Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphisms

M. Faisal Beg[*], Michael I. Miller[†], Alain Trouvé[‡] and Laurent Younes[§]

September 25, 2003

Abstract

This paper examine the Euler-Lagrange equations for the solution of the large deformation diffeomorphic metric mapping problem studied in [9, 17] in which two images I_0, I_1 are given and connected via the diffeomorphic change of coordinates $I_0 \circ \varphi^{-1} = I_1$ where $\varphi = \phi_1$ is the end point at $t = 1$ of curve $\phi_t, t \in [0, 1]$ satisfying $\dot{\phi}_t = v_t(\phi_t)$, $t \in [0, 1]$ with $\phi_0 = id$. The variational problem takes the form

$$\operatorname*{argmin}_{v:\dot{\phi}_t = v_t(\phi_t)} \left(\int_0^1 \|v_t\|_V^2 dt + \|I_0 \circ \phi_1^{-1} - I_1\|_{L^2}^2 \right),$$

where $\|v_t\|_V$ is an appropriate Sobolev norm on the velocity field $v_t(\cdot)$, and the second term enforces matching of the images with $\| \cdot \|_{L^2}$ representing the squared-error norm.

In this paper we derive the Euler-Lagrange equations characterizing the minimizing vector fields $v_t, t \in [0, 1]$ assuming sufficient smoothness of the norm to guarantee existence of solutions in the space of diffeomorphisms. We describe the implementation of the Euler equations using semi-lagrangian method of computing particle flows and show the solutions for various examples. As well, we compute the metric distance on several anatomical configurations as measured by $\int_0^1 \|v_t\|_V dt$ on the geodesic shortest paths.

[*]Center for Imaging Science & the department of Biomedical Engineering, The Johns Hopkins University, 301 Clark Hall, Baltimore, MD 21218

[†]Center for Imaging Science, department of Biomedical Engineering, department of Electrical and Computer Engineering and the department of Computer Science, Whiting School of Engineering, The Johns Hopkins University, 301 Clark Hall, Baltimore, MD 21218

[‡]LAGA, Université Paris 13, France

[§]CMLA, Ecole Normale Supérieure de Cachan, 61, avenue du President Wilson, F-94 235 Cachan CEDEX, France

1

1   Introduction

The last decade has witnessed tremendous developments in medical imaging technologies which are delivering exquisitely detailed pictures of human anatomy. The acquisition of structural imagery via Magnetic resonance imaging (MRI), anisotropic or orientational imagery via Diffusion Tensor Magnetic Resonance Imaging (DTMRI) and functional activation imagery via Functional Magnetic Resonance Imaging (fMRI) among others is routine. These capabilities have in turn spurred the development of mathematics and algorithms for analysis of the information contained in these images. A recurring theme in biomedical image analysis is registration; the variability in human anatomy is not an exception but a rule and hence the need to transform the image data into standard coordinates to facilitate analysis and generalize results to a large subset of the population. The goal of registration is to compute a transformation $\varphi : \Omega \to \Omega$ where $\Omega \subseteq \mathbb{R}^n$ is the domain ($n = 2$ for 2D or $n = 3$ for 3D) on which the data (structural, orientational and functional) are defined. Let images representing this data be functions $I : \Omega \to \mathbb{R}^d$ defined on Ω. Structural images acquired via MRI are scalar valued $d = 1$, color images such as RGB are vector valued $d = 3$ and anisotropy or orientational images acquired from DTMRI are matrix valued. Since the diffusion tensor is symmetric, these can be taken to be vector valued $d = 6$. Let I_0 and I_1 denote the template and target images. The transformation of the template image I_0 under such a transformation is the pullback image defined to be $\varphi.I_0 = I_0 \circ \varphi^{-1} = I_0(\varphi^{-1})$. The transformations increase in dimensionality from low dimensional global transformations specified by a few parameters such as the affine transformations to the high dimensional non-rigid transformations which are specified at each point of the image domain.

The early attempts to compute high-dimensional non-rigid transformations in the medical image setting resulted in the development of the elastic matching strategy by Broit, Bajcsy and co-workers [4, 2, 3] and development on these lines continues as this model is refined and specialized to various applications. In this setting, the transformation φ is linearized around the coordinate system of the exemplar or template coordinate system and generated from a *displacement* vector field $u : \Omega \to \mathbb{R}^n$ such that $\varphi(x) = x + u(x)$ or $\varphi^{-1}(x) = x - u(x)$ for all points $x \in \Omega$. The transformed template image I_0 then becomes $I_0(\varphi^{-1}(x)) = I_0(x - u(x))$, $\forall x \in \Omega$. The "goodness" of the transformation is measured by a cost of the form

$$E_2(I_0, I_1, \varphi) = \|I_0 \circ \varphi^{-1} - I_1\|_2^2$$

where $\| \cdot \|_2$ is the standard L^2 norm of square integrable functions $\|f\|_2^2 = \int_\Omega |f(x)|^2 \mathrm{d}x$. The optimal transformation is the one that minimizes this cost and the optimal vector field u that generates such a transformation is, from among the many possible solutions, chosen

2

to be the one with the highest smoothness. The measurement of smoothness is achieved by specifying the norm on the space of vector fields of the domain Ω to be defined through a differential operator L by:

$$E_1(u) = \|Lu\|_2^2$$

where L is commonly chosen to be a differential operator of the form $L = (-\alpha\Delta + \gamma)I_{n\times n}$. In the variational setting (a very nice discussion is in [1]), the optimal displacement vector field is computed by optimization of the cost

$$\operatorname*{argmin}_{u} \|Lu\|_2^2 + \frac{1}{\sigma^2}\|I_0 \circ \varphi^{-1} - I_1\|_2^2.$$

This approach of generating transformations from displacement vector fields has been an important development in computing non-rigid high dimensional transformations of medical imagery allowing for comparison of anatomy in a standard coordinate system. One of the limitations of this approach is that there are no explicit constraints that ensure that the transformations computed are one-to-one or invertible. Indeed, in some cases [8] folding of the grid over itself can occur thereby destroying the neighbourhood structure which is essential for the study of anatomy. This method of computing transformations is known as the small deformations approach as valid transformations of anatomy using this linearized model via displacement vector fields are computed when the images are seperated by small deformations of the domain.

It is of considerable interest to compute transformations which are not only invertible but also preserve properties such as smoothness of curves, surfaces or other features associated to anatomy. Therefore, diffeomorphic transformations, which are smooth invertible transformations with smooth inverse, are of considerable interest in this regard. Constraining the transformations to be diffeomorphisms is a natural choice in the study of anatomy and properties based on anatomical coordinates as under these transformations, connected sets remain connected, disjoint sets remain disjoint, smoothness of anatomical features such as curves and surfaces is preserved, and coordinates are transformed consistently.

The large deformation model for computing transformations developed by Christensen, Rabbitt and Miller [7] overcomes the limitations of the small deformations model by ensuring that the transformations computed between imagery are diffeomorphic. In this setting, the transformation φ of the domain is generated as the end-point $\varphi = \phi_1$ of the flow of a time-dependent velocity vector field $v_t : \Omega \to \mathbb{R}^n$, $t \in [0,1]$ specified by the ODE $\dot{\phi}_t = v_t(\phi_t)$. This gives a path $\phi_t : \Omega \to \Omega$, $t \in [0,1]$ in the space of transformations starting with $\phi_0 = Id$, where Id is the identity transformation $Id(x) = x$, $\forall x \in \Omega$, and terminating at the end-point

3

$t = 1$ of the flow to the particular transformation $\varphi = \phi_1 = \phi_0 + \int_0^1 v_t(\phi_t)\mathrm{d}t$ matching the given images.

The main contribution of this paper is the estimation of the optimal transformation via the basic variational problem that, in the space of smooth velocity vector fields V on domain Ω, takes the form:

$$\hat{v} = \underset{v : \dot{\phi}_t = v_t(\phi_t)}{\mathrm{argmin}} \left(\int_0^1 \|v_t\|_V^2 dt + \frac{1}{\sigma^2} \|I_0 \circ \phi_1^{-1} - I_1\|_{L^2}^2 \right) . \tag{1}$$

The optimizer of this cost then generates the optimal change of coordinates $\varphi = \phi_1^{\hat{v}}$ upon integration $d\hat{\phi}_t^v / dt = \hat{v}_t(\hat{\phi}_t^v)$, $\phi_0 = id$, where the subscript v in ϕ^v is used to explicitly denote the dependence of ϕ on the associated velocity field v. As shown in [9, 17], enforcing a sufficient amount of smoothness on the elements of the space V of allowable velocity vector fields ensures that the solution to the differential equation $\dot{\phi}_t = v_t(\phi_t)$, $t \in [0,1]$, $v_t \in V$ is in the space of diffeomorphisms. The required smoothness is enforced by defining the norm on the space V of smooth velocity vector fields through a differential operator L of the type $L = (-\alpha\Delta + \gamma)^\alpha I_{n \times n}$ where $\alpha > 1.5$ in 3-dimensional space such that $\|f\|_V = \|Lf\|_2$ where $\| \cdot \|_2$ is the standard L^2 norm for square integrable functions defined on Ω. We term the solution satisfying equation 1 as the large deformations diffeomorphic metric mapping (LDDMM) solution in the sense that (1) similar to the flow approach first introduced by Christensen [7], it provides large deformation coordinate system transformation and (2) as shown in [12, 17, 11], in contrast to [7], the length of the shortest path $\inf \int_0^1 \|v_t\|_V \mathrm{d}t$ connecting images I_0 to I_1 defines a metric on the image orbit $\mathcal{I}$.

In this paper, we present the derivation of the Euler-Lagrange equation for the solution of the variational problem 1 in the space of smooth velocity vector fields. We present the implementation details of a gradient algorithm using the Euler-Lagrange equation for numerical computation of the solution to 1, and in particular a semi-lagrangian method of particle flows to integrate the vector fields being a non-dissipative method as opposed to traditional Eulerian methods which are highly dissipative. The optimizer of 1 generates a geodesic shortest length path in the space of the group of transformations in contrast to the path generated by the method of Christensen et. al [7] which is a locally optimal solution with which we compare the solution to 1.

4

2 Euler equations for the variational minimization on vector fields

The starting point for our approach to the analysis of shape and size in anatomical images is modelling anatomy as a deformable template [10], ie. the observed anatomical imagery $\mathcal{I}$ is an orbit under diffeomorphic transformations $\mathcal{G}$ acting on the coordinate space of a family of exemplars. A homeomorphism on the background space Ω is a bijective (invertible) function $\varphi : \Omega \to \Omega$, which, with its inverse φ^{-1} is continuous. Let the set of homeomorphisms acting on the background space be denoted by $\mathrm{Hom}(\Omega)$. The homeomorphisms form a group for the usual law of composition $\psi \cdot \varphi \doteq \psi \circ \varphi$. Moreover, for any $\varphi \in \mathrm{Hom}(\Omega)$ and any image $I : \Omega \to \mathbb{R}^d$, $\varphi.I \doteq I \circ \varphi^{-1}$ defines an action of $\mathrm{Hom}(\Omega)$ on the set of all images. Let $\mathcal{G}$ be a sub-group of $\mathrm{Hom}(\Omega)$ (for instance the set $\mathrm{Diff}(\Omega)$ of any $\varphi \in \mathrm{Hom}(\Omega)$ which are, with its inverse, continuously differentiable). Given a template I_{template}, the deformable template model of an anatomical ensemble is the orbit

$$\mathcal{I} \doteq \{ \; \varphi I_{\mathrm{template}} = I_{\mathrm{template}} \circ \varphi^{-1} \mid \varphi \in \mathcal{G} \; \}$$

of I_{template} under the action of $\mathcal{G}$. The orbit $\mathcal{I}$ being homogenous under the action of the elements of $\mathcal{G}$ which are bijective mappings implies that images in the anatomical ensemble $\mathcal{I}$ are topologically equivalent i. e. they possess the same sub-structures and for any two images $I_0, I_1 \in \mathcal{I}$, there exists a $\varphi \in \mathcal{G}$ that registers the given images $I_1 = \varphi I_0$.

Given two images, the first task is to find the particular element $\varphi \in \mathcal{G}$ that registers the given images $I_1 = \varphi I_0 = I_0 \circ \varphi^{-1}$. In the large deformations setting, this element is estimated as the end point of the flow associated to a smooth time-dependent vector field. Let $v : [0, 1] \to V$ be a time-dependent velocity vector-field where V is a Hilbert space of smooth, compactly supported vector fields on Ω. Let such a velocity vector field define the evolution of a curve $\phi^v : [0, 1] \to \mathcal{G}$ via the evolution equation

$$\frac{d}{dt} \phi_t^v(x) = v_t(\phi_t^v(x)) \tag{2}$$

where the subscript in ϕ^v is used to explicitly denote the dependence of ϕ on the associated velocity field v. The initial point of the curve ϕ^v at $t = 0$ is $\phi_0^v = Id \in \mathcal{G}$ where Id is the identity transformation $Id(x) = x, \forall x \in \Omega$. The end point of the curve ϕ^v at time $t = 1$ is the particular diffeomorphism $\phi_1^v = \varphi \in \mathcal{G}$ that links the given datasets I_0 and I_1 such that $I_1 = I_0 \circ \varphi^{-1}$ and it is this element φ that we compute as the end point ϕ_1^v associated to a flow v. Thus, we seek a time-dependent velocity vector field v which when integrated via equation 2 generates the particular diffeomorphism matching the given image datasets.

5

Let the notation $\phi_{s,t} : \Omega \to \Omega$ denote the composition $\phi_{s,t} = \phi_t \circ (\phi_s)^{-1}$. The interpretation of $\phi_{s,t}(y)$ is that it is the position at time t of a particle that is at position y at time s. Therefore $\phi_1^v(x) = \phi_{0,1}^v(x)$ is the function that denotes the position at time $t = 1$ of particle that is at position x at time 0. Let the Jacobian of mapping $\phi_{s,t}$, the matrix composed with the space derivatives of $\phi_{s,t}$, be denoted by $D\phi_{s,t}$.

To solve the variational problem 1, we first need to compute the variation of the mapping $\phi_{1,0}^v$ under the perturbation of $v \in L^2([0,1], V)$ by $h \in L^2([0,1], V)$ which we state in the following lemma.

Lemma 2.1. *The variation of mapping $\phi_{s,t}^v$ when $v \in L^2([0,1], V)$ is perturbed along $h \in L^2([0,1], V)$ is given by:*

$$\partial_h \phi_{s,t}^v = \alpha_{s,t} = \lim_{\epsilon \to 0} \frac{\phi_{s,t}^{v+\epsilon h} - \phi_{s,t}^v}{\epsilon} = D\phi_{s,t}^v \int_s^t (D\phi_{s,u}^v)^{-1} h_u \circ \phi_{s,u}^v \mathrm{d}u. \tag{3}$$

Proof. We provide a proof under the assumption that the derivative with respect to ϵ in equation 3 exists and proceed to its identification. The proof of existence can be carried on by standard ordinary differential equations (ODE) arguments. We have

$$\frac{d\phi_{s,t}^{v+\epsilon h}}{dt} = v_t \circ \phi_{s,t}^{v+\epsilon h} + \epsilon h_t \circ \phi_{s,t}^{v+\epsilon h} .$$

Computing the differential in ϵ at $\epsilon = 0$ yields

$$\frac{d}{dt}\partial_h \phi_{s,t}^v = D_{\phi_{s,t}^v} v_t \partial_h \phi_{s,t}^v + h_t \circ \phi_{s,t}^v \tag{4}$$

Thus, $\partial_h \phi_{s,t}^v$ is the solution of a non-homogeneous differential equation, and it suffices to show that the last expression in 3 is a solution of the same equation with the correct condition $\partial_h \phi_{s,s}^v = 0$, but this is an obvious computation, using the fact that

$$\frac{d}{dt} D\phi_{s,t}^v = D_{\phi_{s,t}^v} v_t D\phi_{s,t}^v ,$$

which simply comes from computing the space differential of $d\phi_{s,t}^v/dt = v_t \circ \phi_{s,t}$. Note that this identity also provides the solution of the homogeneous equation associated to 4, so that 3 can also be directly identified by variation of the constant.

$\square$

Existence of the transformations generated via equation 2 depend on the smoothness constraints placed on vector fields allowed in V [9, 17]. One choice to ensure existence of solutions in the space of diffeomorphisms for the ordinary differential equation (ODE) 2 has been to

construct V as the completion of the space of smooth, compactly-supported vector fields for the inner-product defined through a differential operator L (denoting its adjoint as $L^\dagger$) given by:

$$\langle f, g \rangle_V \doteq \langle Lf, Lg \rangle_{L^2} = \langle L^\dagger L f, g \rangle_{L^2}, \tag{5}$$

where $\langle , \rangle_{L^2}$ is the usual L^2-product for square integrable vector-fields on Ω. With V defined in this way, the flow of $v \in L^1([0,1], V)$ generates the sub-group of diffeomorphisms $\mathcal{G} \doteq \{ \varphi \mid \varphi = \phi_1^v, \quad v \in L^1([0,1], V) \}$ that are the end-points of flows associated to elements $v \in L^1([0,1], V)$ (the technical details associated with this construction have been published in [9, 17]). From the assumptions placed in the construction of V, a compact self-adjoint operator operator $K : L^2(\Omega, \mathbb{R}^d) \to V$ is uniquely defined by

$$\langle a, b \rangle_{L^2} = \langle Ka, b \rangle_V. \tag{6}$$

and together with 5, one gets for any smooth vector field $f \in V$ that

$$K(L^\dagger L) f = f. \tag{7}$$

The variational problem for dense image matching is now stated and solved in the space of vector fields V.

Theorem 2.1. *Given a continuously differentiable idealized template image I_0 and a noisy observation of anatomy I_1, then $\hat{v} \in L^2([0,1], V)$ for inexact matching of I_0 and I_1 is given by*

$$\hat{v} = \underset{v \in L^2([0,1],V)}{\operatorname{arginf}} E(v) \doteq \int_0^1 \|v_t\|_V^2 \, \mathrm{d}t + \frac{1}{\sigma^2} \|I_0 \circ \phi_{1,0}^v - I_1\|_{L^2}^2 \tag{8}$$

which satisfies the Euler-Lagrange equation given by

$$2\hat{v}_t - K\left(\frac{2}{\sigma^2} \ |D\phi_{t,1}^{\hat{v}}| \ \nabla J_t^0 \ \left(J_t^0 - J_t^1\right) \right) = 0 \tag{9}$$

where $J_t^0 \doteq I_0 \circ \phi_{t,0}, J_t^1 \doteq I_1 \circ \phi_{t,1}$.

Proof. Let the velocity $v \in L^2([0,1], V)$ be perturbed by an ϵ amount along direction $h \in L^2([0,1], V)$. The Gateaux variation $\partial_h E(v)$ of the energy functional is related to its Fréchet derivative $\nabla_v E$ by

$$\partial_h E(v) = \lim_{\varepsilon \to 0} \frac{E(v + \varepsilon h) - E(v)}{\varepsilon} = \int_0^1 \langle \nabla_v E_t, h_t \rangle_V \, \mathrm{d}t.$$

The variation of $E_1(v) = \int_0^1 \|v_t\|_V^2 \, \mathrm{d}t$ is given by:

$$\partial_h E_1(v) = 2 \int_0^1 \langle v_t, h_t \rangle_V \, \mathrm{d}t.$$

7

Variation of $E_2(v) = \frac{1}{\sigma^2} \|I_0 \circ \phi_{1,0}^v - I_1\|_{L^2}^2$ is

$$
\begin{aligned}
\partial_h E_2(v) &= \frac{2}{\sigma^2} \langle I_0 \circ \phi_{1,0}^v - I_1, DI_0 \circ \phi_{1,0}^v \partial_h \phi_{1,0}^v \rangle_{L^2} \\
&\overset{(a)}{=} \frac{2}{\sigma^2} \langle I_0 \circ \phi_{1,0}^v - I_1, DI_0 \circ \phi_{1,0}^v \left(- D\phi_{1,0}^v \int_0^1 (D\phi_{1,t}^v)^{-1} h_t \circ \phi_{1,t}^v dt \right) \rangle_{L^2} \\
&\overset{(b)}{=} \frac{-2}{\sigma^2} \int_0^1 \langle (I_0 \circ \phi_{1,0}^v - I_1, D(I_0 \circ \phi_{1,0}^v)(D\phi_{1,t}^v)^{-1} h_t \circ \phi_{1,t}^v \rangle_{L^2} dt
\end{aligned}
$$

with (a) where subsitution of $\partial_h \phi_{1,0}^v$ is made using lemma 2.1 and in (b) collecting $D(I_0 \circ \phi_{1,0}^v) = DI_0 \circ \phi_{1,0}^v D\phi_{1,0}^v$. Setting $\phi_{1,t}^v(y) = x$ i. e. $\phi_{t,1}^v(x) = y$, gives the Jacobian change of variables $|D\phi_{t,1}^v(x)| dx = dy$. With this, $\phi_{1,0}^v \mapsto \phi_{1,0}^v \circ \phi_{t,1}^v = \phi_{t,0}^v$ and substituting in above, we get:

$$
\begin{aligned}
\partial_h E_2(v) &= \frac{-2}{\sigma^2} \int_0^1 \langle |D\phi_{t,1}^v| (I_0 \circ \phi_{t,0}^v - I_1 \circ \phi_{t,1}^v), D(I_0 \circ \phi_{t,0}^v) h_t \rangle_{L^2} dt \\
&= \frac{-2}{\sigma^2} \int_0^1 \langle |D\phi_{t,1}^v| (J_t^0 - J_t^1) \nabla J_t^0, h_t \rangle_{L^2} dt \\
&= - \int_0^1 \left\langle K \left(\frac{2}{\sigma^2} |D\phi_{t,1}^v| \left(J_t^0 - J_t^1 \right) \nabla J_t^0 \right), h_t \right\rangle_V dt.
\end{aligned}
$$

Collecting terms, the gradient of the energy functional is thus

$$
(\nabla_v E_t)_V = 2v_t - K \left(\frac{2}{\sigma^2} |D\phi_{t,1}^v| \nabla J_t^0 \left(J_t^0 - J_t^1 \right) \right), \tag{10}
$$

where the subscript V in $(\nabla_v E_t)_V$ indicates the gradient is in the space V. The optimizing velocity field satisfies the Euler-Lagrange equation

$$
\partial_h E(\hat{v}) = \int_0^1 \left\langle 2\hat{v}_t - K \left(\frac{2}{\sigma^2} |D\phi_{t,1}^{\hat{v}}| \nabla J_t^0 \left(J_t^0 - J_t^1 \right) \right), h_t \right\rangle_V dt = 0 \tag{11}
$$

Since h is arbitrary in $L^2([0,1], V)$ we get equality (9). $\qquad\qquad\square$

The optimal vector field for 8 is calculated by using the gradient 10 in a standard gradient descent algorithm. As we shall see in the next section, the optimization vector fields assign a metric $\int_0^1 \|v_t\|_V dt$ on $\mathcal{I}$ and we denote the algorithm based on this gradient as the large deformation diffeomorphic metric mapping (LDDMM) algorithm.

3   Numerical implementation of LDDMM algorithm

This section presents the discretization and implementation details of the estimation of the optimizing vector fields. Let $\Omega = [0,1]^2/[0,1]^3$ represent the 2D/3D background space.

8

Let the discretized flow generating the diffeomorphism be indexed by the discretized index $t_j \in [0, T], j \in [0, N]$, the size of a timestep being δt such that $T = N \times \delta t$. Assume piecewise-constant velocities in the discretized time intervals. Let $v_{t_j}^k(y)$ and $\phi_{t_j}^k(y)$ denote the velocity field and the mapping for the kth iteration of the gradient algorithm and the jth timestep along the discretized flow. Let $I_0 = J_0^0$ be the image at $t = 0$ of the flow being mapped to the image $I_1 = J_T^T$ at time index $t = T$ of the flow.

3.1 Gradient descent scheme based optimization

The variational optimization of the energy functional (Equation 8) is performed in a standard steepest descent scheme

$$v^{k+1} = v^k - \epsilon \nabla_{v_{t_j}^k} E \tag{12}$$

where the discretized version of the gradient $\nabla_v E$ from Equation 10 is given by

$$\nabla_{v_{t_j}^k} E_{t_j}^k(y) = 2v_{t_j}^k(y) - \frac{2}{\sigma^2} K \left(|D\phi_{t_j,T}^k(y)| D J_{t_j}^0(y)^* (J_{t_j}^0(y) - J_{t_j}^T(y)) \right). \tag{13}$$

The discretized energy becomes

$$E(v^k) = \sum_{j=0}^{N-1} \|v_{t_j}^k\|_V^2 \delta t + \frac{1}{N_1 N_2 N_3} \sum_{y \in \Omega} |J_T^0(y) - J_T^T(y)|^2 \tag{14}$$

and the length of the path from the identity to the estimated matching diffeomorphism ϕ_T at simulation index k becomes

$$\text{Length}(Id, \phi_T^{v^k}) = \sum_{j=0}^{N-1} \|v_{t_j}^k\|_V \delta t. \tag{15}$$

3.2 Choice of operator L and evaluation of $L^\dagger L$, $(L^\dagger L)^{-1}$

Let discretized functions $f(x)$ and $g(x)$ be defined $f, g : \Omega \subset \mathbb{Z}^n \to \mathbb{R}^n$ on the discrete lattice $\Omega \subset \mathbb{Z}^n$ and which satisfy equation

$$L^\dagger L f = g. \tag{16}$$

The operator L is chosen to be of the Cauchy-Navier type, $L \doteq -\alpha \nabla^2 + \gamma I$ where I is the identity operator and $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is the Laplacian operator. The coefficient α enforces smoothness, higher values ensure solutions of high regularity, and the coefficient γ is chosen to be positive so that the operator is non-singular. As well, as more anatomies are mapped in future, we will use the statistics of the anatomical mappings to choose these coefficients [10]. We assume periodic boundary conditions, in which case L is self-adjoint

9

$L = L^\dagger$. Given the discretized function f, to calculate the function g via Equation 16, write the discretized version for $Lf(x)$ using finite differences on a periodic domain.

$$
\begin{aligned}
g(x_1, x_2, x_3) &= (-\alpha \nabla^2 + \gamma) f(x_1, x_2, x_3) \\
&= -\alpha \Big(\frac{f(x_1 + \triangle x_1, x_2, x_3) - 2 f(x_1, x_2, x_3) + f(x_1 - \triangle x_1, x_2, x_3)}{\triangle x_1^2} \\
&\quad + \frac{f(x_1, x_2 + \triangle x_2, x_3) - 2 f(x_1, x_2, x_3) + f(x_1, x_2 - dx_2, x_3)}{\triangle x_2^2} \\
&\quad + \frac{f(x_1, x_2, x_3 + \triangle x_3) - 2 f(x_1, x_2, x_3) + f(x_1, x_2, x_3 - \triangle x_3)}{\triangle x_3^2} \Big) \\
&\quad + \gamma f(x_1, x_2, x_3) \tag{17}
\end{aligned}
$$

where $\triangle x_1 = 1/N_1, \triangle x_2 = 1/N_2, \triangle x_3 = 1/N_3$ are the discretization of the domain Ω to the unit-cube. The operator $L^\dagger L = L^2$ may be evaluated by $L(Lf))$ or by explicitly writing out the discretized version of $L^2 f$ on the lines of Equation 17. Given the function $g(x)$, function $f(x)$ is evaluated using the Equation 16 by solving $f = (L^\dagger L)^{-1} g$. This computation is efficiently achieved in the Fourier domain: if F et G are the discrete Fourier transforms of f et g, equation 17 yields $G(k) = A(k)^2 F(k)$ where frequency $k = (k_1, k_2, k_3)$ and $A(k) = \gamma + 2\alpha \sum_{i=1}^{3} \big(\frac{1 - \cos(2\pi \triangle x_i k_i)}{\triangle x_i^2} \big)$. Thus given g, equation 17 is solved by taking the fourier transform, yielding G, dividing each $G(k)$ by $A(k)^2$ and then computing the inverse Fourier transform of the result to yield f.

3.3   Integration of velocity field to generate maps

Given the velocity fields, we need to compute the diffeomorphisms $(\phi_t^v)^{-1}$ that we denote as $\phi_{t,0}^v$. The equation of evolution of this function is derived by differentiation with respect to time $(\phi_t^v)^{-1} \circ \phi_t^v(x) = x$ for all $x \in \Omega$ giving:

$$
\frac{\partial}{\partial t}(\phi_t^v)^{-1} \circ \phi_t^v(x) + D(\phi_t^v)^{-1} \circ \phi_t^v(x) \ \ \frac{d}{dt}\phi_t^v(x) = 0
$$

$$
\frac{\partial}{\partial t}(\phi_t^v)^{-1} \circ \phi_t^v(x) + D(\phi_t^v)^{-1} \circ \phi_t^v(x) \ \ v_t \circ \phi_t^v(x) = 0
$$

$$
\frac{\partial \phi_{t,0}^v}{\partial t}(y) + D\phi_{t,0}^v(y) \ \ v_t(y) \overset{(a)}{=} 0
$$

where the derivative $d/dt(\phi_t^v)$ is given by definition to be $v_t \circ \phi_t^v$, D stands for the Jacobian operator, and (a) follows from change of variables $\phi_t^v(x) = y$ and finally changing back to the notation used throughout the paper that $(\phi_t^v)^{-1}$. This differential equation is in the Eulerian frame of reference, it is an example of a transport equation where the flow is observed at the fixed regular Cartesian mesh (note that the velocity is specified on the grid points $y \in \Omega$ at

10

each time t) by following a new set of particles corresponding to those that are at the regular mesh points at each discretized time instant. In contrast, the Lagrangian equation presented in (Equation 2), flow is observed by following the streamlines of the flow of particles, this leads to computations that are on a regular Cartesian mesh to start with but evolve to become computations on an unstructured mesh. A straightforward computation of the solution to the equation in the Eulerian frame of reference, since one needs to compute the function $\phi^v_{t,o}$ is to approximate the derivative by corresponding finite differences and propogate the solution from time t to time $t + 1$ starting with known solution at time $t = 0$ which is the identity map. This scheme is known as the Eulerian scheme. The solution to (Equation 2) also approximated by standard finite differences is called a Lagrangian scheme. The issue here is the choice of the size of a timestep. On one hand, in the interest of accuracy, the timestep of discretization needs to be small, on the other hand, overly small timesteps lead to large computation times and thus impractical. Eulerian schemes have the disadvantage that the size of timestep is restricted by issues related to CFL stability [13] whereas Lagrangian schemes are computational stable and provide good tolerance to numerical-accuracy related errors for choice of larger size timestep compared to Eulerian schemes. On the other hand, Lagrangian schemes are more involved to implement due to the need to track particles whose trajectories evolve away from their starting positions on a regular grid.

Semi-Lagrangian schemes [15] are a hybrid between these two schemes, they involve following the streamlines for a timestep with the streamlines chosen at each timestep to be those that end on the regular grid at the next timestep. This allows the choice of larger timesteps which are an advantage of Lagrangian schemes and data computation on a regular grid, which are an advantage of the Eulerian numerical schemes. We have implemented a two-step semi-Langrangian scheme with iterative backward trajectory calculation which are described in [15]. Given the velocity field $v^k_{t_j}(y)$ for all discretized timesteps along the flow $t_j, j \in [0, N-1], \delta t = T/N$, the semi-Lagrangian based scheme to solve the transport equations governing the evolution of the mappings $\phi^v_{t_j,0}(y), \phi^v_{t_j,T}(y)$ are the equations:

$$\phi^v_{t_j,0}(y) = \phi^v_{t_{j-1},0}(y - \alpha) \tag{18}$$

$$\phi^v_{t_j,T}(y) = \phi^v_{t_{j+1},T}(y + \alpha). \tag{19}$$

where α is the displacement in one-timestep of following the streamlines with appropriate velocity $v^*_{t_j \pm 0.5}(y)$ at the mid-point of the interval $[t_{j-1}, t_j)$ for Equation 18 and $[t_j, t_{j+1})$ for Equation 19. Various schemes to estimate the velocity at the mid-point are described in [15], and we choose the simple approximation $v^*_{t_j \pm 0.5}(y) = v_{t_j}(y)$ based on the assumption of piece-

wise constant velocity in discretized time intervals. Given this velocity, the displacement α is iteratively calculated using the formula:

$$\frac{\alpha}{2} = \frac{\delta t}{2} v^*_{t_{j \pm 0.5}} (y - \frac{\alpha}{2})$$

(20)

with the data for non-cartesian grid points being obtained by bi/tri-linear interpolation.

3.4 Constant Speed Reparametrization of velocity vector field

The minimizer of 8 is constant speed ($\|v_t\|_V$ = constant) since it is a geodesic. This property can be numerically achieved by time reparametrization ([5]), and doing this in parallel to gradient descent can significantly reduce the convergence time. We use the following procedure. Define normalized length function $s : [0, T] \to [0, T]$ given by

$$s_t = \frac{T}{\text{Length}} \int_0^t \|v_t\|_V \, \mathrm{d}t$$

with Length $= \int_0^T \|v_t\|_V \, \mathrm{d}t$. Define the inverse function $h : [0, T] \to [0, T]$ given by $h_t = s_t^{-1}$ so that $s_{h_t} = t$ and $\dot{h}_t = 1/\dot{s}_{h_t}$ Then, define $\tilde{v}_t = \dot{h}_t v_{h_t}$, which has, by construction constant norm, $\|\tilde{v}_t\|_V = \text{Length}/T$. Since the Schwartz inequality implies

$$\int_0^T \|v_t\|^2 \mathrm{d}t \geq \text{Length}^2/T$$

this operation can only reduce the geodesic energy, whereas $\phi_t^{\tilde{v}} = \phi_{h_t}^v$, implying $\phi_T^{\tilde{v}} = \phi_T^v$ and the second term of $E(v)$, given in equation 8 is left unchanged. Thus the total energy after reparametrization monotonically decreases.

3.5 The Large deformation metric-matching algorithm

The matching algorithm initializes with iteration $k = 0$, $v_{t_j}^k = 0$, $\nabla_{v^k} E_{t_j} = 0$, $\phi_{t_j,0} = Id$, $\phi_{t_j,T} = Id$ $\forall t_j \in [0, T]$. For each iteration $k = 1, 2, \ldots K$ iterate the folowing:

1) Calculate new estimate of velocity $v^{k+1} = v^k - \epsilon \nabla_{v^k} E$.   2) Reparametrize the velocity field to be constant speed. This is done every 10 simulations.   3) Calculate for $j = N - 1$ to $j = 0$ the mapping $\phi_{t_j,T}^{k+1}(y)$ using Equation 19.   4) Calculate for $j = 0$ to $j = N - 1$ the mapping $\phi_{t_j,0}^{k+1}(y)$ using Equation 18.   5) Calculate for $j = 0$ to $j = N - 1$ the image $J_{t_j}^0 = I_0 \circ \phi_{t_j,0}^{k+1}$.   6) Calculate for $j = N - 1$ to $j = 0$ the image $J_{t_j}^1 = I_1 \circ \phi_{t_j,T}^{k+1}$. 7) Calculate for $j = 0$ to $j = N - 1$ the gradient of the image $DJ_{t_j}^0$.   8) Calculate for for $j = 0$ to $j = N - 1$ the Jacobian of the transformation $|D\phi_{t_j}|$.   9) Calculate for $j = 0$ to $j = N - 1$ the gradient $\nabla_{v^{k+1}} E$ for v^{k+1} using Equation 10.   10) Calculate the norm of

the new gradient $\|\nabla_{v^{k+1}}E\|$. Stop if below threshold.   11) Calculate the new Energy using Equation 14.   12) If the number of simulations greater than specified number then Stop. Else re-iterate with $k = k + 1$.   13) Denote the final velocity field as $\hat{v}$ which gives the estimate of the desired optimizer of Equation 8.   14) Calculate the length of the path on the manifold using Equation 15. This is the length of the geodesic and hence the estimated metric between the given images.

3.6 $\int_0^1 \|v_t\|_V \mathrm{d}t$ defines a metric on $\mathcal{I}$

The metric distance between any two points $\varphi_0, \varphi_1 \in \mathcal{G}$ comes from the geodesic length of the curve $\phi_t^v \in \mathcal{G}, t \in [0, 1]$ associated to a vector field $v \in L^2([0, 1], V)$ given by $\int_0^1 \|v_t\|_V \mathrm{d}t$ with end-points $\varphi_0 = \phi_0^v$ and $\varphi_1 = \phi_1^v$. The following construction [9, 12, 17]

$$\rho_{\mathcal{G}}(\varphi_0, \varphi_1) \doteq \inf \left\{ \int_0^1 \|v_t\|_V dt \mid \varphi_1 = \phi_1^v \circ \varphi_0 \right\} \tag{21}$$

defines a distance on $\mathcal{G}$ for which $\mathcal{G}$ is complete. From the definition of the distance $\rho_{\mathcal{G}}$ and the construction of $\mathcal{G}$, we get immediately the following right-invariance property:

$$\rho_{\mathcal{G}}(\varphi_0, \varphi_1) = \rho_{\mathcal{G}}(\varphi_0 \circ \varphi, \varphi_1 \circ \varphi).$$

Proposition 3.1. *The function $\rho_{\mathcal{I}} : \mathcal{I} \times \mathcal{I} \to \mathbb{R}_+$ defined on the anatomical ensemble $\mathcal{I}$ by:*

$$\rho_{\mathcal{I}}(I_0, I_1) \doteq \inf\{ \ \rho_{\mathcal{G}}(Id, \varphi) \mid I_1 = I_0 \circ \varphi^{-1}, \ \varphi \in \mathcal{G} \ \}.$$

is positive, symmetric and satisfies the triangle-inequality.

Proof. The distance $\rho_{\mathcal{I}}(I_0, I_1)$ inherits the properties of positivity and symmetry from the distance $\rho_{\mathcal{G}}$. It also satisfies the triangle-inequality if the distance $\rho_{\mathcal{G}}$ is right-invariant to the stabilizer of the template, which is satisfied here by construction as $\rho_{\mathcal{G}}$ is invariant to the entire group $\mathcal{G}$ [12]. From the results of Dupuis et. al, Trouvé and Younes, it can be shown that the infimum is attained with $\rho_{\mathcal{I}}(I_0, I_1) = 0 \implies I_0 = I_1$ and thus $\rho_{\mathcal{I}}$ is a metric on $\mathcal{I}$ [19]. $\square$

Intuitively, if the images of the given anatomies are *close*, then the diffeomorphic transformation needed to register the images will be closer to the identity transformation and the corresponding distance between images as defined in Equation 22 will be smaller than if the given images are *far* away. The distance presented only deals with the deformation aspects, and is not invariant by rigid transformations, which is not an issue for many application, for which rigid registration is given, or at least may be recovered by standard algorithms.

13

It also turns out that the critical points $v \in L^2([0,1], V)$ of the length functional $\int_0^1 \|v_t\|_V \mathrm{d}t$ are also the critical points of the energy functional $\int_0^1 \|v_t\|_V^2 \mathrm{d}t$ with the additional property that they are of constant speed giving that

$$\rho_G^2(\varphi_0, \varphi_1) = \inf \left\{ \int_0^1 \|v_t\|_V^2 \mathrm{d}t \mid \varphi_1 = \phi_1^v \circ \varphi_0, \ \|v_t\| = \text{constant} \right\}$$

the proof for which is found in in [6]. Therefore, the length of the vector field satisfying the Euler-Lagrange equation 10 being the variational optimizer of 8 gives the metric distance between images I_0 and $I_0 \circ \phi_{1,0}^v$. The problem which is dealt with in this paper, ie. the minimization of

$$\int_0^1 \|v_t\|_V^2 \mathrm{d}t + \frac{1}{\sigma^2} \|I_0 \circ \phi_{1,0}^v - I_1\|_{L^2}^2$$

can be seen as an approximation of the exact matching problem which corresponds to the computation of ρ_G, taking into account possible unaccuracies in the observation of I_1.

4   Numerical results

We first present the application of the LDDMM algorithm to compute a geodesic path between two given images, and the metric estimated for the pair. The algorithm is implemented in C++ to work on 2D images as well as 3D volumes and parallelized using MPI to execute on an IBM RS6000-SP computer to take advantages of larger memory and concurrent parallel processing for computations. The time interval of the flow is discretized, unless stated otherwise, into 20 steps, with each step being of length $\delta t = 0.1$.

The "Parallel Translation" experiment results are shown in Figure 1, where the image pair being mapped are a circular ball and its diagonal translated version. The "Heart Mapping" experiment results are shown in Figure 2, where the images being mapped are 2D MRI sections from a normal canine heart and an abnormal canine heart (Data courtesy of Dr. Raimond Winslow, Center for Computational Biology and Medicine, The Johns Hopkins University). The "Schizophrenia Hippocampus Mapping" (Schiz.) and the "Alzheimer Hippocampus Mapping" (Alzh.) are shown in Figures 3 and 4 respectively, where a section from the hippocampus of a patient diagnosed with Schizophrenia and a patient diagnosed with Alzheimer's disease are being mapped to a corresponding section from a young control (Data courtesy of Dr. John Csernansky, Washington University). The "Macaque Cortex Mapping" (Macaq.) results are shown in Figure 5 where the images being mapped are the cryosection images of the cortex from two Macaque brains (Data courtesy of Dr. David Van Essen,

14

Experiment (Image Size)	L	Image Error	Est. Metric	Iter.	Exec. Time
"Parallel Translation" (32x32)	$-0.01\nabla^2 + 0.1I$	1.97%	0.769	70	0.2min
"Heart Mapping"(80x80)	$-0.01\nabla^2 + I$	7.12%	6.122	159	3.7min
"Schiz."(64x64)	$-0.01\nabla^2 + I$	3.01%	4.891	277	4.2min
"Alzh." (64x64)	$-0.01\nabla^2 + I$	2.69%	5.592	255	3.9min
"Macaq."(80x80)	$-0.01\nabla^2 + I$	3.64%	6.869	134	3.2min

Table 1: Summary of parameters used for the 2D image matching experiments as well as the final error in image overlap after the mapping relative to the error before the mapping, metric distance, number of iterations for convergence and execution time.

Washington University). The 2D experiments were run on a single processor, Table 1 shows parameters and data from these experiments. The "Mitochondria Mapping" experiment results (Figure 8) show the segmented mitochondria shapes from high resolution micrographs that were mapped to a single image and the estimated metrics for operator $L = -0.01\nabla^2 + I$. The results of mapping 3D hand-segmented volumes taken from Alzheimers and Schizophrenia populations are shown in the Figure 9. The 3D images are of dimensions 80 pixels by 128 by 128 pixels and the operator $L = -0.01\nabla^2 + 0.1I$ was used for the mapping. Except for the "Parallel Translation" experiment, all other images were registered using the software "Analyze" [14] to remove rigid rotation and translation prior to the experiment. Execution times depend on the size of images and the number of iterations needed to reach convergence. For 2D images of the size 64×64, the execution time of the algorithm is of the order of a few minutes on a single processor. For 3D images of the size $80 \times 128 \times 128$, execution time ranges from half an hour to a few hours running on 8 processors.

For each of these experiments, the sequence comprising the image I_0 composed with the diffeomorphisms at discretized instants $t_j, j = 6, 12, 19$ on the geodesic path to the image I_1 are shown in Figures 1, 2, 3, and 4. Figures 6 shows the vector plot of the superposition of velocity fields estimated for the flow, and 7 the estimated mapping $\phi_{1,0}$ for each of these experiments.

5 The "Christensen Algorithm" for computing large deformation diffeomorphisms

As described in the introduction, Christensen et. al. [7] proposed flows for large deformations generated paths through the space of diffeomorphisms which match the corresponding images. These paths do not correspond to any global variational problem solution as stated in 1, and will not in general generate the shortest path connecting the images through the space of diffeomorphisms. Therefore it is not an algorithm that can be used to give the orbit a metric distance. The Christensen algorithm exploits the fact that if the operator L does not differentiate in time, then discretize space-time $\Omega \times T$ into a sequence of time-indexed optimizations solving for the locally optimal velocity at each time and then forward integrating the solution. This is only a locally-in-time optimal method reducing the dimension of the optimization and that at each time-step, the algorithm attempts to greedily reach the target. The transformation $\phi_{1,0}$ registering the given anatomical images is generated from velocity fields which are assumed piecewise constant within quantized time increments of size δ, $k = 0, \ldots, K$, $t_k = k\delta, k = 0, \ldots, K = \frac{T}{\delta}$. The locally optimal velocity fields satisfy the partial differential equation

$$L^\dagger L v_t + b_t = 0 \tag{22}$$

where the function $b_t : \Omega \to \mathbb{R}^n$ is given by

$$b_t(x) = -\alpha(J_t^0(x) - J_1^1(x))\nabla J_t^0(x), \tag{23}$$

where $J_t^0 = I_0 \circ \phi_{t,0}$ and $J_1^1 = I_1$ and α is some constant. The time-indexed sequence of locally optimal velocity fields v_{t_j} are integrated to yield the sequence of transformations $\phi_{t_j}^v$, $j = 0, 1, 2, \ldots$, which are points along a path on the manifold of diffeomorphisms from the identity transformation to the point $\phi_{1,0}$, the length of which is $\int_0^1 \|v_t\|_V \, dt$.

One way to interpret this method is to explore the connection to the search for the optimizer of the cost in the space of diffeomorphisms $\mathcal{G}$ [18]. Rewrite Equation 8 as a function on $\mathcal{G}$ to be

$$E(\phi) = \underbrace{\rho_{\mathcal{G}}(id, \phi_1)^2}_{E_1(\phi)} + \underbrace{\frac{1}{\sigma^2} \|I_0 \circ \phi_{1,0} - I_1\|_{L^2}^2}_{E_2(\phi)} \tag{24}$$

and now the matching diffeomorphism is sought by a search in the space of diffeomorphisms $\mathcal{G}$. Suppose that the optimization for the matching diffeomorphism ϕ_1 proceeds on the manifold with the estimate ϕ_t at time t being the candidate as the end-point of a curve $\phi : [0, t] \to \mathcal{G}$ joining all previous estimated points with $\phi_0 = id \in \mathcal{G}$. The length of the curve traversed

is then $\int_0^t \|v_t\|_V \mathrm{d}t \geq \rho_{\mathcal{G}}(id, \phi_t)$ and this provides an upper-bound on the metric distance $\rho_{\mathcal{G}}(id, \phi_t)$ as the curve traversed is not necessarily the geodesic. Let the cost for traversing this curve from the identity to $\phi_t \in \mathcal{G}$ on the manifold be denoted by W_t where

$$W_t = \text{Length}(id, \phi_t)^2 + E_2(\phi_t) \geq \rho_{\mathcal{G}}(id, \phi_t)^2 + E_2(\phi_t) \tag{25}$$

where $\text{Length}(id, \phi_t)$ is the length of the curve ϕ with endpoints $id, \phi_t \in \mathcal{G}$. The local optimization process seeks the velocity v_t such that moving along direction $v_t \circ \phi_t$ at the point ϕ_t on the manifold leads to the lowest increase of the cost W_{t+dt} at $t + dt$. Since $W_{t+dt} = (\text{Length}(id, \phi_t) + \|v_t\|_V dt)^2 + E_2(\phi_{t+dt})$, which we can rearrange and simplify to the first order as

$$W_{t+dt} = W_t + 2\|v_t\|_V \times dt \times \text{Length}(id, \phi_t) + \partial_{v_t \circ \phi_t} E_2(\phi_t), \tag{26}$$

where $\partial_{v_t \circ \phi_t} E_2(\phi_t)$ is the change in E_2 at the point ϕ_t along the direction $v_t \circ \phi_t$. Fixing $\|v_t \circ \phi_t\|_{T_{\phi_t} \mathcal{G}} = 1$ i.e. $\|v_t\|_V = 1$, the local optimizer $v_t \circ \phi_t \in T_{\phi_t} \mathcal{G}$ that minimizes $\partial_{v_t \circ \phi_t} E_2(\phi_t)$ is given by

$$\begin{aligned} v_t \circ \phi_t \quad = \quad & \underset{\|v_t\|_V = 1}{\operatorname{argmin}} \left(\partial_{v_t \circ \phi_t} E_2(\phi_t) = \langle \nabla E_2(\phi_t), v_t \circ \phi_t \rangle_{T_{\phi_t} \mathcal{G}} \right) & (27) \\ = \quad & -\nabla E_2(\phi_t) = \text{Opposite the direction of steepest increase} & (28) \\ \Rightarrow \quad v_t \quad = \quad & -\nabla E_2(\phi_t) \circ \phi_t^{-1} & (29) \end{aligned}$$

More formally, this can interpreted as a gradient descent on the manifold of diffeomorphisms [18] as:

$$\frac{\partial \phi_t}{\partial t} = v_t \circ \phi_t = -\nabla E_2(\phi_t). \tag{30}$$

The computation of $\nabla E_2(\phi_t)$ is done by writing

$$\begin{aligned} \partial_{v_t \circ \phi_t} E_2(\phi_t) \quad \overset{(a)}{=} \quad & \frac{-2}{\sigma^2} \langle I_0 \circ \phi_{t,0} - I_1, D(I_0 \circ \phi_{t,0}) v_t \rangle_{L^2} & (31) \\ \overset{(b)}{=} \quad & \frac{-2}{\sigma^2} \langle K\left((J_t^0 - J_1^1)\nabla J_t^0\right), v_t \rangle_V & (32) \\ \overset{(c)}{=} \quad & \langle \nabla E_2(\phi_t), v_t \circ \phi_t \rangle_{T_{\phi_t} \mathcal{G}} \overset{(d)}{=} \langle \nabla E_2(\phi_t) \circ \phi_t^{-1}, v_t \rangle_V & (33) \end{aligned}$$

where (a) follows from differentiating the cost E_2, (b) from writing the transpose and changing the notation $\nabla J_t^0 = D(J_t^0)^t = D(I_0 \circ \phi_{t,0})^t$ along with transferring the gradient from space L^2 to space V using Equation 6, (c) from definition of the gradient in the tangent space $T_{\phi_t} \mathcal{G}$ at point ϕ_t, (d) from transferring the inner-product to the tangent space $V = T_{id} \mathcal{G}$ at the identity of the manifold. This is a Riemannian (sometimes called "natural") gradient in the

space of diffeomorphisms equipped with the right invariant metric, applied to for the data term or the fitting term E_2. Comparing Equations 29, 32 and 33, we get that

$$v_t = -\nabla E_2(\phi_t) \circ \phi_t^{-1} = \frac{2}{\sigma^2} K\left((J_t^0 - J_1^1)\nabla J_t^0\right) \qquad (34)$$

is the optimizer for the incremental local cost given in Equation 24. Rearrangement of Equation 34 provides the PDE in Equations 22 and the body force function 23 establishing the connection of the PDE formulation and the Riemannian gradient on the fitting term for the local optimization of the matching cost in the space of diffeomorphisms. The regularization provided by the operator K gives this gradient numerically stable behaviour in finite time (of course, the limit behavior is unstable, because the regularization term $\rho(id, \phi^v)^2$ is not taken into account). Henceforth, we denote this algorithm proposed by Christensen et. al as the "Christensen" or the "GEC" algorithm summarized as following.

1) Initialize: $t_j = 0$, $v_{t_j} = 0$, $\phi_{t_j} = Id$.   2) Integrate the velocity field v_{t_j} to compute the map $\phi_{t_{j+1},0}$. Adjust the time-step of integration δt such that the largest displacement is within one pixel in magnitude.   3) Use estimated $\phi_{t_{j+1},0}$ to compute cost $E_2(\phi_{t_{j+1},0})$. This cost quantifies the amount of registration of I_0 and I_1 under the estimated mapping.   4) If $E_2 <$ Threshold, STOP.   5) Using Equation 23, calculate the body force $b_{t_{j+1}}$.   6) Solve Equation 22 for local optimizer $v_{t_{j+1}}$ using FFT-based inversion giving $v_{t_{j+1}} = -K(b_{t_{j+1}})$. 7) Set $j = j + 1$ and go to step (2).   8) On STOP, set time $t_j = 1$, and number of steps taken $N = j$.

Our implementation of the above algorithm also includes the template propogation feature as originally described by Christensen et. al. (see [7] for details) to handle large deformations where the computed transformations approach local singularity on a discretized spatial grid and thus lead to numerical precision errors in interpolating concatenated transformations. Note that the "demons" algorithm, proposed by Thirion [16], carries the same purpose (minimization of E_2), but the GEC algorithm has the advantage of performing a true gradient descent, for a suitable Riemannian structure, and always remain within a space of smooth diffeomorphisms.

5.1 Comparison of the LDDMM algorithm with the GEC algorithm

Figures 10 and 11 show results for three experiments with the "LDDMM" and the "GEC" algorithms. The experiments were terminated when estimated mappings from both made the image mismatch error comparable. Figure 10 shows the image mismatch error as a function of the estimated distance on the manifold of the estimated transformation from the

Experiment	Final Image Error		Distance on Manifold		Time(s)
	LDDMM	GEC	LDDMM	GEC	LDDMM/GEC
"Parallel Translation"	5.648%	5.65%	0.7223	0.722757	12/1
"Heart Mapping"	11.558%	11.799%	4.631	6.7145	120/2.4
"Macaque Cortex Mapping"	9.55%	9.53%	4.636	5.45	98/1.2

Table 2: Comparison of the length of the path on the manifold of diffeomorphisms from the identity to the solution of the metric mapping algorithm and the solution of the "GEC" algorithm.

identity. At the same "distance" from the target image, as measured by comparable values of image mismatch error shown in table 2, the "LDDMM" algorithm (solid line) estimates a transformation whose distance from the identity is shorter than that estimated by the "GEC" algorithm (dotted line).

Figure 11 shows the velocity field and the mappings generated using the two algorithms. The mappings generated by both algorithms, by visual inspection, look very similiar. The velocity fields estimated by the two algorithms are different however. The superposition of velocity fields in time on a single plot for both algorithms reveals that for the "LDDMM" algorithm, the velocity fields are smooth not only in space, but also in time, whereas the velocity fields are smooth in space for the "GEC" algorithm but widely varying from one time-step to another. The distance traversed on the manifold to reach those points is shorter for the "LDDMM" algorithm than the "GEC" algorithm.

6 Discussion

6.1 Hilbert gradient versus L^2 gradient

Notice that the gradient presented in Equation 10 is computed according to the inner product on V, and we refer to it as a Hilbert gradient, which we would like to compare to its variant in the space L^2. Let $L^2(\Omega, \mathbb{R}^n)$ be the space of square-integrable vector fields on Ω with the usual L^2 product $\langle \cdot, \cdot \rangle_{L^2}$. Using Equations 5 and 7, in a weak sense, we can rephrase equation 11 as

$$\partial_h E(\hat{v}) = \int_0^1 \left\langle 2(L^\dagger L)\hat{v}_t - \left(\frac{2}{\sigma^2}|D\phi_{t,1}^{\hat{v}}|\nabla J_t^0 \left(J_t^0 - J_t^1 \right) \right), h_t \right\rangle_{L^2} \mathrm{d}t = 0$$

19

giving the L^2 gradient of the cost to be

$$(\nabla_v E_t)_{L^2} = 2(L^\dagger L)v_t - \left(\frac{2}{\sigma^2}|D\phi_{t,1}^v|\nabla J_t^0\left(J_t^0 - J_t^1\right)\right).$$

Set the term $b_t = \frac{-2}{\sigma^2}|D\phi_{t,1}^v|\nabla J_t^0\left(J_t^0 - J_t^1\right)$ and write the two gradients in space V and space $L^2(\Omega, \mathbb{R}^n)$ with this simplification as being

$$\begin{aligned}(\nabla_v E_t)_V &= 2v_t + Kb_t\\(\nabla_v E_t)_{L^2} &= 2(L^\dagger L)v_t + b_t\end{aligned}$$

From this expression, it is clear that the Hilbert gradient can be deduced from the L^2-gradient by applying the compact operator K, yielding in this way much more stable computations. Another way to interpret this is by an expansion in an orthonormal basis for $L^2(\Omega, \mathbb{R}^n)$. Let the self adjoint operator $(L^\dagger L)$ be diagonalized in the orthonormal basis $(w_i)_{i\in\mathbb{N}}$ with eigenvalues given by $(\lambda_i)_{i\in\mathbb{N}}$. The expression for $(\nabla_v E_t)_{L^2}$ becomes:

$$(\nabla_v E_t)_{L^2} = \sum_{i\in\mathbb{N}}\left(\langle\, 2(L^\dagger L)v_t, w_i\,\rangle_{L^2} + \langle\, b_t, w_i\,\rangle_{L^2}\right)w_i = \sum_{i\in\mathbb{N}}\left(\lambda_i\langle\, 2v_t, w_i\,\rangle_{L^2} + \langle\, b_t, w_i\,\rangle_{L^2}\right)w_i \tag{35}$$

In the other case, expansion of Kb_t in the same basis gives

$$Kb_t = \sum_{i\in\mathbb{N}}\langle\, Kb_t, w_i\,\rangle_{L^2}w_i = \sum_{i\in\mathbb{N}}\langle\, b_t, Kw_i\,\rangle_{L^2}w_i = \sum_{i\in\mathbb{N}}\frac{\langle\, b_t, w_i\,\rangle_{L^2}}{\lambda_i}w_i$$

and the gradient $(\nabla_v E_t)_V$ becomes

$$(\nabla_v E_t)_V = \sum_{i\in\mathbb{N}}\left(\langle\, 2v_t, w_i\,\rangle_{L^2} + \frac{\langle\, b_t, w_i\,\rangle_{L^2}}{\lambda_i}\right)w_i \tag{36}$$

The operator K is compact, therefore the eigenvalues $1/\lambda_i \to 0$ as $i \to \infty$ and hence $\lambda_i \to \infty$ as $i \to \infty$. The difference in behavior of the two gradients come from the multiplication or division of the basis coefficients with the eigen-values λ_i. Consequently, in Equation 35, the high frequency coefficients are amplified whereas in Equation 36 they are smoothed. The numerical instability of the L^2 gradient clearly derives from this high-frequency amplification.

The gradient derived for the presented cost in the space V is compared to the more traditional gradient in the space L^2 and we show that the L^2 gradient is unstable due to its high-frequency amplification. As a result, these gradients differ markedly in their numerical properties. We have implemented both of these gradients numerically and found that the Sobolev gradient is much more stable than the corresponding gradient in the space L^2.

6.2   The metric structure on $\mathcal{I}$

The highlight of LDDMM algorithm presented is the computation of the metric distance between given images coming from the computation of a geodesic path on the manifold of diffeomorphisms connecting the images. This method is based on following a gradient-descent based scheme with a Hilbert gradient in the space V of vector fields for the global variational optimization of the proposed cost 1. We present the derivation and implementation details of this gradient algorithm and denote it as the LDDMM algorithm. The velocity vector field solving this variational problem defines a "geodesic" path on the manifold of diffeomorphisms and the length of this path is a metric distance between the images connected via the diffeomorphism at the end point of the flow i. e. I_0 and $I_0 \circ \phi_{1,0}$.

We compare the matching obtained by our gradient method optimizing over the entire flow in $L^2([0,1], V)$ to the Christensen method of generating locally optimal flows via a viscous-fluid PDE formulation which probably is one of the most efficient greedy methods in this context. We also discuss an interpretation of this method as generating a locally optimal velocity based on the Hilbert gradient of the fitting term. This method generates a path on the manifold of diffeomorphisms, not necessarily the shortest path, but one for which the incremental cost at each step is minimized. Comparing the velocity field generated by these two methods, we see that the field generated by our gradient optimizing over the entire flow produces a field that is smooth in space and time, whereas the Christensen method produces a field that is smooth in space but not as markedly smooth with incremental steps in time, which is to be expected from this local-in-time optimization method. The distance generated by the Christensen method provides an upper bound for the metric distance. We show experiments using the two methods for comparing these properties.

In conclusion, we have presented in this paper the derivation and numerical implementation of a new gradient-based method for computing dense image based mappings solving a global variational problem as stated in 1 and estimating metrics for images. We also present results for applying this method to a several biological shapes in 2D as well as 3D. The metrics estimated for Mitochondrial shapes provide a quantification on relative distance between images which can be compared to human intuition and seems to be in agreement with it. This method provides a quantitative measurement tool that can be used to compute many such metric distances for images mapped to a common anatomical reference image to build the distribution of these metrics in the population. This will allow the comparison of images quantifying relative "close" and "far", providing information that may be of potential use in

aiding clinical diagnosis and treatment. This algorithm is also directly applicable to non-rigid dense registration of color images or DT images and the calculation of metric distance on these orbits.

7 Acknowledgements

References

[1] Yali Amit. A nonlinear variational problem for image matching. *SIAM Journal on Scientific Computing*, 15(1):207–224, 1994.

[2] R. Bajcsy and C. Broit. Matching of deformed images. In *Proc. 6th Int. Joint Conf. Patt. Recog.*, pages 351–353, 1982.

[3] R. Bajcsy, R. Lieberson, and M. Reivich. A computerized system for the elastic matching of deformed radiographic images to idealized atlas images. *Journal of Computer Assisted Tomography*, 7(4):618–625, 1983.

[4] C. Broit. Optimal registration of deformed images. *PhD thesis, University of Pennsylvania*, 1981.

[5] M. P. Do Carmo. *Differential Geometry of Curves and Surfaces*. Prentice-Hall Engineering/Science/Mathematics, 1976.

[6] M. P. Do Carmo. *Riemannian Geometry*. Birkhauser, 1993.

[7] G. E. Christensen, R. D. Rabbitt, and M. I. Miller. Deformable templates using large deformation kinematics. *IEEE Transactions on Image Processing*, 5(10):1435–1447, October 1996.

[8] Gary Christensen. *Deformable shape models for anatomy.* PhD Thesis, Dept. of Electrical Engineering, Sever Institute of Technology, Washington Univ., St. Louis, MO., Aug. 1994.

[9] P. Dupuis, U. Grenander, and M.I. Miller. Variational problems on flows of diffeomorphisms for image matching. *Quarterly of Applied Mathematics*, LVI:587–600, September 1998.

[10] U. Grenander and M. I. Miller. Computational anatomy: An emerging discipline. *Quarterly of Applied Mathematics*, 56:617–694, 1998.

[11] M.I. Miller, A. Trouvé, and L. Younes. On the metrics and Euler-Lagrange equations of computational anatomy. *Annual Review of Biomedical Engineering*, 4:375–405, 2002.

[12] M.I. Miller and L. Younes. Group actions, homeomorphisms, and matching: a general framework. *International Journal of Computer Vision*, 41:61–84, 2001.

[13] K. W. Morton and D. F. Mayers. *Numerical Solution of Partial Differential Equations.* Cambridge University Press, University of Cambridge, 1996.

[14] R. A. Robb. *Biomedical Imaging, Vizualization and Analysis.* John Wiley and Sons, Inc., New York, NY, 1999.

[15] A. Staniforth and J. Côté. Semi-lagrangian integration schemes for atmospheric models-a review. *Monthly Weather Review*, 119:2206–2223, 1991.

[16] J-P Thirion. Image matching as a diffusion process: an analogy with maxwell's demons. *Medical Image Analysis*, 2(3):243–260, 1998.

[17] A. Trouvé. An infinite dimensional group approach for physics based models in patterns recognition. *Preprint*, 1995.

[18] A. Trouvé. Diffeomorphic groups and pattern matching in image analysis. *Int. J. Computer Vision*, 28:213–221, 1998.

[19] L. Younes. Optimal matching between shapes via elastic deformations. *Image and Vision Computing*, 17:381–389, 1999.
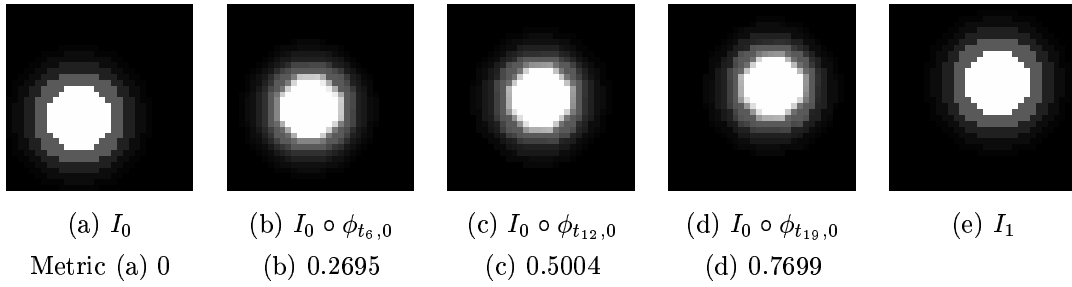
(a) I_0 (b) $I_0 \circ \phi_{t_6,0}$ (c) $I_0 \circ \phi_{t_{12},0}$ (d) $I_0 \circ \phi_{t_{19},0}$ (e) I_1

Metric (a) 0   (b) 0.2695   (c) 0.5004   (d) 0.7699

Figure 1: "Parallel Translation Experiment". Shown is image I_0 composed with the diffeomorphisms at discretized instants $t_j, j = 6, 12, 19$ on the geodesic path to the image I_1 and the corresponding metric distance for the image with respect to image I_0.
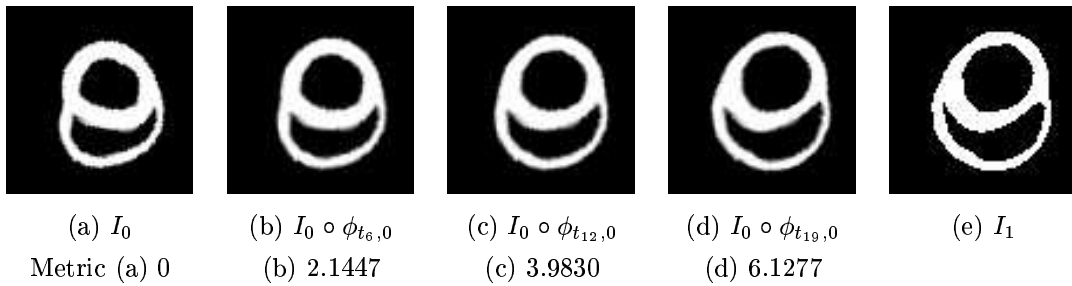

(a) I_0 (b) $I_0 \circ \phi_{t_6,0}$ (c) $I_0 \circ \phi_{t_{12},0}$ (d) $I_0 \circ \phi_{t_{19},0}$ (e) I_1

Metric (a) 0   (b) 2.1447   (c) 3.9830   (d) 6.1277

Figure 2: "Heart Mapping Experiment". Shown is image I_0 composed with the diffeomorphisms at discretized instants $t_j, j = 6, 12, 19$ on the geodesic path to the image I_1 and the corresponding metric distance for the image with respect to image I_0. Data taken from the laboratory of Dr. Raymond Winslow, Johns Hopkins University.
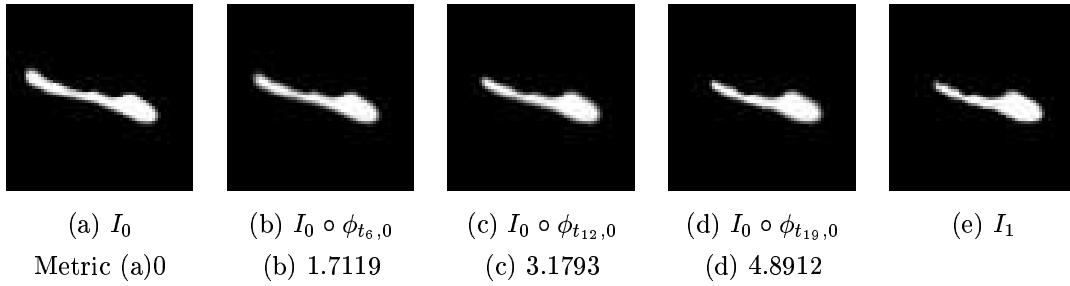
(a) I_0	(b) $I_0 \circ \phi_{t_6,0}$	(c) $I_0 \circ \phi_{t_{12},0}$	(d) $I_0 \circ \phi_{t_{19},0}$	(e) I_1
Metric (a)0	(b) 1.7119	(c) 3.1793	(d) 4.8912	

Figure 3: "Schizophrenia Hippocampus Mapping". Shown is image I_0 composed with the diffeomorphisms at discretized instants $t_j, j = 6, 12, 19$ on the geodesic path to the image I_1 and the corresponding metric distance for the image with respect to image I_0. Data taken from the laboratory of Dr. John Csernansky, Washington University.
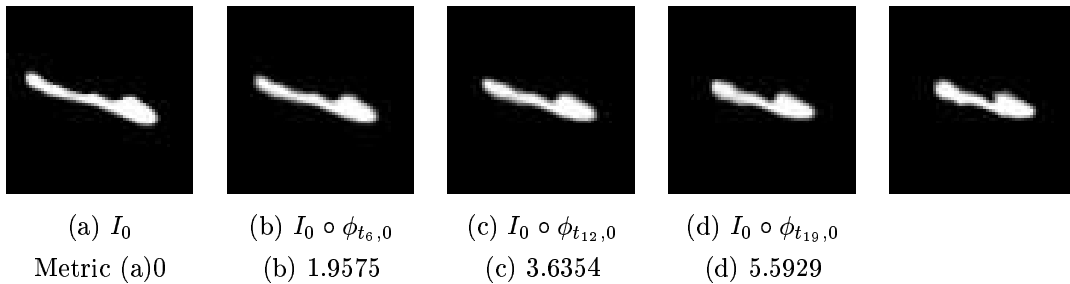

(a) I_0	(b) $I_0 \circ \phi_{t_6,0}$	(c) $I_0 \circ \phi_{t_{12},0}$	(d) $I_0 \circ \phi_{t_{19},0}$
Metric (a)0	(b) 1.9575	(c) 3.6354	(d) 5.5929

Figure 4: "Alzheimer's Hippocampus Mapping". Shown is image I_0 composed with the diffeomorphisms at discretized instants $t_j, j = 6, 12, 19$ on the geodesic path to the image I_1 and the corresponding metric distance for the image with respect to image I_0. Data taken from the laboratory of Dr. John Csernansky, Washington University.
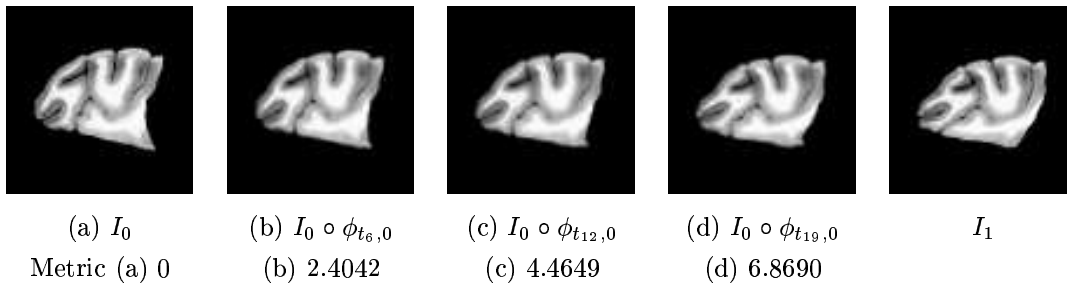
| (a) I_0 | (b) $I_0 \circ \phi_{t_6,0}$ | (c) $I_0 \circ \phi_{t_{12},0}$ | (d) $I_0 \circ \phi_{t_{19},0}$ | I_1 |
| Metric (a) 0 | (b) 2.4042 | (c) 4.4649 | (d) 6.8690 | |

Figure 5: "Macaque Cortex Mapping". Shown is image I_0 composed with the diffeomorphisms at discretized instants $t_j, j = 6, 12, 19$ on the geodesic path to the image I_1 and the corresponding metric distance for the image with respect to image I_0. Data courtesy of Dr. David Van Essen, Washington University.
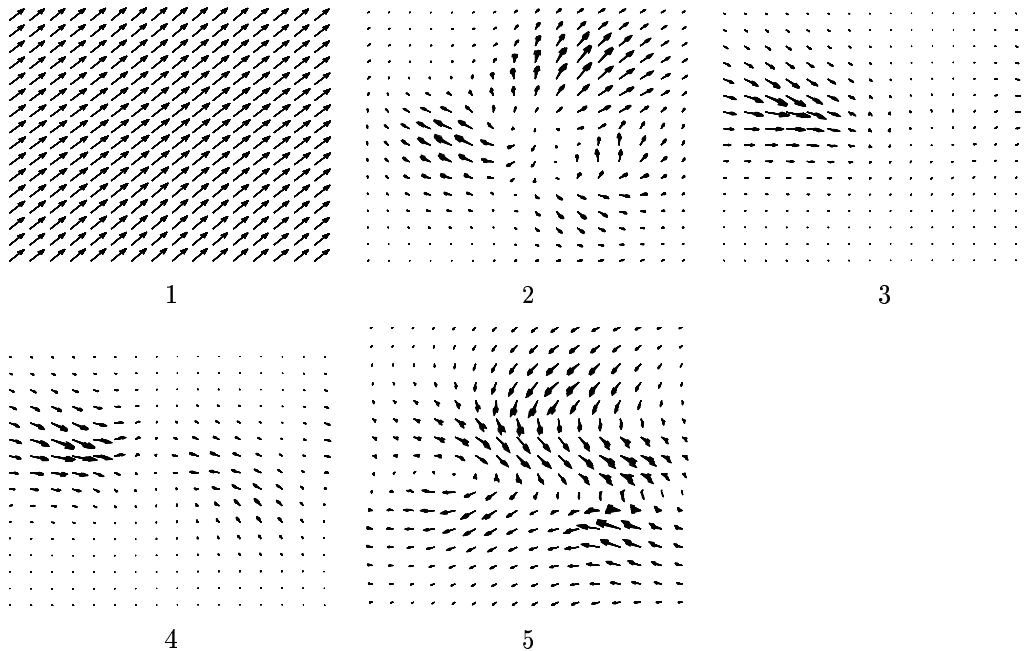

Figure 6: The velocity field along the discretized flow $v_{t_j}, j = 0 \ldots 19$ are superposed on a single figure for each of the "Parallel Translation"(panel 1), "Heart Mapping"(panel 2), "Schizophrenia"(panel 3), "Alzheimer's"(panel 4), "Macaque"(panel 5) experiments respectively. Notice that the velocity is smooth in space. Also interesting to note is that the velocity field is quite smooth in time as well.
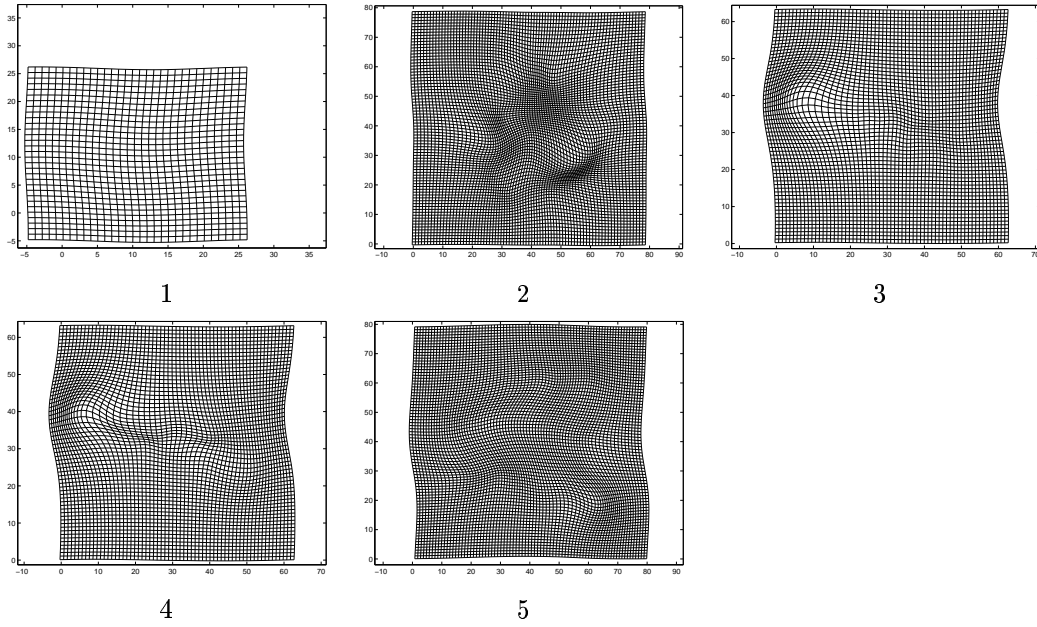
26

Figure 7: Shown here are the estimated matching diffeomorphism $\phi_{1,0}$ for the "Parallel Translation" (panel 1), "Heart Mapping" (panel 2), "Schizophrenia" (panel 3), "Alzheimer's" (panel 4), "Macaque" (panel 5) experiments repsectively.

(I_0) $(I_1)_1$ $(I_1)_2$ $(I_1)_3$ $(I_1)_4$

Distance 0 0.845 1.747 2.568 4.410

(I_0) $(I_1)_5$ $(I_1)_6$ $(I_1)_7$ $(I_1)_8$
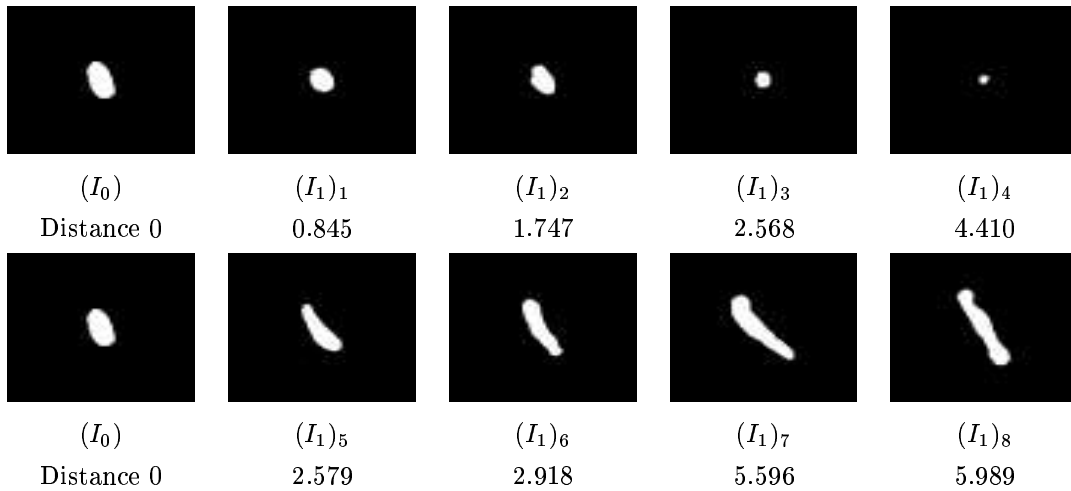
Distance 0 2.579 2.918 5.596 5.989

Figure 8: "Mitochondria Mapping Experiment". Shown in this figure are the results of mapping a Mitochondrial template image I_0 with Mitochondrial target images I_1 subscripted by index $1 \ldots 8$ and the calculated metric distances from I_0 below the respective images. Notice that the estimated metrics agree closely with the human perception of "close" and "far".

(1)I_0	(2) $I_0 \circ \phi_{t_{10},0}$	(3) $I_0 \circ \phi_{t_{19},0}$	(4) I_1
Metric 0	2.541	4.620	

(1)I_0	(2) $I_0 \circ \phi_{t_{10},0}$	(3) $I_0 \circ \phi_{t_{19},0}$	(4) I_1
Metric 0	3.472	6.313	

(1)I_0	(2) $I_0 \circ \phi_{t_{10},0}$	(3) $I_0 \circ \phi_{t_{19},0}$	(4) I_1
Metric 0	2.621	4.766	

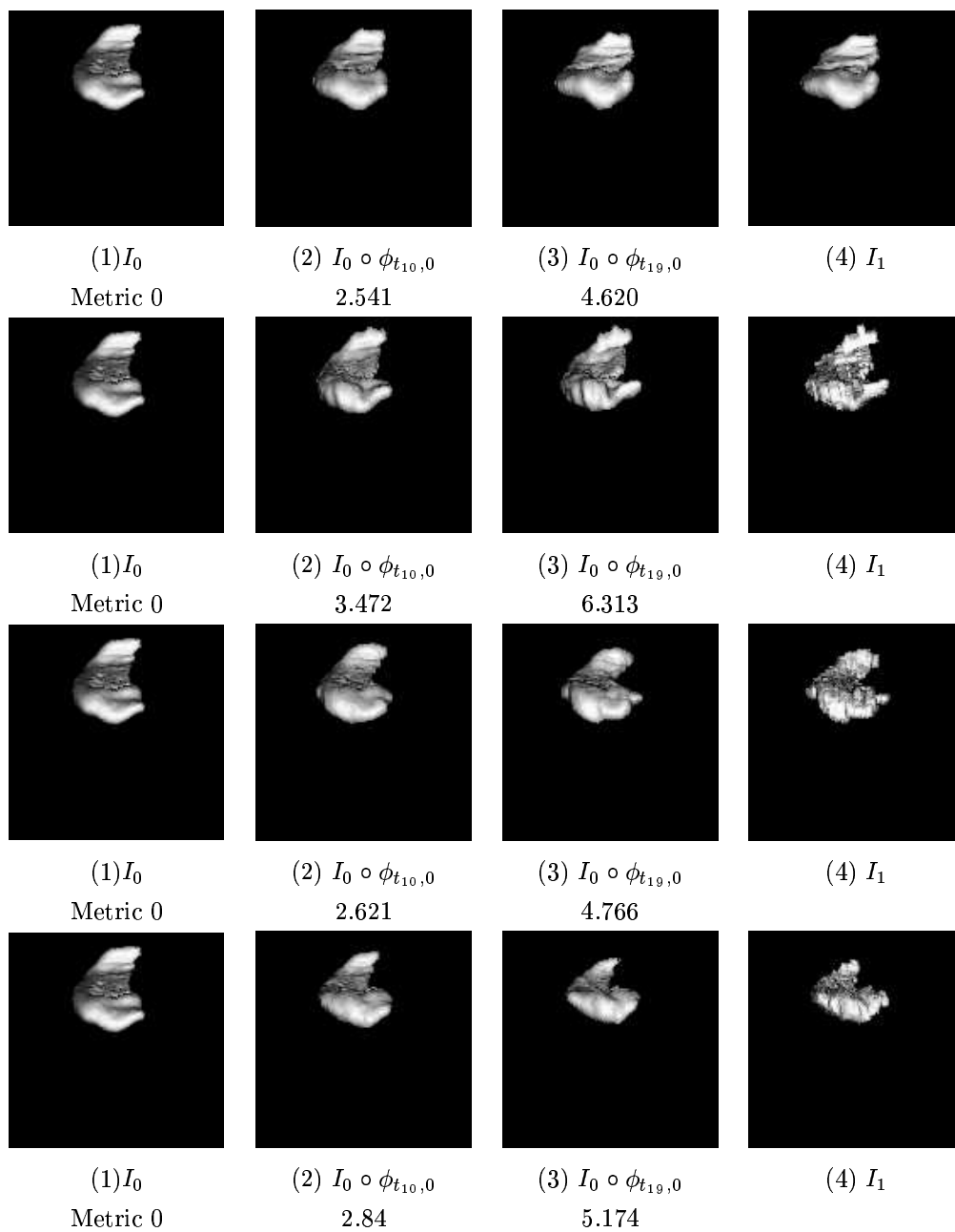| (1)I_0	(2) $I_0 \circ \phi_{t_{10},0}$	(3) $I_0 \circ \phi_{t_{19},0}$	(4) I_1
Metric 0	2.84	5.174	

Figure 9: "3D hippocampus mapping results". The right column of the top two rows show two hippocampus volumes taken from Schizophrenia patients and the right column of the bottom two rows show two hippocampus volumes from DAT patients. Allthe image volumes were mapped to a common template I_0 shown in the first column on the left. The second column from left shows the mapping at discretized instant t_{10} of the flow of 20 discretized timesteps. The third column from left shows the image I_0 deformed with the estimated diffeomorphism $\phi_{1,0}$ at the final discretized instant along the flow. The numbers below each image correspond to the metric distance from the image I_0. Data taken from the laboratory of Dr. John Csernansky, Washington University.
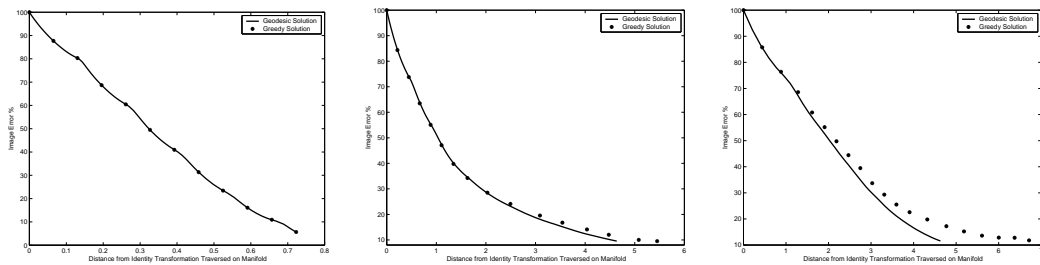
Figure 10: Comparing the "LDDMM" and the "GEC" algorithms. Shown is the Image Error (solid line - "LDDMM" algorithm, dotted line - "GEC" algorithm) as a percentage of error before the transformation via the estimated mapping versus the distance of the path on the manifold joining the identity of the diffeomorphisms to the estimated mapping. In the left panel is the results from the "Parallel Translation" experiment, in the middle panel is the "Macaque Cortex Mapping" experiment and the right panel shows the results from the "Heart Mapping" experiment.
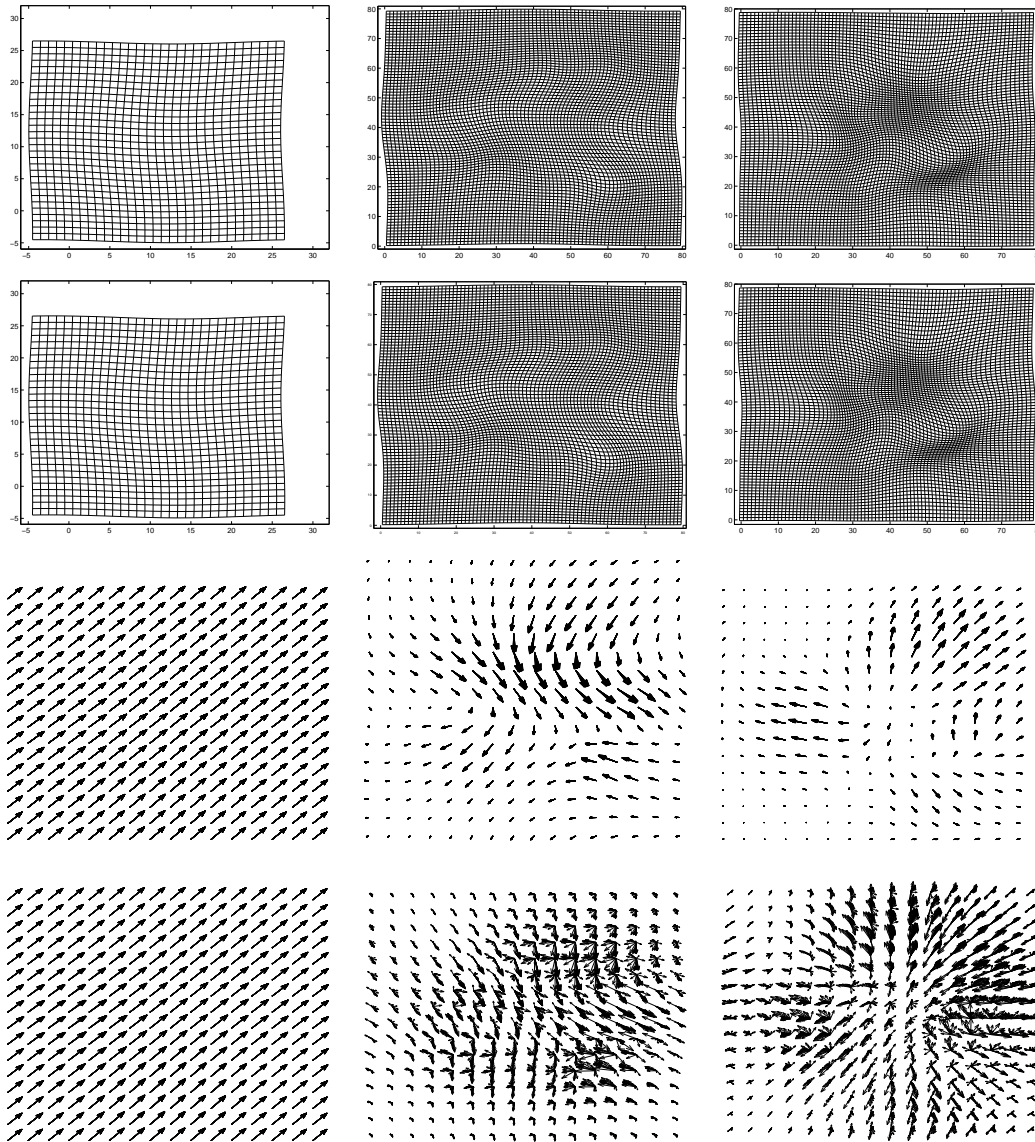
Figure 11: The left column shows data from the "Parallel Translation mapping", the middle column shows the results for "Macaque Cortex mapping" and the right column shows data from the "Heart mapping". In rows 1 and 3 from the top are the results from the "LDDMM" algorithm presented in this paper and in rows 2 and 4 are the results from our implementation of the "GEC" algorithm showing the estimated mappings and the superposed velocity fields that generates these mappings. The velocity field for the "GEC" algorithm is smooth in space but shows marked changes with time-steps.