BRAIN AND SURFACE WARPING
VIA MINIMIZING LIPSCHITZ EXTENSIONS

By

Facundo Mémoli
Guillermo Sapiro
and
Paul Thompson

IMA Preprint Series # 2092
(January 2006)
Based on the notion of minimizing Lipschitz Extensions and its connection with the infinity Laplacian, a computational framework for surface warping and in particular brain warping (the nonlinear registration of brain imaging data) is presented in this paper. The basic concept is to compute a map between surfaces that minimizes a distortion measure based on geodesic distances while respecting the boundary conditions provided. In particular, the global Lipschitz constant of the map is minimized. This framework allows generic boundary conditions to be applied and allows direct surface-to-surface warping. It avoids the need for intermediate maps that flatten the surface onto the plane or sphere, as is commonly done in the literature on surface-based non-rigid brain image registration. The presentation of the framework is complemented with examples on synthetic geometric phantoms and cortical surfaces extracted from human brain MRI scans.
Abstract

Based on the notion Minimizing Lipschitz Extensions and its connection with the infinity Laplacian, a computational framework for surface warping and in particular brain warping (the nonlinear registration of brain imaging data) is presented in this paper. The basic concept is to compute a map between surfaces that minimizes a distortion measure based on geodesic distances while respecting the boundary conditions provided. In particular, the global Lipschitz constant of the map is minimized. This framework allows generic boundary conditions to be applied and allows direct surface-to-surface warping. It avoids the need for intermediate maps that flatten the surface onto the plane or sphere, as is commonly done in the literature on surface-based non-rigid brain image registration. The presentation of the framework is complemented with examples on synthetic geometric phantoms and cortical surfaces extracted from human brain MRI scans.

1 Introduction

Brain warping, a form of brain image registration and geometric pattern matching, is one of the most fundamental and thereby most studied problems in computational brain imaging [37]. Brain images are commonly warped, using 3-dimensional deformation fields, onto a common neuroanatomic template prior to cross-subject comparison and integration of functional and anatomical data. Images of the same subject may be warped into correspondence over time, to help analyze shape changes during development or degenerative diseases. Almost all the active research groups in this area have developed and/or have their favorite brain warping technique. A few representative works can be found at [5, 7, 8, 13, 14, 16, 30, 34, 36, 37, 40, 41], this list being far from complete. In spite of this, the problem is still open and widely studied, since there is not a “ground truth” method to obtain a map between brains. The criteria for matching different features (e.g., geometry or intensity) may also depend on the applications, which range from recovering intraoperative brain change to mapping brain growth, or reducing cross-subject anatomical differences in group functional MRI studies.

The way the brain warping problem is addressed is critical for studies of brain diseases that are based on population comparisons. Examples of this application can be found at [14, 34], although these are a very non-exhaustive account of the rich literature on the subject. The interested reader may also check [35] for numerous applications of brain warping and population studies. As detailed in [37], brain warping approaches can be divided into two classes, those based on volume-to-volume matching and those based on surface-to-surface matching. Our work belongs to the latter of the categories. Surface matching has recently received increasing attention as most functional brain imaging studies focus on the cortex, which varies widely in geometry across subjects. The power of these studies depends on the degree to which the functional anatomy of the cortex can be aligned across subjects, so improved cortical surface registration has become a major goal. In contrast with flow based works such as those in [5, 30, 34], our motivation is as in [1, 16, 17, 18, 21, 38, 39, 42]. That is, we aim to compute a map that preserves certain pre-defined geometric characteristics of the surfaces. While the literature has

\[\text{This includes groups at JHU, UCLA, U. Penn., INRIA, MGH, GATECH, Harvard-BW, and the University of Florida, to name just a few.}\]
mainly attempted to preserve angles and areas, we work with geodesic distances (see also [33]). Our work is inspired by the literature on Lipschitz minimizing maps and in its connection to the infinite Laplacian. The motivation for using these frameworks will be presented after some brief mathematical introduction below.

In this paper we therefore introduce the use of Lipschitz minimizing maps into the area of computational brain imaging, presenting a theoretical and computational framework complemented by examples with artificial and real data. An additional critical contribution of the work here presented is that intermediate distorting maps to the plane or sphere are avoided – these intermediate mappings are common practice in the brain warping literature.

2 Formal statement of the problem

Let \( B_1 \) and \( B_2 \) be two cortical surfaces (2D surfaces in the three dimensional Euclidean space) which we consider smooth and endowed with the metric inherited from \( \mathbb{R}^3 \) so that \( d_{B_1} \) and \( d_{B_2} \) are the geodesic distances measured on \( B_1 \) and \( B_2 \), respectively. Let \( \Gamma_1 \subset B_1 \) and \( \Gamma_2 \subset B_2 \) be subsets which represent features for which a correspondence is already known. In general, the sets \( \Gamma \) are the union of smooth curves traced on the surfaces, e.g., sulcal beds lying between gyri, and/or a union of isolated points. A set of anatomical landmarks that occur consistently in all subjects can be reliably identified using standardized anatomical protocols or automated sulcal labeling techniques (see for example Brain VISA by Mangin and Riviere and SEAL by Le Goualher).

Functional anatomy also varies with respect to sulcal landmarks, but sulci typically lie at the interfaces of functionally different cortical regions so aligning them improves the registration of functionally homologous areas. As commonly done in brain warping [37], we assume that a correspondence between \( \Gamma_1 \) and \( \Gamma_2 \) is pre-specified to the map (boundary conditions of the map). In this correspondence, internal point correspondences may be allowed to relax along landmark curves in the final mappings, e.g., [26].

To fix ideas let’s assume that \( \Gamma_1 = \bigcup_{i=1}^{N} x_i \) and \( \Gamma_2 = \bigcup_{i=1}^{N} y_i \), and that the correspondence is given by \( x_i \mapsto y_i \) for \( 1 \leq i \leq N \).

We want to find a (at least continuous) map \( \phi : B_1 \rightarrow B_2 \) such that \( \phi(x_i) = y_i \) for \( 1 \leq i \leq N \) and such that \( \phi \) produces minimal distortion according to some functional \( J \). One possible way of interpreting this problem is that we are trying to extrapolate or extend the correspondence from \( \Gamma_1 \) to the whole of \( B_1 \) in such a way that we achieve small distortion.

A possible way to measure the distortion produced by a map \( \phi \) is by computing the functionals \( (1 \leq p < \infty) \)

\[
J_p(\phi) = \left( \frac{1}{\mu(B_1)} \int_{B_1} \|D_{B_2} \phi\|^p \mu(dx) \right)^{1/p}
\]

(1)

where \( D_{B_2} \) denotes differentiation intrinsic to the surface \( B_1 \) and \( \mu \) is the area measure on \( B_1 \). One immediate idea is then to consider, for a fixed \( p \in (1, \infty) \), the following variational problem:

**Problem 1 (minimize \( J_p \))** Find \( \phi \in S \) such that \( J_p(\phi) = \inf_{\psi \in S} J_p(\psi) \), where \( S \) is a certain smoothness class of maps \( \phi \) from \( B_1 \) to \( B_2 \) such that they respect the given boundary conditions \( \phi(x_i) = y_i \) for all \( x_i \in \Gamma_1 \).

The case \( p = 2 \) corresponds to the Dirichlet functional and has connections with the theory of (standard) Harmonic Maps. In more generality, it is customary to call the solutions to Problem 1 \( p \)-Harmonic Maps, see for example [10, 20, 11]. It is easy to show, under mild regularity assumptions, that for a fixed \( \phi \), \( J_p(\phi) \) is nondecreasing as a function of \( p \), and that [15]

\[
J_\infty(\phi) := \lim_{p \to \infty} J_p(\phi) = \operatorname{esssup}_{x \in B_1} \|D_{B_2} \phi(x)\|_2,
\]

(2)

which is the Lipschitz constant of \( \phi \).

In this paper we propose to use the functional \( J_\infty \) as a measure of distortion for maps between cortical surfaces and to solve the associated variational problem in order to find a candidate mapping between the cortical surfaces (constrained by the provided boundary conditions).

Let \( \mathcal{L} \) denote the space of all Lipschitz continuous maps \( \psi : B_1 \to B_2 \) such that \( \psi(x_i) = y_i \) for \( 1 \leq i \leq N \). We then propose to solve the following problem:

**Problem 2 (minimize \( J_\infty \))** Find \( \phi \in \mathcal{L} \) such that \( J_\infty(\phi) = \inf_{\psi \in \mathcal{L}} J_\infty(\psi) \).

We now argue in favor of this functional.
2.1 Why use $J_\infty$?

Our first argument is that $J_\infty$ measures distortion in a more global way than any of the $J_p$ for $p \in (1, \infty)$, since instead of computing an averaged integral quantity, we are looking at the supremum of the local distortions, $\|D_{B_2} \phi(x)\|_2$. Note also that as stated above, $J_\infty$ upper-bounds $J_p$ under mild regularity assumptions.

Another element to consider is that this problem is well posed for the kind of general boundary data we want to respect, provided both at curves and isolated points on the surfaces. At least for the case $p \leq 2$, this is not true in general, see [3].

We are then looking for a Lipschitz extension of the map given at $\Gamma_1$ whose Lipschitz constant is as small as possible. Let

$$L(\Gamma_1, \Gamma_2) := \max_{x_i, x_j \in \Gamma_1} \frac{d_{B_2}(y_i, y_j)}{d_{B_1}(x_i, x_j)},$$

that is, the Lipschitz constant of the boundary data. In general, we will have $\inf_{\phi \in \mathcal{L}} J_\infty \geq L(\Gamma_1, \Gamma_2)$. This is related to Kirschbaum’s Theorem which, in one of its many guises states that a Lipschitz map $f : S \to \mathbb{R}^D$, $S \subset \mathbb{R}^d$, has an extension $\tilde{f} : \mathbb{R}^d \to \mathbb{R}^D$ with the same Lipschitz constant as $f$, see [12]. In the same vein, one has Whitney and McShane extensions which apply to the case when the domain is any metric space $X$ and the target is $\mathbb{R}$. These extensions provide functions that agree with $f$ where boundary conditions are given and preserve the Lipschitz constant throughout $X$, see for example [2, 22]. The more general problem of extending $f : S \to Y$ ($S \subset X$, $X$ and $Y$ any metric spaces) to all $X$ with the same Lipschitz constant is not so well understood and only partial results are known, see for example [23, 24, 25].

The idea then is to keep the distortion at the same order as that of the provided boundary conditions. In general there might be many solutions for the Problem (2). One particular class of minimizers which has received a lot of attention is that of absolute minimizers, or absolutely minimizing Lipschitz extensions (AMLE). Roughly speaking, the idea here is to single out those solutions of Problem (2) that also possess minimal local Lipschitz constant, again, see [2] for a general exposition, and [22] for a treatment of the case when the domain is any reasonable metric space and the target is the real line.

3 Proposed computational approach

If we take for example the case of $p$-Harmonic maps, one way of dealing with the computation of the optimal map $\phi_p$ is by implementing the geometric $p$-heat flow associated with the Euler-Lagrange equation of the functional $J_p$, starting from a certain initial condition. As was explained in [29], using an implicit representation for both $B_1$ and $B_2$, we could obtain the partial differential equation $\text{PDE}_p$ we need to solve in order to find $\phi_p$. By taking the formal limit as $p \uparrow \infty$ we would find $\text{PDE}_\infty$, the PDE that characterizes the solution $\phi_\infty$ of the (variational) Problem (2). All of this might work if we had a notion of solution for the resulting PDEs. Whereas this is feasible in the case of $\text{PDE}_p$ for $1 < p < \infty$, to the best of our knowledge, there is no such notion of a solution for $\text{PDE}_\infty$. One could of course still persist and try to solve these equations without the necessary theoretical foundations and call these plausible solutions $\infty$-Harmonic Maps. Nonetheless, this is certainly an interesting line of research.

A different direction is considered in this work. As a guiding example, we first concentrate on the case where $B_1$ is any closed smooth manifold and $B_2$ is replaced by $\mathbb{R}$, as considered in [4] (for scalar data interpolation on surfaces), and in [32]. In [4], the authors propose to follow a similar path to the one we have just described, and they do not obtain a convergent numerical discretization for the resulting PDE. Meanwhile, in [32], the author proposes a convergent discretization of the PDE, basing his construction on the original variational problem. We choose to follow this idea as our guiding principle.

We now explain this alternative approach. The basic idea is simple, instead of first obtaining the Euler-Lagrange equations for $J_\infty$ and then discretizing them, we will first discretize $J_\infty$ and then proceed to solve the resulting discrete problem. Consider that the domain $B_1$ is given discretely as a set of (different) points $\mathbb{B}_1 = \{x_1, \ldots, x_m\}$ together with a neighborhood relation (i.e., a graph). To fix ideas let’s assume the neighborhood relation is a $k$-nearest neighbors one. Denote, for each $1 \leq i \leq m$, by $N_i = \{x_{j_1}, \ldots, x_{j_k}\} \in \mathbb{B}_1$ the set

\footnote{The case when the domain is a subset of $\mathbb{R}^d$ and the target is the real line leads to the so called infinity Laplacian, see [2, 19, 9].}
of $k$ neighbors of the point $x_i$. We consider the discrete local Lipschitz constant of the map $\phi$ at $x_i$:

$$L_i(\phi) := \max_{x_j \in N_i} \frac{d_{B_2}(\phi(x_i), \phi(x_j))}{d_{B_1}(x_i, x_j)} \quad (3)$$

Upon noting that $L_i(\phi)$ serves as a discrete approximation to $\|D_{B_2}\phi(x_i)\|_2$, we see that a possible discretization of the functional $J_\infty(\phi)$ is given by the discrete global Lipschitz constant of $\phi$ given by $\max_{1 \leq i \leq m} L_i(\phi)$. The author of [32] proposed, in the case when $B_2$ is replaced by $B_1$, solving the discrete version of Problem (2) by following the following iterative procedure (here described for $B_2$ a surface as in our problem):

- Let $\phi_0$ be an initial guess of the map.
- For each $n \geq 1$, if $x_i \not\in \Gamma_1$, let
  \[
  \phi_n(x_i) = \arg\min_{y \in B_2} \max_{j \in N_i} \frac{d_{B_2}(y, \phi_{n-1}(x_j))}{d_{B_1}(x_i, x_j)} \quad (4)
  \]
- $\phi_n(x_i) = y_i$ for all $n \geq 0$ for $x_i \in \Gamma_1$.

With computational efficiency related modifications described below, this is the approach we follow in general. The intuition behind this iterative procedure is that, at each point of the domain, we are changing the value of the map in order to minimize the local Lipschitz constant, that is, the local distortion produced by the map. This is in agreement with the notion of AMLEs briefly explained in §2.1. We should remark that since we are using intrinsic distances for the matching, we can let $L_i(\phi)$ play the role of (the norm of) the displacement field for analyzing the deformation.\footnote{One can imagine a situation in which two isometric surfaces are matched by our algorithm such that $L_i(\phi) = 1$ for all $i$, but the displacement field $|x_i - \phi(x_i)|$ is large since there may be no rigid motion that aligns the two surfaces. One simple example is a flat sheet of paper and the same sheet slightly bent.}

## 4 Implementation details

In addition to discretizing the domain $B_1$, we also use a discretization $B_2 = \{y_1, \ldots, y_{m'}\}$ of the target space $B_2$ for our implementation. We endow $B_2$ with a neighborhood relation given by the $k$-nearest neighbors of each point. For computational efficiency, we work at all times with two different scales in the discrete domain $B_1$. We choose a subset $F_1$ of $B_1$ such that $|F_1| \ll m$ but still $F_1$ is an efficient (well separated) covering of $B_1$ with small covering radius. We do this by using the well known (geodesic) Farthest Point Sampling (FPS) procedure, see [28, 31], which can be efficiently constructed based on optimal computational techniques. Roughly speaking, we apply the iterative procedure on this subset of points only and then extend the map to the rest of the points in the domain $B_1$. We now show how to obtain a reasonable initial condition $\phi_0$ and then discuss additional details regarding the implementation of the iterative procedure described in the previous section.

### Building the initial condition:
We compute, for all $x_r \in F_1 \setminus \Gamma_1$, $\phi_0(x_r) = \arg\min_{y \in B_2} \max_{x_i \in \Gamma_1} \frac{d_{B_1}(y, \phi_0(x_i))}{d_{B_1}(x_r, x_i)}$. For this step we use the classical Dijkstra’s algorithm for approximating the distances $d_{B_1}$ and $d_{B_2}$ since they might be evaluated at faraway points. This is of course run on the graphs obtained from connecting each point to its $k$-nearest neighbors.

### The iterative procedure:
After $\phi_0$ is computed for all points in the set $F_1$, we run the iterative procedure from §2 on this set of points. The main modification here is that whereas we still use Dijkstra’s algorithm for approximating $d_{B_2}$ in the target surface, since in the domain we must compute $d_{B_1}$ only for neighboring points ($F_1$ was chosen to be dense enough), for computationally efficiency we can approximate $d_{B_1}(x_i, x_j) \approx ||x_i - x_j||$ for $x_i \in \Gamma_1$. We should also point out that for points in $F_1$, the neighborhood relation is defined to be that of $k$-nearest neighbors with respect to the metric on $B_1$ defined by the adjacency matrix of $B_1$. Let $\phi_* : F_1 \rightarrow B_2$ denote the map obtained as the output of this stage.

### Extension to the whole domain:
After we have iterated over points in $F_1$ until convergence, we extend the map $\phi_*$ to all points $x_i$ in $B_1 \setminus \{F_1 \cup \Gamma_1\}$. This is done by computing $\phi_* (x_i) = \arg\min_{y \in B_2} \max_{x \in F_1 \cup \Gamma_1} \frac{d_{B_1}(y, \phi_*(x))}{d_{B_1}(x_i, x)}$. For this step, and since we have already obtained the map for a relatively dense subset, we approximate both $d_{B_1}$ and $d_{B_2}$ by the Euclidean distance. Once again, the motivation for this is just computational efficiency.
5 Examples

In this section we present some computational examples of the ideas presented in previous sections. First, in Figure 1 the domain $B_1$ is a cube ($n = 10086$) and the target $B_2$ is a sphere ($n' = 17982$). For the purposes of visualizing the map, we assigned the clown texture (which can be thought of as a function $I : \mathbb{B}_2 \to \mathbb{R}^3$) to the sphere, which can be seen on the bottom-right corner of the figure. The sphere and the cube were concentric and of approximately the same size. We selected $F_1$ on the cube consisting of 1000 well separated points using the FPS procedure alluded to in §4. Also, we set $k = 6$ (number of neighbors). We then chose $\Gamma_1$ to be the first 100 points of the set and then projected them onto the sphere, obtaining in this way, the corresponding set $\Gamma_2$ to use as boundary conditions. We then followed the computational procedure detailed before. The top-left figure shows the composition $I \circ \phi_s : \text{cube} \to \mathbb{R}$ as a texture on the cube. Finally, the top-right and the bottom-left images show the histogram of $L_1(\phi_s)$ and its spatial distribution in the domain (we paint the cube at each point $x_i$ with the color corresponding to $L_1(\phi_s)$, respectively. Ideally, we would like to obtain a $\delta$-type histogram, meaning that the distances have been constantly scaled. Of course, this is not possible (unless one of the surfaces is isometric to a scaled version of the other), and we attempt to obtain histograms as concentrated as possible. This is quite nicely obtained for this and the additional examples in this paper.

Figure 2 shows the construction of a map from the unit sphere $S^2$ into a cortical hemisphere $\mathbb{B}$ ($\mathcal{B}$). The boundary conditions consisted of 6 pairs of points. We first took the following 6 points on the sphere $\Gamma_1 = \{ (\pm 1, 0, 0), (0, \pm 1, 0), (0, 0, \pm 1) \}$. We then constructed the intrinsic distance matrix $[d_{ij}^e(p_i, p_j)]$ for all $p_i, p_j \in \Gamma_1$. Finally, we chose 6 points $\{ q_1, \ldots, q_6 \} = \Gamma_2$ in $\mathbb{B}$ such that $\max_{i \neq j} d_{ij}^e(q_i, q_j)$ was as close as possible to $\frac{1}{2}\text{diam}(\mathbb{B})$. We painted $\mathbb{B}$ with a texture $I_\mathcal{B}$ depending on its mean curvature so as to more easily visualize the sulci/crests: If $H(x)$ stands for mean curvature of $\mathcal{B}$ at $x$, then $I_\mathcal{B}(x) = (H(x) - \min_x H(x))^2$. See the caption for more details.

The example in Figure 3 is about computing a map $\Phi$ from a subject’s left hemisphere $\mathbb{B}_1$ to another subject’s left hemisphere $\mathbb{B}_2$. The boundary conditions were constructed in a way similar to the one used for the previous example, but in this case, 300 points were chosen. Note that, if available, hand traced curves could be used as commonly done in the literature (in other words, more anatomical/functional oriented boundary conditions). In the first two rows we show 4 different views of each cortical surface, and in the third row we show $\mathbb{B}_2$ colored with the values of $L_1(\Phi)$ which we interpret as a measure of the local deformation of the map needed to match $\mathbb{B}_1$ to $\mathbb{B}_2$. See the caption for more details.

![Figure 1: Artificial example of the proposed warping algorithm. From top to bottom and left to right: The domain surface, with a picture painted on it to help in visualizing the computed map; histogram of the Lipschitz constant (note how it is concentrated around a single value); color coded distribution of the Lipschitz constant for the computed map; and mapped texture following the computed map.](image)

6 Concluding remarks

In this paper we have introduced the notions of minimizing Lipschitz extensions into the area of surface and brain warping. These maps provide a more global constraint than ordinary p-harmonic ones, and allow for more general boundary conditions. The proposed computational framework leads to an efficient surface-to-surface warping algorithm that avoids distorting intermediate steps that are common in the brain warping literature. We are currently investigating the use of this new warping technique...
for creating population averages and applying it to disease and growth studies. In earlier work, the Jacobian of a deformation mapping over time has been used to map the profile of brain tissue growth and loss in a subject scanned serially (tensor-based morphometry [6, 34]). The discrete local Lipschitz constants of our computed mappings also provide a useful index of deformation that can be analyzed statistically across subjects. The framework here introduced can also be applied in 3D for volumetric warping and with weighted geodesic distances instead of natural ones to include additional geometric characteristics in the matching. As recently shown in [27], the use of pairwise distances is of importance of other matching and computer vision tasks. Results in these directions will be reported elsewhere.

Acknowledgments

This work is partially supported by DARPA, the National Science Foundation, the National Institutes of Health, and the National Geospatial-Intelligence Agency.

References


Figure 2: An example of mapping between the cortex and a sphere. The order is the same as in the previous figure, but now the domain and target surface are colored with a curvature-based color code. Note once again the concentration of the Lipschitz constant for the computed map. On the top, the texture map corresponding to $l_H(x)$ as described in the text is used. On the bottom, the texture is $\max(l_H(x), \delta)$ for a user selected value of the threshold $\delta$, which highlights the gyral crests.
Figure 3: Warping between the cortical surfaces of two brains. In the first row we show 4 views of $B_1$: posterior, medial, lateral and directly viewing the occipital cortex. The corresponding 4 views of $B_2$ are shown in the second row. In the third row, we show $B_3$ with texture $I(x_i) = L(x_i)$ which can be interpreted as a measure of local deformation needed to match $x_i \in B_1$ to $\Phi(x_i) \in B_2$. Relatively little deformation (blue colors) is required to match features across subjects on the flat interhemispheric surface (second image in the second row). This is consistent with the lower variability of the gyral pattern in the cingulate and medial frontal cortices. By contrast, there is significant expansion required to match the posterior occipital cortices of these two subjects, especially in the occipital poles which are the target of many functional imaging studies of vision. The final panel in the figure shows the corresponding histogram for $L(x_i)$, the local Lipschitz constants of the map.