Landmark Matching via Large Deformation Diffeomorphisms

Sarang C. Joshi and Michael I. Miller

***Abstract*—This paper describes the generation of large deformation diffeomorphisms $\phi : \Omega = [0,1]^3 \leftrightarrow \Omega$ for landmark matching generated as solutions to the transport equation $d\phi(x,t)/dt = v(\phi(x,t),t), t \in [0,1]$ and $\phi(x,0) = x$, with the image map defined as $\phi(\cdot,1)$ and therefore controlled via the velocity field $v(\cdot,t), t \in [0,1]$. Imagery are assumed characterized via sets of landmarks $\{x_n, y_n, n = 1,2,\ldots,N\}$. The optimal diffeomorphic match is constructed to minimize a running smoothness cost $\|Lv\|^2$ associated with a linear differential operator L on the velocity field generating the diffeomorphism while simultaneously minimizing the matching end point condition of the landmarks.**

Both inexact and exact landmark matching is studied here. Given noisy landmarks x_n matched to y_n measured with error covariances Σ_n, then the matching problem is solved generating the optimal diffeomorphism $\hat{\phi}(x,1) = \int_0^1 \hat{v}(\hat{\phi}(x,t),t)\,dt + x$ where

$$\hat{v}(\cdot) = \arg\min_{v(\cdot)} \int_0^1 \int_\Omega \|Lv(x,t)\|^2\,dx\,dt + \sum_{n=1}^{N} [y_n - \phi(x_n,1)]^T \Sigma_n^{-1} [y_n - \phi(x_n,1)]. \quad (1)$$

Conditions for the existence of solutions in the space of diffeomorphisms are established, with a gradient algorithm provided for generating the optimal flow solving the minimum problem. Results on matching two-dimensional (2-D) and three-dimensional (3-D) imagery are presented in the the macaque monkey.

***Index Terms*—Deformable templates, medical imaging, pattern theory.**

I. Introduction

RECENT revolutionary advances in the field of medical imaging have facilitated digital imaging modalities such as magnetic resonance (MR), X-ray computed tomography (CT), and cryosection imaging (CI), to name a few, and are enabling extremally detailed study of anatomy. Although the study of anatomical variability can be traced back to the beginnings of modern science, the exquisite resolution and the three-dimensional (3-D) and four-dimensional (4-D) capabilities of these imaging modalities, combined with the advances in digital computation, is only now enabling the detailed and precise *computational* study of the infinite biological variability of anatomy. This is emerging as the discipline that we are coming to call *computational anatomy* [1], [2] in which the main aim of our own work is to develop mathematical and software tools specialized to the understanding of the variability of brain anatomy in humans and primate monkeys [3]–[12].

The transformations $h \in \mathcal{H} : \Omega \to \Omega$ are constructed from the group of diffeomorphisms of the coordinate system Ω, $\mathcal{H} \equiv \mathrm{Diff}(\Omega)$, defined by vector fields of displacements

$$h \in \mathcal{H} : x = (x_1,x_2,x_3) \in \Omega \mapsto h(x) = (h_1(x),h_2(x),h_3(x) \in \Omega.$$

Within the framework of computational anatomy, the single most important component is the generation of large deformation diffeomorphisms. Given any two anatomical images I_1, I_2 assumed to be in the same homogeneous anatomy (see [2]) compute a diffeomorphism from one anatomy to the other: $I_1 \overset{h}{\longleftarrow\!\!\!\!\longrightarrow}_{\phi = h^{-1}} I_2$. In our work, not only are the global structures of *deep nuclei* important but as well we study the differential geometric features associated with the finest geometric structures including *sulcal trajectories and cortical folds*. Notions such as Riemannian length, Gaussian curvature, and surface area measures of highly complex folded structures are at the heart of our investigative work and others [13]. Methods which allow for the quantitative study of shape associated with Riemannian lengths, curvatures, and surface area measures are of crucial importance. It is therefore natural to organize the transformations around the continuum, emphasizing the properties of diffeomorphisms as they map the various tangent spaces and curvature features of the embedded submanifolds.

The approach is motivated by the image matching problem formulated via flows previously by Rabbitt and Christensen and posed as a control formulation in Dupuis et al. [12]. The basic protocol [5] followed for generation of the registration maps employs a hierarchy of transformations "increasing" the dimensionality successively via the fusion of landmark and image matching. Low-dimensional, nonrigid, coarse registration including the affine motions of global scale, rotation and translation, proceeds to high-dimensional, fine image matching registration resulting in the final maps [1]. One of the fundamental limitations of our previous application of the compositional approach is that the landmark matching was based on small deformation kinematics, implying it will not necessarily produce a diffeomorphism. Therefore, the ultimate goal of forming a group cascade as championed by Matejic [14] might not succeed as landmark matching will not in general provide diffeomorphisms. This is the principal focus of this paper, to extend

Manuscript received June 16, 1998; revised November 22, 1999. This work was supported by the NIH under Grants RO1-MH525158-01A1, RO1-NS35368-02, and NSF BIR-9 424264. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Glenn Healey.

S. C. Joshi is with the University of North Carolina, Chapel Hill, NC 27599 USA.

M. I. Miller is with the Center for Imaging Science, The Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: mim@cis.jhu.edu).

Publisher Item Identifier S 1057-7149(00)05337-9.

1057–7149$10.00 © 2000 IEEE

landmark matching to the large deformation setting insuring the generation of diffeomorphisms.

A. Large Deformation Landmark Matching Problem

Our approach is to construct diffeomorphisms $\phi : \Omega = [0,1]^3 \leftrightarrow \Omega$ in terms of the solutions to the ordinary differential equation (ODE) defined by the transport equation as first used in [7]

$$\frac{d\phi(x,t)}{dt} = v(\phi(x,t),t), \quad \phi(x,0) = x, \quad t \in [0,1]. \tag{2}$$

The final time diffeomorphism $\phi(\cdot,1)$ mapping the anatomy is therefore controlled via the velocity field $v(\cdot,t), t \in [0,1]$

$$\phi(x,1) = x + \int_0^1 v(\phi(x,s),s)\,ds, \quad x \in \Omega. \tag{3}$$

We assume the *targets* are characterized via sets of landmarked imagery $\{x_n, y_n, n = 1,2,\ldots,N\}$. A quadratic registration distance is defined between the measurements from the various target anatomies. The transformation is generated which minimize the distance while at the same time being a diffeomorphism.

Following the recent work of Dupuis *et al.* [12], we formulate the landmark matching problem as a control problem, with the optimal diffeomorphic match constructed to minimize a running smoothness cost on the velocity field generating the diffeomorphism while simultaneously minimizing the matching end point condition of the landmarks. If $v(x,t), x \in \Omega, t \in [0,1]$ is a sufficiently smooth vector field on $[0,1]$, then, by the existence and uniqueness theorem of ODE's [15], the solution exists and is uniquely determined by the velocity field $v(x,t)$ and the initial condition $\phi(x,0) = x$. Furthermore, it defines a unique diffeomorphism of $\Omega \leftrightarrow \Omega$ via the solution to the above ODE. By a continuously differentiable vector field, following [15] we mean that each of the coordinate functions $v_i(x,t), i = 1,2,3$ are continuously differentiable with respect to $x_i, i = 1,2,3$ and t.

Diffeomorphic landmark transformations are constructed by forcing the velocity fields to minimize quadratic energetics on $\Omega \times [0,1] = [0,1]^4 \subset I\!R^4$ defined via a matrix constant coefficient differential operator of the form

$$E(v) = \int_\Omega \int_{[0,1]} \|Lv(x,t)\|^2\,dx\,dt$$

where L is a constant-coefficient, differential operator. The matching problem for the N-landmarks $\{x_n, n = 1,\ldots,N\}$ becomes

$$\begin{aligned} &\hat{v}(x,t) = \arg\min_{v(x,t)} \int_0^1 \int_\Omega \|Lv(x,t)\|^2\,dx\,dt \\ &\text{subject to } \phi(x_n,1) = y_n \quad n = 1,\ldots,N \end{aligned} \tag{4}$$

where

$$\frac{d\phi(x,t)}{dt} = v(\phi(x,t),t), \quad \phi(x,0) = x.$$

The large deformation setting reduces in the small deformation setting to the small deformation landmark matching problem of Bookstein [16], [17] and others [18], [19]. Although these methods have proved to be very powerful in the study of biological shape [16]–[18], the deformations are not constrained to be *diffeomorphic* transformations as they are based on quadratic penalties derived from differential operators motivated by small deformation kinematics. They do not allow for large deformations that maintain the geometry and topology of the template [7]. To illustrate, assume the anatomies are deformed one to another according to $\phi(x,1) = x + u(x)$. This essentially removes the kinematic nonlinearity. Then the minimization of the small deformation transformation mapping the template to the target is chosen to minimize the thin-plate "bending energies" while constraining the transformation at the known landmark points. The energy minimization problem solved is of the form

$$\begin{aligned} &\hat{u} = \arg\min_u \int_\Omega \|Lu(x)\|^2\,dx \\ &\text{subject to } u(x_n) = y_n - x_n, \quad n = 1,\ldots,N. \end{aligned}$$

The basic difficulty for curved trajectories is illustrated in Fig. 1, in which the small deformation solution forces the grid lines to cross thus destroying the geometric properties of the maps.

II. Generating Large Deformation Diffeomorphisms

We force the transformations mapping the landmarks to be diffeomorphisms by associating with them an energy $\|Lv\|^2$ which gives them sufficient differentiability to insure the existence and uniqueness of the solution of the ODE. We have been using operators motivated by mechanics such as $L = -a\nabla^2 + b\nabla\nabla\cdot + c$. Throughout we shall assume the compact setting $\Omega = [0,1]^3$ with the operator and boundary conditions chosen so that the Green's function is nonsingular and continuous in both variables. For landmark matching we have used powers of the Laplacian with $a = 1, b = 0$ with zero boundary conditions which differentiates only in space, implying L is a diagonal operator giving energetics in the form

$$E(v) = \int_0^1 \int_\Omega \sum_{i=1}^3 |(-\nabla^2 + c)v_i(x,t)|^2\,dx\,dt.$$

Our results on generating diffeomorphisms apply for general constant coefficient matrix differential operators L as long as the 3×3 matrix Green's function $G(x,y)$ is continuous in both x and y and the 3×3 matrix function $K(x,y) = GG^\dagger(x,y)$ is positive definite as an operator. This is true for the diagonal Laplacian plus identity case

$$L = \begin{pmatrix} -\nabla^2 + c & 0 & 0 \\ 0 & -\nabla^2 + c & 0 \\ 0 & 0 & -\nabla^2 + c \end{pmatrix} \tag{5}$$

and is assumed throughout for the theorems.

Assuming that the velocity fields are Gaussian random fields induced by the differential operator L [5], [18], we define the spatial $3N \times 3N$ covariance matrix $K(\phi(t))$ according to

$$K(\phi(t)) = \underbrace{\begin{pmatrix} K(\phi(x_1,t),\phi(x_1,t)) & \cdots & K(\phi(x_1,t),\phi(x_N,t)) \\ \vdots & \vdots & \vdots \\ K(\phi(x_N,t),\phi(x_1,t)) & \cdots & K(\phi(x_N,t),\phi(x_N,t)) \end{pmatrix}}_{3N\times 3N} \tag{6}$$

with $(K(\phi(t))_{ij}$ denoting the $ij, 3\times 3$ block entry $(K(\phi(t))_{ij} = K(\phi(x_i,t),\phi(x_j,t))$.

A. Inexact Landmark Matching

Assume the measured landmark points are defined to within some covariance $\Sigma_n, n = 1,\ldots,N$. Then the matching problem has the following form.

Theorem 1: With the squared error distance for landmark matching given by

$$D(\phi(1)) = \sum_{n=1}^{N} [y_n - \phi(x_n,1)]^T \Sigma_n^{-1} [y_n - \phi(x_n,1)]$$

let L be a constant coefficient matrix differential operator with 3×3 matrix Green's function $G(x,y)$ continuous in both x and y and $K = GG^{\dagger}$ positive definite, on compact domain $\Omega = [0,1]^3$. Then the optimization

$$\hat{v}(\cdot) = \arg\min_{v} \int_0^1 \int_\Omega ||Lv(x,t)||^2 \, dx \, dt + D(\phi(1)) \tag{7}$$

where

$$\phi(x,1) = \int_0^1 v(\phi(x,t),t)\, dt + x$$

with the minimizer $\hat{v}$ defined by (7) is of the form

$$\hat{v}(x,t) = \sum_{i=1}^{N} K(\phi(x_i,t),x) \sum_{j=1}^{N} (K(\phi(t))^{-1})_{ij} \dot{\hat{\phi}}(x_j,t) \tag{8}$$

where

$$\begin{matrix} \dot{\hat{\phi}}(x_n,\cdot) \\ n = 1,\ldots,N \end{matrix} = \arg \min_{\substack{\dot{\phi}(x_n,\cdot) \\ n=1,\ldots,N}} \int_0^1 \sum_{ij} \dot{\phi}(x_i,t)(K(\phi(t))^{-1})_{ij} \times \dot{\phi}(x_j,t) + D(\phi(1)). \tag{9}$$

The minimizing velocity fields $\dot{\hat{\phi}}(x_n,\cdot)$ are continuous and $\hat{\phi}(x,1) = \int_0^1 \hat{v}(\hat{\phi}(x,t),t)\, dt + x$ is a diffeomorphism of $\Omega \leftrightarrow \Omega$.

Proof: The proof has two parts. The first demonstrates that the optimal velocity field is of the form given by (8). The second part demonstrates that the resulting optimization (9) is a classic nonlinear quadratic control problem in the Bolza form with the continuity properties. With the uniqueness of the solution of the ODE the diffeomorphism property results. See Appendix, Theorem I-A for proof.

Equation (9) demonstrates that the optimization is reduced to optimization of the N, landmark trajectories $\phi(x_n,\cdot), n = 1,\ldots,N$. The Euler–Lagrange conditions for these optimizers are given as follows. □

Theorem 2: The minimizer $\hat{\phi}(t)$ for the landmark matching problem corresponding to $\hat{\phi}(x_n,0) = x_n, n = 1,\ldots,N$ with

$$\begin{matrix} \dot{\hat{\phi}}(x_n,\cdot) \\ n = 1,\ldots,N \end{matrix} = \arg \min_{\substack{\dot{\phi}(x_n,\cdot) \\ n=1,\ldots,N}} \int_0^1 \sum_{ij} \dot{\phi}(x_i,t)^T \times (K(\phi(t))^{-1})_{ij} \dot{\phi}(x_j,t)\, dt + D(\phi(1)) \tag{10}$$

satisfies the Euler–Lagrange conditions for $l = 1,2,3, n = 1,\ldots,N$

$$0 = 3\sum_{j=1}^{N} \dot{\phi}(x_j,t)^T \left(\frac{\partial K(\phi(t))^{-1}}{\partial \phi_l(x_j,t)} \right)_{nj} \dot{\phi}(x_j,t) + \left(2\sum_{j=1}^{N} (K(\phi(t))^{-1})_{nj} \ddot{\phi}(x_j,t) \right)_l ; \tag{11}$$

$$0 = 2\sum_{j=1}^{N} (K(\phi(1))^{-1})_{nj} \dot{\phi}(x_j,1) + \Sigma_n^{-1}(\phi(x_n,1) - y_n). \tag{12}$$

Proof: See Appendix, Section V-B for the proof. □

Remark 1 (Relation to Eulerian Setting): In [7], [12], and [20], we have studied the quadratic image matching in the inverse Eulerian frame $h = \phi^{-1}$. Then

$$D(h(1)) = \sum_{n=1}^{N} [x_n - h(y_n,1)]^T \Sigma_n^{-1} [x_n - h(y_n,1)] \tag{13}$$

where

$$h(x,1) = \int_0^1 -\nabla h(x,t) v(x,t) + x.$$

Then the optimal diffeomorphism is given by the matching problem stated in Dupuis [12] *et al.*; for landmark matching it becomes

$$\hat{v}(x,t) = \arg\min_{v} \int_0^1 \int_\Omega ||Lv(x,t)||^2 \, dx \, dt + D(h(1))$$

with optimal diffeomorphism $\hat{h}(x,1) = \int_0^1 -\nabla \hat{h}(x,t)\hat{v}(x,t)\, dt + x$.

The power of working in the Lagrangian frame via (7)–(9) is that we have reduced the optimization from velocity fields $v(\cdot)$ on $\Omega \times [0,1]$ to N velocity fields $\dot{\phi}(x_n,\cdot), n = 1,\ldots,N$ on $[0,1]$.

B. Including the Affine Motions

We include the affine motions by allowing the target points to carry the affine motions

$$(A,a): y_n \in I\!R^3 \mapsto Ay_n + a \in R^3, \quad A \in \mathbf{GL}(3), \quad a \in R^3. \tag{14}$$

Then the distance function is straightforwardly modified to include the affine motions

$$D(\phi(1),A,a) = \sum_{n=1}^{N} [Ay_n + a - \phi(x_n,1)]^T \Sigma_n^{-1} [Ay_n + a - \phi(x_n,1)].$$

Then, Theorem 1 holds for the solution with the affine motion added to the estimation. Let L be the constant coefficient matrix differential operator as above with 3×3 matrix Green's function $G(x,y)$ continuous in both x and y and $K = GG^{\dagger}$ positive definite, on compact domain Ω. The optimizing diffeomorphism is given by

$$(\hat{v}(\cdot),\hat{A},\hat{a}) = \arg\min_{v,A,a} \int_0^1 \int_\Omega \|Lv(x,t)\|^2 \, dx \, dt + D(\phi(1),A,a) \tag{15}$$

where $\dot{\phi} = v(\phi(x,t),t)$, $\phi(x,0) = x$, with $\hat{v}(\cdot)$ satisfying (8) and

$$\dot{\hat{\phi}}(x_n,\cdot)_{n=1,\ldots,N}, \hat{A}, \hat{a} = \arg \min_{\dot{\phi}(x_n,\cdot)_{n=1,\ldots,N},A,a} \int_0^1 \sum_{ij} \dot{\phi}(x_i,t)(K(\phi(t))^{-1})_{ij}\dot{\phi}(x_j,t). + D(\phi(1),A,a). \tag{16}$$

C. Exact Landmark Matching

As we will be interested in generating flows that match exactly a given set of landmarks in the template and the target we now state the associated energy minimization problem. Notice that care must be taken as $\Sigma \to 0$. The problem of exact matching of a set of landmarks in the template and the target is formulated as follows.

Corollary 1: Given landmarks $\{x_n, n = 1,2,\ldots,N\} \subset \Omega = [0,1]^3$ that can be identified exactly with points $\{y_n, n = 1,\ldots,N\}$ in the data, with operator L as in theorem 1, the solution to the energy minimization problem

$$\hat{v}(\cdot) = \arg\min_v \int_0^1 \int_\Omega \|Lv(x,t)\|^2 \, dx \, dt \quad \text{subject to } \phi(x_n,1) = y_n, \quad n = 1,\ldots,N \tag{17}$$

where

$$\phi(x,t) = x + \int_0^T v(\phi(x,s),s)\, ds \tag{18}$$

exists and defines a diffeomorphism $\phi(\cdot,1)$ of $\Omega \leftrightarrow \Omega$. The optimum velocity field $\hat{v}$ and diffeomorphism $\hat{\phi}$ are given by

$$\hat{v}(x,t) = \sum_{i=1}^{N} K(\phi(x_i,t),x) \sum_{j=1}^{N} (K(\phi(t))^{-1})_{ij} \dot{\hat{\phi}}(x_j,t) \tag{19}$$

where

$$\hat{\phi}(x_n,\cdot)_{n=1,\ldots,N} = \arg \min_{\phi(x_n,\cdot)_{n=1,\ldots,N}} \int_0^1 \sum_{ij} \dot{\phi}(x_i,t)^T (K(\phi(t))^{-1})_{ij} \dot{\phi}(x_j,t)\, dt \quad \text{subject to } \phi(x_n,1) = y_n, \quad n = 1,\ldots,N \tag{20}$$

with the optimal diffeomorphism given by $\hat{\phi}(x,1) = \int_0^1 \hat{v}(\hat{\phi}(x,t),t)\, dt + x$.

Proof: As in Theorem 1, the above minimization problem of (17) is is equivalent to that of a finite dimensional optimal control with fixed end point conditions. Following the same reasoning as in theorem 1, it is equivalent to that of finding the optimum paths of the N landmarks points satisfying the minimization of (20). It satisfies all the conditions required for existence in theorem 1. However we must also prove that the set of exact landmark matches is not empty, i.e. that their exists a velocity field having finite cost which carries one set of landmarks to the other. This is proven in Theorem 3 in the Appendix V-C. The remainder of the proof is identical to theorem 1, ■

D. Implementation Algorithm

We now state the algorithm for the inexact landmark matching case (see remark below for exact landmark matching algorithm). The algorithm for landmark matching reduces the problem to a finite dimensional problem by defining the flows on the finite grid of fixed times of size δ, $t_k = k\delta$, $k = 0,1,\ldots,K = 1/\delta$. Assume velocities piecewise constant within the quantized time intervals, so that for $t \in [t_{k-1},t_k)$, $\dot{\phi}(x_n,t) = (\phi(x_n,t_k) - \phi(x_n,t_{k-1})/\delta)$. It will be helpful for the 3×1 vectors A to define the notation $(A)_l$ to mean the lth component of the vector A.

Algorithm 1: The finite dimensional minimization over $I\!R^{3NK}$ becomes $\hat{\phi}(x_n,0) = x_n$, $n = 1,\ldots,N$ and

$$\hat{\phi}(x_n,t_k)_{\substack{n=1,\ldots,N,\\ k=1,\ldots,K}} = \arg \min_{\substack{\phi(x_n,t_k)\\ n=1,\ldots,N\\ k=1,\ldots,K}} \frac{1}{\delta^2} \sum_{k=1}^{K} \sum_{ij=1}^{N} [\phi(x_i,t_k) - \phi(x_i,t_{k-1})]^T \times \left(\int_{t_{k-1}}^{t_k} (K(\phi(t))^{-1})_{ij}\, dt \right) [\phi(x_j,t_k) - \phi(x_j,t_{k-1})] + \sum_{n=1}^{N} [y_n - \phi(x_n,1)]^T \Sigma_n^{-1} [y_n - \phi(x_n,1)]. \tag{21}$$

The gradient algorithm for minimizing (21) initializes with $m = 0$ and $\phi^{(0)}(x_n, t_k) = x_n$, $n = 1, \ldots N$, $k = 1, \ldots, K$, and iterates for $m = 0, 1, \ldots$.

1) Calculate gradient perturbation for each

$$\phi(x_n, t_k) = \begin{pmatrix} \phi_1(x_n, t_k) \\ \phi_2(x_n, t_k) \\ \phi_3(x_n, t_k) \end{pmatrix}, n = 1, \ldots, N, \quad k = 1, \ldots, K$$

$$\vdots$$

$$\phi^{(m+1)}(x_n, t_k) = \begin{pmatrix} \phi_1^{(m)}(x_n, t_k) \\ \phi_2^{(m)}(x_n, t_k) \\ \phi_3^{(m)}(x_n, t_k) \end{pmatrix} - \Delta \begin{pmatrix} \frac{\partial}{\partial \phi_1(x_n,t_k)} P\left(\phi^{(m)}(1)\right) + \frac{\partial}{\partial \phi_1(x_n,t_k)} D\left(\phi^{(m)}(1)\right) \\ \frac{\partial}{\partial \phi_2(x_n,t_k)} P\left(\phi^{(m)}(1)\right) + \frac{\partial}{\partial \phi_2(x_n,t_k)} D\left(\phi^{(m)}(1)\right) \\ \frac{\partial}{\partial \phi_3(x_n,t_k)} (P\left(\phi^{(m)}(1)\right) + \frac{\partial}{\partial \phi_3(x_n,t_k)} D\left(\phi^{(m)}(1)\right) \end{pmatrix} \tag{22}$$

where for $l = 1, 2, 3$

$$\frac{\partial D(\phi(1))}{\partial \phi_l(x_n, t_k)} = \delta[t_k - 1] \left(2\Sigma_n^{-1}[y_n - \phi(x_n, 1)]\right)_l$$

$$\frac{\partial P^{(m)}(\phi)}{\partial \phi_l(x_n, t_k)} = 2\sum_{j=1}^{N} \left(\int_k^{t_{k+1}} (K(\phi(t))^{-1})_{nj}\, dt\, [\phi(x_j, t_k) - \phi(x_j, t_{k+1})] \right)_l + 2\sum_{j=1}^{N} \left(\int_{t_{k-1}}^{t_k} (K(\phi(t))^{-1})_{nj}\, dt\, [\phi(x_j, t_k) - \phi(x_j, t_{k-1})] \right)_l + \sum_{j=1}^{N} (\phi(x_j, t_{k+1}) - \phi(x_j, t_k))^T \times \frac{\partial \int_{t_k}^{t_{k+1}} (K(\phi(t))^{-1})_{nj}\, dt}{\partial \phi_l(x_j, t_k)} \times [\phi(x_j, t_{k+1}) - \phi(x_j, t_k)]$$

where

$$\frac{\partial \int_{t_k}^{t_{k+1}} (K(\phi(t))^{-1})_{nj}\, dt}{\partial \phi_l(x_j, t_k)} = \int_{t_k}^{t_{k+1}} \left(K(\phi(t))^{-1} \frac{\partial K(\phi(t))}{\partial \phi_l(x_j, t_k)} K(\phi(t))^{-1} \right)_{nj} dt \tag{23}$$

and $\delta[t_k - 1] = 1$ for $t_k = 1$, and 0 otherwise.

After stopping, define the final iterate as $\hat{\phi} = \phi^{(m+1)}$, and

$$\dot{\hat{\phi}}(x_n, t) = \frac{\hat{\phi}^{(m+1)}(x_n, t_k) - \hat{\phi}^{(m+1)}(x_n, t_{k-1})}{\delta}, \quad t \in [t_{k-1}, t_k), \quad k = 1, \ldots, K, \tag{24}$$

with

$$\hat{v}(x, t) = \sum_{n=1}^{N} K(\hat{\phi}(x_n, t), x) \sum_{j=1}^{N} (K(\hat{\phi}(t))^{-1})_{nj} \dot{\hat{\phi}}(x_j, t), \tag{25}$$

$$\hat{\phi}(x, 1) = \int_0^1 \hat{v}(\hat{\phi}(x, t), t)\, dt + x, \quad \text{for all } x \in \Omega.$$

For including the affine motions, within each step of the gradient on the velocity field we fix the diffeomorphism from the previous iteration $\phi^{(m)}(\cdot)$ and solve the linear landmark matching problem

$$\hat{A}^{(m+1)}, \hat{a}^{(m+1)} = \arg\min_{A,a} \sum_{n=1}^{N} \left[Ay_n + a - \phi^{(m)}(x_n, 1) \right]^T \times \Sigma_n^{-1} \left[Ay_n + a - \phi^{(m)}(x_n, 1) \right] \tag{26}$$

giving

$$\hat{A}^{(m+1)} = \left(\sum_{n=1}^{N} y_n y_n^T \Sigma_n^{-1} \right)^{-1} \times \left(\sum_{n=1}^{N} y_n \left(\phi^{(m)}(x_n, 1) - \hat{a}^{(m+1)} \right)^T \right) \tag{27}$$

$$\hat{a}^{(m+1)} = \left(\sum_{n=1}^{N} \Sigma_n^{-1} \right)^{-1} \times \left(\sum_{n=1}^{N} \Sigma_n^{-1} \hat{A}^{(m+1)} y_n - \phi^{(m)}(x_n, 1) \right). \tag{28}$$

These are linear equations for fixed $\phi^{(m)}$ which are solved via matrix inversion.

Remark 2: For choosing initial conditions for the inexact landmark matching the identity map given by $v(\cdot) = 0$ is used for the initial condition. For exact landmark matching an initial condition is generated from the inexact landmark matching solution following an approach suggested by Younes [21]. Construct an inexact landmark match to within an ϵ-ball of the target landmarks. The initial condition for exact matching is generated by linearly interpolating the inexact landmark match onto the target points. This insures an initial condition which does not cross and maps the landmarks exactly.

III. Results

For speed of implementation we have implemented the algorithm corresponding to the infinite domain so that the Green's functions and resulting covariance $K(\cdot, \cdot)$ have analytic closed form expressions. For the Laplacian operator $L = \nabla^2$ this gives the Green's functions and covariances of Kent [18], of the form $K(x, y) = \|x - y\|, x, y \in I\!R^3$.

We now derive the covariance induced by the self-adjoint operator $L = -\nabla^2 + c$ with infinite boundary conditions. As the covariance operator $K(x, y)$ is given by the Green's function for the operator L^2 which is shift invariant, we use the Fourier transform method for computing the Green's function at the point

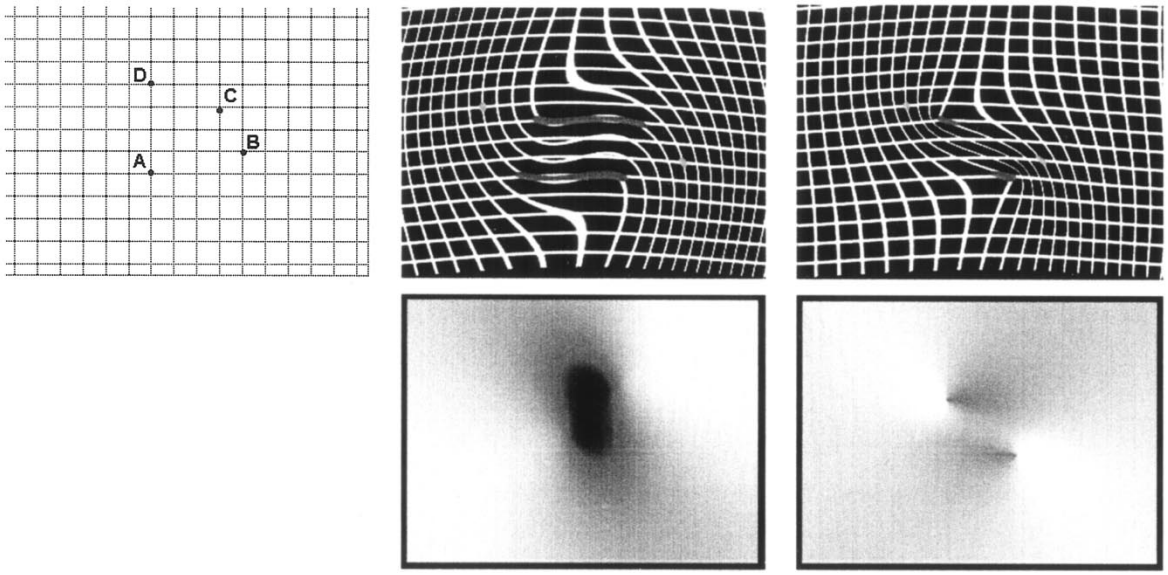

Fig. 1. Top row: Left panel shows phantom of two landmark point matching, A matching to B, and C matching to D, with fixed corners used for the first experiment. Middle and right panels compare the small deformation thin plate landmark deformation (middle panel) and the large deformation solution (right panel). Bottom row: Left panel shows the determinant of the Jacobian of the thin plate landmark deformation. The right panel shows the determinant of the Jacobian associated with the the *diffeomorphic* transformation. Notice that the determinant is negative in the region where the grid lines in the thin plate landmark deformation cross over while the determinant of the *diffeomorphic* transformation is strictly positive everywhere.

$y = 0$ with the shift invariance used to deduce it every where. In the Fourier domain the spectrum of the the operator L^2 centered at the origin is given by $(||\omega||^2 + c)^2$. Using rotational invariance and spherical coordinates gives the space domain representation via the Fourier integral

$$K(x,0) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{(||\omega||^2+c)^2}e^{-j2\pi\langle\omega,x\rangle}\,d\omega \tag{29}$$

$$= \frac{2}{\sqrt{c(2\pi)}}e^{-\sqrt{c}||x||}. \tag{30}$$

This gives the entire kernel according to $K(x,y) = (2/\sqrt{c(2\pi)})e^{-\sqrt{c}||x-y||}$.

A. Small Deformation Solution

Large deformations present a significant challenge to small deformation landmark matching such as in Bookstein [16], [17] or Kent and Mardia [18], or Wahba [19] on spline methods. Although these methods have proved to be very powerful in the study of brain structures [16]–[18], the deformations are not constrained to be *diffeomorphic* transformations. For small deformations, assume the anatomies are deformed one to another according to $\phi(x,1) = x + u(x)$. Then the minimization of the small deformation transformation mapping the template to the target is chosen to minimize the thin-plate "bending energies" while constraining the transformation at known landmark points $x_n, n = 1,\ldots,N$. We include the affine motions by allowing the target points to carry the affine motions

$$(A,a) : y_n \in I\!R^3 \mapsto Ay_n + a \in R^3,$$
$$A \in \mathbf{GL}(3), \quad a \in R^3. \tag{31}$$

Following Joshi [5], [22] the small deformation cost is minimized according to

$$\hat{u},\hat{A},\hat{a} = \arg\min_{u,A,a}\int_{\Omega}||Lu(x)||^2\,dx + \sum_{n=1}^{N}[Ay_n + a - (x_n + u(x_n))]^T \times \Sigma_n^{-1}[Ay_n + a - (x_n + u(x_n))]. \tag{32}$$

This gives the optimal small deformation shown in (33) and (34) at the bottom of the next page.

The large deformation landmark matching algorithm was implemented in 2-D initially. Fig. 1 shows a set of results based on a simple test pattern containing 2 matching points with fixed corners requiring a nonlinear twisting motion for the matching. The left panel of Fig. 1 shows the test grid pattern containing four points A, B, C, and D which where chosen as landmark points on the grid. Point A was mapped to point B and point C was mapped to point D while the four corners were mapped onto themselves.

Shown in the top row, middle panel of Fig. 1 is the small deformation solution of equations (33) and (34). Notice how the grid lines cross in generating the mapping.

The large deformation matching algorithm 1 was implemented for the 2-D grid of points. Shown in the top right panel of Fig. 1 is the result of matching A to B and C to D. Notice that no grid lines cross. The bottom row shows the determinant of the Jacobian of the small deformation (left panel) with the *diffeomorphic* transformation (right panel). The color scale shows black where the determinant of the Jacobian is negative, and shows white where the Jacobian has positive determinant. Notice that the determinant (left panel) flips sign and is negative in the region where the grid lines in the small deformation

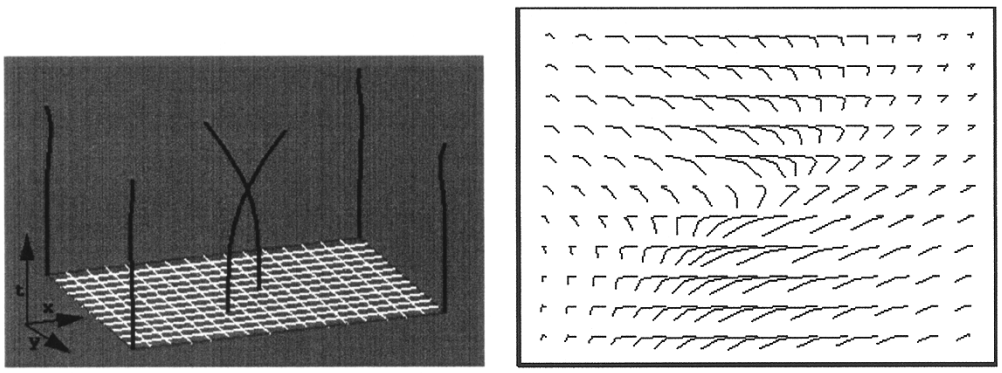

Fig. 2. Left panel shows the optimum paths $\hat{\phi}(x_n, t), n = 1, 2$ traced out by the landmark points A, C, and the four corners of the image. Right panel shows paths $\hat{\phi}(x, t)$ traced out by the grid points under the optimal velocity field $\hat{v}(x, t)$.

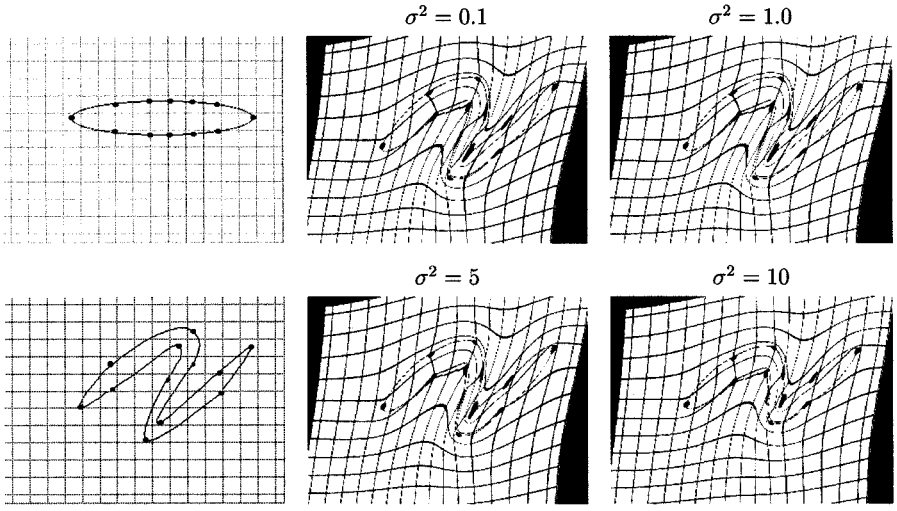

Fig. 3. Left column shows the two test patterns. Columns 2 and 3 show results of the deformation process using variances of $\sigma^2 = 0.1, 1, 5$, and 10 respectively.

landmark deformation crosses over; the determinant of the *diffeomorphic* transformation is strictly positive everywhere.

To illustrate the complete flow of points associated with the diffeomorphic transformation, the left panel of Fig. 2 shows the optimum paths traced out by the landmark points A, C and the four corners. The right panel of Fig. 2 shows the paths traced out by all the grid points under the optimum velocity field $\hat{v}(x,t), x \in [0,1]^2, t \in [0,1]$. Notice again that the mapping is one-to-one as none of the paths cross each other.

Shown in Fig. 3 are results from two other test patterns, the "OVAL" and "S" (left column). The corresponding points shown in the figure were used as landmark points and the covariances $\Sigma_n, n = 1, \ldots, N$ associated with them was successively varied. We chose $\Sigma_n = \sigma^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ using variances of

$$\hat{u}(x) = \sum_{n=1}^{N} K(x_n, x)\beta_n$$
$$= \sum_{n=1}^{N} \underbrace{\frac{2}{\sqrt{c2\pi}} \begin{pmatrix} e^{-\sqrt{c}\|x_n - x\|} & 0 & 0 \\ 0 & e^{-\sqrt{c}\|x_n - x\|} & 0 \\ 0 & 0 & e^{-\sqrt{c}\|x_n - x\|} \end{pmatrix}}_{K(x_n, x)} \underbrace{\begin{pmatrix} \hat{\beta}_{n1} \\ \hat{\beta}_{n2} \\ \hat{\beta}_{n3} \end{pmatrix}}_{\beta_n} \quad \text{with} \tag{33}$$

$$\begin{pmatrix} \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_N \end{pmatrix} = \begin{pmatrix} (K(x_1, x_1) + \Sigma_1 & \cdots & K(x_1, x_N) \\ & \vdots & \\ K(x_N, x_1) & \cdots & K(x_N, x_N) + \Sigma_N \end{pmatrix}^{-1} \begin{pmatrix} x_1 - (\hat{A}y_1 + \hat{a}) \\ \vdots \\ x_N - (\hat{A}y_N + \hat{a}) \end{pmatrix}$$
$$0 = \sum_n \beta_n, \quad 0 = \sum_n \beta_n y_n^T. \tag{34}$$

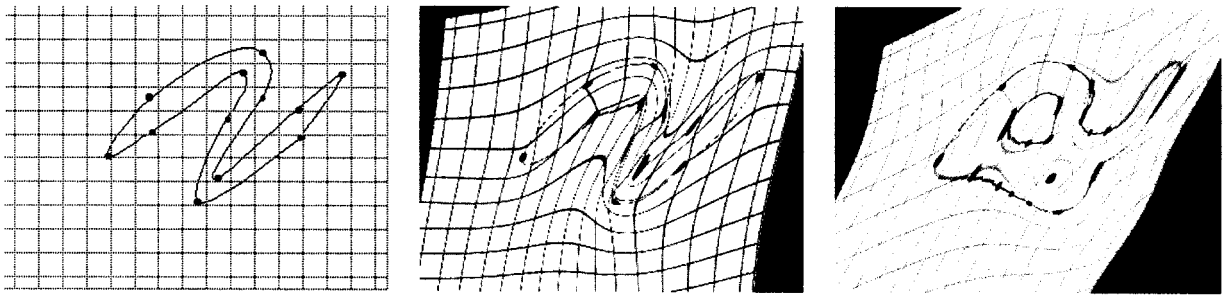

Fig. 4. Left panel shows the target oval test pattern. Middle panel shows the $\sigma^2 = 0.1$ large deformation solution. Right panel shows the small deformation solution with $\sigma^2 = 0.1$.

$\sigma^2 = 0.1, 1, 5, 10$. Shown in the middle and right columns of Fig. 3 are the results of the large deformation process. Notice that as the variance associated with the landmark identification process is increased the OVAL does not deform exactly to the target pattern.

Fig. 4 shows a comparison of the large deformation solution (middle panel) compared to the small deformation solution (right panel). Because the small deformation landmark matching does not generate a diffeomorphism, it creates several image landmarks in the mapped target.

B. Application to the Study of the Cortical Cartography

The mammalian cerebral cortex has the form of a layered, highly convoluted thin shell of grey matter surrounding white matter, and is one of the most striking features of the brain. The cortex contains a complex mosaic of anatomically and functionally distinct areas which play a tremendously important role in understanding brain functions [23]. As championed by Van Essen *et al.* [24], to aid in the understanding of the geometry and the functional topography of the cortex the convoluted cerebral cortex is mapped to a plane to generate a cortical flat map. The cortical flat map, although inherently induces distortions, allows for convenient visualization.

To understand individual variation in the cortical topography the Van Essen group has been using the large deformation tools to establish correspondences between the flat maps of various individual cortical surfaces. Shown in Fig. 5 are the geometrical features associated with the two flat maps of the cortical surfaces depicting the cortical geometry and the associated partitioning schemes by Brodmann on the left, and Fellman and Van Essen on the right. Notice the large variation in the geometry and the shape of the cortical surfaces as depicted by the difference in the positions of the geometrical features associated with the deep folds of the sulci and the fundi that were identified on both the flat maps. These features are used as the landmarks for driving the deformation algorithms herein assumed with known predefined correspondence (see [25] and [26] for automated sulcus generation and matching based on the Frenet frame supporting automated landmark correspondence generation). Shown in Fig. 6 is the result of the deformation process. For this $\Omega = [0,1]^2$ with $L = \begin{pmatrix} -\nabla^2 + c & 0 \\ 0 & -\nabla^2 + c \end{pmatrix}$, a 2×2 matrix operator defined on $\mathbb{R}^2$ valued vector fields deforming in the unit-square $[0,1]^2$. The left panel of Fig. 6 shows the deformed flat map corresponding to the partitioning schemes by Brodmann. Shown in the right panel is the overlay of the deformed partitioning schemes by Brodmann on the partitioning scheme by Felleman and Van Essen.

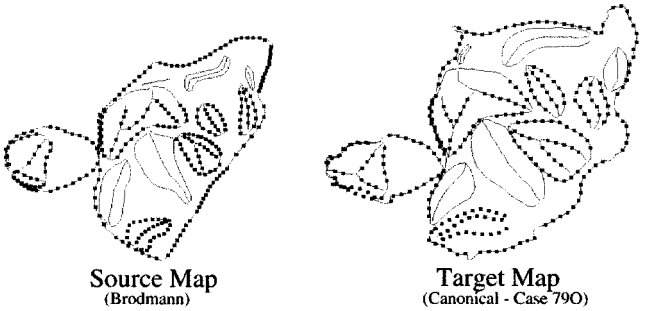

Fig. 5. Figure shows the geometrical features such as the sulci and the fundi that were identified on the flat maps and used as landmarks for the deformation process. Data taken from the laboratory of Dr. D. Van Essen.

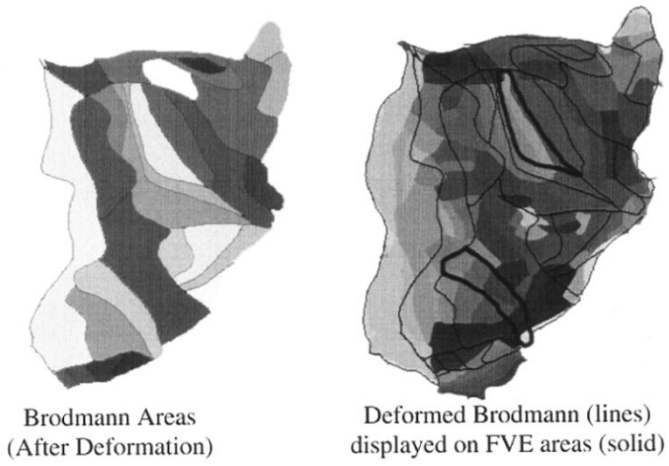

Fig. 6. Shown on the left is the deformed flat map corresponding to the partitioning schemes by Brodman. Shown on the right the overlay of the deformed partitioning schemes by Brodman on the partitioning scheme by Fellman and Van Essen. Data taken from the laboratory of Dr. D. Van Essen.

C. Three-Dimensional Macaque and Human Hippocampus Results

Examine the whole macaque cryosection brains shown in Fig. 7 in which the gyri and associated sulci have been labeled. The sulci and gyri are defined precisely in terms of the geometrical properties of the cortical surface using the notions of ridge curves and crest lines (extremal points of curvature) following

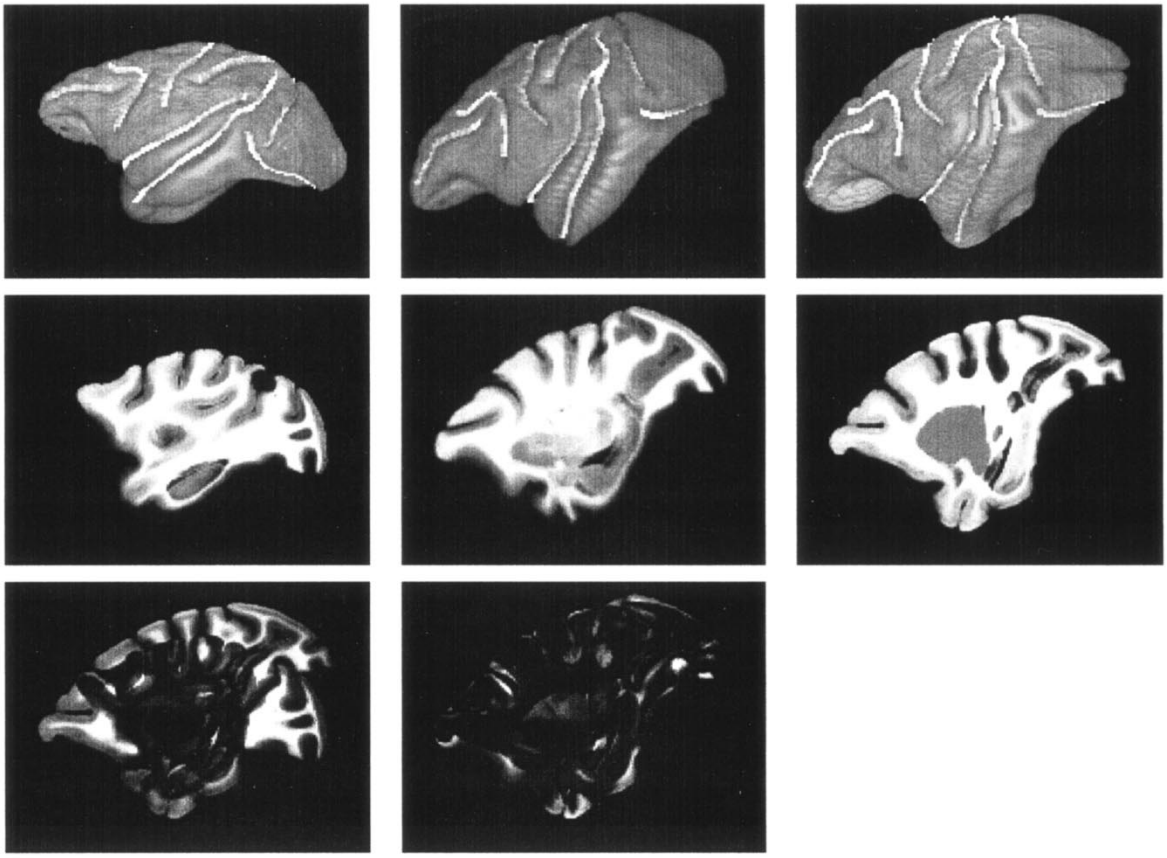

Fig. 7. Top row shows the volume rendering of the template 87A (left panel) and the template mapped to the target 90C (middle) using only the sulcal line constraints to define the transformation. The right panel shows the target brain 90C. Middle row shows corresponding sections through the template 87A (left) the target 90C (right) and the deformed template (middle). Bottom row shows in the squared error between the respective sections and the target. Data taken from the laboratory of Dr. D. Van Essen.

[27]. Using the dynamic programming algorithm for tracking geometrically significant contours on the brain developed in [27] the gyri and associated sulci were labeled in several whole brains. The sulcal maps constrain the transformation from one brain to the other. The top row of Fig. 7 shows fundus curves of the major sulci which have been identified and placed into the whole brains. The left and right panels show brains 87A and 90C with the fundus curves placed on the map. The deformation field was constrained so that the corresponding points were mapped on to each other in $I\!R^3$.

Our group has extensively studied the mapping of the hippocampus in MRI of human brains [8], [9]. All of this previous work has relied on the use of small deformation landmark matching, assuming that template and target landmarks are similar. Fig. 7 shows large deformation landmark maps of the macaque. The top panels show the volume rendering of the template 87A (left panel) and the template mapped to the target 90C using only the sulcal curve constraints to define the transformation. The right panel shows the target brain 90C. The middle row of Fig. 7 shows corresponding sections through the template 87A (left) the target 90C (right) and the deformed template (middle). Shown in the bottom row is the squared error between the respective sections and the target. Notice that there is a large difference in the shape and positions of the major subvolumes (the thalamus and the cortical folds) between the undeformed template and the target. While there has been an alignment of the major subvolumes in the deformed template with the target, the strong differences that still remain are associated with the fact that a relatively small number of landmarks are being used. To generate more complete matches the landmark matching must be coupled to the image matching as described in [1], [28].

Fig. 8 shows large deformation landmark maps of the hippocampus studied via MRI. Shown in Fig. 8 are results on the landmark matches of the hippocampus compared to hand segmentations generated by Dr. J. Haller and Dr. L. Wang at Washington University.

IV. Conclusion

This paper describes the generation of large deformation diffeomorphisms $\phi : \Omega = [0,1]^3 \leftrightarrow \Omega$. Both inexact and exact landmark matching is studied here. Given noisy landmarks x_n matched to y_n measured with error covariances Σ_n, then the matching problem is solved generating the optimal diffeomorphism $\hat{\phi}(x,1) = \int_0^1 \hat{v}(\hat{\phi}(x,t),t)\,dt + x$ where

$$\hat{v}(\cdot) = \arg\min_{v(\cdot)} \int_0^1 \int_\Omega ||Lv(x,t)||^2\,dx\,dt + \sum_{n=1}^{N} [y_n - \phi(x_n,1)]^T \Sigma_n^{-1} [y_n - \phi(x_n,1)]. \quad (35)$$

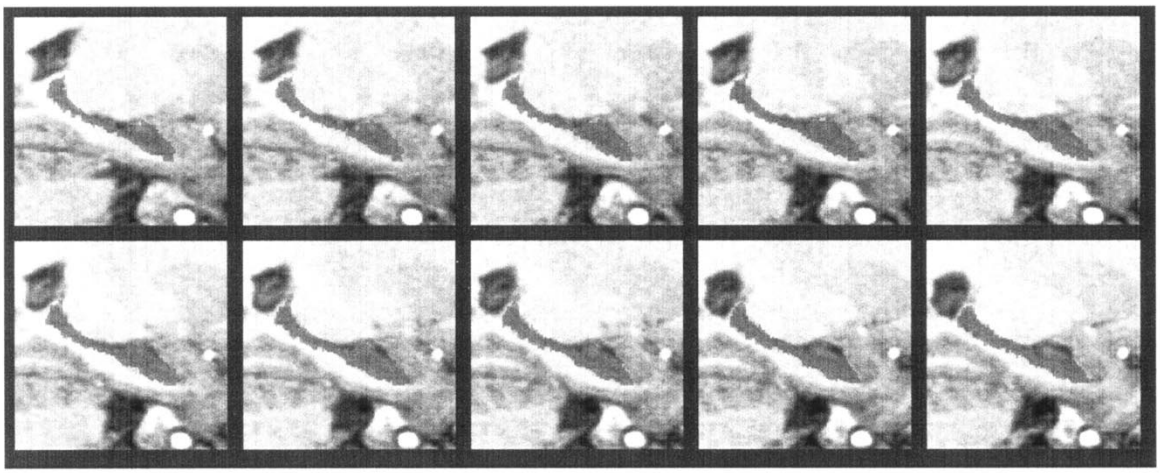

Fig. 8. Figure showing matching of the hippocampus based on landmarks showing the comparison of hand segmentations and automated landmark based matches.

Conditions for the existence of solutions in the space of diffeomorphisms are established, with a gradient algorithm provided for generating the optimal flow solving the minimum problem.

Results demonstrate that large deformation solutions are appropriate in the setting that the anatomical maps are required to follow curved trajectories. This corresponds to the kinematic nonlinearity introduced fundamentally via the transport equation. As shown for large curved deformations, the small deformation solutions are inadmissible as they result in maps which deform the lattice in nonphysical ways. Inconsistencies result corresponding to noninvertibility of the maps.

Results on matching 2-D and 3-D imagery are presented in the macaque monkey. We have found that for the understanding of the curved geometries associated with cortical folding [29] small deformation mapping cannot work. On the other hand, results are provided within demonstrating that by introducing the kinematic nonlinearity matches can be generated consistently on cortical folded surfaces. Perhaps the most fundamental limitation of the work is the assumption of known correspondence between curvilinear landmarks such as defined by sulcus and gyrus curves. In [25] and [26], automated methods for sulcus and gyrus curve generation are described; as well diffeomorphic curve matching in $\mathbb{R}^3$ is described based on the Frenet frame. This approach is similar to the geometry based curve matching proposed in [30] and [31]. As demonstrated in [26] this allows for the matching across cortical surfaces based on automatic correspondence generation between geometric features such as the deepest cortical folds.

Appendix A
Proof of Landmark Minimization Theorem

Proof of Theorem 1: Let us prove the optimizing v field of (7) is of the form

$$\hat{v}(x) = \sum_{n=1}^{N} K(\phi(x_n,t),x) \underbrace{\sum_{j=1}^{N} (K(\phi(t))^{-1})_{nj} \dot{\hat{\phi}}(x_j,t)}_{\beta_n(t)}. \tag{36}$$

Choose $v \in \mathcal{H}$ the Hilbert space with finite energy $\|Lv\|^2 < \infty$ defined in terms of a second velocity field $\tilde{v}$ according to

$$\begin{aligned} v(x,t) &= \tilde{v}(x,t) + \psi(x,t) \\ &= \underbrace{\sum_{n=1}^{N} K(\phi(x_n,t),x)\beta_n(t)}_{\tilde{v}(x,t)} + \psi(x,t) \end{aligned} \tag{37}$$

with the choice of $\beta_n(\cdot)$ chosen to satisfy

$$v(\phi(x_n,t),t) = \sum_{j=1}^{N} K(\phi(x_j,t),\phi(x_n,t))\beta_j(t), \quad n = 1,\ldots,N \tag{38}$$

implying $\psi(\phi(x_n,t),t)) = 0, n = 1,\ldots,N$. That (38) can be made true follows from the fact that L is chosen so that it is positive definite with Green's function that has finite energy, so that $\|L\tilde{v}(\cdot,t)\|^2 < \infty$ for each t, and $K = GG^{\dagger}$ is invertible (full rank covariance). Then $\tilde{v} \in \mathcal{H}$ since $\|L\tilde{v}\|^2 < \infty$ following from the fact that

$$\begin{aligned} \|L\tilde{v}(\cdot,t)\|^2 &= \int_{\Omega} \left\| L \sum_{n=1}^{N} K(\phi(x_n,t),x)\beta_n(t) \right\|^2 dx \\ &= \int_{\Omega} \left\| \sum_{n=1}^{N} G(\phi(x_n,t),x)\beta_n(t) \right\|^2 dx \\ &\le \sum_{n=1}^{N} \sum_{k=1}^{3} \int_{\Omega} \|G_k(\phi(x_n,t),x)\beta_n(t)\|^2 \, dx \\ &< \infty \end{aligned}$$

if $\|G_k\|^2 < \infty$ for each of the matrix Green's function components. This is true for the diagonal Laplacian case.

Hence, it is sufficient to show that $\|Lv\|^2 \ge \|L\tilde{v}\|^2$ (see (39)–(40) at the bottom of the next page),with (a) following since $\mathrm{TERM1} = 0$ implied by the facts that $\psi(\phi(x_n,t),t) = 0$ for $n = 1,\ldots,N$ and $K = GG^{\dagger}$ giving $L^{\dagger}LK(\phi(x_n,t),x) = \delta(\phi(x_n,t) - x)$.

The optimal diffeomorphism satisfying $\phi(x,1) = \int_0^1 v(\phi(x,\sigma),\sigma)\,d\sigma + x$ for the matching problem becomes (41)–(42), shown at the bottom of the next page. Then,

choosing $\beta_n(\cdot)$ from (38) and taking the inverse covariance and substituting for $\beta(\cdot)$ gives the equivalent optimization stated in terms of the N-Lagrangian landmark flow velocities according to

$$\hat{v}(x,t) = \sum_{n=1}^{N} K(\phi(x_n,t),x) \sum_{j=1}^{N} (K(\phi(t))^{-1})_{nj} \dot{\hat{\phi}}(x_j,t) \tag{43}$$

where

$$\begin{aligned} &\dot{\hat{\phi}}(x_n,\cdot) \\ &n = 1,\ldots,N \\ &\equiv \arg \min_{\substack{\dot{\phi}(x_n,\cdot) \\ n=1,\ldots,N}} \int_0^1 \sum_{ij} \dot{\phi}(x_i,t)(K(\phi(t))^{-1})_{ij} \dot{\phi}(x_j,t) \\ &\quad + D(\phi(1)). \end{aligned} \tag{44}$$

This completes the first part of the proof.

To establish continuity of the minimizer, and the diffeomorphism property we notice that this is a classic quadratic control problem in Bolza form ([32, ch. 3]), which we prove via the following lemma. □

Lemma 1 (Bolza Lemma): There exists continuously differentiable set of optimal paths $\phi(x_n,\cdot), n = 1,\ldots,N$ satisfying $\phi(x_n,t) = \int_0^T v(\phi(x_n,s),s)\,ds + x_n$ minimizing

$$\begin{aligned} &\dot{\hat{\phi}}(x_n,\cdot) \\ &n = 1,\ldots,N \\ &\equiv \arg \min_{\substack{\dot{\phi}(x_n,\cdot) \\ n=1,\ldots,N}} \int_0^1 \underbrace{\sum_{ij} \dot{\phi}(x_i,t)^T (K(\phi(t))^{-1})_{ij} \dot{\phi}(x_j,t)}_{H(\phi,\dot{\phi},t)}\, dt \\ &\quad + \underbrace{\sum_{n=1}^{N} [y_n - \phi(x_n,1)]^T \Sigma_n^{-1} [y_n - \phi(x_n,1)]}_{D(\phi(1))}. \end{aligned}$$

Proof: Let $\phi(t) = [\phi(x_1,t),\ldots,\phi(x_N,t)]^T \in \mathbb{R}^{3N}$, with velocity $\dot{\phi}(t) \in \mathbb{R}^{3N}$. The canonical *Bolza* form becomes

$$\hat{\dot{\phi}} = \arg \min_{\{\dot{\phi}\}} J(\dot{\phi}) = \arg \min \int_0^1 H(\phi,\dot{\phi},t)\,dt + D(\phi(1)). \tag{45}$$

We now use the existence of optimal control to prove the existence of a minimum to the optimization of (45). For this, define the following sets. Let $\mathcal{F}$ be the set of all Lebesgue-integrable functions $\dot{\phi}(\cdot)$ with values in $\mathbb{R}^{3N}$ and $J(\dot{\phi}) < \infty$. By the existence theorem of optimal control (see [32, ch. 3]) if 1) $\mathcal{F}$ is not empty, 2) Ω is compact and D is continuous on Ω, 3) $H(\cdot,\dot{\phi})$ is convex in the control variable $\dot{\phi}$, and 4) $H(\phi,\dot{\phi},t)) \geq c_1 \|\dot{\phi}(t)\|^\beta, c_1 > 0, \beta > 1$, then there exists $\hat{\dot{\phi}}(t)$ minimizing $J(\dot{\phi})$ on $\mathcal{F}$. If in addition H is of class $C^r, r \geq 2$ and the matrix $H_{\dot{\phi}\dot{\phi}}(\phi(t),\dot{\phi}(t),t)$ is positive definite for all $\dot{\phi}$ then $\hat{\dot{\phi}}(t)$ is of class C^{r-1} and $\hat{\phi}(t)$ is of class C^r on $[0,1]$. The above minimization of (45) satisfies all the above conditions as follows.

1) $\mathcal{F}$ is not empty as the identity mapping with the trivial control $\dot{\phi}(t) = 0$ is in $\mathcal{F}$ as $J(0) = \sum_{n=1}^{N} [x_n - y_n]^T \Sigma_n^{-1} [x_n - y_n] < \infty$.
2) $\Omega = [0,1]^3$ is compact and D is continuous as it is a quadratic function of the endpoint $\phi(1)$.
3) $K(\phi(t))$ is a positive definite covariance matrix and $H_{\dot{\phi}\dot{\phi}}(\phi(t),\cdot t) = 2K^{-1}(\phi(t))$ is positive definite, implying that H is a convex function in the control variable $\dot{\phi}$.
4) We now prove that $H(\phi,\dot{\phi},t) \geq c_1 \|\dot{\phi}(t)\|^2$. As $K^{-1}(\phi(t))$ is symmetric positive definite

$$\begin{aligned} H(\phi,\dot{\phi},t) &= \dot{\phi}(t)^T K^{-1}(\phi(t)) \dot{\phi}(t)^T \\ &\geq \frac{1}{\|K(\phi(t))\|_s} \|\dot{\phi}(t)\|^2 \\ &\geq \frac{1}{\text{Trace}(K(\phi(t)))} \|\dot{\phi}(t)\|^2 \end{aligned}$$

$$\begin{aligned} \|Lv\|^2 &= \|L(\tilde{v}+\psi)\|^2 \\ &= \int_0^1 \int_\Omega \left\| L\left[\sum_{n=1}^{N} K(\phi(x_n,t),x)\beta_n(t) \right] \right\|^2 dx\,dt + \|L\psi\|^2 + \underbrace{2 \int_0^1 \sum_{n=1}^{N} \beta_n(t)^T \int_\Omega LK(\phi(x_n,t),x) L^\dagger \psi(x,t)\,dx\,dt}_{\text{TERM1}} \end{aligned} \tag{39}$$

$$\stackrel{(a)}{=} \|L\tilde{v}\|^2 + \|L\psi\|^2 \tag{40}$$

$$\hat{v} = \arg \min_v \int_0^1 \int_\Omega \|Lv(x,t)\|^2\,dx\,dt + D(\phi(1)), \tag{41}$$

$$= \arg \min_{\beta(\cdot): v(\cdot) = \sum_{i=1}^{N} K(\phi(x_i,t),\phi(\cdot,t))\beta_i(t)} \int_0^1 \sum_{ij} \beta_i^T(t) K(\phi(x_i,t),\phi(x_j,t))\beta_j(t) + D(\phi(1)). \tag{42}$$

where $\|\cdot\|_s$ is the operator norm of the matrix K. Since the Green's operator G is assumed to be continuous and Ω is compact $K(x,y)$ is bounded and hence $\mathrm{Trace}(K(\phi(t)) < \infty$ implying that $H(\phi,\dot\phi,t) \geq c_1\|\dot\phi(t)\|^2$. Furthermore, $K(\phi(t))$ is of class C^2 as the Green's operator of L is assumed to be continuous, implying that $H(\phi,\dot\phi,t)$ is of class C^2.

This implies that there exists optimal paths $\hat\phi(t)$ of class C^2 and $\hat{\dot\phi}(t) = \dot{\hat\phi}(t)$ of class C^1 minimizing (45). Lemma 1 ■

With the existence and differentiability of the solution of (45) we see that the velocity field given by $\hat v(x,t) = \sum_{j=1}^N K(\hat\phi(x_j,t),x)\hat\beta_j(t)$ minimizes (7). Furthermore as $\hat\beta_j(t)$ satisfy the system of linear equations

$$\frac{d}{dt}\hat\phi(x_n,t) = \sum_{j=1}^N K(\phi(x_n,t),\phi(x_j,t))\hat\beta_j(t)$$

and as $\dot{\hat\phi}(t)$ is of class C^1, it implies that $\dot{\hat\phi}(x,t)$ is of class C^1 and hence by Lemma 2 below defines a diffeomorphism $\hat\phi(x,t)$.

Lemma 2: Let Ω be $[0,1]^3$ subset of $\mathbb{R}^3$. Let $v : (x,t) \in \Omega\times[0,1] \to \dot v(x,t) \in \mathbb{R}^3$ be a continuously differentiable vector field with support, $\mathrm{supp}(v(x,t))$, contained in Ω for each t. Let ϕ be a solution to the system of ordinary differential equations

$$\frac{d\phi(x,t)}{dt} = v(\phi(t,x),t)$$

with initial condition $\phi(x,0) = x$. Then for each t, $\phi(t,\cdot)$ is a diffeomorphism of $\Omega \leftrightarrow \Omega$.

Proof: As the vector field $v(x,t)$ has compact support it is a *complete* vector field and hence has a unique solution $\phi(x,t)$ for all t (see [15]). We prove that $\phi(x,t)$ is a diffeomorphism by showing that for each t

1) $\phi(t,\cdot)$ is one-to-one and onto;
2) Both $\phi(t,\cdot)$ and $\phi^{-1}(t,\cdot)$ are differentiable.

For all $x \in \partial\Omega$, as $\mathrm{supp}(v) \subset \Omega$, $v(x,t) = 0$, it is immediate that for all t, $\phi(x,t) = x$. For a fixed t, consider the velocity field $\tilde v(x,s) = -v(x,t-s)$ with solution $\tilde\phi(x,t)$ to the ODE $d\tilde\phi(x,s)/ds = \tilde v(\tilde\phi(x,s),s)$. By the uniqueness and existence theorem of ordinary differential equations it is immediate that for all $x \in \Omega$, $\phi(t,\tilde\phi(x,t)) = x$ and $\tilde\phi(\phi(x,t),t) = x$. This implies that for all $x \in \Omega$, $\phi(x,t) \in \Omega$ as if for some $t, \phi(x,t) = y \notin \Omega$ then $\tilde\phi(y,s) = y$ for all s as $d\tilde\phi(y,s)/ds = 0$. This is a contradiction, as by the above remark $\tilde\phi(t,y) = \tilde\phi(\phi(x,t),t) = x$. Similarly for all $x \in \Omega$, $\tilde\phi(x,t) \in \Omega$. This implies that $\phi(t,\cdot)$ is one-to-one and onto and $\phi^{-1}(t,\cdot)$ is given by $\tilde\phi(t,\cdot)$. By smooth dependence on initial conditions, both ϕ and $\tilde\phi$ are continuously differentiable mappings for each t implying that, $\phi(t,\cdot)$ is a diffeomorphism. Theorem ■

Appendix B
Euler–Lagrange

Proof of Theorem 2: The minimizer $\hat\phi(x_n,t), n = 1,\ldots,N$ for the landmark matching problem with $\phi(x_n,0) = x_n, n = 1,\ldots,N$ and minimizing

$$\begin{aligned}&\hat\phi(x_n,\cdot)\\ &n=1,\ldots,N\\ &= \arg\min_{\substack{\phi(x_n,\cdot)\\ n=1,\ldots,N}} \int_0^1 \underbrace{\sum_{ij=1}^N \dot\phi(x_i,t)^T(K(\phi(t))^{-1})_{ij}\dot\phi(x_j,t)}_{H(\phi,\dot\phi,t)}\\ &\quad+\sum_{n=1}^N [y_n-\phi(x_n,1)]^T\Sigma_n^{-1}[y_n-\phi(x_n,1)]^T\end{aligned}$$

satisfies the Euler–Lagrange conditions of (11) and (12).

Proof: Let $\hat\phi(x_n,t), n = 1,\ldots,N$ be the minimizer and consider the one parameter family of functions $\phi(x_n,t) = \hat\phi(x_n,t) + \epsilon\eta(x_n,t), n = 1,\ldots,N$ is an allowable perturbation around the minimizer implying that $\eta(x_n,0) = 0$ as $\phi(x_n,0) = x_n, n = 1,\ldots,N$. Thus the objective functional

$$\begin{aligned}J(\epsilon) = &\int_0^1 H(\phi,\dot\phi,t)\,dt\\ &+\sum_{n=1}^N [y_n-\phi(x_n,1)]^T\Sigma_n^{-1}[y_n-\phi(x_n,1)]^T\end{aligned}$$

satisfies $J'(\epsilon)|_{\epsilon=0} = 0$ for all allowable perturbations. Differentiating and integrating by parts gives

$$\begin{aligned}0 = &\int_0^1\sum_{n=1}^N\sum_{l=1}^3 \frac{\partial H}{\partial\phi_l(x_n,t)}\eta_l(x_n,t)\,dt\\ &+\int_0^1\sum_{n=1}^N\sum_{l=1}^3 \frac{\partial H}{\partial\dot\phi_l(x_n,t)}\dot\eta_l(x_n,t)\,dt\\ &+\sum_{n=1}^N \eta(x_n,1)^T\Sigma_n^{-1}[y_n-\phi(x_n,1)],\\ 0 = &\sum_{n=1}^N\sum_{l=1}^3 \frac{\partial H}{\partial\dot\phi_l(x_n,1)}\eta_l(x_n,1)\\ &+\int_0^1\sum_{n=1}^N\sum_{l=1}^3 \left[\left(\frac{\partial H}{\partial\phi_l(x_n,t)}\right)-\frac{d}{dt}\left(\frac{\partial H}{\partial\dot\phi_l(x_n,t)}\right)\right]\\ &\times\eta_l(x_n,t)\,dt\\ &+\sum_{n=1}^N \eta_l(x_n,1)^T\Sigma_n^{-1}[y_n-\phi(x_n,1)],\end{aligned}\tag{46}$$

implying for all allowable perturbations yielding the Euler–Lagrange conditions, for all $n = 1,\ldots,N, l = 1,2,3$

$$0 = \left(\frac{\partial H}{\partial\phi_l(x_n,t)}\right) - \frac{d}{dt}\left(\frac{\partial H}{\partial\dot\phi_l(x_n,t)}\right)\tag{47}$$

$$0 = \frac{\partial H}{\partial\phi_l(x_n,1)} + \left(\Sigma_n^{-1}[y_n-\phi(x_n,1)]\right)_l.\tag{48}$$

As $H(\phi,\dot{\phi},t) = \sum_{ij}\dot{\phi}(x_i,t)^T(K(\phi(t))^{-1})_{ij}\dot{\phi}(x_j,t)$, gives for $n=1,\ldots,N, l=1,2,3$

$$\begin{aligned}
&\frac{\partial H}{\partial \phi_l(x_n,t)} + \frac{d}{dt}\left(\frac{\partial H}{\partial \dot{\phi}_l(x_n,t)}\right)\\
&\quad = \sum_j \dot{\phi}(x_j,t)^T\left(\frac{\partial K(\phi(t))^{-1}}{\partial \phi_l(x_n,t)}\right)_{nj}\dot{\phi}(x_j,t)\\
&\qquad + 2\sum_j \dot{\phi}(x_j,t)^T\left(\frac{\partial K(\phi(t))^{-1}}{\partial \phi_l(x_n,t)}\right)_{nj}\dot{\phi}(x_j,t)\\
&\qquad + \left(2\sum_{j=1}^{N}(K(\phi(t))^{-1})_{nj}\ddot{\phi}(x_j,t)\right)_l;\\
&\frac{\partial H}{\partial \dot{\phi}_l(x_n,t)} = \left(2\sum_{j=1}^{N}(K(\phi(1))^{-1})_{nj}\dot{\phi}(x_j,1)\right)_l.
\end{aligned}$$

■

APPENDIX C
HOMOGENEITY LEMMA FOR EXACT LANDMARK MATCHING

As we will be interested in generating flows that match exactly a given set of landmarks in the template and the target we now state the associated energy minimization problem and prove existence of a solution. Notice that care must be taken as $\Sigma \to 0$. Also, the trivial velocity field used in the proof of part 1 showing that $\mathcal{F}$ is not empty will not be satisfied. We first prove that given N distinct landmark points, there exists a diffeomorphism which maps the points on to each other exactly. For this define the concept of isotopy. Let $\mathrm{Diff}(\Omega)$ be the space of all compactly supported diffeomorphisms of $\Omega \leftrightarrow \Omega$.

Definition 1: Elements $\phi^0,\phi^1 \in \mathrm{Diff}(\Omega)$ are said to be isotopic, ϕ^0–ϕ^1, if there is a C^∞ map $I : \Omega\times[0,1]\to\Omega$ such that $\phi^0(x) = I(x,0)$, $\phi^1(x) = I(x,1)$. The map I is called an isotopy between ϕ_0 and ϕ_1.

We use the Homogeneity Lemma from differential geometry to extend from one landmark for the existence of a flow to that of exactly matching the N given landmarks.

Let $x,y \in U \subset \Omega, U$ open and connected, then there is $\phi \in \mathrm{Diff}(U)$ and a compactly supported isotopy I such that $\phi(x) = y$ and $I(\cdot,0) = Id, I(\cdot,1) = \phi(\cdot)$.

Lemma 3: Let $x_i, y_i, i = 1,\ldots,N$ be distinct points in Ω then there is a $\phi \in \mathrm{Diff}(\Omega)$ and a compactly supported isotopy I such that for all N, $\phi(x_i) = y_i$ and $I(\cdot,0) = Id, I(\cdot,1) = \phi(\cdot)$.

Proof: We prove the above statement by induction. For this, we use the Homogeneity Lemma for $N = 1$. That is, let $x,y \in U \subset \Omega, U$ open and connected, then there is $\phi \in \mathrm{Diff}(U)$ and a compactly supported isotopy I such that $\phi(x) = y$ and $I(\cdot,0) = Id, I(\cdot,1) = \phi(\cdot)$.

Assume that it is true for $L-1$. There exists disjoint open connected sets $U^1, U^2 \subset \Omega, U^1 \cap U^2 = \emptyset$ such that $x_i, y_i \in U^1$ for $i = 1,\ldots,N-1$ and $x_N, y_N \in U^2$. By theorem above there exists $\phi^1 \in \mathrm{Diff}(U^1)$ and $\phi^2 \in \mathrm{Diff}(U^2)$ and compactly supported isotopes I^1 and I^2. Let $\phi = \phi^1 \circ \phi^2$. It is immediate that ϕ has the desired properties. Similarly the isotopy $I = I^1 \circ I^2$ also has the desired properties $I(\cdot,0) = Id, I(\cdot,1) = \phi$. ■

We now show that flows can generate all diffeomorphisms isotopic to the identity and hence particularly diffeomorphisms that map the given landmarks.

Theorem 3: If $f \in \mathrm{Diff}(\Omega)$ is isotopic to the Id then there exists a vector field $v(x,t)$ and an associated flow $\phi(x,t)$ such that $f(x) = \phi(x,1)$.

Proof: As $f \sim Id$ there exists a map $I : \Omega\times[0,1]\to\Omega$ such that $x = I(x,0)$, $f(x) = I(x,1)$. Let $v(I(x,t),t) = \partial I(x,t)/\partial t$. As for each $t \in [0,1]$ $I(\cdot,t) \in \mathrm{Diff}(\Omega)$, it implies that $\mathrm{supp}(v(x,t)) \subset \Omega$ is compactly supported and hence complete. By existence and uniqueness theorem there exists a unique flow $\phi(x,t)$ such that $\phi(x,0) = x$ that satisfies the differential equation $\partial\phi(x,t)/\partial t = v(\phi(x,t),t)$. This implies that the flow is in fact given by the isotopy $I(x,t)$. ■

REFERENCES

[1] M. Miller, A. Banerjee, G. Christensen, S. Joshi, N. Khaneja, U. Grenander, and L. Matejic, "Statistical methods in computational anatomy," *Statist. Meth. Med. Res.*, vol. 6, pp. 267–299, 1997.

[2] U. Grenander and M. I. Miller, "Computational anatomy: An emerging discipline," *Quart. Appl. Math.*, vol. 56, pp. 617–694, 1998.

[3] M. I. Miller, G. E. Christensen, Y. Amit, and U. Grenander, "Mathematical textbook of deformable neuroanatomies," *Proc. Nat. Acad. Sci.*, vol. 90, no. 24, Dec. 1993.

[4] U. Grenander and M. I. Miller, "Representations of knowledge in complex systems," *J. R. Statist. Soc. B*, vol. 56, no. 3, pp. 549–603, 1994.

[5] S. C. Joshi, M. I. Miller, G. E. Christensen, A. Banerjee, T. A. Coogan, and U. Grenander, "Hierarchical brain mapping via a generalized dirichlet solution for mapping brain manifolds," in *Proc. SPIE 1995 Int. Symp. Optical Science, Engineering, Instrumentation*, vol. 2573, Aug. 1995, pp. 278–289.

[6] S. C. Joshi, J. Wang, M. I. Miller, D. Van Essen, and U. Grenander, "On the differential geometry of the cortical surface," in *Proc. SPIE 1995 Geometric Methods Applied Imaging*, San Diego, CA, July 9–14, 1995.

[7] 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, Oct. 1996.

[8] J. W. Haller, G. E. Christensen, S. Joshi, J. W. Newcomer, M. I. Miller, J. C. Csernansky, and M. W. Vannier, "Hippocampal MR imaging morphometry by means of general pattern matching," *Radiology*, vol. 199, pp. 787–791, June 1996.

[9] J. W. Haller, A. Banerjee, G. E. Christensen, S. Joshi, M. I. Miller, M. W. Vannier, and J. C. Csernansky, "Three-dimensional hippocampal volumetry by high dimensional transformation of a neuroanatomical atlas," *Radiology*, vol. 202, pp. 504–510, Feb. 1997.

[10] S. Joshi, M. I. Miller, and U. Grenander, "On the geometry and shape of brain sub-manifolds," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 11, no. 8, 1997.

[11] G. E. Christensen, S. C. Joshi, and M. I. Miller, "Volumetric transformation of brain anatomy," *IEEE Trans. Med. Imag.*, vol. 16, pp. 864–877, Dec. 1997.

[12] P. Dupuis, U. Grenander, and M. I. Miller, "Variational problems on flows of diffeomorphisms for image matching," *Quart. Appl. Math.*, vol. LVI, pp. 587–600, Sept. 1998.

[13] L. D. Griffin, "The intrinsic geometry of the cerebral cortex," *J. Theor. Biol.*, vol. 166, pp. 261–273, 1994.

[14] L. Matejic, "Group cascades for representing biological variability,", Ph.D. dissertation, Brown Univ., Providence, RI, 1997.

[15] W. M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*. New York: Academic, 1986.

[16] F. L. Bookstein, *The Measurement of Biological Shape and Shape Change*. New York: Springer-Verlag, 1978, vol. 24.

[17] ——, *Morphometric Tools for Landmark Data*. New York: Cambridge Univ. Press, 1991.

[18] J. T. Kent and K. V. Mardia, "The link between kriging and thin-plate splines," in *Probability, Statistics and Optimization*, F. P. Kelly, Ed. New York: Wiley, 1994.

[19] G. Wahba, *Spline Models for Observational Data*. Philadelphia, PA: SIAM, 1990.

[20] G. E. Christensen, R. D. Rabbitt, M. I. Miller, S. C. Joshi, U. Grenander, and T. A. Coogan, *Topological Properties of Smooth Anatomic Maps*. Norwell, MA: Kluwer, 1995.
[21] L. Younes, "Discussion of mathematics for object recognition shape, invariance and deformations," in *Proc. Workshop Mathematics Object Recognition Shape, Invariance, Deformations*, Nov. 1997.
[22] S. Joshi, "Large deformation diffeomorphisms and Gaussian random fields for statistical characterization of brain submanifolds," Ph.D. dissertation, Dept. Elect. Eng., Sever Inst. Technol., Washington Univ., St. Louis, MO, Aug. 1997.
[23] D. J. Felleman and D. C. Van Essen, "Distributed hierarchical processing in the primate cerebral cortex," *Cerebral Cortex*, vol. 1, no. 1, pp. 1–47, 1991.
[24] H. A. Drury, D. C. Van Essen, C. H. Anderson, C. H. Lee, T. A. Coogan, and J. W. Lewis, "Computerized mappings of the cerebral cortex. A multiresolution flattening method and a surface-based coordinate system," *J. Cogn. Neurosci.*, vol. 8, pp. 1–28, 1996.
[25] M. I. Miller, N. Khaneja, and U. Grenander, "Dynamic programming generation of curves on brain surfaces," *Pattern Anal. Machine Intell.*, vol. 20, pp. 1260–1264, Nov. 1998.
[26] M. Bakircioglu, U. Grenander, N. Khaneja, and M. Miller, "Curve matching on brain surfaces using frenet distances," *Hum. Brain Map.*, vol. 6, no. 5, pp. 329–332, 1998.
[27] N. Khaneja, "Statistics and geometry of cortical features," M.S. Thesis, Dept. Elect. Eng., Sever Inst. Technol., Washington Univ., St. Louis, MO, Dec. 1996.
[28] M. I. Miller, S. C. Joshi, and G. E. Christensen, "Large deformation fluid diffeomorphisms for landmark and image matching," in *Brain Warping*, A. W. Toga, Ed. New York: Academic, 1999, pp. 115–131.
[29] D. C. Van Essen, H. Drury, S. Joshi, and M. I. Miller, "Functional and structural mapping of human cerebral cortex: Solutions are in the surfaces," *Proc. Nat. Acad. Sci.*, vol. 95, pp. 788–795, Feb. 1998.
[30] L. Younes, "Computable elastic distances between shapes," *SIAM J. Applied Math.*, 1998.
[31] H. Tagare, D. O'Shea, and A. Rangarajan, "A geometric criterion for shape based nonrigid correspondence," in *Proc. 5th Int. Conf. Computer Vision*, 1995, pp. 434–439.
[32] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Control*. Berlin, Germany: Springer-Verlag, 1975.

Sarang C. Joshi received the B.Sc., M.Sc., and D.Sc. degrees in electrical engineering from Washington University, St. Louis, MO.

In November 1997, he joined IntellX, L.L.C., (now part of Medtronic), Boulder, CO, as Director of Technology Development. In May 2000, he joined the University of North Carolina, Chapel Hill, as an Assistant Professor of radiation oncology and biomedical engineering. His research interests include pattern recognition, computational geometry, image understanding and deformable templates, and computational anatomy.

Michael I. Miller received the B.S.E.E. degree from the State University of New York, Stony Brook, in 1976, and the M.S.E.E. and Ph.D.B.M.E. degrees from The Johns Hopkins University, Baltimore, MD, in 1978 and 1983, respectively.

He has been affiliated with Washington University, St. Louis, MO, since 1984 as a Professor of electrical engineering, with joint appointments in the Institute for Biomedical Computing and Mallinckrodt Institute of Radiology. In 1995, he was named to the Newton R. and Sarah L. Wilson Professorship in Biomedical Engineering, which he held until 1998. In July 1998, he joined The Johns Hopkins University as a Professor of biomedical engineering and electrical and computer engineering, where he currently directs the Center for Imaging Science. His research and teaching interests are in the areas of pattern theory, computational linguistics, computational neuroscience, and computational anatomy.