A Hierarchical Bayesian Model for Multi-Site Diffeomorphic Image Atlases

Michelle Hromatka[1], Miaomiao Zhang[1], Greg M. Fleishman[2,3], Boris Gutman[2], Neda Jahanshad[2], Paul Thompson[2], and P. Thomas Fletcher[1]

[1] School of Computing, University of Utah, Salt Lake City, UT, USA
[2] Imaging Genetics Center, University of Southern California, Marina del Rey, CA, USA
[3] Dept. of Bioengineering, University of California, Los Angeles, CA, USA

Abstract. Image templates, or atlases, play a critical role in imaging studies by providing a common anatomical coordinate system for analysis of shape and function. It is now common to estimate an atlas as a deformable average of the very images being studied, in order to provide a representative example of the particular population, imaging hardware, protocol, etc. However, when imaging data is aggregated across multiple sites, estimating an atlas from the pooled data fails to account for the variability of these factors across sites. In this paper, we present a hierarchical Bayesian model for diffeomorphic atlas construction of multi-site imaging data that explicitly accounts for the inter-site variability, while providing a global atlas as a common coordinate system for images across all sites. Our probabilistic model has two layers: the first consists of the average diffeomorphic transformations from the global atlas to each site, and the second consists of the diffeomorphic transformations from the site level to the individual input images. Our results on multi-site datasets, both synthetic and real brain MRI, demonstrate the capability of our model to capture inter-site geometric variability and give more reliable alignment of images across sites.

1 Introduction

Recent years have seen a movement towards combining neuroimaging data collected across multiple sites. Such multi-site data has the potential to accelerate scientific discovery by providing increased sample sizes, broader ranges of participant demographics, and publicly available data. Different approaches include large, coordinated multi-site neuroimaging studies, such as the Alzheimer's Disease Neuroimaging Initiative (ADNI) [7], as well as data sharing initiatives that combine multiple independent single-site studies, such as the Autism Brain Imaging Data Exchange (ABIDE) [3]. Larger sample sizes are especially critical in genome-wide association studies (GWAS), in order to provide sufficient statistical power to test millions of genetic variants. This requires aggregation across a broad range of neuroimaging studies, such as those involved in the Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA) consortium [10].

Analysis using these multi-site data sets, however, is not straightforward. Styner et al. [9] compared intra- and inter-site variability by analyzing variability of tissue and structural volumes of the same subject imaged twice at each of five sites. This analysis showed that, regardless of segmentation approach, there was always higher inter-site variability, most notably when those sites used different brands of MRI scanner. In large multi-site studies, there are possibly multiple confounding factors across sites, including different MRI scanners, protocols, populations, and diagnosis techniques.

A common first step in neuroimage analysis is to register all images to a common coordinate system, or atlas, often generated as an unbiased deformable average of the input images themselves. When dealing with multi-site image data, one possibility is to pool the images from all of the sites and estimate an atlas. However, treating multi-site data as a single, homogeneous dataset fails to account for the variability across sites, which can be detrimental to the statistical power and counteract the gains made by increasing the sample size. Another option is to carry out individual image processing and statistical analyses at each site, and combine the statistical results in a post hoc meta-analysis. While this can be an effective way to combine statistical tests of low-dimensional summary measures, it is not applicable to problems on high-dimensional data, such as images, without first learning and establishing a common coordinate frame between the pooled datasets.

We present a hierarchical Bayesian model to estimate atlases on multi-site imaging data, which controls for the inter-site variability, while providing a common coordinate system for analysis. This builds on methods presented in [12], where the large deformation diffeomorphic metric mapping (LDDMM) problem is formulated as a probabilistic model, and the transformations between the atlas and individual images are considered random variables. To build an atlas from multi-site image data, we propose to add an additional layer in this probabilistic model to account for systematic geometric differences between imaging sites, resulting in concatenated transformations, one site-specific, one subject-specific, which describe the deformation of a subject's image to the estimated atlas. Bayesian inference is performed through an iterative, *maximum a posteriori* (MAP) estimate of the random variables until convergence conditions are met. We demonstrate that the resulting model reduces the confounding inter-site variability, and results in improved statistical power in a statistical analysis of brain shape.

2 Background

Single Site Bayesian Atlas Building We will work within the framework of large deformation diffeomorphic metric mapping (LDDMM) [1], as it provides a rigorous setting for defining a distance metric on deformations between images. The atlas building problem in LDDMM can be phrased as a minimization of the sum-of-squared distances function from the atlas to the input images. Images are treated as L^2 functions on a compact image domain Ω. The diffeomorphism

registering the atlas image $A \in L^2(\Omega, \mathbb{R})$ to the input image $I_k \in L^2(\Omega, \mathbb{R})$ will be denoted $\phi_k \in \mathrm{Diff}(\Omega)$. It is given by the flow of a time-varying velocity field, $v_k(t) \in C^\infty(T\Omega)$. In the geodesic shooting framework, this velocity field is governed by the geodesic shooting equation on $\mathrm{Diff}(\Omega)$. As such, it is sufficient to represent a geodesic by its velocity at $t = 0$. Thus, we will simplify notation by excluding the time variable, and write $v_k = v_k(0)$. Given this setup, the atlas building problem seeks to minimize the energy

$$E(v_k, A) = \sum_{k=1}^{N} (Lv_k, v_k) - \frac{1}{\sigma^2} \sum_{k=1}^{N} \|A \circ \phi_k^{-1} - I_k\|_{L^2}^2, \tag{1}$$

where L is a Riemannian metric on velocities, given by a self-adjoint differential operator that controls the regularity of the transformations, and σ^2 is the image noise variance.

Zhang et al. [12] describe a Bayesian interpretation of this diffeomorphic atlas building problem, where the initial velocities, v_k, become latent random variables. In the Bayesian setting, the regularization term on v_k is the log-prior and the image match term is the log-likelihood. We will adopt this probabilistic interpretation, as it allows us to define a hierarchical Bayesian model for multi-site atlas building.

Geodesic Shooting of Diffeomorphisms In the interest of space, we will give a very brief description of geodesic shooting in the space of diffeomorphisms. Given an initial velocity $v \in C^\infty(T\Omega)$, the evolution of the velocity along a geodesic path is given by the Euler-Poincaré equations (EPDiff) [6],

$$\frac{\partial v}{\partial t} = -K \operatorname{ad}_v^* m = -K \left[(Dv)^T m + Dm\, v + m \operatorname{div} v \right], \tag{2}$$

where D denotes the Jacobian matrix, and $K = L^{-1}$. The operator ad^* is the dual of the negative Lie bracket of vector fields,

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

The EPDiff equation (2) results in a time-varying velocity $v_t : [0, 1] \to V$, which is integrated in time by the rule $(d\phi_t/dt) = v_t \circ \phi_t$ to arrive at the geodesic path, $\phi_t \in \mathrm{Diff}^s(\Omega)$.

3 Hierarchical Bayesian Model

We now present our hierarchical Bayesian model for multi-site atlas estimation. The input are images I_{ik}, where $i = 1, \ldots, S$ represents the site index and $k \in 1, \ldots, N_i$ represents the subject index at site i. The goal is to estimate a common atlas image for all sites, $A \in L^2(\Omega, \mathbb{R})$, and simultaneously capture the geometric differences between sites. We represent the inter-site variability as a set of site-specific deformations, ϕ_i, which, when composed with A, describe

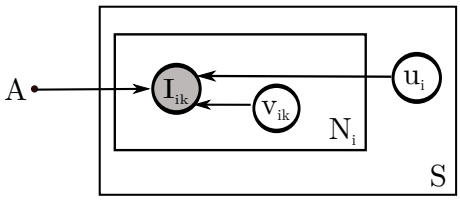

Fig. 1. The hierarchical Bayesian model for multi-site atlas estimation.

site-specific atlases, $A \circ \phi_i^{-1}$. The site-specific atlas is then deformed by the individual-level diffeomorphism, ψ_{ik}, to arrive at the final registration of the atlas, $A \circ \phi_i^{-1} \circ \psi_{ik}^{-1}$, to the input image I_{ik}. As above, we will use the initial geodesic velocity to capture a diffeomorphism: denote u_i for the intial velocity of ϕ_i and v_{ik} for that of ψ_{ik}. This approach gives rise to a hierarchical Bayesian model, shown in Figure 1, which decomposes the diffeomorphic transformation from the atlas A to the individual image I_{ik} into the site-specific initial velocity, u_i, and the subject-specific initial velocity, v_{ik}.

In order to estimate the u_i, v_{ik} in the Bayesian setting, we need to define priors on our random variables:

$$u_i \sim N(0, L^{-1})$$
$$v_{ik} \sim N(0, L^{-1}),$$

where $L = (-\alpha\Delta + I)^c$, the same value discussed in Section 2.

Our model assumes i.i.d. Gaussian image noise, yielding the likelihood:

$$p(I_{ik}|v_{ik}, u_i, A) \sim \prod_{k \in N_i} N(A \circ \phi_i^{-1} \circ \psi_{ik}^{-1}, \sigma^2). \tag{3}$$

While it is not readily apparent that v_{ik}, u_i are related in the priors, these two variables are conditionally dependent, as seen in the head-to-head nature of the sub-model involving v_{ik}, I_{ik}, u_i. When I_{ik} are also random variables, v_{ik} and u_i are independent; when I_{ik} is observed, this introduces a link between v_{ik}, u_i and they are now dependent in their respective posterior distributions:

$$\ln\ p(v_{ik}|I_{ik}, A, u_i; \sigma) \propto -(Lv_{ik}, v_{ik}) - \frac{1}{2\sigma^2}\|A \circ \phi_i^{-1} \circ \psi_{ik}^{-1} - I_{ik}\|^2, \tag{4}$$

$$\ln\ p(u_i|I_{ik}, A, v_{ik}; \sigma) \propto -(Lu_i, u_i) - \frac{1}{2\sigma^2}\sum_{k=1}^{N_i}\|A \circ \phi_i^{-1} - I_{ik} \circ \psi_{ik}\|^2 |D\psi_{ik}|. \tag{5}$$

Performing Bayesian inference on the system described above could be done in several ways. We have chosen a *maximum a posteriori* (MAP) optimization that alternates between maximizing the posterior equations given in (4) and (5) over v_{ik}, u_i, and the closed-form maximization for A given in (6). Alternatively, we could sample the v_{ik}, u_i using Gibbs sampling on the posteriors in a Monte

Carlo expectation maximation (MCEM) procedure. However, the MAP estimate is a mode approximation to the full EM algorithm and sufficient to show the utility of our hierarchical Bayesian model for estimating multi-site image atlases.

We compute the MAP estimates through a gradient ascent procedure in which we alternate between updating the v_{ik} and the u_i, then updating the atlas, A, using an adaptation of the closed form MAP update derived in [12]:

$$A = \frac{\sum_{i \in S} \sum_{k \in N_i} (I_{ik} \circ \psi_{ik} \circ \phi_i) |D(\psi_{ik} \circ \phi_i)|}{\sum_{i \in S} \sum_{k \in N_i} |D(\psi_{ik} \circ \phi_i)|}. \tag{6}$$

This is done iteratively until the energy, $E(v_{ik}, u_i, A)$ converges within reasonable tolerance. In our model, we fix the parameters $\sigma = 0.05$, $\alpha = 3$ and $c = 3$. To compute the gradient w.r.t. the initial velocity v_{ik}, we follow the geodesic shooting algorithm and reduced adjoint Jacobi fields from Bullo [2]. We first forward integrate the geodesic evolution equation (2) along time points $t \in [0, 1]$, and generate the diffeomorphic deformations by $(d/dt)\phi(t, x) = v(t, \phi(t, x))$. The gradient at $t = 1$ is computed as

$$\nabla_{v_{ik}} \ln p(v_{ik} | I_{ik}, A, u_i; \sigma) = -K \left[\frac{1}{\sigma^2} (A \circ \phi_i^{-1} \circ \psi_{ik}^{-1} - I_{ik}) \cdot \nabla (A \circ \phi_i^{-1} \circ \psi_{ik}^{-1}) \right].$$

We then integrate the gradient above backward to time point $t = 0$ by reduced adjoint Jacobi fields to update the gradient w.r.t. the initial velocity v_{ik}. Similarly for u_i, we compute the gradient at $t = 1$ by

$$\nabla_{u_i} \ln p(u_i | I_{ik}, A, v_{ik}; \sigma) = -K \left[\frac{1}{\sigma^2} (A \circ \phi_i^{-1} - I_{ik} \circ \psi_{ik}) \cdot |D\psi_{ik}| \cdot \nabla (A \circ \phi_i^{-1}) \right].$$

4 Results

We present results from a multi-site neuroimaging dataset as a demonstration of the hierarchical model's ability to capture inter-site variability. Furthermore, we show that controlling for this inter-site variability results in improved statistical characterization of the underlying shape variability due to factors of interest, such as age and diagnosis.

Data The Autism Brain Imaging Data Exchange (ABIDE) database is an online consortium of MRI and resting-state fMRI data from 17 international sites, resulting in brain imaging data for 539 individuals with an autism spectrum disorder (ASD) and 573 typically developing (TD) controls [3]. These sites vary in scanner type and imaging protocol, among many other variables. From three sites with different scanner brands, we selected 45 age and group matched subjects (15 from each site), including 27 TD and 18 ASD subjects. The MRIs for these subjects were then skull stripped, motion corrected, bias field corrected, rigidly registered, and intensity normalized prior to analysis.

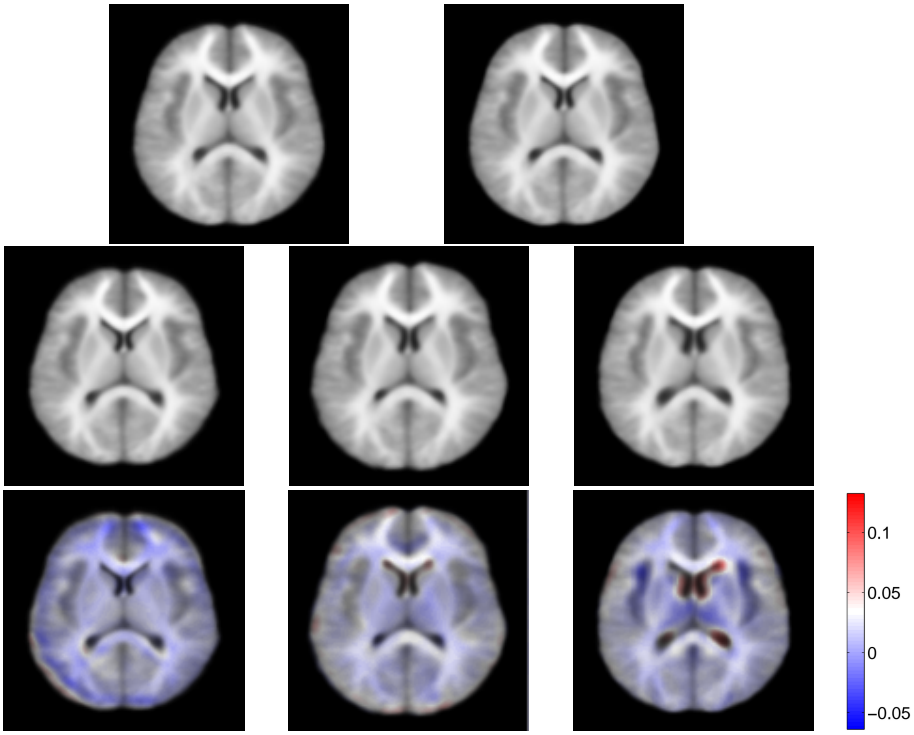

Fig. 2. Axial slices of the estimated atlases: pooled (top left), multi-site (top right), site-specific (middle) and the site-specific atlases overlaid with the log-determinant Jacobian of the site-specific deformations (bottom). A negative value (blue) denotes shrinkage and positive value (red) denotes expansion from the site level to the estimated atlas.

Statistical Comparison We compared our hierarchical multi-site atlas to a single atlas computed from the pooled data. The results are shown in Figure 2. Notice that the top-level atlas A estimated in the multi-site model is similar to the single atlas from the pooled data. However, our model captures differences between the site-specific atlases, as can be seen in the log-determinant Jacobian maps of the site-specific diffeomorphisms. Similar analysis was performed on the pooled results, using the estimated initial velocities within each site for the estimated pooled atlas to generate the site deformations. This analysis showed that just pooling the data and estimating site-specific deformations yields smaller differences across sites in terms of magnitude and average deformation.

In order to focus on biological shape variability, we can control for the confounding inter-site variability by "removing" the estimated site-specific transformations from our multi-site approach. We do this by adjoint transport of the initial velocity fields, v_{ik}, from the site atlas back to the top-level atlas coordinates, giving transported velocities, $\tilde{v}_{ik} = \mathrm{Ad}_{\phi_i^{-1}} v_{ik}$.

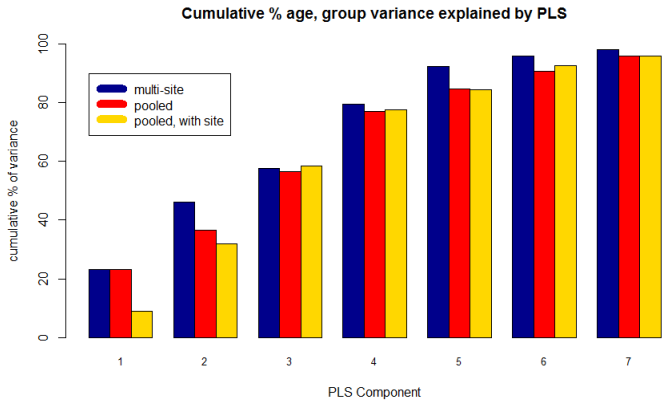

Fig. 3. The cumulative age and group variance (blue, red) or age, group and site (gold) accounted for by each PLS component of the velocity fields.

To test if our multi-site atlas is better able to capture age-related and diagnosis-related shape variability, we use partial least squares (PLS) regression on the diagnosis (ASD or TD) and age versus the initial velocity fields $\tilde{v}_{ik}$. PLS seeks to predict a set of dependent variables from a set of independent variables, yielding, among other analyses, what percentage of the variance with respect to the independent variables is accounted for by the dependent variables [11]. In our case, the dependent variables for analyzing the multi-site method are age and diagnosis, and for the pooled method are age, diagnosis and site versus the initial velocity fields. Figure 3 shows that variance in the multi-site velocity field data is able to explain age and diagnosis in fewer PLS components than in the velocities estimated by pooling the data, even when site is included as a dependent variable. Of course with enough dimensions, both methods are able to fully explain the responses, but our hierarchical model explains the variance more efficiently. Additionally, the total variance of the system described by PLS on the multi-site data was nearly 14% lower than the total variance on pooled data, for totals of 1129056 and 1305985 respectively. This shows that our proposed hierarchical model is able to more efficiently capture biological variability in shape of the multi-site ABIDE dataset by reducing the overall inter-site variance.

5 Conclusion

We have presented a novel approach to the problem of multi-site atlas estimation and shown its utility in reducing the variability of high-dimensional, multi-site imaging data. Our hierarchical Bayesian model captures inter-site variability in site-specific diffeomorphisms which, when composed with diffeomorphisms at the individual level, achieve the final diffeomorphic transformations between the atlas and input images. Our multi-site model was able to reduce overall

variability and capture the relevant variability of the data, i.e., that due to age or diagnosis, in fewer PLS components than its pooled atlas counterpart.

We chose to compare our approach to a single atlas on the pooled data due to the prevalence of this approach in the literature, e.g., [4, 8]. An alternative comparison would be to use a multi-atlas [5] on the pooled data to determine if our decrease in variance is due solely to the fact that we are estimating multiple atlases. The difference is that our multi-site atlas directly models and controls for the inter-site variance. We leave comparison to a multi-atlas as future work.

Acknowledgements This work was supported by NIH Grant R01 MH084795, NSF Grant 1251049, and NSF CAREER Grant 1054057.

References

1. 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)
2. Bullo, F.: Invariant affine connections and controllability on lie groups. Tech. rep., Geometric Mechanics, California Institute of Technology (1995)
3. Di Martino, A., Yan, C.G., Li, Q., Denio, E., Castellanos, F.X., Alaerts, K., Anderson, J.S., Assaf, M., Bookheimer, S.Y., Dapretto, M., et al.: The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Molecular Psychiatry (2013)
4. Jahanshad, N., Kochunov, P.V., Sprooten, E., Mandl, R.C., Nichols, T.E., et al.: Multi-site genetic analysis of diffusion images and voxelwise heritability analysis: a pilot project of the enigma–dti working group. Neuroimage 81, 455–469 (2013)
5. Leporé, N., Brun, C., Chou, Y.Y., Lee, A., Barysheva, M., De Zubicaray, G.I., et al.: Multi-atlas tensor-based morphometry and its application to a genetic study of 92 twins. In: 2nd MICCAI Workshop on Mathematical Foundations of Computational Anatomy. pp. 48–55 (2008)
6. Miller, M.I., Trouvé, A., Younes, L.: Geodesic shooting for computational anatomy. Journal of Mathematical Imaging and Vision 24(2), 209–228 (2006)
7. Mueller, S.G., Weiner, M.W., Thal, L.J., Petersen, R.C., Jack, C.R., Jagust, W., Trojanowski, J.Q., Toga, A.W., Beckett, L.: Ways toward an early diagnosis in Alzheimer's disease: The Alzheimer's Disease Neuroimaging Initiative (ADNI). Alzheimer's & Dementia 1(1), 55–66 (2005)
8. Qiu, A., Miller, M.I.: Multi-structure network shape analysis via normal surface momentum maps. NeuroImage 42(4), 1430–1438 (2008)
9. Styner, M.A., Charles, H.C., Park, J., Gerig, G.: Multisite validation of image analysis methods: assessing intra-and intersite variability. In: Medical Imaging 2002. pp. 278–286. International Society for Optics and Photonics (2002)
10. Thompson, P.M., Stein, J.L., Medland, S.E., Hibar, D.P., Vasquez, A.A., Renteria, M.E., Toro, R., Jahanshad, N., Schumann, G., Franke, B., et al.: The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. Brain Imaging and Behavior 8(2), 153–182 (2014)
11. Wold, H.: Partial least squares. Encyclopedia of statistical sciences (1985)
12. 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 (2013)