi_{nj} - I_n\|_2^2. \quad (13)$$

The closed-form solutions for hyperparameters k and β are

$$k = \psi^{-1} \left(\frac{1}{NS} \sum_{i=1}^N \sum_{j=1}^S \ln \alpha_{nj} - \ln \beta \right), \quad \beta = \frac{1}{NSk} \sum_{n=1}^N \sum_{j=1}^S \alpha_{nj}. \quad (14)$$

Here ψ is a digamma function, which is the logarithmic derivative of the gamma function $\Gamma(\cdot)$. The inverse of digamma function ψ^{-1} is computed by using a fixed-point iteration algorithm [20].

As there is no closed-form update for initial velocities, we employ a gradient ascent algorithm to estimate \tilde{v}_{nj} . The gradient $\nabla_{\tilde{v}_{nj}} Q$ is computed by a forward-backward sweep approach. Details are introduced in the FLASH algorithm [30].

4 Experimental Evaluation

We compare the proposed model with LDDMM atlas building algorithm that employs single-penalty regularization with manually tuned parameters on 3D

brain images [30]. In HMC sampling, we draw 300 samples for each subject, with initialized value of $\alpha = 10$, $k = 9.0$, $\sigma = 0.05$, and $\beta = 0.1$. An averaged image of all image intensities is used for atlas initialization.

Data. We include 100 3D brain MRI scans with segmentation maps from a public released resource Open Access Series of Imaging Studies (OASIS) for Alzheimer’s disease [10]. The dataset covers both healthy and diseased subjects, aged from 55 to 90. The MRI scans are resampled to 128^3 with the voxel size of $1.25mm^3$. All MRIs are carefully preprocessed by skull-stripping, intensity normalization, bias field correction, and co-registration with affine transformation.

Experiments. We estimate the atlas of all deformed images by using our method and compare its performance with LDDMM atlas building [30]. Final results of atlases estimated from both our model and the baseline algorithm are reported. We also compare the time and memory consumption of proposed model with the baseline that performs HMC sampling in a full spatial domain [31]. To measure the sharpness of estimated atlas I , we adopt a metric of normalized standard deviation computed from randomly selected 3000 image patches [15]. Given $N(i)$, a patch around a voxel i of an atlas I , the local measure of the sharpness at voxel i is defined as $\text{sharpness}(I(i)) = \text{sd}_{N(i)}(I) / \text{avg}_{N(i)}(I)$, where sd and avg denote the standard deviation and the mean of N_i .

To further evaluate the quality of estimated transformations, we perform atlas-based segmentation after obtaining transformations from our model. For a fair comparison, we fix the atlas for both methods and examine the registration accuracy by computing the dice similarity coefficient (DSC) [8] between the propagated segmentation and the manual segmentation on six anatomical brain structures, including cerebellum white matter, thalamus, brain stem, lateral ventricle, putamen, caudate. The significance tests on both dice and sharpness between our method and the baseline are performed.

Results. Fig. 1 visualizes a comparison of 3D atlas on real brain MRI scans. The top panel shows that our model substantially improves the quality of atlas with sharper and better details than the baseline with different values of manually set regularization parameters, e.g., $\alpha = 0.1, 3.0, 6.0, 9.0$. Despite the observation of a smaller value of $\alpha = 0.1$ produces sharper atlas, it breaks the smoothness constraints on the transformation fields hence introducing artifacts on anatomical structures (outlined in purple boxes). The mean and standard deviation of our estimated hyperprior parameters k and β in Eq. (7) over 30 pairwise image registrations are 47.40/7.22, and 0.036/0.005. The bottom panel quantitatively reports the sharpness metric of all methods. It indicates that our algorithm outperforms the baseline by offering a higher sharpness score while preserving the topological structure of brain anatomy.

Fig. 2 reports results of fixed-atlas-based segmentation by performing the baseline with various regularization parameters and our algorithm. It shows the dice comparison on six anatomical brain structures of all image pairs. Our algorithm produces better dice coefficients without the need of parameter tuning.

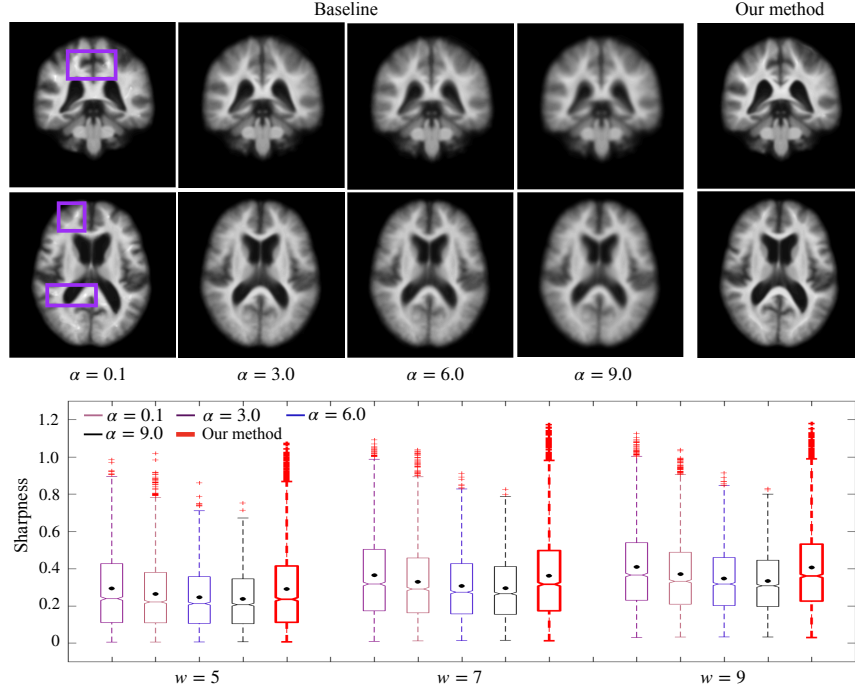


Fig. 1: Top: atlases estimated by baseline with different α and our model (artifacts introduced by small regularization are outlined in purple boxes). Bottom: sharpness measurement of atlas for all methods with different patch size w . The mean of the sharpness metric of **our method** vs. the best performance of baseline without artifacts ($\alpha = 3$) is **0.290/0.264**, **0.362/0.323**, **0.405/0.360**.

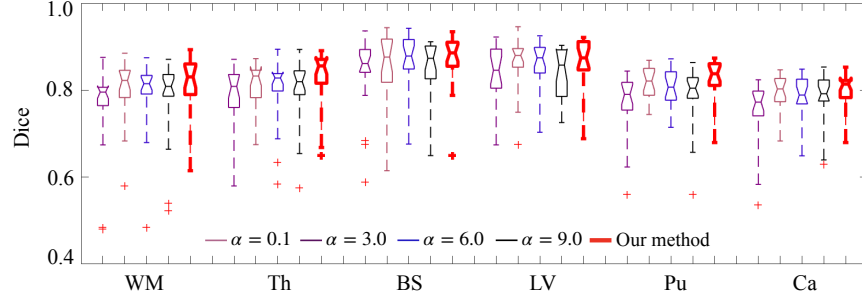


Fig. 2: A comparison of dice evaluation for fixed-atlas-based segmentation on six brain structures (cerebellum white matter (WM), thalamus (Th), brain stem (BS), lateral ventricle (LV), putamen(Pu), caudate (Ca)).

The runtime of our atlas building on 100 3D brain MR images are 4.4 hours with 0.89GB memory consumption. The p-values of significance differences test on both dice ($p = 0.002$) and sharpness ($p = 0.0034$) reject the null hypothesis that there's no differences between our model estimation and baseline algorithms.

5 Conclusion

This paper presents a novel hierarchical Bayesian model for unbiased diffeomorphic atlas building with subject-specific regularization. We design a new parameter choice rule that allows adaptive regularization to control the smoothness of image transformations. We introduce a hierarchical prior that provides prior information of regularization parameters at multiple levels. The developed MCEM inference algorithm eliminates the need of manual parameter tuning, which can be tedious and infeasible in multi-parameter settings. Experimental results show that our proposed algorithm yields a better registration model as well as an improved quality of atlas. While our algorithm is presented in the setting of LDDMM, the theoretical development is generic to other deformation models, e.g., stationary velocity fields [4]. In addition, this model can be easily extended to multi-atlas building where a much higher degree of variations exist in the population studies. Our future work will focus on conducting subsequent statistical shape analysis in the resulting atlas space.

References

1. Allasonnière, S., Amit, Y., Trouvé, A.: Towards a coherent statistical framework for dense deformable template estimation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **69**(1), 3–29 (2007)
2. Allasonnière, S., Kuhn, E.: Stochastic algorithm for parameter estimation for dense deformable template mixture model. *arXiv preprint arXiv:0802.1521* (2008)
3. Arnol'd, V.I.: 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. Arsigny, V., Commowick, O., Pennec, X., Ayache, N.: A log-euclidean framework for statistics on diffeomorphisms. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 924–931. Springer (2006)
5. Ashburner, J., Friston, K.J.: Diffeomorphic registration using geodesic shooting and gauss-newton optimisation. *NeuroImage* **55**(3), 954–967 (2011)
6. Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C.: Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical image analysis* **12**(1), 26–41 (2008)
7. Beg, M.F., Miller, M.I., Trouvé, A., Younes, L.: Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International journal of computer vision* **61**(2), 139–157 (2005)
8. Dice, L.R.: Measures of the amount of ecologic association between species. *Ecology* **26**(3), 297–302 (1945)
9. Duane, S., Kennedy, A.D., Pendleton, B.J., Roweth, D.: Hybrid monte carlo. *Physics letters B* **195**(2), 216–222 (1987)

10. Fotenos, A.F., Snyder, A., Girton, L., Morris, J., Buckner, R.: Normative estimates of cross-sectional and longitudinal brain volume decline in aging and ad. *Neurology* **64**(6), 1032–1039 (2005)
11. Gori, P., Colliot, O., Marrakchi-Kacem, L., Worbe, Y., Poupon, C., Hartmann, A., Ayache, N., Durrleman, S.: A bayesian framework for joint morphometry of surface and curve meshes in multi-object complexes. *Medical image analysis* **35**, 458–474 (2017)
12. Hong, Y., Golland, P., Zhang, M.: Fast geodesic regression for population-based image analysis. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 317–325. Springer (2017)
13. Iglesias, J.E., Sabuncu, M.R., Van Leemput, K.: Incorporating parameter uncertainty in bayesian segmentation models: Application to hippocampal subfield volumetry. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 50–57. Springer (2012)
14. Joshi, S., Davis, B., Jomier, M., Gerig, G.: Unbiased diffeomorphic atlas construction for computational anatomy. *NeuroImage* **23**, S151–S160 (2004)
15. Legouhy, A., Commowick, O., Rousseau, F., Barillot, C.: Online atlasing using an iterative centroid. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 366–374. Springer (2019)
16. Liao, R., Turk, E.A., Zhang, M., Luo, J., Adalsteinsson, E., Grant, P.E., Golland, P.: Temporal registration in application to in-utero mri time series. *arXiv preprint arXiv:1903.02959* (2019)
17. Lorenzo-Valdés, M., Sanchez-Ortiz, G.I., Mohiaddin, R., Rueckert, D.: Atlas-based segmentation and tracking of 3d cardiac mr images using non-rigid registration. In: *International conference on medical image computing and computer-assisted intervention*. pp. 642–650. Springer (2002)
18. Ma, J., Miller, M.I., Trouvé, A., Younes, L.: Bayesian template estimation in computational anatomy. *NeuroImage* **42**(1), 252–261 (2008)
19. Miller, M.I., Trouvé, A., Younes, L.: Geodesic shooting for computational anatomy. *Journal of mathematical imaging and vision* **24**(2), 209–228 (2006)
20. Minka, T.: Estimating a dirichlet distribution (2000)
21. Pohl, K.M., Fisher, J., Grimson, W.E.L., Kikinis, R., Wells, W.M.: A bayesian model for joint segmentation and registration. *NeuroImage* **31**(1), 228–239 (2006)
22. Rohlfing, T., Brandt, R., Menzel, R., Maurer Jr, C.R.: Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. *NeuroImage* **21**(4), 1428–1442 (2004)
23. Simpson, I.J., Cardoso, M.J., Modat, M., Cash, D.M., Woolrich, M.W., Andersson, J.L., Schnabel, J.A., Ourselin, S., Initiative, A.D.N., et al.: Probabilistic non-linear registration with spatially adaptive regularisation. *Medical image analysis* **26**(1), 203–216 (2015)
24. Simpson, I.J., Woolrich, M.W., Andersson, J.L., Groves, A.R., Schnabel, J.A.: A probabilistic non-rigid registration framework using local noise estimates. In: *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*. pp. 688–691. IEEE (2012)
25. Vialard, F.X., Risser, L., Rueckert, D., Cotter, C.J.: Diffeomorphic 3d image registration via geodesic shooting using an efficient adjoint calculation. *International Journal of Computer Vision* **97**(2), 229–241 (2012)
26. Vialard, F.X., Risser, L., Holm, D.D., Rueckert, D.: Diffeomorphic atlas estimation using karcher mean and geodesic shooting on volumetric images. In: *MIUA*. pp. 55–60 (2011)

27. Wang, J., Xing, W., Kirby, R.M., Zhang, M.: Data-driven model order reduction for diffeomorphic image registration. In: International conference on information processing in medical imaging. pp. 694–705. Springer (2019)
28. Wells III, W.M., Viola, P., Atsumi, H., Nakajima, S., Kikinis, R.: Multi-modal volume registration by maximization of mutual information. *Medical image analysis* **1**(1), 35–51 (1996)
29. Yeo, B.T., Sabuncu, M.R., Desikan, R., Fischl, B., Golland, P.: Effects of registration regularization and atlas sharpness on segmentation accuracy. *Medical image analysis* **12**(5), 603–615 (2008)
30. Zhang, M., Fletcher, P.T.: Fast diffeomorphic image registration via fourier-approximated lie algebras. *International Journal of Computer Vision* **127**(1), 61–73 (2019)
31. Zhang, M., Singh, N., Fletcher, P.T.: Bayesian estimation of regularization and atlas building in diffeomorphic image registration. In: Information Processing in Medical Imaging. pp. 37–48. Springer (2013)