ANTS: Open-Source Tools for Normalization And Neuroanatomy

Brian B. Avants, Nicholas J. Tustison, Gang Song, and James C. Gee

Abstract— **Computational anatomy (CA) seeks to quantify natural variation in biological shape and function with roots that reach back to the seminal works of Charles Darwin and D'arcy Thompson. CA is currently applied to study health, disease and epidemiology and uses deformable mappings between images as a basic technique. However, there is a lack of standards and reproducibility in the field that is due, in part, to the use of proprietary software and private data that is difficult to openly evaluate. To facilitate reproducibility in CA measurements and advancement of imaging sciences, NIH has recently committed significant support to open-source data and software resources. Here, we report a recent product of this commitment: Advanced Normalization Tools (ANTS), an ITK-based toolkit for CA and related areas. The ANTS open-source library consists of a well-evaluated suite of state-of-the-art normalization, segmentation and template-building tools for quantitative morphometric analysis. We highlight the prominent features of ANTS, including diffeomorphic normalization methods, and demonstrate its utility by performing a detailed analysis on openly-available anatomically labeled brain data from the non-rigid image registration evaluation project (NIREP). The results from this analysis evidences the high level of accuracy achievable with ANTS using intensity-based registration and segmentation. In addition, we show the significant performance gains may be achieved by coupling intensity-based image metrics and label-based metrics from specific, sensibly selected cortical structures. Additional features are highlighted in the appendix.**

I. Introduction

The rapid advancement of biological and medical imaging technologies has caused a proliferation in the development of quantitative tools for computational anatomy. The principal tools of this emerging field are deformable mappings between images whether they be driven by similarity metrics which are intensity-based, point-set based, or both. Several categories of mappings have been proposed in the literature. Of particular recent interest are diffeomorphic transformations which, by definition, preserve topology. Topology preservation is fundamental to making comparisons between objects in the natural world that are thought to change in such a way that local neighborhoods are preserved. Cytoarchitectonic brain mapping studies also suggest that the layout of cell types throughout the brain is generally preserved [1], further motivating diffeomorphic mapping in the context of the brain.

Despite the number of proposed algorithms, our limited assessment of published research mirrors the experience of many others who prefer a working paradigm of what has been referred to as *reproducible research*. As described by Dr. Kovacevic "[reproducible research] refers to the idea that, in 'computational' sciences, the ultimate product is not a published paper but, rather, the entire environment used to produce the results in the paper (data, software,etc.)." After an informal survey of 15 published papers, she found "none had code available" and "in only about half the cases were the parameters [of the algorithm] specified" [2]. Recent discussions within the computational sciences research community, particularly among advocates of "open science," have also voiced similar concerns [3], [4]. In this paper, we discuss our contribution to the open-source medical image analysis research community which we call ANTS (Advanced Normalization Tools). Built on an ITK framework, this software package comprises a suite of tools for image normalization and template building based on previously published research.

Perhaps the most persuasive evidence motivating the use of our contributions discussed in this paper is the recent outcome of a large-scale comparative image registration algorithm assessment [5]. In the largest evaluation study to date involving 14 popular non-linear registration algorithms, our Symmetric Normalization (SyN) transformation model [6] discussed below, was consistently one of the top two performers across all tests. Overlap and distance measures used for assessment employed three completely independent analysis methods (permutation tests, one-way ANOVA tests, and indifference zone ranking). Unlike some of the other algorithms used in this brain registration evaluation study, all of our methods (not just SyN) are available as open-source.

We first provide an overview of the different components included in ANTS, such as the available transformation models and similarity metrics offered. This is followed by an extensive experimental analysis that builds upon the results from the recent image normalization evaluation of [5] which was limited to a single configuration of ANTS. A set of ANTS tools are evaluated on the sixteen subjects from the public, expert-labeled neuroanatomy in the NIREP dataset. This will provide a useful benchmark for evaluating other possible ANTS configurations. Importantly, this work also suggests—to our knowledge, for the first time—the clear practical benefit of time-parameterized diffeomorphic mapping over the greedy and exponential mapping approaches.

The paper is organized as follows. Section II gives an overview of the transformation models and similarity metrics available in ANTS, along with optimization schemes. Section III connects the theory with the user interface. Section IV, first, illustrates the performance of different transformation models on a pair of "classic" image registration examples. Second, we report results on a series of large-scale experiments that profile ANTS registration schema against a cortical labeling problem using the NIREP evaluation dataset. Finally, we close with discussion of the toolkit and our findings.

II. Theoretical Overview of ANTS

A useful classification schema of normalization techniques is based upon the following three principal components [7], [8]:

- the *transformation model*, which includes the regularization kernels,
- the *similarity (or correspondence) measures*, and
- the *optimization strategy*.

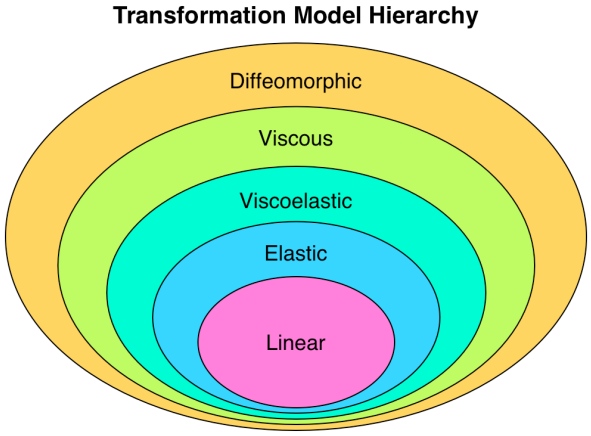
In general, image normalization is the process of finding the optimal transformation, ϕ, within a specified transformation space which maps each $\mathbf{x}$ of image $\mathcal{I}(\mathbf{x})$ to a location in image $\mathcal{J}(\mathbf{z})$ such that a specified cost function, $\mathcal{C}$, describing the similarity between $\mathcal{I}$ and $\mathcal{J}$, is minimized. A summary of available transformation models and similarity measures are provided in Table I. Details are given in subsequent sections.

A. ANTS Transformation Models

A variety of transformation models have been proposed in the literature with varying degrees of freedom (illustrated in Figure 1). For deformable transformations, one approach is to optimize within the space of non-topology preserving, yet physically motivated transformations—an approach pioneered by Bajcsy [9]. Elastic and related models, such as HAMMER [10], statistical parametric mapping (SPM) [11], free-form deformations (FFD) [12], and Thirion's demons [13] operate in the space of vector fields, which does not preserve topology. In other words, barring ad hoc constraints to prevent otherwise, these algorithms may allow the topology to change in an uncontrolled way which makes the deformable mappings difficult to interpret in functional or anatomical studies.

Thus, in addition to shape-based, biological motivation, diffeomorphic mapping is motivated by the desire for deformable transformations that provide well-behaved solutions with mathematical guarantees about distances and regularity. Furthemore, the diffeomorphic space has group structure [14]. Optimizing directly within this space has shown remarkable success in various computational anatomy studies involving longitudinal [15], [16], functional [17], and population data [15]. We include three such diffeomorphic algorithms in the ANTS toolbox based on previously reported research and a new time-parameterized extension to the standard symmetric normalization algorithm [6].

Regardless of current research trends, however, we recognize that selection of the transformation model is ultimately application-specific, that no single choice is optimal for all scenarios [18], and, therefore, the transformation model must be chosen in a principled fashion. In fact, several non-diffeomorphic algorithms performed quite well in Klein's comparative study of nonrigid registration algorithms [5]. For this reason, ANTS also includes elastic-related methods for a panoply of transformation model options. As we show below, all of these methods may be implemented, when using gradient descent optimization, in a generic framework that generates transformations (when necessary) and updates transformations according to the needs of each model. ANTS implementation follows this general form for each model.



1: Diagrammatic illustration of the transformation model hierarchy where the encompassing transformation spaces are characterized by increasing degrees of freedom.

1) Rigid and Affine Linear Transformations : Image registration strategies often begin with a linear transformation for initial global alignment which precedes a deformable transformation with increased degrees of freedom. The linear transformations available within ANTS optimize either a mean-squared difference (MSQ) or mutual information (MI) similarity metric which are optimized with respect to translation, rotation, and, in the case of affine transformations, scaling and shearing. The successive optimization of each component of the linear transformation allows for careful control over increasing degrees of freedom. ANTS also composes the affine transformation with the deformable transformation field before performing any interpolation or downsampling. In this way, ANTS normalization never requires more than a single image interpolation step and is able to always refer back to the original full-resolution image sources. The ANTS implementation of rigid mapping is quaternion-based with additional scaling and shearing terms when affine mapping is desired (via the `--rigid-affine true` flag).

2) Vector Field Operators for Regularization: Deformable normalization strategies typically invoke a deformation regularization step which smooths the displacement field, $\mathbf{u}$, or velocity field, $\mathbf{v}$, or both by a linear operator such as the Laplacian or Navier-Stokes operator. One may write this regularization as a variational minimization in terms of its linear operator or in terms of a kernel function operating on the field itself, e.g., $\mathbf{u}_{smooth} = K \star \mathbf{u}_{not\ smooth}$, where $K\star$ denotes convolution with the Green's kernel, K, for the linear operator, L. Regularization models operate on either the whole mapping ϕ or the gradient of the similarity term or both. The same regularization schema is available for both diffeomorphic and the recently proposed directly manipulated free-form deformation (DMFFD) [19], allowing regularization of both total deformation and deformation update. Viewed from this perspective, hybrid configurations incorporating discretized FFD strategies and diffeomorphisms can be combined for novel image normalization approaches. ANTS enables a variety of choices for K including the Gaussian with varying

Category	Transformation, ϕ	Similarity Measures	Brief Description
Linear	Rigid†	MI, MSQ	translation and rotation
	Affine†	MI, MSQ	rigid, scaling, and shear
Elastic	Deformable	CC, PR, MI, MSQ, JHCT, PSE	Demons-like algorithm
	DMFFD	CC, PR, MI, MSQ, JHCT, PSE	FFD variant
Diffeomorphic	Exponential	CC,PR, MI, MSQ, JHCT, PSE	minimizes $\boldsymbol{v}(\mathbf{x})$
	Greedy SyN†	CC, PR, MI, MSQ, JHCT, PSE	minimizes $\boldsymbol{v}(\mathbf{x},\mathbf{t})$ locally in time
	Geodesic SyN†	CC, PR, MI, MSQ, JHCT, PSE	minimizes $\boldsymbol{v}(\mathbf{x},\mathbf{t})$ over all time

I: Available transformations and similarity metrics available in ANTS. Similarity metric acronyms: CC = cross correlation, PR = probabilistic matching, MSQ = mean squared difference, MI = mutual information, JHCT = Jensen-Havrda-Charvat-Tsallis divergence, PSE = point-set expectation. ANTS also provides the inverse of those transformations denoted by the '†' symbol.

σ and a variety of B-spline functions, both of which induce adequate regularity for normalization models used in ANTS. While additional physical operators will be included in future releases, current B-spline options provide many orders of flexibility [20].

3) Diffeomorphic Transformations: Diffeomorphisms form a group of differentiable maps with differentiable inverse [21], [22] that is closed under composition. In contrast, the vector space that most deformable image registration methods use is closed under *addition*, an operation that cannot guarantee topology preservation. That is, two topology-preserving vector fields added together are still a vector field but the result may no longer preserve topology. Furthermore, most regularization models for vector spaces use a quadratic penalty, thus making large deformations difficult to realize. Modeling transformations with diffeomorphisms, on the other hand, ensures a flexible, linear penalty on deformation while guaranteeing topology preservation. Additionally, distance metrics in the space of diffeomorphisms allow geodesic properties to be explored [23], [24].

ANTS assumes the diffeomorphism, ϕ, is defined on the image domain, Ω, and maintains an affine transform at the boundary such that $\phi(\partial\Omega) = A(\mathbf{Id})$ where $A(\mathbf{Id})$ is an affine mapping applied to the identity transformation. The map ϕ, over time, parameterizes a family of diffeomorphisms, $\phi(\mathbf{x},t) : \Omega \times t \rightarrow \Omega$, which can be generated by integrating a (potentially) time-dependent, smooth velocity field, $\boldsymbol{v} : \Omega \times t \rightarrow \mathbb{R}^d$, through the ordinary differential equation (o.d.e.)

$$\frac{d\phi(\mathbf{x},t)}{dt} = \boldsymbol{v}(\phi(\mathbf{x},t),t), \;\; \phi(\mathbf{x},0) = \mathbf{x}. \tag{1}$$

The existence and uniqueness theorem for o.d.e.'s implies that integrating Equation (1) generates a diffeomorphism. The deformation field yielded by ϕ is $\mathbf{u}(\mathbf{x}) = \phi(\mathbf{x},1) - \mathbf{x}$.

Dupuis et al. [25] motivated the usage of diffeomorphisms for CA by showing that the variational form

$$D(\mathcal{I},\mathcal{J}) = \int_0^1 ||L\boldsymbol{v}||dt, \; \mathcal{I}(\phi(\mathbf{x},1)) = \mathcal{J}(\mathbf{z}) \tag{2}$$

represents a true mathematical metric between anatomical instances $\mathcal{I}$ and $\mathcal{J}$ given an appropriate norm, L, on the velocity field, $\boldsymbol{v}$. An optimal solution, $\boldsymbol{v}^*$, minimizes the metric $D(\mathcal{I},\mathcal{J})$ with respect to L. Dupuis [25] also showed that such a solution is guaranteed to be well-posed. Intuitively, Equation (2) provides a sense of distance between two anatomical shapes. It also illustrates that the optimal diffeomorphic solution is analogous to finding the geodesic curve between two points in a curved space. [1]

In most real-world applications, however, a diffeomorphic path connecting the anatomical instance $\mathcal{J}$ with $\mathcal{I}$ is non-existent (due, for example, to photometric variation, idiosyncratic cortical folding or the presence/absence of a tumor in neuroanatomical images). Therefore, the following minimizing variational form is used for optimization in diffeomorphic normalization to accommodate inexact matching [25], [26]

$$\boldsymbol{v}^* = \underset{\boldsymbol{v}}{\operatorname{argmin}} \left\{ \int_0^1 ||L\boldsymbol{v}||^2 dt + \lambda \int_\Omega ||\mathcal{I} \circ \phi(\mathbf{x},1) - \mathcal{J}||d\Omega \right\}. \tag{3}$$

The Euler-Lagrange equations characterizing the optimizing time-varying velocity field, $\boldsymbol{v}^*$, were derived in [26] and later used in formulating the gradient-descent optimization scheme known as *large deformation diffeomorphic metric-matching* (LDDMM) [24] with the similarity metric, or data term, for LDDMM being the squared intensity difference with weight λ.

To accommodate a variety of medical image normalization tasks, one typically encounters more complex intensity transfers between one anatomical instance $\mathcal{J}$ and another instance $\mathcal{I}$. Thus, ANTS enables not only a variety of similarity metric possibilities beyond the conventional squared difference metric but it also permits any number of different similarity metrics for a particular image normalization task. This leads to the following generalization of Equation (3):

$$\boldsymbol{v}^* = \underset{\boldsymbol{v}}{\operatorname{argmin}} \left\{ \int_0^1 ||L\boldsymbol{v}||^2 dt + \lambda \int_\Omega \Pi_\sim(\mathcal{I}, \phi(\mathbf{x},1), \mathcal{J}) d\Omega \right\} \tag{4}$$

where $\Pi_\sim$ is a similarity metric depending on the images and the mapping and λ controls the degree of exactness in the matching. If $\Pi_\sim$ is selected as cross-correlation, then one is estimating the diffeomorphism under more robust illumination constraints, as described in [6].

[1] It is important to note the similarity between the definition of curve length, $\int ||\mathcal{C}'(t)||dt$, for the parametric curve $\mathcal{C}(t)$ and Equation (2). In this sense, the solution for Equation (2) is the geodesic diffeomorphism, where $\boldsymbol{v}$ is the tangent vector of the diffeomorphism, such that the shape distance, D, between $\mathcal{I}$ and $\mathcal{J}$ is minimized.

Exploiting the fact that the diffeomorphism, ϕ, can be decomposed into two components ϕ_1 and ϕ_2, one may construct a *symmetric* alternative to Equation (4). Now define, in $t \in [0, 0.5]$, $\boldsymbol{v}(x,t) = \boldsymbol{v}_1(x,t)$ and $\boldsymbol{v}(x,t) = \boldsymbol{v}_2(x, 1-t)$ when $t \in [0.5, 1]$. This leads to the symmetric variant of Equation (4),

$$\{\boldsymbol{v}_1^*, \boldsymbol{v}_2^*\} = \underset{\boldsymbol{v}_{1,2}}{\operatorname{argmin}} \Bigg\{ \int_0^{0.5} ||L\boldsymbol{v}_1(x,t)||^2 \, dt + \int_0^{0.5} ||L\boldsymbol{v}_2(x,t)||^2 \, dt + \lambda \int_\Omega \Pi_\sim \left(\mathcal{I} \circ \phi_1(\mathbf{x}, 0.5), \mathcal{J} \circ \phi_2(\mathbf{x}, 0.5)\right) d\Omega \Bigg\}. \tag{5}$$

Note that the regularization term, here, is equivalent to that in equation 3. The only change is the splitting of the integral into two time intervals reflecting the underlying optimized components of the velocity field. The corresponding symmetric Euler-Lagrange equations are similar to [26]. The difference, here, is that in finding $\boldsymbol{v}^*$, we minimize the variational energy from either end-point towards the mid-point of the transformation, as indicated by the data term. This strategy "splits" the optimization dependence equally between both images. Thus, gradient-based iterative convergence deforms $\mathcal{I}$ and $\mathcal{J}$ along the geodesic diffeomorphism, ϕ, to a fixed point midway (intuited by the notion of shape distance) between $\mathcal{I}$ and $\mathcal{J}$ thus motivating the denotation of the solution strategy as Symmetric Normalization (SyN).

Other diffeomorphic algorithms have since been reported in the research literature e.g. DARTEL [27] and Diffeomorphic Demons [28], [29], both of which use the constant velocity, exponential model for generating diffeomorphisms. We thus include three diffeomorphic transformation models for parameterizing $\phi(\cdot)$. These include Geodesic SyN, Greedy SyN, and exponential mapping. As summarized in Table I, each of these transformation models can utilize a host of similarity measures both individually and in mutual combination.

a) Geodesic SyN: Using a gradient-based optimization strategy for minimizing equation (5) first requires a specified discretization of t, s.t. $t \in [0,1] \approx \{0, 1/k, 2/k, \cdots, 1\}$ where integer $k > 1$ is the desired number of discretized intervals. Calculation of the gradient of Π with respect to the diffeomorphisms ϕ_1 and ϕ_2 is performed at each of these t values, t_k,

$$\nabla \Pi_i(\boldsymbol{x}, t_k) = \frac{\partial}{\partial \phi_i} \Pi_\sim (\mathcal{I}(\phi_1^{-1}(\boldsymbol{x}, t_k)), \mathcal{J}(\phi_2^{-1}(\boldsymbol{x}, 1 - t_k)) \tag{6}$$

for $i \in \{1, 2\}$. The total velocity field is then updated from the previous iteration according to, for each i,

$$\boldsymbol{v}(\mathbf{x}, t) = \boldsymbol{v}(\mathbf{x}, t) + \delta K \star \nabla \Pi_i(\mathbf{x}, t), \tag{7}$$

where δ is the user-specified gradient descent parameter. Note that the update, here, is a $N+1$-dimensional vector field and K is a $N+1$-dimensional operator, when the images are N dimensions. We generate $\phi_i(\mathbf{x}, t)$ for each $t \in [0, 1]$ and $i \in \{1, 2\}$ by integrating Equation (1) using Runge-Kutta methods. We cycle through these steps until convergence or iterative exhaustion.

b) Greedy SyN: Although the Geodesic SyN algorithm conforms most closely to the theoretical diffeomorphic foundations culminating with Equation (5), the computational and memory cost is significant due to the dense-in-time gradient calculations and requisite reintegration of the diffeomorphisms after each iterative update. As a lower-cost alternative, we offer a greedy variant which performs quite well for most medical image normalization problems we have encountered. Additionally, this was the strategy used in the large-scale comparative image registration algorithm assessment of [5].

Greedy optimization of Equation (5) calculates the gradient only at the mid-point of the full diffeomorphism, i.e. at $t = 0.5$ in equation 6

$$\nabla \Pi = \frac{\partial}{\partial \phi_i} \Pi_\sim (\mathcal{I}(\phi_1^{-1}(\mathbf{x}, 0.5)), \mathcal{J}(\phi_2^{-1}(\mathbf{x}, 0.5))) \tag{8}$$

for $i \in \{1, 2\}$. $\phi_1(\mathbf{x}, 0.5)$ and $\phi_2(\mathbf{x}, 0.5)$ are then updated from the previous iteration according to

$$\phi_i(\mathbf{x}, 0.5) = \phi_i(\mathbf{x}, 0.5) + (\delta K \star \nabla \Pi_i(\mathbf{x}, 0.5)) \circ \phi_i(\mathbf{x}, 0.5). \tag{9}$$

That is, the gradient at the mid-point is mapped back to the origin of each diffeomorphism. We then update the full mapping by explicitly enforcing $\phi^{-1}(\phi(\mathbf{x}, 1)) = \mathbf{x}$ in the discrete domain, as described in [6].

c) Exponential Mapping: Ashburner introduced DARTEL (Diffeomorphic Anatomical Registration using Exponentiated Lie algebra) as a rapidly computed alternative to time parameterized diffeomorphic schemes [27]. The key difference between a time-varying diffeomorphism and a diffeomorphism generated by an exponential mapping [27] is that the exponential mapping maintains only a single vector field that is constant in time. A diffeomorphism can be generated by exponentiation of a constant velocity field from the following o.d.e (cf Equation (1))

$$\frac{d\phi(\mathbf{x}, t)}{dt} = \boldsymbol{v}(\phi(\mathbf{x}, t)), \;\; \phi(\mathbf{x}, 0) = \mathbf{x}. \tag{10}$$

Note that there is no explicit time parameter in the velocity field. Theoretically, restricting the velocity field to be constant in time reduces the size of the space that may be generated [14] in a way that is similar to the difference between real and rational numbers, the latter of which are sparsely distributed through the reals. ANTS registration via exponential mapping optimizes the functional,

$$||L\boldsymbol{v}(x)||^2 + \lambda \int_\Omega \Pi_\sim \left(\mathcal{I}, \mathcal{J} \circ \phi(\mathbf{x}, 1)\right) d\Omega. \tag{11}$$

ANTS currently updates the velocity field based on the gradient only at the end-point. That is, the gradient is computed after the image $\mathcal{J}$ is mapped to the "fixed" image $\mathcal{I}$. In contrast, DARTEL optimizes over the time discretization, computing the gradient at each discrete time point, as in the full ANTS geodesic optimization. However, ANTS also allows a greedy exponential mapping strategy that is analogous to

the method described in [29] (`GreedyExp`). An important difference between Diffeomorphic Demons and exponential mapping is that Diffeomorphic Demons composes exponential maps together. It is therefore closer to a greedy descent strategy similar to that used in Christensen's classic approach [30].

4) Vector Space Transformations: Potential mapping solutions to the image matching problem operating in vector spaces are constructed in a similar variational form as that for the diffeomorphic formulation. We write this general variational energy, Π, as

$$\Pi(\mathcal{I},\mathcal{J},\phi) = \int_{\Omega} (\Pi_{\sim}(\mathcal{I},\mathcal{J},\phi)(\mathbf{x}) + \lambda\ \Pi_R(\phi)(\mathbf{x}))\, d\Omega, \quad (12)$$

where $\mathcal{I}$ and $\mathcal{J}$ are, again, the moving and fixed images, respectively, and ϕ is the transformation which maps between $\mathcal{I}$ and $\mathcal{J}$. $\Pi_{\sim}$ is the similarity metric and Π_R is the explicit regularization term.

a) Gaussian-Regularized Elastic Deformation: A simple and efficient, yet powerful image normalization algorithm is the approach known as Thirion's demons [13]. Using an optical flow based similarity, the solution is obtained by iterating between the calculation of image forces and subsequent Gaussian regularization. In ANTS, we extend this basic approach to include the similarity metrics available for deformable registration (see Table I). This leads to an update form that is performed as follows:

$$\begin{aligned} \mathbf{U} &= \frac{\partial}{\partial\phi}\Pi_{\sim}(\mathcal{I}(\mathbf{x}),\mathcal{J}(\phi^{-1}(\mathbf{x}))), \\ \mathbf{u}(\mathbf{x}) &= K_{\mathbf{u}} \star (\mathbf{u}(\mathbf{x}) + \delta K_{\mathbf{U}} \star \mathbf{U}), \end{aligned} \quad (13)$$

where $K_{\mathbf{u}}$ regularizes the total deformation field and $K_{\mathbf{U}}$ regularizes the gradient field. [2] Thus, ANTS vector field transformation models default to a Demons style optimization. Like the Demons algorithm, this method uses a greedy descent on the energy and does not truly optimize the total variational form. The update strategy for the exponential mapping is very similar to the above, but with an additional composition on the gradient that maps the gradient back to time zero. See [29] for a nice discussion of the distinction between additive and compositive updates in these methods.

B. ANTS Intensity-Based Similarity Metrics

Several intensity-based image metrics have been proposed in the literature with varying levels of performance dependent upon specific applications. We have included several of the most popular similarity metrics within ANTS. In addition, our software framework facilitates the development of other image metrics. Both mutual information [31] and mean-squared difference similarity metrics are available for the linear transformations. Also included are the cross-correlation (the PR metric) [15], local mutual information [12], [32], and mean squared difference similarity metrics for the non-linear transformation models. The parameters for the different metrics are discussed in the ANTS documentation [33]. The cross-correlation uses a local neighborhood in a spatially varying correspondence model which provides robustness to illumination and inhomogeneity. Therefore, we choose the correlation metric for most practical image registration problems in real imagery.

[2] Any ANTS transformation model may be used with both update and total deformation regularization.

C. ANTS Label or Point-Based Similarity Metrics

In addition to intensity-based metrics, ANTS also contains similarity metrics for registering labeled point-sets or label images. These include a landmark matching metric and two point-set metrics which can accommodate point-sets of different cardinality. These point-set metrics can be used alone for strict point-set registration or in conjunction with intensity-based metrics for dual intensity/point-set registration. Exact matching and partial (or incompletely labeled) point set matching are available.

1) Exact Landmark Matching: Generalizing the B-spline fitting algorithm of [34], we developed a scattered data approximation algorithm [35] and contributed the code to the ITK library [20]. This code is included in ANTS and forms the basis of our exact landmark matching where this metric seeks to minimize the weighted sum of distances between corresponding landmarks using a hierarchical approach. One can also associate relative confidence values with each landmark for fine-tuning exact landmark matching results.

2) Point-Set Expectation: In [36], the point-set matching problem was formulated in the context of incomplete label matching but is equally applicable to the general scenario of registering point-sets not necessarily of equivalent cardinality. Given two point-sets, X and Y, the essential idea underlying the point-set expectation matching algorithm is that the optimal solution minimizes the distance between each point $y \in Y$ with its corresponding *expected* point in X. When discussing these point-based metrics, we abuse notation slightly by "hiding" the application of the deformation to the points to simplify the notation. However, in practice, we apply the same mapping to the points as we do to the imagery that is associated with them and that defines their spatial domains. Thus, the transformations that appear in the image similarity term are also used in the point similarity for all below.

We calculate the expected point using a Bayesian formulation and a non-parametric Parzen windowing scheme. This allows one to define the probability of the point $x \in X$ given a point $y \in Y$ as

$$\mathbf{P}(X = x|Y = y) = G(y; x, \sigma_X) \quad (14)$$

where $G(y; x, \sigma_X)$ is a normalized Gaussian with mean x and standard deviation σ_X. The expected point $\mathrm{E}(X|y)$ is then calculated to be

$$\begin{aligned} \mathrm{E}(X|Y = y) &= \sum_{j=1}^{|X|} \mathbf{P}(X = x_j|Y = y)x_j \\ &= \frac{1}{|X|}\sum_{j=1}^{|X|} G(y; x_j, \sigma_X)x_j \end{aligned} \quad (15)$$

where $|\cdot|$ denotes cardinality. The weighted sum of distances between the points in Y and their corresponding expected

points in X is calculated from Equation (15), i.e.

$$\mathrm{PSE}(X,Y) = \frac{1}{|Y|}\sum_{i=1}^{|Y|}\left\| y_i - \frac{1}{|X|}\sum_{j=1}^{|X|} G(y_i; x_j, \sigma_X)x_j \right\|^2 . \tag{16}$$

The ANTS parameters for this metric require one to choose the percentage of points to use from the input data (subset selection increases efficiency), the Parzen window sigma (σ_X), the number K for the number of nearest neighbors used in the Parzen window calculation (again for efficiency) and, finally, the number of iterations over which one optimizes the similarity symmetrically. This latter option is useful for partial matching problems as in [36]. A related similarity metric, based on maximizing the Jensen-Havrda-Charvat-Tsallis Divergence between the point sets, is described in the appendix.

III. ANTS Implementation and Usage

ANTS, built upon an ITK foundation, maintains the same coding style as its base. For much of its functionality, ANTS requires ITK, necessitating the installation of ITK prior to installing ANTS. All ANTS source code is available via the online source code repository SourceForge.[3] Binaries for Windows, OSX, 32 and 64-bit Linux are also available from the same online location. For quality assurance and maintenance purposes we have established an ANTS test reporting open-source "dashboard"[4] on our lab website[5] to monitor compilation and testing of the ANTS program. A screenshot from a daily testing period is given in Figure 2. Such a configuration facilitates reporting of problems encountered by users on a multitude of computing platforms.

Based on our experience with standard command line argument parsing packages (e.g. `getopt`), we developed our own set of classes for an intuitive command line interface. A summary of command line arguments are given in Table II. These ANTS argument parsing classes provide an intuitive compromise between parsers where every variable requires a unique flag and strict ordering requirements on the command line. A multivariate command line call to ANTS is given by:

```
>ANTS 3
  --metric MSQ[fixedImage.nii,movingImage.nii,1]
  --metric PSE[fixedImage.nii,movingImage.nii,
     fixedLabels.nii,movingLabels.nii,0.25,0.1,100,0,10]
  --transformation SyN[1,2,0.1] --geodesic 2
  --regularization Gauss[6.0,0.25]
  --iterations 50x20x10x5
  --output-naming results.nii
```

The MSQ metric uses weight 1 while the PSE metric has weight 0.25. In the PSE metric, ten percent of points are selected from the Label images (param 0.1). The point set variance is 100, points are selected densely (not just from the boundaries of the labels) and 10 nearest neighbors are used. The transformation model is the full time optimization of the symmetric SyN transformation (chosen by including –geodesic 2) with gradient step 1, two time points in the spatiotemporal discretization and a time-step of 0.1 in the Runge-Kutta integration that generates the diffeomorphisms. The correspondence between the ANTS command line specification and the image normalization formulation with images I and J and their corresponding labeled images X and Y illustrates the motivation for our command line interface,

$$\begin{aligned}\Pi(\mathcal{I},\mathcal{J},X,Y,\phi) = & \underbrace{\int_0^1 ||L\boldsymbol{v}(x,t)||^2 dt}_{\texttt{-t SyN[}\cdot\texttt{], -r Gauss[}\cdot,\cdot\texttt{]}} + \\ & \underbrace{\lambda_1 \int_\Omega ||\mathcal{I}\circ\phi_1(\mathbf{x},0.5) - \mathcal{J}\circ\phi_2(\mathbf{x},0.5)||d\Omega}_{\texttt{-m MSQ[}\mathcal{J}\texttt{,}\mathcal{I}\texttt{,}\lambda_1\texttt{]}} + \\ & \underbrace{\frac{\lambda_2}{|Y|}\sum_{i=1}^{|Y|}\left\| y_i - \frac{1}{|X|}\sum_{j=1}^{|X|} G(y_i; x_j, \sigma_X)x_j \right\|^2}_{\texttt{-m PSE[}Y\texttt{,}X\texttt{,}\lambda_2\texttt{]}} .\end{aligned} \tag{17}$$

Here, we apply the expectation-based point set registration method for mapping labeled points sets, as described in [36]. ITK-SNAP may be used to label images and exported segmentation images may be input to the PSE metric below, as labeled data. The Frown and Smile data is used as example and is shown in figure 3. This data is available in the ANTS/Examples/Data/ directory. This example should run on the downloaded ANTS data so you may see the results. The ANTS CMakeLists.txt file contains the command that runs this data in automated testing (via the CMake ctest command) and allows the user to evaluate whether s/he is getting the expected performance from their own installation. In Table II we give a brief summary of the arguments available for the normalizations offered by the ANTS package. This includes the corresponding variable specification. More information can be found on the ANTS website.

IV. Experimental Evaluation

The first part of our experimental evaluation will assess ANTS performance on qualitative, "classic" examples of large deformation image registration. Two examples will suffice, both of which are based on the letter "**C**" examples pioneered by [30] and also used in [29]. The purpose of these examples is to show large deformation capabilities with different transformation options. We follow this example with experiments that assess whether, in the cortical labeling problem with normal control data, the added flexibility of true large deformation implementations conveys a performance advantage.

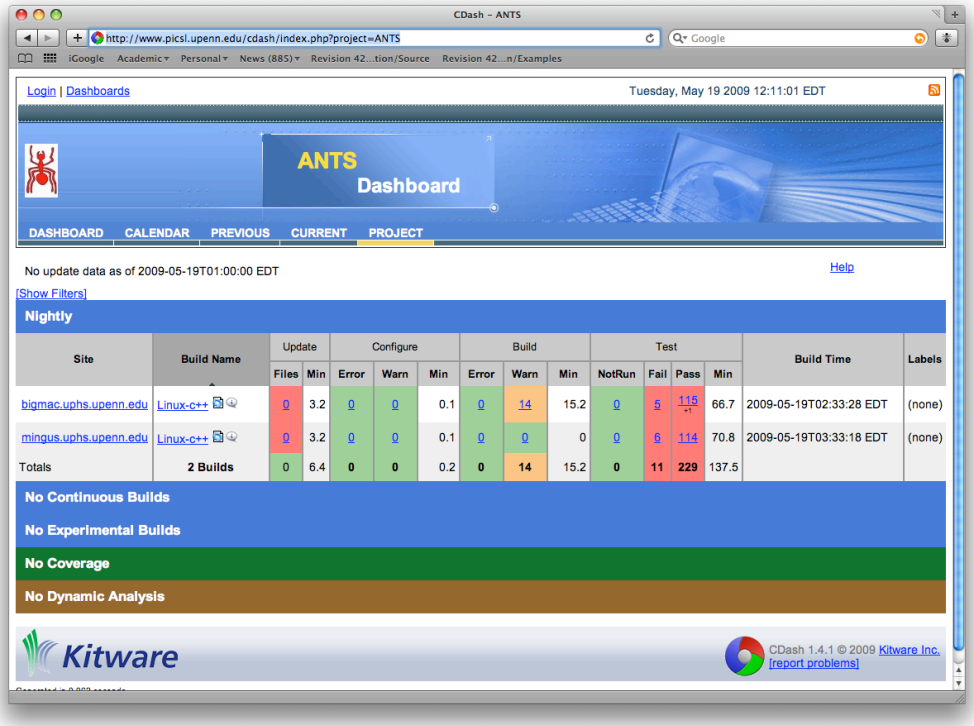
A. Classic Registration Examples

The classic C examples are used to illustrate differences in the ability of image registration transformation models to achieve "large-deformation" mappings between images. While good performance in these examples provides little insight about normalization of real brain imagery, one is able to evaluate the relative accuracy and flexibility of the transformation model when topology and similarity concerns are minimized. An example is shown in figure 4. (which uses the LDDMM like `--geodesic 1` flag). Here, we will

[3] http://sourceforge.net/projects/advants/

[4] http://www.cdash.org

[5] http://www.picsl.upenn.edu/cdash/index.php?project=ANTS



2: The ANTS dashboard, which is hosted on the PICSL website, reports daily building and testing of the ANTS software. It also allows any user to submit their own building and testing configurations to help with debugging issues and maintenance for a variety of computing platforms.

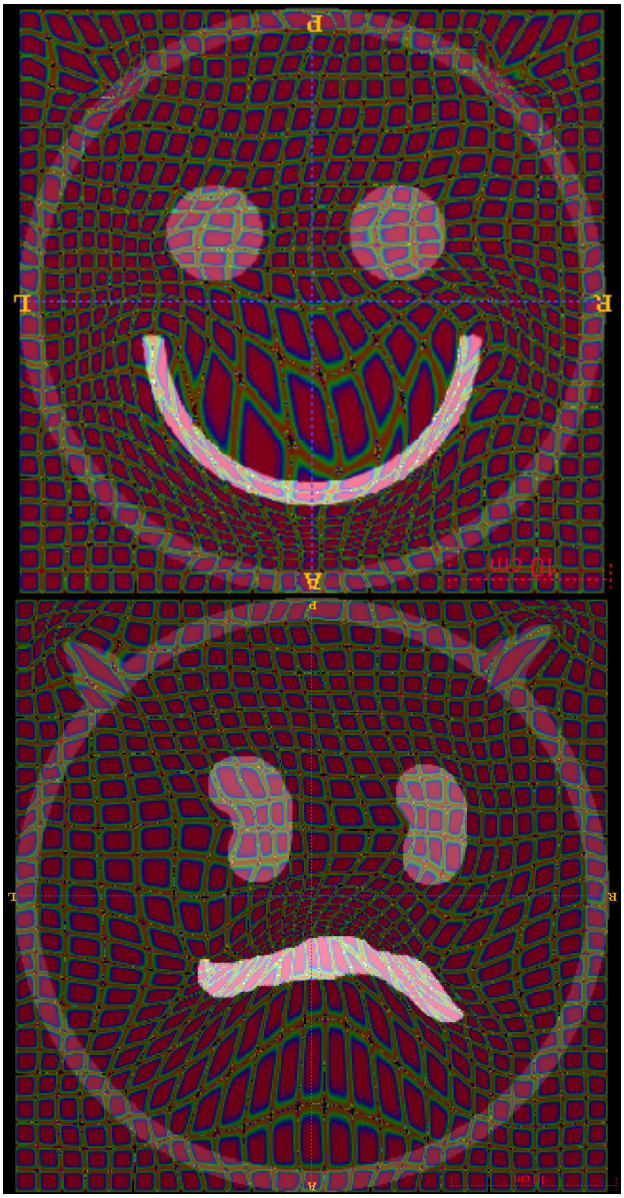
	Argument	Flag	Variables	Sample Parameters
Linear	Iterations	`--linear-iterations`		N_1xN_2xN_3x...
	Similarity	`--linear-metric`	`MI,MSQ`	[N_{bins},$N_{samples}$]
	Affine or Rigid	`--do-rigid`		`true / false`
Deform.	Image Similarity	`--metric,-m`	`MI,CC,PR,MSQ`	[$\mathcal{I}$,$\mathcal{J}$,radius]
	Point-Set Similarity	`--metric,-m`	`PSE,JHCT`	[$\mathcal{I}$,$\mathcal{J}$,X,Y]
	Iterations/Level	`--iterations,-i`		N_1xN_2xN_3x...
	Regularization	`--regularization,-r`	`Gauss,DMFFD`	[$\sigma^2_{gradient}$,σ^2_{total}], `[1x1,3x3]`
	Transformation	`--transformation,-t`	`Elast,SyN,Exp`	[$\Delta_{gradient}$]
	Transformation	`--large-deformation`	`SyN`	[$\Delta_{gradient}$,# time points,dT]
Misc.	Histogram Match $\mathcal{I},\mathcal{J}$	`--use-histogram-matching`		1
	NN Interpolation	`--use-NN`		0
	Mask Image	`--mask,-x`		`mask.nii`
	Output Naming	`--output-naming,-o`		`filename.nii`

II: The various flags and variables for a variety of image registration possibilities. Additional information can be found on the ANTS website [33].

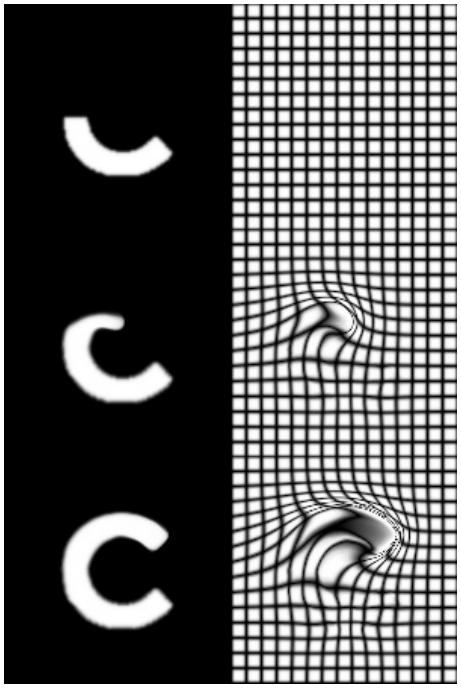
compare advantages and disadvantages of Elastic mapping, Diffeomorphic Demons, greedy SyN and geodesic SyN on this example. The commands used for the example are available in supplementary material. The images are available in the ANTS Examples/Data directory.

B. Cortical Labeling

There are many avenues for exploration of the various components of ANTS. However, due to space constraints, we limit experimental analysis within this paper to an extension of the normalization assessment carried out by [5]. As previously mentioned, this large-scale assessment encompassed evaluation of 14 popular registration algorithms which were optimized, in terms of their parameters, by their respective authors before a thorough brain image normalization study. Although the Greedy SyN algorithm, outlined in an earlier section, was consistently one of the top two performers in Klein's study, for the benefit of the users of ANTS, we explore the other transformation model possibilities within ANTS



3: Expected output for the frown to smile shows a smooth, though large deformation. The grids are overlaid on the deformed images. In this case, the label images and the similarity metric images are identical. The labelings are important for guiding this image mapping to a good local minimum. Without the label guidance, an accurate mapping between the "smile" and the "frown" cannot be achieved as the image similarity term only maps part of the image correctly. Users may run the example themselves as all data is available in the ANTS toolkit Examples/Data directory.



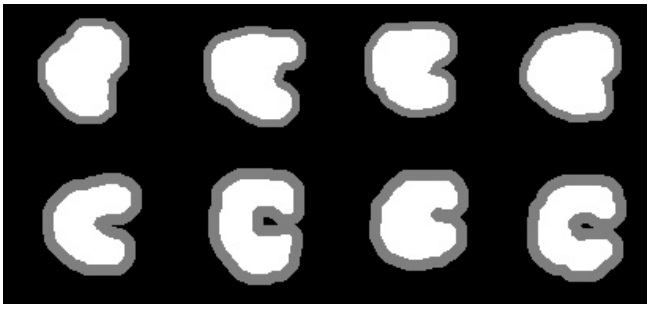
4: ANTS Large Deformation: The original goal of ANTS was to develop public, open source large deformation image registration. This is a classic example showing the progress of deforming a half C to a full C along a geodesic diffeomorphism. The deforming grid accompanies each deformed image.

and compare them with Greedy SyN. In terms of data, we utilize the NA0 evaluation database of the Non-Rigid Image Registration Evaluation Project (NIREP) for future comparison with evaluation studies that have been proposed by the NIREP initiative.[6]

As outlined in the introduction, in addition to the transformation, the optimization strategy and similarity metric form the image normalization scheme. Since our optimization strategy is limited to gradient descent, experimental analysis includes an exploration of optimal gradient steps within a sensible window where the steps are scaled according to the voxel spacing. In terms of similarity metric, we limit exploration to cross-correlation while varying the radius within reasonable values. Other metrics were not explored since the labeled brain images entail a simple intensity relationship between image pairs obviating the need for similarity metrics for more complex intensity relationships (e.g. MI) in addition to the fact that the consistently top two performers in [5] used cross-correlation.

Briefly, two experiments were performed. The first experiment consisted of a more extensive parameter search over both the transformation model space and the cross-correlation metric radius using the exhaustive pairwise combination of the eight 2-D images illustrated in Figure 5. Based on the results of the first experiment, the parameter space was pruned and subsequently used for the registrations performed during the second experiment involving 10 randomly selected image pairs

[6]http://www.nirep.org/



5: The eight 2-D simulated images used for the initial parameter search. These images are available with the ANTS source distribution.

from the 16 labeled NA0 NIREP brain images.

C. 2-D Simulated Image Normalization Evaluation

Eight simulated 2-D images were created to model the types of deformations one would encounter in brain image normalization. Each image was created with isotropic spacing and of size 102×95. The foreground of each image was comprised of two labels representing the white matter and grey matter. Since the 2-D simulated image experiments were used to prune the transformation model space (and not the image similarity metric space) the foreground intensities were not created to model the intensity variation normally seen in the grey/white matter. Note that these images are distributed with the ANTS open source package.

D. 3-D NIREP Brain Image Normalization Evaluation

The Non-Rigid Image Registration Project is a large-scale evaluation resource for deformable registration algorithms headed by Gary Christensen at the University of Iowa. In addition to the development and distribution of the necessary software tools for algorithmic validation, this project includes the distribution of appropriate image data and corresponding segmentations. One such database that has been made available is referred to as the "NA0" database consisting of 16 MR image volumes of normal adult human volunteers. A brief demographic sketch of the NA0 database is as follows: 8 males with mean age of 32.5 $\pm$ 8.4 years (range of 25 to 48 years) and 8 females with mean age of 29.8 $\pm$ 5.8 years (range of 24 to 41).

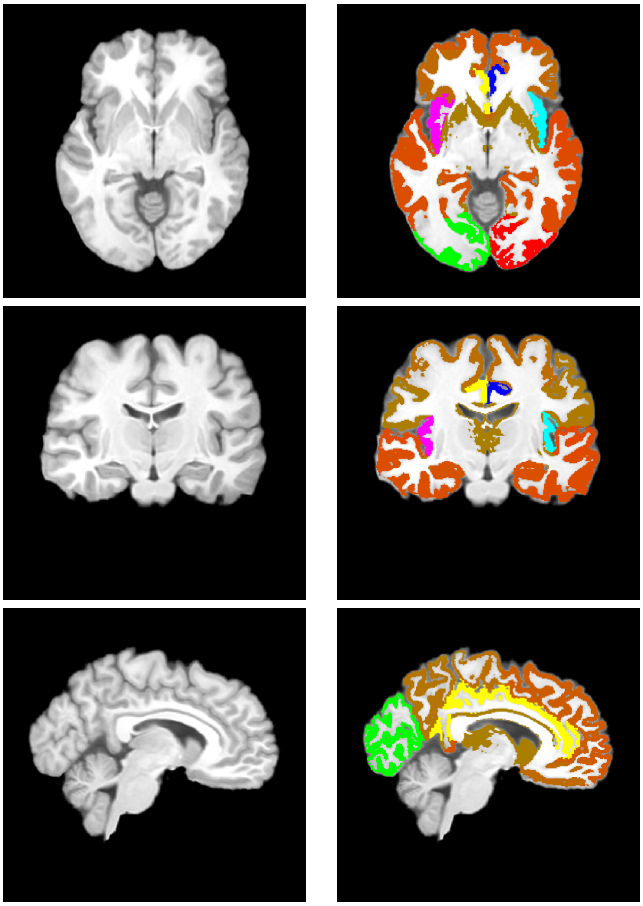
Following acquisition, each image was resampled and padded to isotropic voxel spacing of $0.7 \times 0.7 \times 0.7$ mm^3 and total size of $256 \times 300 \times 256$ voxels. The cortex of each of the 16 MR image volumes was segmented into 32 regions [37], [38] using Brainvox which were later refined using manual editing.

Build template parallel with affine update .... scripts etc .

V. Discussion

Acknowledgment

ANTS is supported by Grant 1R01EB006266-01 from the National Institute Of Biomedical Imaging and Bioengineering and administered through the UCLA Center for Computational Biology.



6: Canonical image views of NA01 image from the NIREP data base (left column) and the corresponding segmentations (right column).

VI. Additional ANTS Features

a) Directly Manipulated Free-Form Deformation: Another top performer in Klein's study [5] was the Image Registration Toolkit (IRTK) based on the research originally reported in [12] in which mutual information and a free-form deformation (FFD) transformation model were used to analyze breast deformation. In ANTS we provide an implementation of a variant of the well-known FFD transformation model for image registration known as *directly manipulated free-form deformation* [19]. The DMFFD model replaces the standard FFD gradient used in [12] with an intuitive preconditioned gradient to overcome problematic energy topographies intrinsic with the traditional approach. DMFFD, in ANTS, is a regularization option as opposed to a specific registration method. That is, we use the same transformation models and gradient descent strategies detailed above, but allow the DMFFD model to regularize the update or total deformation.

For n-D images, the FFD (and DMFFD) transformational model, ϕ_{FFD}, is defined as

$$\phi_{FFD} = \sum_{i_1=1}^{M_1} \cdots \sum_{i_n=1}^{M_n} \mathbf{P}_{i_1,\ldots,i_n} \prod_{j=1}^{n} B_{i_j,d_j}(u_j) \qquad (18)$$

where $\mathbf{P}_{i_1,\ldots,i_n}$ is an n-D grid of control points and $B_{i_j,d_j}(u_j)$ is the B-spline in the i_j^{th} direction of order d_j. The gradient of

the image normalization energy, Π, with respect to the control points used during gradient-based optimization is easily calculated to be:

$$\frac{\partial \Pi}{\partial \mathbf{P}_{i_1,\ldots,i_n}} = \sum_{c=1}^{N_\Omega} \left(\frac{\partial \Pi_\sim}{\partial \phi} + \frac{\partial \Pi_R}{\partial \phi} \right)_c \prod_{j=1}^{n} B_{i_j,d_j}(u_j^c), \quad (19)$$

which is the gradient used in [12]. In contrast, the DMFFD approach uses a preconditioned gradient given by:

$$\begin{aligned}\frac{\partial \Pi}{\partial \mathbf{P}_{i_1,\ldots,i_n}} =& \left(\sum_{c=1}^{N_\Omega} \left(\frac{\partial \Pi_\sim}{\partial \phi} + \frac{\partial \Pi_R}{\partial \phi} \right)_c \prod_{j=1}^{n} B_{i_j,d_j}(u_j^c) \right. \\ & \left. \cdot \frac{\prod_{j=1}^{n} B_{i_j,d_j}^2(u_j^c)}{\sum_{k_1=1}^{d_1+1} \cdots \sum_{k_n=1}^{d_n+1} \prod_{j=1}^{n} B_{k_j,d_j}^2(u_j^c)} \right) \\ & \cdot \left(\frac{1}{\sum_{c=1}^{N_\Omega} \prod_{j=1}^{n} B_{i_j,d_j}^2(u_j^c)} \right). \end{aligned} \quad (20)$$

The difference between the two gradients is seen to reside strictly in terms of the B-spline shape functions which serve to normalize the DMFFD gradient in a unique fashion so as to minimize its susceptibility to hemstitching during the course of optimization. The DMFFD approach, while available in ANTS, has not yet been exploited due to its relatively increased computational demands relative to the default regularization which is a fast Gaussian regularizer as in the Demons method.

b) Jensen-Havrda-Charvat-Tsallis Divergence: Recent information theoretic approaches have been proposed for point-set registration. A previous open-source contribution [19] generalizes the Jensen-Shannon divergence to the Jensen-Havrda-Charvat-Tsallis (JCHT) divergence which permits a fine-tuning of the divergence measure such that emphasis can vary between robustness and sensitivity for application-specic tailoring [39].

Each point-set is represented as a PDF via a Gaussian mixture model (GMM). Assuming K point-sets denoted by $\{X_k, k \in \{1,\ldots,K\}\}$, the k^{th} point-set is denoted by $\{x_1^k,\ldots,x_{|X_k|}^k\}$. The corresponding k^{th} PDF is calculated from the k^{th} point-set as

$$\mathbf{P}_k(s) = \frac{1}{|X_k|} \sum_{i=1}^{|X_k|} G(s; x_i^k, C_i^k) \quad (21)$$

where $G(s; x_i^k, C_i^k)$ is a normalized Gaussian with mean x_i^k and covariance C_i^k evaluated at s. For each point, x_i, the associated weighted covariance matrix, $C_{\mathcal{K}_i}$, is given by

$$C_{\mathcal{K}_i} = \frac{\sum_{x_j \in \mathcal{N}_i, x_j \neq x_i} \mathcal{K}(x_i; x_j)(x_i - x_j)^{\mathrm{T}}(x_i - x_j)}{\sum_{x_j \in \mathcal{N}_i, x_j \neq x_i} \mathcal{K}(x_i; x_j)} \quad (22)$$

where $\mathcal{N}_i$ is the local neighborhood of the point x_i and $\mathcal{K}$ is a user-selected neighborhood weighting kernel. We use an isotropic Gaussian for $\mathcal{K}$ with variance $\sigma_{\mathcal{K}_i}^2$ as well as a k-d tree structure for efficient determination of $\mathcal{N}_i$ [40]. Calculation of the gradient requires the inverse of each covariance matrix. To avoid ill-conditioned covariance matrices, we use the modified covariance $C_i = C_{\mathcal{K}_i} + \sigma_n^2 I$ where I is the identity matrix and σ_n is a parameter denoting added isotropic Gaussian noise.

We designate the number of sample points generated for each of the K probability density functions as $\{M_1,\ldots,M_K\}$ and the k^{th} set of points as $\{s_1^k,\ldots,s_{M_k}^k\}$. The JHCT divergence is then calculated using the K sets of points and the formula

$$\begin{aligned}\mathrm{JHCT}_\alpha(\mathbf{P}_1,\ldots,\mathbf{P}_K) =& \frac{1}{1-\alpha} \\ & \left[\frac{1}{M} \left(\sum_{k=1}^{K} \sum_{j=1}^{M_k} \left[\mathbf{P}^*(s_j^k) \right]^{\alpha-1} - 1 \right) \right. \\ & \left. + \frac{1}{N} \sum_{k=1}^{K} \frac{|X_k|}{M_k} \left(\sum_{j=1}^{M_k} \left[\mathbf{P}_k(s_j^k) \right]^{\alpha-1} - 1 \right) \right] \end{aligned} \quad (23)$$

where

$$\mathbf{P}^*(X) = \frac{1}{N} \sum_{k=1}^{K} \sum_{i=1}^{|X_k|} G(x; x_i^k, C_i^k), \quad (24)$$

$N = \sum_{k=1}^{K} |X_k|$, and $M = \sum_{k=1}^{K} M_k$. The prior weighting values are calculated from $\gamma_k = |X_k|/N$ such that the larger point-sets are weighted more heavily.

REFERENCES

[1] A. Schleicher, P. Morosan, K. Amunts, and K. Zilles, "Quantitative architectural analysis: A new approach to cortical mapping." *J Autism Dev Disord*, Jul 2009. [Online]. Available: http://dx.doi.org/10.1007/s10803-009-0790-8

[2] J. Kovacevic, "From the editor-in-chief," *IEEE Transactions on Image Processing*, vol. 15, p. 12, 2006.

[3] T. S. Yoo and D. N. Metaxas, "Open science–combining open data and open source software: medical image analysis with the insight toolkit." *Med Image Anal*, vol. 9, no. 6, pp. 503–506, Dec 2005. [Online]. Available: http://dx.doi.org/10.1016/j.media.2005.04.008

[4] L. Ibanez, R. S. Avila, and S. R. Aylward, "Open source and open science: how it is changing the medical imaging community," in *Proc. of the International Symposium on Biomedical Imaging*, 2006.

[5] A. Klein, J. Andersson, B. A. Ardekani, J. Ashburner, B. Avants, M.-C. Chiang, G. E. Christensen, L. D. Collins, J. Gee, P. Hellier, J. H. Song, M. Jenkinson, C. Lepage, D. Rueckert, P. Thompson, T. Vercauteren, R. P. Woods, J. J. Mann, and R. V. Parsey, "Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration." *Neuroimage*, Jan 2009. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2008.12.037

[6] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee, "Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain." *Med Image Anal*, vol. 12, no. 1, pp. 26–41, Feb 2008.

[7] L. G. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, no. 4, pp. 325–376, December 1992.

[8] L. Ibanez, L. Ng, J. C. Gee, and S. Aylward, "Registration patterns: The generic framework for image registration of the Insight Toolkit," in *IEEE International Symposium on Biomedical Imaging*, July 2002, pp. 345–348.

[9] R. Bajcsy and S. Kovacic, "Multiresolution elastic matching," *Computer Vision, Graphics, and Image Processing*, vol. 46, pp. 1–21, 1989.

[10] D. Shen and C. Davatzikos, "Hammer: Hierarchical attribute matching mechanism for elastic registration," *IEEE Transactions on Medical Imaging*, vol. 21, no. 11, pp. 1421–1439, 2002.

[11] J. Ashburner and K. Friston, "Voxel-based morphometry—The methods," *Neuroimage*, vol. 11, pp. 805–821, 2000.

[12] D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, and D. Hawkes, "Nonrigid registration using free-form deformations: application to breast MR images." *IEEE Trans Med Imaging*, vol. 18, no. 8, pp. 712–21, Aug 1999.

[13] J.-P. Thirion, "Image matching as a diffusion process: an analogy with Maxwell's demons," *Medical Image Analysis*, vol. 2, no. 3, pp. 243–260, 1998.
[14] V. I. Arnold, *Ordinary Differential Equations*. Springer-Verlag, 1991.
[15] B. Avants, C. Anderson, M. Grossman, and J. C. Gee, "Spatiotemporal normalization for longitudinal analysis of gray matter atrophy in frontotemporal dementia." *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv*, vol. 10, no. Pt 2, pp. 303–310, 2007.
[16] N. Fox, W. Crum, R. Scahill, J. Stevens, J. Janssen, and M. Rossor, "Imaging of onset and progression of alzheimer's disease with voxel-compression mapping of serial magnetic resonance images," *Lancet*, vol. 358, pp. 201–205, 2001.
[17] M. I. Miller, A. Trouvè, and L. Younes, "Geodesic shooting for computational anatomy," *J. Mathematical Imaging and Vision*, 2005, submitted.
[18] D. Wolpert and W. Macready, "No free lunch theorems for optimization," *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, pp. 67–82, April 1997.
[19] N. J. Tustison, B. B. Avants, and J. C. Gee, "Directly manipulated free-form deformation image registration." *IEEE Trans Image Process*, vol. 18, no. 3, pp. 624–635, Mar 2009. [Online]. Available: http://dx.doi.org/10.1109/TIP.2008.2010072
[20] N. J. Tustison and J. C. Gee, "N-D C^k B-spline scattered data approximation," *Insight Journal*, p. published online, 2005. [Online]. Available: http://hdl.handle.net/1926/140
[21] D. G. Ebin and J. Marsden, "Groups of diffeomorphisms and the motion of an incompressible fluid," *Annals of Mathematics*, no. 92, pp. 102–163, 1970.
[22] D. Mumford, "Pattern theory and vision," *Questions Matheematiques En Traitement Du Signal et de L'Image, Institut Henri Poincare*, vol. 3, pp. 7–13, 1998.
[23] M. I. Miller and L. Younes, "Group actions, homeomorphisms, and matching: A general framework," *Internation Journal of Computer Vision*, vol. 41, no. 1/2, pp. 61–84, 2001.
[24] M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes, "Computing large deformation metric mappings via geodesic flows of diffeomorphisms," *Int. J. Comput. Vision*, vol. 61, no. 2, pp. 139–157, 2005.
[25] P. Dupuis, U. Grenander, and M. I. Miller, "Variational problems on flows of diffeomorphisms for image matching," 1998. [Online]. Available: citeseer.comp.nus.edu.sg/67701.html
[26] M. I. Miller, A. Trouve, and L. Younes, "On the metrics and euler-lagrange equations of computational anatomy." *Annu Rev Biomed Eng*, vol. 4, pp. 375–405, 2002. [Online]. Available: http://dx.doi.org/10.1146/annurev.bioeng.4.092101.125733
[27] J. Ashburner, "A fast diffeomorphic image registration algorithm." *Neuroimage*, vol. 38, no. 1, pp. 95–113, Oct 2007. [Online]. Available: http://dx.doi.org/j.neuroimage.2007.07.007
[28] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Diffeomorphic demons using itk's finite difference solver hierarchy," *Insight Journal—2007 MICCAI Open Science Workshop*, 2007.
[29] ——, "Diffeomorphic demons: efficient non-parametric image registration." *Neuroimage*, vol. 45, no. 1 Suppl, pp. S61–S72, Mar 2009. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2008.10.040
[30] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinematics," *IEEE Transactions on Image Processing*, vol. 5, no. 10, pp. 1435–1447, 1996.
[31] P. Viola and W. M. Wells, "Alignment by maximization of mutual information," *International Journal of Computer Vision*, vol. 24, no. 2, pp. 137–154, 1997.
[32] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: a survey." *IEEE Trans Med Imaging*, vol. 22, no. 8, pp. 986–1004, Aug 2003.
[33] B. B. Avants, N. J. Tustison, G. Song, and J. C. Gee, *ANTS: Advanced Open-Source Normalization Tools for Neuroanatomy*, Penn Image Computing and Science Laboratory, 2009. [Online]. Available: http://www.picsl.upenn.edu/ANTS
[34] S. Lee, G. Wolberg, and S. Y. Shin, "Scattered data interpolation with multilevel b-splines," *IEEE Transactions on Visualization and Computer Graphics*, vol. 3, no. 3, pp. 228–244, 1997.
[35] N. J. Tustison and J. C. Gee, "Generalized n-D C^k B-spline scattered data approximation with confidence values," in *Proc. Third International Workshop Medical Imaging and Augmented Reality*, 2006, pp. 76–83.
[36] J. Pluta, B. B. Avants, S. Glynn, S. Awate, J. C. Gee, and J. A. Detre, "Appearance and incomplete label matching for diffeomorphic template based hippocampus segmentation." *Hippocampus*, vol. 19, no. 6, pp. 565–571, Jun 2009. [Online]. Available: http://dx.doi.org/10.1002/hipo.20619
[37] J. S. Allen, H. Damasio, and T. J. Grabowski, "Normal neuroanatomical variation in the human brain: an mri-volumetric study." *Am J Phys Anthropol*, vol. 118, no. 4, pp. 341–358, Aug 2002. [Online]. Available: http://dx.doi.org/10.1002/ajpa.10092
[38] J. S. Allen, H. Damasio, T. J. Grabowski, J. Bruss, and W. Zhang, "Sexual dimorphism and asymmetries in the gray-white composition of the human cerebrum." *Neuroimage*, vol. 18, no. 4, pp. 880–894, Apr 2003.
[39] N. J. Tustison, S. P. Awate, G. Song, T. S. Cook, and J. C. Gee, "A new information-theoretic measure to control the robustness-sensitivity trade-off for dmffd point-set registration," in *Proceedings of Information Processing in Medical Imaging*, 2009.
[40] M. de Berg, M. van Kreveld, M. Overmars, and O. Scharzkopf, *Computational Geometry: Algorithms and Applications*. Springer, 2000.