A Computational Model of Multidimensional Shape

**Xiuwen Liu · Yonggang Shi · Ivo Dinov ·
Washington Mio**

Abstract We develop a computational model of shape that
extends existing Riemannian models of curves to multidi-
mensional objects of general topological type. We construct
shape spaces equipped with geodesic metrics that measure
how costly it is to interpolate two shapes through elastic de-
formations. The model employs a representation of shape
based on the discrete exterior derivative of parametrizations
over a finite simplicial complex. We develop algorithms to
calculate geodesics and geodesic distances, as well as tools
to quantify local shape similarities and contrasts, thus ob-
taining a formulation that accounts for regional differences
and integrates them into a global measure of dissimilarity.
The Riemannian shape spaces provide a common frame-
work to treat numerous problems such as the statistical mod-
eling of shapes, the comparison of shapes associated with
different individuals or groups, and modeling and simulation
of shape dynamics. We give multiple examples of geodesic
interpolations and illustrations of the use of the models in
brain mapping, particularly, the analysis of anatomical vari-
ation based on neuroimaging data.

X. Liu
Department of Computer Science, Florida State University,
Tallahassee, FL 32306, USA

Y. Shi · I. Dinov
Laboratory of Neuro Imaging, UCLA School of Medicine, Los
Angeles, CA, 90095, USA

W. Mio (✉)
Department of Mathematics, Florida State University,
Tallahassee, FL 32306, USA
e-mail: mio@math.fsu.edu

1 Introduction

The development of computational models of shape is mo-
tivated by core problems in computer vision, pattern analy-
sis and computational anatomy, spanning a breadth of ap-
plications such as the statistical analysis of shapes, map-
ping the anatomy of the human brain, characterizing normal
anatomical variation and pathological changes, and morph-
ing shapes to estimate dynamics from observations. Rie-
mannian models based on the construction of shape spaces
are very appealing because they provide a framework to treat
many of these problems in a unified manner. These mod-
els are of particular interest in problems involving analysis,
estimation, and modeling of full geometries, as opposed to
specific features that reflect shape properties, but do not let
us to fully recover shape. For example, to model variation of
the shape of an organ within a population, it is desirable to
have models that allow us to analyze and infer full anatomies
from which one may extract local or global features for more
specific purposes.

In this paper, we construct a computational model of mul-
tidimensional shape equipped with geodesic metrics that al-
low quantification of shape similarity and difference. We de-
velop algorithms to calculate shape geodesics and geodesic
distances, as well as a tool to analyze regional shape sim-
ilarity and difference between a pair or group of shapes.
This type of tool allows us to address an important prob-
lem: *where do the most significant morphological contrasts
occur?* Our primary goals are to develop a framework for al-
gorithmic analysis of multidimensional shape and the basic
algorithms and tools needed in applications. We also demon-
strate the computational tractability of the model through

many examples and carry out experiments to illustrate some of its potential uses. However, we emphasize the development of the model and tools, leaving more extensive applications for future work.

Our approach leads to several variants: models that are sensitive or invariant to scale, as well as to orientation or orthogonal transformations. All of these variants are of potential interest and the most adequate choice for a problem will likely be guided by the context and application. For example, to model variation of the anatomy of an organ within a population with imaging techniques, one might want to employ a scale-invariant model in order to minimize or eliminate the influence of image acquisition parameters, or to disregard global size. On the other hand, to analyze atrophy of, say, the hippocampus using longitudinal magnetic resonance data, one might prefer to scale the whole brain to a template and use a scale-dependent model of the shape of the hippocampus to make it more sensitive to localized tissue loss. Figure 1 gives a preliminary illustration of a shape geodesic interpolating the contour surfaces of two octopuses constructed with one of the models of this paper.

The literature on shape is vast and comprises many different approaches. They include morphometric analysis based on specific shape descriptors, deformable templates, as well as use of a variety of representations and metrics, not necessarily derived from Riemannian structures. The choice is largely determined by the application and some of the non-Riemannian alternatives are well suited for problems such as shape indexing, classification, or retrieval. However, as we target problems that can be treated more effectively with the Riemannian approach, such as the development of statistical models of shape, we focus our discussion on this class of models. Pioneering work on shape due to Kendall and Bookstein led to the development of the first formal shape model using a representation by a finite indexed set of landmark points and a metric based on Procrustes alignment (Kendall 1984, 1999). Kendall's model has been employed in a variety of applications and also triggered numerous developments in statistics on manifolds (Dryden and Mardia 1998; Le and Kendall 1993; Bhattacharya and Patrangenaru 2003; Fletcher et al. 2004). As it is natural to perceive the shape of the contour of an object as a continuum, more recently, shape spaces of curves and surfaces are receiving a great deal of attention. A shape space of plane elastic curves was first proposed by Younes (1998, 1999); a reformulation of the model by Younes et al. (2007) yields a framework for efficient analysis and registration of both open and closed pla-

nar curves. A model that uses arc-length parametrizations of plane curves was developed by Klassen et al. (2004); as the curves are traversed with constant speed, the shape representation only uses directional information. This model can be computed quite efficiently (Schmidt et al. 2006), however, the lack of elasticity sometimes forces the model to rely on somewhat unnatural curve correspondences. Mio et al. (2007, 2009) constructed a shape space of elastic strings equipped with a 1-parameter family of geodesic metrics; planar curves were studied in Mio et al. (2007) and an extension to curves in Euclidean space of any dimension in Mio et al. (2009), where an alternative, more robust computational approach is employed. These metrics are indexed by a parameter that controls how much resistance a curve offers to stretching or compression relative to deformation by bending. The elastic model of Younes (1998), Younes et al. (2007) is isometric to a special case of the model of Mio et al. (2007, 2009), however, the formulation of Younes et al. (2007) for this special case has several computational advantages. The shape space of Joshi et al. (2007) is also equivalent to a special case of the general model of Mio et al. (2007, 2009), but in this case the computational costs for closed curves are comparable. In Mio et al. (2009), the arclength model of Klassen et al. (2004) is also interpreted as a limit case of elastic curves, as the resistance to stretching and compression becomes progressively larger. Other metrics in the space of curves have been investigated by Michor and Mumford (2007), including their organization and hierarchy. A model of shape of curves based on principles of linear elasticity that accounts for deformations interior to the curve has been developed by Fuchs et al. (2009); however, in the current formulation, the computations appear to be costly.

Progress on the study of shape of curves has been brisk, however, the development of Riemannian models of multidimensional shape is still rather incipient. Kilian et al. (2007) have proposed a model for parametric surfaces in 3D space, but the emphasis is on morphing and animating shapes in computer graphics. Liu et al. (2008, 2009) developed continuous analogues of the classical landmark model for parametric surfaces replacing Procrustes alignment in Euclidean space with alignment based on variants of the Sobolev metric. Thus far, the main applications of the model have been to analysis of magnetic resonance images of the human brain. The linear elasticity model of Fuchs et al. (2009) also applies to surfaces in 3D space; some examples of surface geodesics are presented in Fuchs et al. (2009).

The computational model developed in this paper employs a parametric representation of shape over an abstract simplicial complex K. A simplicial complex is a combinatorial structure that allows us to extend to more general topologies the familiar discretizations of curves as polygons and surfaces as triangle meshes. Thus, built-in to the parametric representation is the assumption that the connectivity of the meshes underlying the shapes to be compared are the same. Another point to be explained is related to shape registration. If $\alpha, \beta \colon K \to \mathbb{R}^p$ are parametric shapes in Euclidean space of dimension p, then for each vertex z of K, we think of the points $\alpha(z)$ and $\beta(z)$ as being in correspondence. Shape models based on parametric representations may be used in at least two different ways: (i) to quantify shape dissimilarity via the metric and also as a shape registration tool by selecting parametrizations that minimize the shape distance; (ii) to quantify shape dissimilarity using pre-computed parametrizations and correspondences. The latter is of interest in many applications such as problems in neuroimaging where registration techniques for specific anatomies that account for geometrical and/or biological features already have been developed and the data is fundamentally discrete, with no clear underlying continuous structure, for example, as in Morra et al. (2009), Liu et al. (2009). More details on this are provided in the examples of Sect. 5.5. In its present form, our computational model falls in the second category, with shape registration and analysis treated as related, but complementary problems. However, the discussion in Appendix C supports the view that the methods developed in this paper may also be interpreted as a step towards a new technique of shape registration.

We represent a parametric shape by its discrete exterior derivative because first-order representations tend to provide a good balance of robustness, computational tractability, and geometrical accuracy. Representations and metrics beyond order zero also can be more effective in capturing non-linear deformations such as bends, folds, regional expansions and contractions. The space of mappings $K \to \mathbb{R}^p$ is equipped with a family of Riemannian metrics indexed by a parameter that controls how much resistance an edge offers to deformation via stretching/compression relative to directional deformations. The choice of a parameter value may be based on experimentation with specific data, but an alternative is to consider "all" metrics. Philosophically, this viewpoint is analogous to a scale-space approach, which oftentimes encodes and reveals structure in more effective ways. The computational model we propose was motivated by a continuous model of shape of multidimensional Riemannian manifolds. As a detailed discussion of this model would take us too far afield, we only sketch the basic ideas with broad strokes in Appendix C.

The paper is organized as follows. In Sect. 2, we discuss the shape representation and examine invariance under shape-preserving transformations. In Sect. 3, we define

the shape metrics. The algorithm to calculate geodesics via energy minimization is developed in Sects. 4 and 5, with numerous illustrations and applications to the computation of mean shapes. Methods to detect and quantify regional morphological differences are discussed in Sect. 6. We thank Ying Wang for helping with the illustrations.

2 Shape Spaces

An *abstract finite simplicial complex K* is a classical abstraction of a more geometric notion of simplicial complex, which includes familiar objects such as polygonal representations of curves, triangle meshes that represent surfaces, and their higher dimensional analogues that are built out of vertices, edges, triangles, and higher simplexes, according to well established assembly rules. Formally, K is a finite collection of finite nonempty sets with the property that if A is in K, so is every nonempty subset of A. Intuitively, a vertex corresponds to an element of K containing a single point, an edge to an element of K containing 2 points, and a triangle to an element with 3 points. In general, an element A of K with $k + 1$ points is referred to as a k-simplex, and any nonempty subset of A is called a *face* of A. The requirement that every nonempty subset of a simplex A also be in K simply means that all faces of A should be in K. This also implies that the intersection of any two simplexes of K is either empty or a common face. The reader may consult (Munkres 1984) for more details. The vertex set of K is the union of all elements of K with a single point. We write the vertex set as $V = \{z_1, \ldots, z_r\}$. It follows that any k-simplex is of the form $\{z_{i_0}, \ldots, z_{i_k}\}$, with $z_{i_j} \in V$. We often orient an edge $\{z_i, z_j\}$ of K either as $e_{ij} = (z_i, z_j)$ or $e_{ji} = (z_j, z_i)$.

A parametric shape (modeled on K) in p-dimensional Euclidean space $\mathbb{R}^p$ is a mapping $\alpha \colon V \to \mathbb{R}^p$ defined on the vertex set. Although α is only defined on V, if $\{z_{i_0}, \ldots, z_{i_k}\}$ is a k-simplex of K, we think of the convex hull of $\{\alpha(z_{i_0}), \ldots, \alpha(z_{i_k})\}$ in $\mathbb{R}^p$ as a geometric realization of the k-simplex under α. For example, if $\{z_i, z_j\}$ is a 1-simplex (edge) of K, the line segment in $\mathbb{R}^p$ connecting $\alpha(z_i)$ and $\alpha(z_j)$ is a geometric edge of α. Likewise, abstract 2-simplexes give rise to triangles in $\mathbb{R}^p$.

2.1 Shape Representation

We employ a first-order representation of $\alpha \colon V \to \mathbb{R}^p$ so that the Riemannian model of shape will be more sensitive to non-linear deformations such as bends, folds, local expansions and contractions. This is done through the *discrete exterior derivative* also known as the *coboundary operator* in the simplicial homology literature (cf. Munkres 1984). The usual derivative of a mapping measures its infinitesimal variations. Since α is defined on the vertices of K, the natural discrete analogue is the variation measured as we move

from a vertex to an adjacent one. In other words, variations along the (oriented) edges of K. Thus, the discrete derivative $d\alpha$ along an oriented edge e is defined as

$$d\alpha(e) = \alpha(e^+) - \alpha(e^-), \tag{1}$$

where e^- and e^+ are the initial and terminal vertices of e, respectively. If we reverse the orientation of e, the variation of α gets multiplied by -1. Hence, it suffices to record the value of $d\alpha$ for one orientation of each edge. Therefore, we fix an orientation for each edge of K and let

$$E = \{e_1, \ldots, e_m\} \tag{2}$$

be the resulting set of oriented edges. The discrete derivative may be viewed as a mapping $d\alpha \colon E \to \mathbb{R}^p$. We only consider mappings α such that $d\alpha(e_i) \neq 0$, for every $e_i \in E$, and refer to them as *immersions*. This just means that no edge gets crushed to a single point under α. For each i, $1 \le i \le m$, we write the modular and directional components of $d\alpha(e_i)$ as

$$r_i = \log\|d\alpha(e_i)\| \quad \text{and} \quad v_i = d\alpha(e_i)/\|d\alpha(e_i)\|, \tag{3}$$

respectively, with the modular part expressed in logarithmic scale. Thus, the variation of α along the (oriented) ith edge of K is given by the vector $d\alpha(e_i) = e^{r_i} v_i$. Writing each v_i as a row vector $v_i = [v_{i1} \ \ldots \ v_{ip}]$, we represent α by the pair $(r, v) \in \mathbb{R}^m \times \mathbb{R}^{m \times p}$, where

$$r = \begin{bmatrix} r_1 \\ \vdots \\ r_m \end{bmatrix} \quad \text{and}$$

$$v = \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix} = \begin{bmatrix} v_{11} & \ldots & v_{1p} \\ \vdots & \ddots & \vdots \\ v_{m1} & \ldots & v_{mp} \end{bmatrix}. \tag{4}$$

Since the Riemannian metrics on $\mathbb{R}^m \times \mathbb{R}^{m \times p}$ to be used in the development of our shape model differs from the standard Euclidean metric, we will use the notation L for this space to emphasize this distinction. Thus far, the only constraints imposed on (r, v) are that the rows of v be unit vectors, which we write as

$$J_i(r, v) = \|v_i\|^2 = 1, \tag{5}$$

$1 \le i \le m$. We denote by N the subspace of L formed by all pairs (r, v) satisfying these m conditions.

2.2 Normalization of Scale

The (r, v)-representation of an immersion α is clearly invariant under translations. As discussed in the Introduction, both scale-invariant and scale-sensitive models of shape are of interest, the choice depending on the particular application. A simple way of normalizing scale is to fix the total edge length, say, to be unitary. In other words, to require that (r, v) satisfy

$$F(r, v) = \sum_{i=1}^{m} e^{r_i} = 1. \tag{6}$$

Remark We adopt normalization based on total edge length because it is simplest and it applies to arbitrary finite simplicial complexes. However, if K represents a manifold of dimension d, an alternative is to normalize scale by fixing the total d-volume (length if $d = 1$, area if $d = 2$, etc.) of the shape α. This normalization is more stable with respect to refinement of K to a simplicial complex of "higher resolution", say, via barycentric subdivision.

2.3 Integrability

An important step in our construction is the characterization of the pairs (r, v) that arise as the exterior derivatives of immersions. The integrability conditions will reveal the further constraints to be imposed on (r, v). If γ is an oriented path in K formed by a sequence of oriented edges, let the integer n_i denote the net number of times that the oriented edge $e_i \in E$ is traversed by γ, where a negative sign indicates reversal of orientation. The variation of a mapping α along γ may be expressed as

$$\sum_{i=1}^{m} n_i d\alpha(e_i) = \sum_{i=1}^{m} n_i e^{r_i} v_i. \tag{7}$$

The variation along any (oriented) cycle c in K clearly must vanish. Conversely, given (r, v) and an oriented path γ, define the integral of (r, v) along γ to be

$$I_\gamma(r, v) = \sum_{i=1}^{m} n_i e^{r_i} v_i. \tag{8}$$

If $I_c(r, v)$ vanishes along every cycle c, then (r, v) represents the exterior derivative of an immersion, unique up to translations. This can be seen as follows: fix a vertex z_i of K and a point $x_i \in \mathbb{R}^p$, and define $\alpha(z_i) = x_i$. For any vertex z_j, choose a path γ from z_i to z_j and let $\alpha(z_j) = x_i + I_\gamma(r, v)$. The vanishing condition over cycles ensures that $\alpha(z_j)$ is independent of the path chosen.

We need a computationally tractable way of checking whether $I_c(r, v) = 0$, for every cycle c. Clearly, there is a high amount of redundancy built-in to these conditions because the integral over many cycles can be calculated as a combination of integrals over more "primitive" ones. Our next goal is to identify such primitives. We first consider the cycles of length 3 formed by the oriented boundaries of the

2-simplexes (triangles) of K, where the orientation of each triangle is chosen arbitrarily. Standard cohomology theory arguments, cf. Munkres (1984), show that if $|K|$ is simply connected (or more generally, if the first Betti number of K is zero), then it is enough to check the vanishing of the integral over these cycles. One should note, however, that there is some redundancy even among the integrals along the boundaries of triangles. For example, consider the simple case where K is equivalent to a tetrahedron. If the integrals along the boundaries of 3 faces are zero, then the integral automatically vanishes on the boundary of the fourth face. Minimal families of triangles can be determined explicitly, but we omit further discussion of this because it is rather intuitive for the complexes K that arise most often in computer vision and neuroimaging. Several examples are given below.

Besides a minimal set of boundaries of triangles, we need to consider β_1 additional cycles, $\gamma_1, \ldots, \gamma_{\beta_1}$, where β_1 is the first Betti number of K (Munkres 1984). These cycles should be chosen to wrap around all "essential 1-dimensional holes" of K. (More precisely, they should represent a basis of the first simplicial homology group of K with real coefficients.) The integrability of (r, v) is equivalent to the vanishing of the integral along the boundaries of a minimal set of triangles and a collection of β_1 essential cycles.

Examples

(i) If the topology underlying K is that of a line segment, there are no triangles and $\beta_1 = 0$. As expected, in this case, there are no integrability conditions to be enforced because any (r, v) can be integrated to an arc.

(ii) If K has the topology a circle, then there are no triangles and $\beta_1 = 1$. Thus, we need a single cycle γ, the simplest choice being a cycle that goes once around the circle. In this case, the shape model of this paper coincides with the model of elastic strings studied in Mio et al. (2009).

(iii) If K has the topology of a closed orientable surface of genus g and K has T triangles, then we must consider the boundaries of $(T - 1)$ triangles and $2g$ additional cycles because the first Betti number is $\beta_1 = 2g$. Figure 2(a) shows two essential cycles of a surface of genus 1.

(iv) If K represents an orientable surface of genus g with $h > 0$ holes punctured, then all triangles are needed and $\beta_1 = 2g + h - 1$. Figure 2(b) shows an essential cycle on a surface of genus zero with two holes punctured.

(v) To construct the "robocat" of Fig. 2(c), we first punctured two holes on a surface of genus zero and then glued the two feet creating a non-manifold point. This produces an additional essential cycle in the leg area.
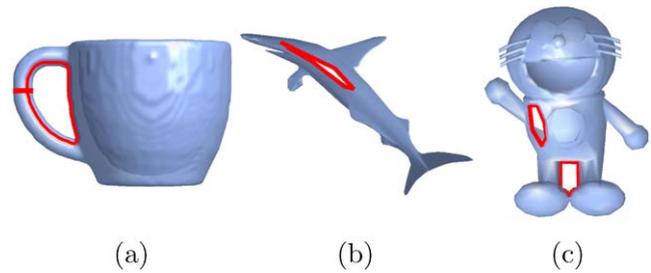

Fig. 2 Examples of minimal sets of essential cycles for different topologies

Let $F = \{T_1, \ldots, T_n\}$ be a minimal set of triangles of K needed to guarantee the vanishing of the integral along the boundaries of all 2-simplexes. Fix an orientation for each T_ℓ, $1 \le \ell \le n$, so that F becomes an oriented-face set. For each $T_\ell \in F$, let $\epsilon(\ell, i)$ be the number of times that the oriented edge $e_i \in E$ appears on the oriented boundary of T_ℓ. Most coefficients vanish, except for three, which correspond to the edges $e_{\ell_1}, e_{\ell_2}, e_{\ell_3}$ that form the boundary of T_ℓ. Clearly, $\varepsilon(\ell, \ell_i) = \pm 1$, $1 \le i \le 3$. We make the convention that $\ell_1 < \ell_2 < \ell_3$. With this notation, the integrability condition associated with T_ℓ may be expressed as

$$I_{\partial T_\ell}(r, v) = \sum_{i=1}^{m} \varepsilon(\ell, i) e^{r_i} v_i = 0. \tag{9}$$

For each $1 \le \ell \le n$, we rewrite this vector equation as a set of p scalar conditions

$$G_{\ell, j}(r, v) = \sum_{i=1}^{m} \varepsilon(\ell, i) e^{r_i} v_{ij} = 0, \tag{10}$$

$1 \le j \le p$, obtaining pn independent integrability conditions that (r, v) should satisfy. Likewise, fixing a family of essential cycles $\gamma_1, \ldots, \gamma_{\beta_1}$, let $\zeta(k, i)$ be the net number of times that $e_i \in E$ is traversed by γ_k. We obtain an additional set of $p\beta_1$ independent integrability conditions

$$H_{k, j}(r, v) = \sum_{i=1}^{m} \zeta(k, i) e^{r_i} v_{ij} = 0, \tag{11}$$

$1 \le k \le \beta_1$, $1 \le j \le p$. In conclusion, (r, v) represents the exterior derivative of an immersion α if and only if it satisfies the $p(n + \beta_1)$ integrability conditions described in (10) and (11).

2.4 Pre-Shape and Shape Spaces

We refer to pairs $(r, v) \in N$ satisfying conditions (6), (10) and (11) as *pre-shapes* and denote by P the space of all pre-shapes. A pre-shape gives a representation of immersions that is invariant to scale and translations. For a scale-sensitive variant, we simply drop condition (6) in the definition of pre-shapes. Thus, we have a nested sequence

$$P \subset N \subset L \tag{12}$$

of spaces, where P is the space of primary interest. The others will be useful at various stages of our computation of geodesics.

In some applications, we may want to consider mappings that differ by a rotation, or a more general orthogonal transformation, as representing the same shape. In such cases, we consider the action of the group $O(p)$ of $p \times p$ orthogonal matrices on pairs (r, v). If $U \in O(p)$ and α is an immersion, the result of applying the rigid transformation U to α is the composition $U \circ \alpha \colon V \to \mathbb{R}^p$. The induced action on (r, v) is $(r, v) \mapsto (r, vU^T)$. The action on the modular component is trivial because edge lengths do not change under U. Preshapes clearly get transformed into pre-shapes and the orbit of (r, v) is the set

$$[r, v] = \{(r, vU^T) : U \in O(p)\}. \tag{13}$$

We define the shape space $\mathcal{S}$ as the orbit space of P under the action of the orthogonal group. In other words, as the quotient space $\mathcal{S} = P/O(p)$.

3 Riemannian Metrics

We introduce Riemannian structures on the space L of all pairs (r, v) that will induce a family of pre-shape and shape metrics. The basic idea underlying the model is that the edges of a shape offer resistance to two different types of deformations, namely: (i) stretching and compression, which correspond to changes in edge length; (ii) bending, which is associated with directional changes. The energy needed to perform such deformations will be determined by two parameters, which we call the *tension* and *rigidity* coefficients. The geodesic distance is related to the minimum energy required to morph a shape into another through immersions.

Recall that a Riemannian structure on a manifold consists of a collection of inner products, one for each tangent space, that varies smoothly over the manifold. For each $(r, v) \in L$, we define

$$\langle (h, w), (\tilde{h}, \widetilde{w}) \rangle_{(r,v)} = a \sum_{i=1}^{m} h_i \tilde{h}_i e^{r_i} + b \sum_{i=1}^{m} (w_i \cdot \widetilde{w}_i) e^{r_i}, \tag{14}$$

where $a, b > 0$ are the tension and rigidity coefficients. If we think of (h, w) as an infinitesimal deformation of (r, v), the quantity

$$\|(h, w)\|^2_{(r,v)} = \langle (h, w), (h, w) \rangle_{(r,v)}$$
$$= a \sum_{i=1}^{m} h_i^2 e^{r_i} + b \sum_{i=1}^{m} \|w_i\|^2 e^{r_i} \tag{15}$$

may be interpreted as the energy cost of the deformation. The first and second terms of (15) account for the stretching/compression and bending energies, respectively. Henceforth, the submanifolds $P, N \subset L$ will have the Riemannian structure induced by (36). If (r, v) and $(\bar{r}, \bar{v})$ are pre-shapes, we let $\delta((r, v), (\bar{r}, \bar{v}))$ be the geodesic distance between them on the pre-shape submanifold P. If s and $\bar{s}$ are the shapes represented by the $O(p)$-orbits of (r, v) and $(\bar{r}, \bar{v})$, the shape distance is defined as

$$d(s, \bar{s}) = \inf_{U, V \in O(p)} \delta((r, vV^T), (\bar{r}, \bar{v}U^T))$$
$$= \inf_{U \in O(p)} \delta((r, v), (\bar{r}, \bar{v}U^T)). \tag{16}$$

The last equality follows from the fact that $O(p)$ acts by isometries; that is, $\delta((r, v), (\bar{r}, \bar{v})) = \delta((r, vU^T), (\bar{r}, \bar{v}U^T))$, for any $U \in O(p)$. As already pointed out, the use of the pre-shape or shape distance will depend on the particular application.

Much of the remaining work is devoted to the calculation of geodesics in P and the geodesic distance δ. The minimization over $O(p)$ appearing in (16) is a relatively simpler calculation. We briefly indicate how this minimization problem can be treated. We may obtain an initial estimate of

$$\hat{U} = \operatorname*{argmin}_{U \in O(k)} \delta((r, v), (\bar{r}, \bar{v}U^T)), \tag{17}$$

say, by minimizing $\|(r, v) - (\bar{r}, \bar{v}U^T)\|^2_{(r,v)}$. This is the same as minimizing $\sum_{i=1}^{m} \|v_i - \bar{v}_i U^T\|^2 e^{r_i}$, which is equivalent to maximizing $\sum_{i=1}^{m} (e^{r_i/2} v_i) \cdot (\bar{v}_i U^T)$. This problem is analogous to the one that arises in Procrustes alignment of shapes (Kendall 1984) and admits a closed form solution. Once this initial estimate is obtained, a gradient search over the orthogonal group $O(p)$ can be used to locally refine the estimation. As in Mio et al. (2009), the $O(p)$-gradient can be calculated numerically via finite differences in the Lie algebra of $O(p)$, which is formed by the $p \times p$ antisymmetric matrices.

4 Energy Minimization

We adopt an energy minimization approach to the estimation of minimal length geodesics in pre-shape space. Although this interpretation of geodesics is classical, the algorithmic aspects of the calculation pose numerous problems, both in terms of efficiency and robustness. We describe a computational strategy in the setting of more general Riemannian manifolds, not only pre-shape manifolds. This is done not just for the sake of generality, but because the method becomes more transparent. The proposed strategy is an extension of the energy minimization techniques used in the special case of shape of curves in Mio et al. (2009).

Let M be a Riemannian manifold with each tangent space T_pM, $p \in M$, equipped with the inner product $\langle, \rangle_p$. The associated norm is denoted $\|v\|_p = \langle v, v \rangle_p^{1/2}$, for $v \in T_pM$. Given a path $\gamma: I \to M$ defined over the interval $I = [0, 1]$, our goal is to develop an algorithm to deform γ to a geodesic maintaing the endpoints $\gamma(0) = p$ and $\gamma(1) = q$ fixed throughout the deformation. The energy of γ is defined as

$$E(\gamma) = \int_0^1 \langle \partial_t \gamma(t), \partial_t \gamma(t) \rangle_{\gamma(t)} \, dt = \int_0^1 \|\partial_t \gamma(t)\|_{\gamma(t)}^2 \, dt. \tag{18}$$

We are interested in paths of minimal energy as they correspond to minimal length geodesics (do Carmo 1994).

4.1 Path Variation

An infinitesimal deformation of a path $\gamma: I \to M$ can be described by a vector field $x(t) \in T_{\gamma(t)}M$, $t \in I$, along γ. We calculate the gradient of E with respect to the metric structure given by the following inner product of vector fields:

$$\langle x, y \rangle_\gamma = \langle x(0), y(0) \rangle_{\gamma(0)} + \int_0^1 \langle D_t x(t), D_t y(t) \rangle_{\gamma(t)} \, dt, \tag{19}$$

where D_t denotes *covariant differentiation* in M along γ (cf. do Carmo 1994). Covariant differentiation is differentiation of vector fields from a viewpoint intrinsic to the Riemannian manifold M. As we are interested in geodesics with prescribed endpoints, our goal is to minimize the energy on the subspace of paths satisfying the boundary conditions $\gamma(0) = p$ and $\gamma(1) = q$. The use of (19) leads to a computationally robust expression for the gradient of E and also let us easily enforce the boundary conditions, as further explained below. In pattern analysis, this type of metric in path space has been used in the construction of elastica (Mio et al. 2004), the evolution of Sobolev active contours (Sundaramoorthi et al. 2007), and the study of shape of curves (Joshi et al. 2007; Mio et al. 2009).

4.2 The Unconstrained Gradient

Given a path γ, we first calculate the unconstrained gradient of E at γ, that is, without enforcing the boundary conditions. Consider a variation $\gamma(t; \mu)$ of γ along a direction $x(t)$, where $\mu \in (-\epsilon, \epsilon)$ is the variation parameter. This means that $\gamma(t; 0) = \gamma(t)$ and

$$x(t) = \frac{\partial}{\partial \mu} \gamma(t; \mu)\big|_{\mu=0}. \tag{20}$$

Differentiating (18) at $\mu = 0$, the directional derivative of E at γ in the direction x may be expressed as

$$dE_\gamma(x) = 2 \int_0^1 \langle D_t x(t), (\partial_t \gamma(t))_{\gamma(t)} \, dt. \tag{21}$$

To rewrite this expression in terms of the inner product (19), let F be a vector field along γ such that $D_t F(t) = 2 \partial_t \gamma(t)$ and $F(0) = 0$. In other words, a covariant integral of the field $f(t) = 2 \partial_t \gamma(t)$ that vanishes at $t = 0$. Then, (21) becomes $dE_\gamma(x) = \langle x, F \rangle_\gamma$, which means that the gradient of E at γ is

$$\nabla E(\gamma) = F. \tag{22}$$

Numerically, the fact that $\nabla E(\gamma)$ is obtained via integration (not differentiation) of the tangent field $\partial_t \gamma$ is one of the key advantages of using the proposed metric structure on the space of paths. This leads to a more stable calculation and robust algorithm as integration may be viewed as a smoothing operation.

4.3 The Boundary Conditions

To calculate the constrained gradient $\nabla^* E(\gamma)$, we need to understand which vector fields along γ represent variations of the curve that are orthogonal to those that respect the boundary conditions. Then, $\nabla^* E(\gamma)$ can be calculated from $\nabla E(\gamma)$ by subtracting the component orthogonal to the subspace of variations that are compatible with the boundary conditions.

Clearly, vector fields that represent variations that keep the endpoints of γ fixed are those that vanish at $t = 0$ and $t = 1$. A simple (covariant) integration-by-parts argument, which is presented in Appendix B for completeness, shows that the vector fields that are orthogonal to those that vanish at the endpoints are precisely the covariantly linear fields. In other words, the smooth fields with trivial second covariant derivative along γ.

By construction, the field $F = \nabla E(\gamma)$ is zero at $t = 0$. Thus, to obtain $\nabla^* E(\gamma)$, we just need to subtract from F the covariantly linear field that vanishes at $t = 0$ and coincides with F at $t = 1$. To construct such a field, again, we use covariant integration, this time applied to the reverse of γ. First, construct a parallel field by integrating the everywhere zero field with initial condition $F(1)$ along the reverse of the path γ. (Parallel fields are those with zero covariant derivative so the construction of parallel fields is a special case of covariant integration.) Reversing the path again, after integration, we obtain a parallel field G along γ whose value at $t = 1$ is $F(1)$. The field $t \mapsto tG(t)$, $t \in I$, is covariantly linear with the desired properties implying that

$$\nabla^* E(\gamma) = \nabla E(\gamma) - tG. \tag{23}$$

4.4 Algorithm

The arguments of Sects. 4.2 and 4.3 show that, in practice, it is possible to calculate $\nabla^* E(\gamma)$ provided that we can perform covariant integration of vector fields, as this is needed for: (i) the calculation of $\nabla E(\gamma)$; (ii) the construction of the parallel field used for the orthogonal projection of $\nabla E(\gamma)$ onto $\nabla^* E(\gamma)$. To implement gradient descent, we also need an ϵ-update rule for paths in M along directions prescribed by a vector field. These issues will be discussed in detail below for the pre-shape space P. For a general Riemannian manifold M for which these ingredients are available, we obtain the following algorithm to deform a path $\gamma: I \to M$ to a geodesic.

Let $\epsilon, \delta > 0$ be (small) real numbers:

(i) Integrate the field $f(t) = 2\partial_t \gamma(t)$ covariantly along the path γ with zero initial condition. The integral field F gives the gradient $\nabla E(\gamma)$.

(ii) Construct a parallel field G along γ that coincides with $F(1) \in T_{\gamma(1)}M$ at $t = 1$. By (23), $\nabla^* E(\gamma) = \nabla E(\gamma) - tG$.

(iii) ϵ-update the path γ in the direction of the vector field $-\nabla^* E(\gamma)$ to obtain a path γ_{new}.

(iv) Iterate the process until $\|\gamma_{\text{new}} - \gamma\|_\gamma < \delta$, where $\|\cdot\|_\gamma$ denotes the norm associated with the inner product defined in (19).

5 Pre-Shape Geodesics

The calculation of covariant integrals in the pre-shape manifold P can become costly as the complexity of the reference simplicial complex K increases. This happens in many important instances, for example, high resolution meshes often used in medical imaging and computer vision. To be more specific, if a path γ in P is discretized by subdividing it into T parts and the number of independent constraints that characterize pre-shapes is c, then each covariant integration along γ requires setting up and solving T linear systems with coefficient matrices of size $c \times c$. Even though the matrices of these systems are symmetric and sparse, this could limit the applicability of the algorithm of Sect. 4.4 in a significant way. The number c of independent constraints that characterize pre-shapes was discussed in Sect. 2.1 and is typically large. For example, for surface meshes, c is of the order of the number of vertices of K. To overcome this obstacle, we propose a variant of the algorithm that replaces gradient descent for the energy of paths in P with its counterpart for paths in N followed by a projection of the path onto P. Recall that, in the space N, we only require that each row of v be a unit vector (see (5)). Since we will use first-order integration of the differential equations that govern covariant integration, to justify this approach, we would

need the two operations to be interchangeable to first order. However, the cost of computing a projection map with this property is comparable to the cost of covariant integration in P, so we employ a greedy strategy by relaxing the condition on the projection map. In the special case of curves in the plane or 3D space, where both procedures can be implemented efficiently, in previous work we experimented with both approaches obtaining nearly identical results (Mio et al. 2009). Experimentation with low resolution surface meshes also leads to much the same results. For high-resolution or multi-dimensional shapes, however, the computational gains are substantial. We begin the "greedy" calculation of pre-shape geodesics with a description of the projection map.

5.1 Projection Map

We introduce a projection that maps a pair $(r, v) \in N$ near P onto P. For the present purposes, it suffices to consider pairs (r, v) in the vicinity of P because we will start with a path γ in P, perturb it infinitesimally in N, and project it back onto P. In (6), (10) and (11), we introduced the constraint functions F, $G_{\ell,j}$, and $H_{k,j}$, $1 \le j \le p$, $1 \le \ell \le n$, $1 \le k \le b_K$, that define pre-shapes. Clearly, the residual functions $\rho(r, v) = 1 - F(r, v)$, $\rho_{\ell,j}(r, v) = -G_{\ell,j}(r, v)$, and $\sigma_{k,j}(r, v) = -H_{k,j}(r, v)$ have the property that their simultaneous vanishing is equivalent to (r, v) being a pre-shape. Given (r, v) near P, we use Newton's method to find a nearby zero of

$$J(r, v) = \frac{1}{2}\rho^2(r, v) + \frac{1}{2}\sum_{\ell,j}\rho_{\ell,j}^2(r, v) + \frac{1}{2}\sum_{k,j}\sigma_{k,j}^2(r, v)$$

(24)

and project (r, v) to the pre-shape represented by that zero.

A calculation shows that if we do not impose the condition that each row of v must be a unit vector, then the negative (unconstrained) gradient of J is

$$-\nabla J(r, v) = \rho(r, v)\nabla F(r, v) + \sum_{\ell,j}\rho_{\ell,j}(r, v)\nabla G_{\ell,j}(r, v)$$

$$+ \sum_{k,j}\sigma_{k,j}(r, v)\nabla H_{k,j}(r, v).$$

(25)

Explicit formulae for $\nabla F(r, v)$, $\nabla G_{\ell,j}$, and $\nabla H_{k,j}(r, v)$ are given in Appendix A.

We write the modular and directional parts of the (negative) gradient as $-\nabla J(r, v) = (h, w^*)$. To enforce the constraint $\|v_i\|^2 = 1$, we need to make w_i^* tangential to the unit sphere in $\mathbb{R}^p$ at v^i. For this, we replace w^* with w, where the ith row of w is $w_i = w_i^* - (w_i^* \cdot v_i)v_i$. Letting

$\epsilon(r, v) = J(r, v)/\|(h, w)\|^2_{(r,v)}$, we update (r, v) as follows:

$$\begin{cases} r = r + \epsilon h, \\ v_i = \cos(\epsilon \|w_i\|)v_i + \sin(\epsilon \|w_i\|)\frac{w_i}{\|w_i\|}, & \text{if } w_i \neq 0, \end{cases} \tag{26}$$

and v_i stays unchanged, otherwise. This spherical update of v_i ensures that v_i remains a unit vector. We iterate the process until $J(r, v)$ becomes small.

5.2 Covariant Integration

A path γ in N will be denoted $\gamma(t) = (r(t), v(t))$. Similarly, we separate the modular and directional components of a vector field along γ and write it as $(f(t), x(t))$. In Appendix B, we show that covariant integrals of $(f(t), x(t))$ are vector fields $(F(t), X(t))$ along γ that satisfy the system of differential equations

$$\begin{cases} \partial_t F_i(t) = f_i(t) - \frac{1}{2}[\partial_t r_i(t)]F_i(t) \\ \qquad + \frac{1}{2}\frac{b}{a}[X_i(t) \cdot \partial_t v_i(t)] \\ \partial_t X_i(t) = x_i(t) - \frac{1}{2}[X_i(t)\partial_t r_i(t) + F_i(t)\partial_t v_i(t)] \\ \qquad - [X_i(t) \cdot \partial_t v_i(t)]v_i(t), \end{cases} \tag{27}$$

for $1 \le i \le p$. Given $(f(t), x(t))$ along γ and the initial condition $(F(0), X(0)) = (F_0, X_0)$, we use Euler's method to estimate the covariant integral field $(F(t), X(t))$. Note that parallel transport, which is needed in the computation of geodesics, corresponds to the special case of covariant integration with $(f(t), x(t))$ everywhere zero.

5.3 Pre-Shape Geodesic Algorithm

We now modify the algorithm of Sect. 4.4 to estimate a geodesic between given pre-shapes (r, v) and (r^*, v^*). To initialize the process, let α, α^* be immersions of K into $\mathbb{R}^p$ represented by (r, v) and (r^*, v^*), respectively. We linearly interpolate α and α^* to obtain a 1-parameter family of mappings $\alpha_t \colon K \to \mathbb{R}^p$. If necessary, we gently deform α_t to make $d\alpha_t(e) \neq 0$ for every oriented edge e and $0 < t < 1$. After normalizing scale (if considering the scale-invariant model), the path $\gamma(t) = (r(t), v(t))$ obtained from the log-polar representation of $d\alpha_t$ is a path in pre-shape space and can be used to initialize the search.

Let $\epsilon, \delta > 0$ be (small) real numbers:

(i) Initialize the search with a path $\gamma(t) = (r(t), v(t))$ in pre-shape space constructed, for example, as described above.

(ii) Let $(f(t), x(t)) = 2(\partial_t r(t), \partial_t v(t))$. Using (27), integrate this field covariantly along the path γ with zero initial condition. The integral field $(F(t), X(t))$ gives the gradient $\nabla E(\gamma)$ of the energy as a functional on paths in N.

(iii) Using (27), construct a parallel field $(G(t), Y(t))$ along γ that agrees with $(F(1), X(1))$ at $t = 1$. By (23), the constrained gradient of E is $\nabla^* E(\gamma) = (F(t), X(t)) - (tG(t), tY(t))$.

(iv) Write the modular and the directional components of $\nabla E(\gamma)$ as $(h(t), w(t))$, $0 \le t \le 1$. Update $\gamma(t) = (r(t), v(t))$ as a path in N according to

$$\begin{cases} \tilde{r}(t) = r(t) - \epsilon h(t); \\ \tilde{v}_i(t) = v_i(t), & \text{if } w_i(t) = 0; \\ \tilde{v}_i(t) = \cos(\epsilon \|w_i(t)\|)v_i(t) - \sin(\epsilon \|w_i(t)\|)\frac{w_i(t)}{\|w_i(t)\|}, \\ \qquad \text{otherwise.} \end{cases}$$

Note that the update of $v_i(t)$ takes place along great circles of the unit sphere in $\mathbb{R}^k$ ensuring that each $\tilde{v}_i(t)$ is a unit vector.

(v) Project each $(\tilde{r}(t), \tilde{v}(t))$ onto the pre-shape space P to obtain a path $(r_{\text{new}}(t), v_{\text{new}}(t))$ in P.

(vi) Iterate the process until $\|(r_{\text{new}}(t), v_{\text{new}}(t)) - (r(t), v(t))\|_\gamma < \delta$, where $\|\cdot\|_\gamma$ denotes the norm associated with (19).

5.4 Discretization

We are employing a discrete representation of pre-shapes derived from the exterior derivative of a mapping $\alpha \colon V \to \mathbb{R}^p$. Thus, the only parameter that remains to be discretized is the "time" parameter of a path $\gamma(t) = (r(t), v(t))$. This is done in a standard way by sampling the unit interval $I = [0, 1]$ uniformly at points $0 = t_0 < t_1 < \cdots < t_k = 1$. To simplify notation, we write the various points $\gamma(t_j)$ on the discrete path as $\gamma(j) = (r(j), v(j))$. The partial derivatives $\partial_t \gamma(j) = (\partial_t r(j), \partial_t v(j))$ needed in the calculation of geodesics are computed via finite differences and defined for $0 \le j \le k - 1$.

The covariant integral of a discrete vector field $(f(j), x(j))$ along γ, $0 \le j \le k - 1$, with prescribed initial condition (F_0, X_0) is calculated by integrating (27) with Euler's method. The integral field $(F(j), X(j))$ is defined for $0 \le j \le k$. All other steps in the algorithms only involve fixed values of t so they are identical in the discrete version.

5.5 Examples of Geodesics

Figure 3 shows two examples of shape geodesics calculated with coefficients $a = b = 0.5$. In both cases, we started with a single spherical shape, punctured two holes, and for the "robocat" we also attached the two feet at a point to make the topology more complex. Then, we "manually" deformed the shape to obtain a second shape with the same underlying mesh structure. These are shown as the first and last frames of each row. In each case, the reference complex K is the abstract simplicial complex underlying the shapes, so that

they admit a natural parametrization over K. The topology of the modified robocat has two (classes of) essential cycles and the shark only one, as shown in Fig. 2.

Figure 4 shows a shape geodesic between cortical surfaces and a geodesic interpolating hippocampal surfaces. The surface meshes were extracted from magnetic resonance images and correspondences calculated with the direct mapping techniques of Shi et al. (2007a, 2007b). In each case, one of the surfaces was re-meshed by transferring the mesh structure of the other surface via the correspondence so that they both have the same underlying simplicial structure. The mesh representing the cortices have 122,880 edges and the hippocampal meshes have 30,720 edges. The geodesics were calculated with parameter values $a = 0.3$, $b = 0.7$ for the cortices and $a = 0.15$, $b = 0.85$ for the hippocampal surfaces.

5.6 Mean Shape

To further illustrate the methods of this paper, we apply the shape model to the construction of anatomical atlases of the cortex and hippocampus as sample mean shapes. We begin with a brief review of the definition of sample Fréchet mean shape and a gradient descent strategy to compute the mean of samples on Riemannian manifolds. Let $s_1, \ldots, s_n$ be parametric shapes represented by the pairs (r_i, v_i), $1 \le i \le n$. For a shape s represented by (r, v), the total spread of the data with respect to s is

$$V(s) = \sum_{i=1}^{n} d^2((r, v), (r_i, v_i)). \tag{28}$$

A shape that minimizes V is called a sample Fréchet mean shape. Karcher showed that if (h_i, w_i) represents the initial velocity of the geodesic that starts at (r, v) and reaches (r_i, v_i) in unit time, then the gradient of V at (r, v) is $\sum (r_i, v_i)$ (Karcher 1977). Thus, our algorithm to compute geodesics allows us to estimate mean shapes via gradient descent. This approach has been used to calculate mean shape in numerous different contexts, for example, Dryden and Mardia (1998), Srivastava et al. (2005), Liu et al. (2009). We also should mention that the method of Fletcher et al. (2008) for estimating median shapes readily adapts to this context, as well.

Figure 5 shows the mean shape of 7 cortical surfaces of the left hemisphere extracted from magnetic resonance images. One of the shapes was chosen as a reference and the remaining six were registered with it by the method of Shi et al. (2007a). As in the computation of geodesics discussed above, the six surfaces were re-meshed via the correspondence, so that all shapes have the same underlying spherical mesh structure with 122,880 edges. The mean was calculated with $a = b = 0.5$. Figure 6 shows a hippocampal atlas computed in a similar manner as the mean shape of 50 samples, using meshes with 30,720 edges. For registration of the hippocampal surfaces, we used the algorithm of Shi et al. (2007b).

6 Energy Profile and Energy Density

A (parametric) geodesic $\gamma(t) = (r(t), v(t))$, $0 \le t \le 1$, has parallel velocity field. In particular, it is traversed with constant speed ω, where ω is the length of γ. Thus, the energy

Fig. 3 Two shape geodesics calculated with $a = b = 0.5$ with a model sensitive to scale
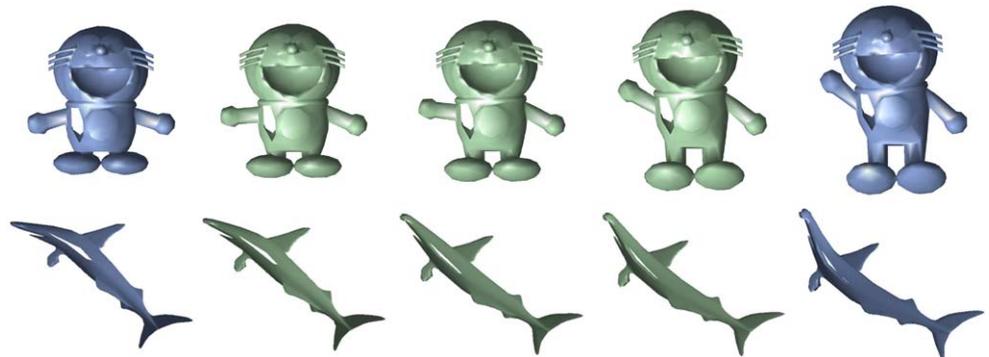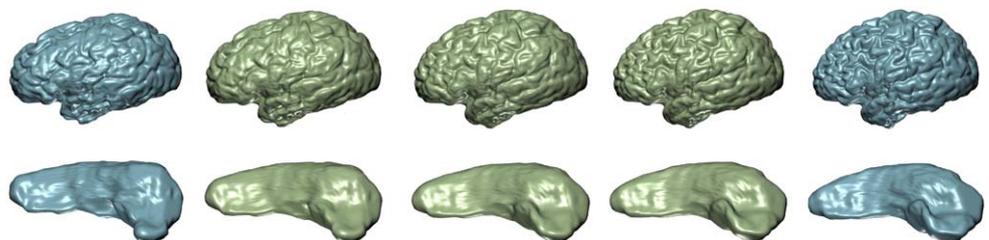

Fig. 4 Scale-invariant geodesic interpolations of cortical surfaces with $a = 0.3$ and $b = 0.7$ and hippocampal surfaces with $a = 0.15$ and $b = 0.85$

of γ is

$$E(\gamma) = \int_0^1 \|\partial_t \gamma(t)\|_{\gamma(t)}^2 \, dt = \omega^2. \tag{29}$$

On the other hand, we may write the energy as

$$E(\gamma) = \sum_{i=1}^m \int_0^1 (a|\partial_t r_i(t)|^2 e^{r_i(t)} + b\|\partial_t v_i(t)\|^2 e^{r_i(t)}) \, dt, \tag{30}$$

which expresses $E(\gamma)$ as a sum of the contributions of the individual edges. Thus, the *energy profile function* (EPF) defined as

$$\psi(e_i) = \int_0^1 (a|\partial_t r_i(t)|^2 e^{r_i(t)} + b\|\partial_t v_i(t)\|^2 e^{r_i(t)}) \, dt \tag{31}$$

quantifies the contribution of each edge to the total geodesic deformation energy. Similarly, we define the *energy density function* (EDF) by

$$\rho(e_i) = \frac{1}{\omega^2} \int_0^1 (a|\partial_t r_i(t)|^2 e^{r_i(t)} + b\|\partial_t v_i(t)\|^2 e^{r_i(t)}) \, dt, \tag{32}$$

which measures the fraction of the total deformation energy associated with the ith edge. Although geodesic distance is a global quantifier of shape difference, the EPF and EDF provide a means to measure local shape differences and identify the regions where shape similarity and divergence are most pronounced. Despite the fact that, for a single shape geodesic, the EPF and EDF just differ by a multiplicative constant, there is a qualitative difference between them. To explain this point, suppose the EPF of a geodesic indicates

that the main deformations of a shape are localized to a region R. If this trend is present in the early stages of the deformation, although the difference between the EPFs of the full geodesic and the geodesic that describe the initial deformation can be significant, the EDFs will bear much higher similarity and better capture the early trends of a deformation that takes place primarily over the region R. This characteristic of the EDF might become useful, for example, in applications such as uncovering morphological signatures that can help in early detection of tissue loss characteristic of neurodegenerative diseases. The EPF, in turn, is likely to be more useful in applications where it is important to quantify regional morphological changes in more absolute terms, for example, to compare observations of a population of shapes. One may further decompose the local energy into its tension and rigidity components to separately quantify the local shape differences due to stretching and bending. One may also modify ψ and ρ to a function defined on the vertex set by letting the value on a vertex be the average value of ρ on the edges incident with that vertex.

As a first illustration, Fig. 7 displays the initial shapes of the geodesics in Fig. 3 overlaid with intensity maps of the respective energy profile functions. The plots are consistent with the main localized differences observed in the data used in the calculation of the geodesics of Fig. 3. The main visible differences in the robocats are the position of the arms and length of the legs and these are the areas where the EPF has highest intensity. A more localized difference in the nose
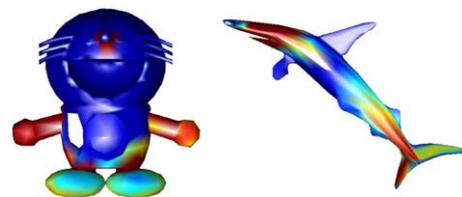

Fig. 7 Plots of the energy density functions for the geodesics in Fig. 3

Fig. 5 The mean of 7 cortical surfaces of the left hemisphere calculated with $a = b = 0.5$. The mean shape is highlighted on the last panel
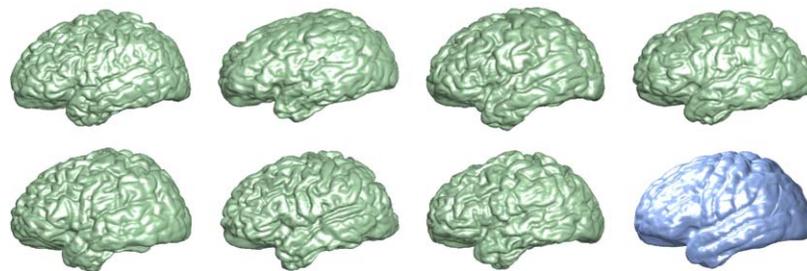

Fig. 6 The last panel highlights the mean of 50 right hippocampal surfaces calculated with $a = b = 0.5$. The other panels display 9 samples from the group of 50 shapes
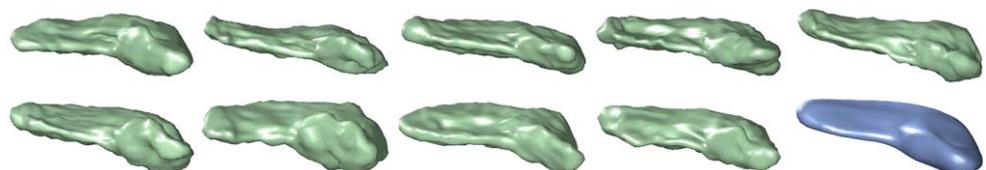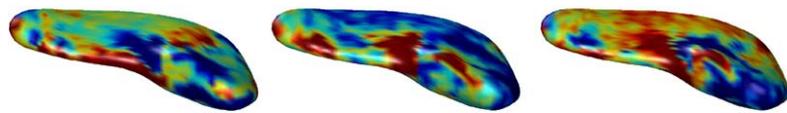
area that is not so apparent to the naked eye is also captured but the EPF. Similarly, the EPF for the shark is strongest near the head and tail because in one the shapes these regions are bent downward and bent upward on the other.

Now, we use energy profile functions to quantify regional shape variation observed in the 50 hippocampal surfaces of Sect. 5.6. For each of the 50 samples, we calculated the EPF associated with the geodesic from the group atlas (mean shape) to the sample. These energy profile functions may be viewed as functions defined on the surface of the atlas. This functional data can be analyzed with standard techniques such as principal component analysis. Figure 8 shows the first three principal modes of variation plotted over the atlas, which allow us to visualize the regions where shape variation about the mean is most pronounced as measured by the model.

7 Summary and Discussion

We developed a computational model of multidimensional shape, within the framework of Riemannian geometry, using a parametric representation of shape over a finite simplicial complex. We constructed a shape space equipped with a 1-parameter family of geodesic metrics derived from Riemannian structures that decouple the energy cost of an infinitesimal deformation of an edge of a mesh by changes in magnitude and direction. The parameter in the model controls how much change in direction is penalized relative to change in magnitude. The choice of parameter value can be based on experimentation with specific data, but we also suggested the possibility of working with a summary of "all" parameter values, as this may help to elucidate and reveal finer aspects of shape difference. We proposed the use of the energy profile and energy density functions of a geodesic as tools to identify regions where shape difference and similarity are most salient. We also developed algorithms to calculate geodesic paths, geodesic distances, and the associated energy profile and density functions.

We illustrated the applicability of the methods with the calculation of geodesics between contour surfaces of multiple objects, as well as between pairs of cortical and hippocampal surfaces represented by high resolution meshes segmented from magnetic resonance images of the human brain. Geodesic distances, geodesic interpolations, and models of shape variation developed in shape space enable us to quantify and visualize anatomical resemblance and divergence between individuals and across populations. We applied the methods developed to the construction of a cortical

atlas, as well as an atlas of the right hippocampus as sample Fréchet mean shapes.

The characterization and quantification of global and regional morphological differences are important problems in many contexts, for example, in the study of anatomical changes in brain structures due to neurodegenerative diseases. In forthcoming work, the methodology developed will be applied to large comparative studies of morphological changes in the brain associated with normal aging and progression of neurodegenerative disease, as well as other problems in neuroimaging.

The model of this paper employs a parametric representation of shape over a finite simplicial complex. However, the method also applies to other combinatorial structures such as parametrizations over cubical meshes. Such discrete representations may be more natural, for example, in the study of shape of solids in 3D space such as the entire volume of the hippocampus. A sketch of a continuous model that motivated the constructions of this paper is presented in Appendix C. A continuous limit is of interest for many reasons. For example, it integrates discrete models developed over different meshes as approximations to a common model and it might help to interpret our shape metrics as a step towards a new approach to registration of multidimensional shape. As the detailed investigation of these aspects of the problem falls beyond the intended scope of this paper, these problems are left for future work.

Appendix A: Pre-Shape Constraints

Differentiating (6), (10) and (11), one can show that the gradients with respect to the inner product (19) of the functions F, $G_{\ell,j}$, and $H_{k,j}$ that define pre-shapes are as follows:

(i) The modular component of $\nabla F(r, v)$ is the constant vector $[1/a \ldots 1/a]^T$ and the directional component is the zero matrix.

(ii) The modular part of $\nabla G_{\ell,j}(r, v)$ is $\frac{1}{a}[\varepsilon(\ell, 1)v_{1,j} \ldots \varepsilon(\ell, m)v_{m,j}]^T$ and the jth column of the directional component is $\frac{1}{b}[\varepsilon(\ell, 1) \ldots \varepsilon(\ell, m)]^T$. The other columns of the directional component are zero. These two vectors may have nonzero entries only in rows ℓ_1, ℓ_2 and ℓ_3.

(iii) The modular part of $\nabla H_{k,j}(r,v)$ is $\frac{1}{a}[\zeta(k,1)v_{1,j} \dots$
$\zeta(k,m)v_{m,j}]^T$ and the jth column of the directional
component is $\frac{1}{b}[\zeta(k,1) \dots \zeta(k,m)]^T$. The other
columns of the directional component are zero.

Appendix B: Covariant Integration

For $(z,y) = (z,y_1,\dots,y_p) \in \mathbb{R} \times \mathbb{R}^p$, we use the sub-
script 0 to identify the z-coordinate and the subscript j,
$1 \le j \le p$, for the coordinate y_j. Given $a, b > 0$, define a
Riemannian metric on $\mathbb{R} \times \mathbb{R}^p$ whose metric tensor at (z,y)
is $g_{00}(z,y) = ae^z$, $g_{jj}(z,y) = be^z$, and 0, otherwise. Then,
the Riemannian structure on L, defined in Sect. 3, is isomet-
ric to the Cartesian product of these $(p+1)$-dimensional
models over $1 \le i \le m$. Thus, to derive the differential equa-
tion that governs covariant integration along a path in N, it
suffices to derive the corresponding differential equation for
covariant integration in $\mathbb{R} \times \mathbb{S}^{p-1}$ with respect to the induced
metric. The Christoffel symbols of the Levi-Civita connec-
tion on $\mathbb{R} \times \mathbb{R}^p$ are $\Gamma_{00}^0 = \Gamma_{0j}^j = \Gamma_{j0}^j = 1/2$, $\Gamma_{jj}^0 = -b/2a$,
$1 \le j \le p$, and zero otherwise. Therefore (cf. do Carmo
1994), the covariant derivative of a vector field (F^t, X^t)
along a path (z^t, y^t) in $\mathbb{R} \times \mathbb{R}^k$ is given by

$$\begin{cases} D_t F(t) = \partial_t F(t) + \frac{1}{2}(\partial_t z(t)) F(t) - \frac{1}{2}\frac{b}{a}[X(t) \cdot \partial_t y(t)] \\ D_t X(t) = \partial_t X(t) + \frac{1}{2}(X(t)\partial_t z(t) + F(t)\partial_t y(t)). \end{cases}$$
(33)

If $(z(t), y(t))$ is a path in $\mathbb{R} \times \mathbb{S}^{p-1}$, then fields that are tan-
gential to $\mathbb{R} \times \mathbb{S}^{p-1}$ are those that satisfy the additional or-
thogonality condition $F(t) \cdot y(t) = 0$. Thus, if $(f(t), x(t))$
and $(F(t), X(t))$ are both tangential to $\mathbb{R} \times \mathbb{S}^{p-1}$, the fact
that the covariant derivative of $(F(t), X(t))$ in the submani-
fold $\mathbb{R} \times \mathbb{S}^{p-1}$ is $(f(t), x(t))$ may be rephrased as $D_t F(t) =$
$f(t)$ and $D_t X(t) = x(t) + \tau(t)y(t)$, where $\tau(t)$ is a scalar
field to be determined. Substituting in (33), we obtain

$$\begin{cases} \partial_t F(t) = f(t) - \frac{1}{2}\partial_t z(t) F(t) + \frac{1}{2}\frac{b}{a}[X(t) \cdot \partial_t y(t)] \\ \partial_t X(t) = x(t) - \frac{1}{2}(X(t)\partial_t z(t) + F(t)\partial_t y(t)) + \tau(t)y(t). \end{cases}$$
(34)

Differentiating $X(t) \cdot y(t) = 0$, we get $\partial_t X(t) \cdot y(t) =$
$-X(t) \cdot \partial_t y(t)$. From (34), it follows that $\tau(t) = -X(t) \cdot$
$\partial_t y(t)$, where we used the facts that $X(t) \cdot y(t) = 0$ and
$\partial_t y(t) \cdot y(t) = 0$. Substituting this value of $\tau(t)$ in (34), we
obtain (27).

Appendix C: Sketch of a Continuous Model

Let M be a closed, n-dimensional Riemannian manifold;
that is, a compact n-manifold without boundary such that

each tangent space $T_x M$, $x \in M$, is equipped with an inner
product $\langle , \rangle_x$ that varies smoothly over M. If $\phi : M \to \mathbb{R}^p$ is
a smooth map, the differential $d\phi$ may be viewed as an $\mathbb{R}^p$-
valued differential 1-form on M. This 1-form is the assem-
bly of the pointwise differentials $d\phi_x : T_x M \to \mathbb{R}^p$, $x \in M$.
Each $d\phi_x$ is a linear mapping $T_x M \to \mathbb{R}^p$. If $h \in T_x M$,
$d\phi_x(h)$ measures the rate of variation of ϕ at x in the di-
rection h. We assume that $n < p$ and ϕ is an immersion;
that is, each $d\phi_x$ is non-singular. Using a polar decompo-
sition, we express $d\phi_x$ uniquely as a composite mapping
$d\phi_x = U_x \circ S_x$, where $S_x : T_x M \to T_x M$ is a positive, self-
adjoint operator and $U_x : T_x M \to \mathbb{R}^p$ is a linear isometric
embedding. In other words, $U_x^* \circ U_x = I_x$, where U^* is the
adjoint of U and I_x is the identity map of $T_x M$. Letting
$A_x = \log S_x$ be the operator logarithm, we identify $d\phi_x$ with
the pair (A_x, U_x). Under this representation,

$$d\phi_x = U_x \circ e^{A_x},$$
(35)

where A_x is a self-adjoint (not necessarily positive) operator.
Let $\text{Sym}(M)$ be the total space of the vector bundle over M
whose fibre over $x \in M$ is the space of self-adjoint operators
$T_x M \to T_x M$ and $\text{Iso}(M, p)$ be the total space of the bundle
whose fibre over $x \in M$ is the space $\text{Iso}(T_x M, \mathbb{R}^p)$ of lin-
ear isometric embeddings $T_x M \to \mathbb{R}^p$. Then, the differen-
tial $d\phi$ may be identified with the pair (A, U), where A and
U are the sections of the bundles $\text{Sym}(M)$ and $\text{Iso}(M, p)$
given by $A(x) = A_x$ and $U(x) = U_x$, respectively.

Next, we introduce a Riemannian structure on the space
of pairs (A, U). Let $H, \tilde{H}$ and $W, \tilde{W}$ represent infinitesimal
deformations of A and U. Define

$$\langle (H, W), (\tilde{H}, \tilde{W}) \rangle_{(A,U)} = a \int_M \langle H_x e^{A_x}, \tilde{H}_x \rangle dV(x)$$
$$+ b \int_M \langle W_x e^{A_x}, \tilde{W}_x \rangle dV(x), \quad (36)$$

where $a, b > 0$, dV is the volume form of M and the in-
ner products on the integrands are operator inner products.
For simplicity, we consider the scale-dependent model, in
which case, the only constraints on (A, U) are the integra-
bility conditions. Motivated by (35), let ω be the $\mathbb{R}^p$-valued
1-form on M whose value at $x \in M$ is $\omega(x) = U_x \circ e^{A_x}$.
The first integrability condition is $d\omega = 0$, where d denotes
exterior differentiation. The vanishing of $d\omega$ translates to
ω being a closed 1-form. If the first Betti number of M is
zero, this is the only constraint needed on (A, U). Other-
wise, let $c_1, \dots, c_{\beta_1}$ denote 1-cycles that represent a basis of
$H_1(M; \mathbb{R})$, the first singular homology group of M with real
coefficients. The additional independent conditions are the
vanishing of the integral of ω along each of these 1-cycles
to ensure that the form ω be exact.

Note that if $\{v_1(x), \dots, v_n(x)\}$ is an orthonormal basis
of $T_x M$, the operator inner product appearing on the first

integrand in (36) may be expressed as

$$\langle H_x e^{A_x}, \tilde{H}_x \rangle = \sum_{i=1}^{n} \langle H_x e^{A_x}(v_i(x)), \tilde{H}_x(v_i(x)) \rangle_x. \quad (37)$$

In particular, if we choose an orthonormal basis of eigenvectors of A_x with eigenvalues $\lambda_i(x)$, we get

$$\langle H_x e^{A_x}, \tilde{H}_x \rangle = \sum_{i=1}^{n} e^{\lambda_i(x)} \langle H_x(v_i(x)), \tilde{H}_x(v_i(x)) \rangle_x. \quad (38)$$

Similarly for the other integrand. Letting $h_i(x) = U_x \circ H_x(v_i(x))$ and $w_i(x) = W_x(v_i(x))$, we may rewrite (36) as

$$\langle (H, W), (\tilde{H}, \tilde{W}) \rangle_{(A,U)}$$

$$= a \sum_{i=1}^{n} \int_M e^{\lambda_i(x)} h_i(x) \cdot \tilde{h}_i(x) \, dV(x)$$

$$+ b \sum_{i=1}^{n} \int_M e^{\lambda_i(x)} w_i(x) \cdot \tilde{w}_i(x) \, dV(x), \quad (39)$$

where we used the fact that U_x is an isometric embedding so that

$$\langle H_x(v_i(x)), \tilde{H}_x(v_i(x)) \rangle_x = U_x(H_x(v_i(x))) \cdot U_x(\tilde{H}_x(v_i(x)))$$

$$= h_i(x) \cdot \tilde{h}_i(x). \quad (40)$$

Thus, the energy of (H, W) is

$$\|(H, W)\|^2_{(A,U)} = a \sum_{i=1}^{n} \int_M e^{\lambda_i(x)} \|h_i(x)\|^2 \, dV(x)$$

$$+ b \sum_{i=1}^{n} \int_M e^{\lambda_i(x)} \|w_i(x)\|^2 \, dV(x). \quad (41)$$

If H_x and A_x commute, one can show that $\|h_i(x)\|$ is the magnitude of the perturbation of $\|A(v_i(x))\|$ due to H. This is the interpretation of the inner product that motivated (15) and the Riemannian structure used on the computational model of Sect. 3. Although commutativity does not hold in general, this interpretation is reasonable because the edges of a discrete mesh can be perturbed somewhat independently; that is, the perturbations only need to be compatible with the integrability conditions. Note that we employed a naïve discretization of the integral over a mesh representing M. One can refine the discrete integral by allowing different weights at the vertices, a modification whose implementation is straightforward.

References

Bhattacharya, R., & Patrangenaru, V. (2003). Large sample theory of intrinsic and extrinsic sample means on manifolds. I. *Annals of Statistics*, *31*(1), 1–29.

do Carmo, M. P. (1994). *Riemannian geometry*. Basel: Birkhauser.

Dryden, I. L., & Mardia, K. V. (1998). *Statistical shape analysis*. New York: Wiley.

Fletcher, P., Lu, C., Pizer, S., & Joshi, S. (2004). Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions on Medical Imaging*, *23*(8), 995–1005.

Fletcher, P. T., Venkatasubramanian, S., & Joshi, S. (2008). Robust statistics on Riemannian manifolds via the geometric median. In *IEEE international conference on computer vision and pattern recognition*.

Fuchs, M., Jüttler, B., Scherzer, O., & Yang, H. (2009). Shape metrics based on elastic deformations. *Journal of Mathematical Imaging and Vision*, *35*(1), 86–102.

Joshi, S., Klassen, E., Srivastava, A., & Jermyn, I. (2007). An efficient representation for computing geodesics between *n*-dimensional elastic shapes. In *IEEE conference on computer vision and pattern recognition*.

Karcher, H. (1977). Riemann center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics*, *30*, 509–541.

Kendall, D. G. (1984). Shape manifolds, Procrustean metrics and complex projective spaces. *Bulletin of London Mathematical Society*, *16*, 81–121.

Kendall, D. G., Barden, D., Carne, T. K., & Le, H. (1999). *Shape and shape theory*. New York: Wiley.

Kilian, M., Mitra, N., & Pottmann, H. (2007). Geometric modeling in shape space. *ACM Transactions on Graphics*, *26*(3), 64–3–64–8.

Klassen, E., Srivastava, A., Mio, W., & Joshi, S. (2004). Analysis of planar shapes using geodesic paths on shape manifolds. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *26*, 372–383.

Le, H. L., & Kendall, D. G. (1993). The Riemannian structure of euclidean shape spaces: a novel environment for statistics. *Annals of Statistics*, *21*(3), 1225–1271.

Liu, X., Mio, W., Shi, Y., Dinov, I., Liu, X., Lepore, N., Lepore, F., Fortin, M., Voss, P., Lassonde, M., & Thompson, P. M. (2008). Models of normal variation and local contrasts in hippocampal anatomy. In *MICCAI*.

Liu, X., Shi, Y., Morra, J., Liu, X., Thompson, P. M., & Mio, W. (2009). Mapping hippocampal atrophy with a multi-scale model of shape. In *ISBI*.

Michor, P., & Mumford, D. (2007). An overview of the Riemannian metrics on spaces of curves using the Hamiltonian approach. *Applied and Computational Harmonic Analysis*, *23*, 74–113.

Mio, W., Srivastava, A., & Klassen, E. (2004). Interpolations with elasticae in Euclidean spaces. *Quarterly of Applied Mathematics*, *62*, 359–378.

Mio, W., Srivastava, A., & Joshi, S. (2007). On shape of plane elastic curves. *International Journal of Computer Vision*, *73*(3), 307–324.

Mio, W., Bowers, J. C., & Liu, X. (2009). Shape of elastic strings in Euclidean space. *International Journal of Computer Vision*, *82*, 96–112.

Morra, J., Tu, Z., Apostolova, L. G., Green, A., Avedissian, C., Madsen, S., Parikshak, N., Toga, A., Jack, C., Schuff, N., Weiner, M. W., & Thompson, P. M., & the ADNI (2009). Automated mapping of hippocampal atrophy in 1-year repeat MRI data from 490 subjects with Alzheimer's disease, mild cognitive impairment, and elderly controls. *NeuroImage*, *45*(1), S3–S15. Supplement 1.

Munkres, J. (1984). *Elements of algebraic topology*. New York: Perseus Books.

Schmidt, F. R., Clausen, M., & Cremers, D. (2006). Shape matching by variational computation of geodesics on a manifold. In *LNCS: Vol. 4174. Pattern Recognition (Proc. DAGM)* (pp. 142–151). Berlin: Springer.

Shi, Y., Thompson, P., Dinov, I., Osher, S., & Toga, A. (2007a). Direct cortical mapping via solving partial differential equations on implicit surfaces. *Medical Image Analysis*, *11*(3), 207–223.

Shi, Y., Thompson, P., Zubicaray, G., Ross, S., Tu, Z., Dinov, I., & Toga, A. (2007b). Direct mapping of hippocampal surfaces with intrinsic shape context. *NeuroImage*, *36*(3), 792–807.

Srivastava, A., Joshi, S., Mio, W., & Liu, X. (2005). Statistical shape analysis: Clustering, learning and testing. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, *27*, 590–602.

Sundaramoorthi, G., Yezzi, A., & Mennucci, A. (2007). Sobolev active contours. *The International Journal of Computer Vision*, *73*, 345–366.

Younes, L. (1998). Computable elastic distance between shapes. *SIAM Journal of Applied Mathematics*, *58*, 565–586.

Younes, L. (1999). Optimal matching between shapes via elastic deformations. *Journal of Image and Vision Computing*, *17*(5/6), 381–389.

Younes, L., Michor, P., Shah, J., & Mumford, D. (2007). A metric on shape space with explicit geodesics. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei, Mat. Appl.*, *9*.