Shape Analysis of Elastic Curves in Euclidean Spaces

Anuj Srivastava, Eric Klassen, Shantanu H. Joshi, and Ian Jermyn

Abstract

This paper introduces a square-root velocity (SRV) representation for analyzing shapes of curves in Euclidean spaces using an elastic metric. The SRV representation has several advantages: the well-known elastic metric simplifies to the $\mathbb{L}^2$ metric, the re-parameterization group acts by isometries, and the space of unit length curves becomes the familiar unit sphere. The shape space of closed curves is a submanifold of the unit sphere, modulo rotation and re-parameterization groups, and one finds geodesics in that space using a path-straightening approach. Several experiments are presented to demonstrate these ideas: (i) Shape analysis of cylindrical helices for studying structures of protein backbones, (ii) Shape analysis of facial curves for use in recognition, (iii) A wrapped probability distribution to capture shapes of planar closed curves, and ii) Parallel transport of deformations from one shape to another.

I. Introduction

Shape is an important feature for characterizing objects in several branches of science, including computer vision, bioinformatics, and biometrics. The variability exhibited by shapes within and across classes are often quite structured and there is a need to capture these variations statistically. One of the earliest works in statistical analysis and modeling of shapes of objects

A. Srivastava is with the Department of Statistics, Florida State University, Tallahassee, USA.

E. Klassen is with the Department of Mathematics, Florida State University, Tallahassee, USA.

S. H. Joshi is with the Laboratory of Neuroimaging, University of California, Los Angeles, USA.

I. H. Jermyn is with the Ariana project-team, INRIA, Sophia Antipolis, France.

came from Kendall [11] and his school of researchers [5], [30]. While this formulation took major strides in this area, its limitation was the use of landmarks in defining shapes. Since the choice of landmarks is often subjective, and also because objects in images or in imaged scenes are more naturally viewed as having continuous boundaries, there has been a recent focus on shape analysis of curves and surfaces, albeit in the same spirit as Kendall's formulation. Consequently, there is now a significant literature on shapes of continuous curves as elements of infinite-dimensional Riemannian manifolds, called shape spaces. This highly-focused area of research started with the efforts of Younes who first defined shape spaces of closed curves and imposed Riemannian metrics on them [35]. In particular, he computed geodesic paths between open curves under these metrics and projected them to obtain deformations between closed curves denoting boundaries of objects in 2D images. A related approach on analyzing shapes of sulcal curves in a brain waspresented in [12]. Klassen et al. [14] restricted to arc-length parameterized planar curves and derived numerical algorithms for computing geodesics between closed curves, the first ones to directly do so on the space of closed curves. Among other things, they applied it to statistical modeling and analysis using large databases of shapes [32]. Michor and Mumford [18] have exhaustively studied several choices of Riemannian metrics on spaces of closed, planar curves for the purpose of comparing their shapes. Mio et al. [20] presented a family of elastic metrics that quantified the bending and stretching needed to deform shapes into each other. Similarly, Shah [28], [29] derived geodesic equations for planar closed curves under different elastic metrics and different representations of curves. In these formulations, a shape space is typically constructed in two steps. First, a mathematical representation of curves with appropriate constraints leads to a *pre-shape space*. Then, one identifies elements of the pre-shape space that are related by the actions of shape-preserving transformations (rotations, translations, and scalings, as well as re-parameterizations). The resulting quotient space, *i.e.* the set of orbits under the respective group actions, is the desired shape space. If a pre-shape space is a Riemannian (Hilbert) manifold, then the shape space inherits this Riemannian structure and can be viewed as a quotient manifold or orbifold.

The choice of a shape representation and a Riemannian metric are critically important - for improved understanding, physical interpretations, and efficient computing. This paper introduces a particularly convenient representation that enables simple physical interpretations of the resulting deformations. This representation is motivated by the well-known Fisher-Rao metric, used previously for imposing a Riemannian structure on the space of probability densities; taking the positive square-root of densities results in a simple Euclidean structure where geodesics, distances, and statistics are straightforward to compute. A similar idea was introduced by Younes [35] and later used in Younes et al. [36] for studying shapes of **planar** curves under an elastic metric. The representation used in the current paper is similar to these earlier ideas, but is sufficiently different to be **applicable to curves in arbitrary** $\mathbb{R}^n$. The main contributions of this paper are as follows:

(1) Presentation of a square-root velocity (SRV) representation for studying shapes of elastic closed curves in $\mathbb{R}^n$, first introduced in the conference papers [7], [8]. This has several advantages as discussed later.

(2) The use of a numerical approach, termed *path-straightening*, for finding geodesics between shapes of closed elastic curves. It uses a gradient-based iteration to find a geodesic where, using the Palais metric on the space of paths, the gradient is available in a convenient analytical form.

(3) The use of a gradient-based solution for optimal re-parameterization of curves when finding geodesics between their shapes. This paper compares the strengths and weaknesses of this gradient solution versus the commonly used Dynamic Programming (DP) algorithm.

(4) The application and demonstration of this framework to: (i) shape analysis of cylindrical helices in $\mathbb{R}^3$ for use in studies of protein backbone structures, (ii) shape analysis of 3D facial curves, (iii) development of a wrapped normal distribution to capture shapes in a shape class, and (iv) parallel transport of deformations from one shape to another. The last item is motivated by the need to predict individual shapes or shape models for novel objects, or novel views of the objects, using past data. A similar approach has recently been applied to shape representations using deformable templates [37] and for studying shapes of 3D triangulated

meshes [13].

The paper is organized as follows. Section II introduces the proposed elastic shape framework, while Section III discusses its merits relative to existing literature. Section IV describes a path-straightening approach for finding geodesics and a gradient-based approach for elastic curve registration. Section V presents four applications of this framework.

II. Shape Representation and Riemannian Structure

In order to develop a formal framework for analyzing shapes of curves, one needs a mathematical representation of curves that is natural, general and efficient. We describe one such representation that allows a simple framework for shape analysis.

A. Square-root Velocity Representation and Pre-Shape Space

Let β be a parameterized curve ($\beta : D \rightarrow \mathbb{R}^n$), where D is a certain domain for the parameterization. We are going to restrict to those β that are differentiable and their first derivative is in $\mathbb{L}^2(D, \mathbb{R}^n)$. In general D will be $[0, 2\pi]$, but for closed curves it will be more natural to have $D = \mathbb{S}^1$. We define a mapping: $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ according to $F(v) \equiv v/\sqrt{\|v\|}$ if $\|v\| \neq 0$ and 0 otherwise. Here, $\|\cdot\|$ is the Euclidean 2-norm in $\mathbb{R}^n$ and note that F is a continuous map. For the purpose of studying the shape of β, we will represent it using the square-root velocity (SRV) function defined as $q : D \rightarrow \mathbb{R}^n$, where $q(t) \equiv F(\dot{\beta}(t)) = \dot{\beta}(t)/\sqrt{\|\dot{\beta}(t)\|}$. This representation includes those curves whose parameterization can become singular in the analysis. Also, for every $q \in \mathbb{L}^2(D, \mathbb{R}^n)$ there exists a curve β (unique up to a translation) such that the given q is the SRV function of that β. In fact, this curve can be obtained using the equation: $\beta(t) = \int_0^t q(s)\|q(s)\|ds$. The motivation for using this representation and comparisons with other such representations are described in the Section III.

To remove scaling variability, we rescale all curves to be of length 2π. (One can use any finite value here, including one, and we have chosen 2π simply to include the unit circle when $n = 2$.) We remark that this restriction to a "slice" of the full space of curves is identical to Kendall's [11] approach for removing scale variability. The remaining transformations (rotation,

translation, and re-parameterization) will be dealt with differently. This difference is due to the difference in the actions of scaling and other groups on the representation space of curves, as described later. The restriction that β is of length 2π translates to the condition that $\int_D \|q(t)\|^2 dt = \int_D \|\dot{\beta}\| dt = 2\pi$. Therefore, the SRV functions associated with these curves are elements of a hypersphere in the Hilbert manifold $\mathbb{L}^2(D, \mathbb{R}^n)$; we will use the notation $\mathcal{C}^o$ to denote this hypersphere. According to Lang [15] pg. 27, $\mathcal{C}^o$ is a Hilbert submanifold in $\mathbb{L}^2(D, \mathbb{R}^n)$.

For studying shapes of closed curves, we impose an additional condition that $\beta(0) = \beta(2\pi)$. In view of this condition, it is natural to have the domain D be the unit circle $\mathbb{S}^1$ for closed curves. With pre-determined (fixed) placement of the origin on $\mathbb{S}^1$, it can be identified with $[0, 2\pi)$ using the function $t \mapsto (\cos(t), \sin(t))$. We will use them according to convenience. In terms of the SRV function, this closure condition is given by: $\int_{\mathbb{S}^1} q(t)\|q(t)\| dt = 0$. Thus, we have a space of fixed length, closed curves represented by their SRV functions:

$$\mathcal{C}^c = \{q \in \mathbb{L}^2(\mathbb{S}^1, \mathbb{R}^n) | \int_{\mathbb{S}^1} \|q(t)\|^2 dt = 2\pi, \int_{\mathbb{S}^1} q(t)\|q(t)\| dt = 0\}.$$

The superscript c implies that we have imposed the closure condition. With the earlier identification of $[0, 2\pi)$ with $\mathbb{S}^1$, $\mathcal{C}^c \subset \mathcal{C}^o \subset \mathbb{L}^2(D, \mathbb{R}^n)$. What is the nature of the set $\mathcal{C}^c$? We are going to sketch a proof that $\mathcal{C}^c$ is a codimension-n submanifold of $\mathcal{C}^o$; this proof is based on pages 25-27 of [15]. Let $G : \mathcal{C}^o \to \mathbb{R}^n$ be a map defined as $G(q) = \int_{\mathbb{S}^1} q(t)\|q(t)\| dt$. First, we need to check that its differential, $dG_q : T_q(\mathcal{C}^o) \to \mathbb{R}^n$, is surjective at every $q \in G^{-1}(\mathbf{0})$; $\mathbf{0} \in \mathbb{R}^n$ is the origin. One easily verifies that this is true, except in cases where the vector $q(t)$ lies in the same one-dimensional subspace of $\mathbb{R}^n$ for every t; in these cases it is not true. These exceptional functions correspond to curves that lie entirely in a straight line in $\mathbb{R}^n$. This collection of curves is a "very small" (measure zero) subset of $\mathcal{C}^o$, and we conclude that G is a submersion at the remaining points of $G^{-1}(\mathbf{0})$. Therefore, using [15], $\mathcal{C}^c$ is a codimension-n submanifold of $\mathcal{C}^o$, for all points except those in this measure zero subset. We will ignore this subset since there is essentially a zero probability of encountering it in real problems. We conclude that $\mathcal{C}^c$, with the earlier proviso, is a submanifold of the Hilbert space $\mathcal{C}^o$ and, thus,

$\mathbb{L}^2(\mathbb{S}^1, \mathbb{R}^n)$.

Now we have two submanifolds – $\mathcal{C}^o$ and $\mathcal{C}^c$ – associated with open and closed curves in $\mathbb{R}^n$, respectively. They are called *pre-shape spaces* for their respective cases. We will call $\mathcal{C}^o$ the pre-shape space of open curves since the closure constraint is not enforced here, even though it *does* contain closed curves also, while $\mathcal{C}^c$ is purely the pre-shape space of closed curves. To impose Riemannian structures on these pre-shape spaces, we consider their tangent spaces.

1) **Open Curves**: Since $\mathcal{C}^o$ is a sphere in $\mathbb{L}^2([0, 2\pi], \mathbb{R}^n)$, its tangent space at a point q is given by:

$$T_q(\mathcal{C}^o) = \{v \in \mathbb{L}^2([0, 2\pi], \mathbb{R}^n) | \langle v, q \rangle = 0\} \ .$$

Here $\langle v, q \rangle$ denotes the inner product in $\mathbb{L}^2([0, 2\pi], \mathbb{R}^n)$: $\langle v, q \rangle = \int_0^{2\pi} \langle v(t), q(t) \rangle dt$.

2) **Closed Curves**: The tangent space to $\mathcal{C}^c$ at a point q is, of course, a subset of $\mathbb{L}^2(\mathbb{S}^1, \mathbb{R}^n)$. Since $\mathcal{C}^c$ is a submanifold, this subset is often defined using the differential of the map G. In fact, the tangent space $T_q(\mathcal{C}^c)$ at a point $q \in \mathcal{C}^c$ is given by the kernel of the differential of G at that point [19]. Therefore, it is often easier to specify the normal space, *i.e.* the space of functions in $\mathbb{L}^2(\mathbb{S}^1, \mathbb{R}^n)$ perpendicular to $T_q(\mathcal{C}^c)$. This normal space is found using the differential of G as follows: for the function $G_i(t) = \int_{\mathbb{S}^1} q_i(t) \|q(t)\| dt$, $i = 1, 2, \ldots, n$, its directional derivative in a direction $w \in \mathbb{L}^2(\mathbb{S}^1, \mathbb{R}^n)$ is given by:

$$dG_i(w) = \int_{\mathbb{S}^1} \langle w(t), \frac{q_i(t)}{\|q(t)\|} q(t) + \|q(t)\| \mathbf{e}_i(t) \rangle dt \ ,$$

where $\mathbf{e}^i$ is a unit vector in $\mathbb{R}^n$ along the i^{th} coordinate axis. This specifies the normal space:

$$N_q(\mathcal{C}^c) = \text{span } \{q(t), (\frac{q_i(t)}{||q(t)||} q(t) + ||q(t)|| \mathbf{e}_i), \ i = 1, \ldots, n\} \ , \tag{1}$$

Hence, $T_q(\mathcal{C}^c) = \{v \in \mathbb{L}^2(\mathbb{S}^1, \mathbb{R}^n) | \langle v, w \rangle = 0, \quad \forall w \in N_q(\mathcal{C}^c)\}$.

The standard metric on $\mathbb{L}^2(D, \mathbb{R}^n)$ restricts to Riemannian structures on the two manifolds $\mathcal{C}^o$ and $\mathcal{C}^c$. These structures can then be used to determine geodesics and geodesic lengths between elements of these spaces. Let $\mathcal{C}$ be a Riemannian manifold denoting either $\mathcal{C}^o$ or $\mathcal{C}^c$, and let $\alpha : [0, 1] \to \mathcal{C}$ be a parameterized path such that $\alpha(0) = q_0$ and $\alpha(1) = q_1$. Then, the length

of α is defined to be: $L[\alpha] = \int_0^1 \langle \dot{\alpha}(t), \dot{\alpha}(t) \rangle^{1/2} dt$, and α is a said to be a *length-minimizing geodesic* if $L[\alpha]$ achieves the infimum over all such paths. The length of this geodesic becomes a distance:

$$d_c(q_0, q_1) = \inf_{\alpha:[0,1]\to\mathcal{C} | \alpha(0)=q_0, \alpha(1)=q_1} L[\alpha] \ .$$

The computation of geodesics in $\mathcal{C}^o$ is straightforward, since it is a sphere, but the case of $\mathcal{C}^c$ is more complicated and requires numerical methods as described later.

B. Shape Space as Quotient Space

By representing a parameterized curve $\beta(t)$ by its SRV function $q(t)$, and imposing the constraint $\int_D \langle q(t), q(t) \rangle dt = 2\pi$, we have taken care of the translation and the scaling variability, but the rotation and the re-parameterization variability still remain. A rotation is an element of $SO(n)$, the special orthogonal group of $n \times n$ matrices, and a re-parameterization is an element of Γ, the set of all orientation-perserving diffeomorphisms of D. In the following discussion, $\mathcal{C}$ stands for either $\mathcal{C}^o$ or $\mathcal{C}^c$.

The rotation and re-parameterization of a curve β are denoted by the actions of $SO(n)$ and Γ on its SRV. While the action of $SO(n)$ is the usual: $SO(n) \times \mathcal{C} \to \mathcal{C}, \quad (O, q(t)) = Oq(t)$, the action of Γ is derived as follows. For a $\gamma \in \Gamma$, the composition $\beta \circ \gamma$ denotes its re-parameterization; the SRV of the re-parameterized curve is $F(\dot{\beta}(\gamma(t))\dot{\gamma}(t)) = q(\gamma(t))\sqrt{\dot{\gamma}(t)}$, where q is the SRV of β. This gives use the action $\Gamma \times \mathcal{C} \to \mathcal{C}, \quad (q, \gamma) = (q \circ \gamma)\sqrt{\dot{\gamma}}$. In order for shape analysis to be invariant to these transformations, it is important for these groups to act by isometries. We note the following properties of these actions.

Lemma 1: The actions of $SO(n)$ and Γ on $\mathcal{C}$ commute.

Proof: This follows directly from the definitions of the two group actions.

Therefore, we can form a joint action of the product group $SO(n) \times \Gamma$ on $\mathcal{C}$ according to $((O, q), \gamma) = O(q \circ \gamma)\sqrt{\dot{\gamma}}$.

Lemma 2: The action of the product group $\Gamma \times SO(n)$ on $\mathcal{C}$ is by isometries with respect to the chosen metric.

Proof: For a $q \in \mathcal{C}$, let $u, v, \in T_q(\mathcal{C})$. Since $\langle Ou(t), Ov(t)\rangle = \langle u(t), v(t)\rangle$, for all $O \in SO(n)$ and $t \in D$, the proof for $SO(n)$ follows.

Now, fix an arbitrary element $\gamma \in \Gamma$, and define a map $\phi : \mathcal{C} \to \mathcal{C}$ by $\phi(q) = (\gamma, q)$. A glance at the formula for (γ, q) confirms that ϕ is a linear transformation. Hence, its derivative $d\phi$ has the same formula as ϕ. In other words, the mapping $d\phi : T_q(\mathcal{C}) \to T_{(\gamma,q)}(\mathcal{C})$ is given by: $u \mapsto \tilde{u} \equiv (u \circ \gamma)\sqrt{\dot{\gamma}}$. The Riemannian metric after the transformation is:

$$\begin{aligned}\langle \tilde{u}, \tilde{v}\rangle &= \int_D \langle \tilde{u}(t), \tilde{v}(t)\rangle dt \\ &= \int_D \langle u(\gamma(t))\sqrt{\dot{\gamma}(t)}, v(\gamma(t))\sqrt{\dot{\gamma}(t)}\rangle dt = \int_D \langle u(\tau), v(\tau)\rangle d\tau, \quad \tau = \gamma(t)\ .\end{aligned}$$

Putting these two results together, the joint action of $\Gamma \times SO(n)$ on $\mathcal{C}$ is by isometries with respect to the chosen metric. □

Therefore, we can define a quotient space of $\mathcal{C}$ modulo $\Gamma \times SO(n)$. The orbit of a function $q \in \mathcal{C}$ is given by:

$$[q] = \{O(q \circ \gamma)\sqrt{\dot{\gamma}}) | (\gamma, O) \in \Gamma \times SO(n)\}\ .$$

In this framework, an orbit is associated with a shape and comparisons between shapes are performed by comparing the orbits of the corresponding curves and, thus, the need for a metric on the set of orbits. In order for this set to inherit the metric from $\mathcal{C}$, we need the orbits to be closed sets in $\mathcal{C}$. Since these orbits are not closed in $\mathcal{C}$, we replace them by their closures in $\mathbb{L}^2(D, \mathbb{R}^n)$. With a slight abuse of notation, we will call these orbits $[q]$. Then, define the quotient space $\mathcal{S}$ as the set of all such closed orbits associated with the elements of $\mathcal{C}$, *i.e.* $\mathcal{S} = \{[q] | q \in \mathcal{C}\}$.

Since we have a quotient map from $\mathcal{C}$ to $\mathcal{S}$, its differential induces a linear isomorphism between $T_{[q]}(\mathcal{S})$ and the normal space to $[q]$ at any point $\tilde{q}$ in $[q]$. The Riemannian metric on $\mathcal{C}$ (*i.e.* the $\mathbb{L}^2$ inner product) restricts to an inner product on the normal space which, in turn, induces an inner product on $T_{[q]}(\mathcal{S})$. The fact that $\gamma \times SO(n)$ act by isometries implies that the resulting inner product on $T_{[q]}(\mathcal{S})$ is independent of the choice of $\tilde{q} \in [q]$. In this manner, $\mathcal{S}$ inherits a Riemannian structure from $\mathcal{C}$. Consequently, the geodesics in $\mathcal{S}$ correspond to those

geodesics in $\mathcal{C}$ that are perpendicular to all the orbits they meet in $\mathcal{C}$ and the geodesic distance between any two points in $\mathcal{S}$ is given by:

$$d_s([q]_0,[q]_1) = \min_{\tilde{q}_1 \in [q_1]} d_c(q_0,\tilde{q}_1) = \inf_{(\gamma,O)\in\Gamma\times SO(n)} d_c(q_0, O(q_1\circ\gamma)\sqrt{\dot{\gamma}}) \ . \tag{2}$$

III. Motivation & Comparisons

We first motivate the choice of SRV and the elastic metric for shape analysis and then compare our choice with previous ideas.

A. Motivation for the SRV Representation

Let $\beta : D \to \mathbb{R}^n$ be an open curve in $\mathbb{R}^n$. Assume that for all $t \in D$, $\dot{\beta}(t) \neq 0$.[1] We then define $\phi : D \to \mathbb{R}$ by $\phi(t) = \ln(\|\dot{\beta}(t)\|)$, and $\theta : D \to \mathbb{S}^{n-1}$ by $\theta(t) = \dot{\beta}(t)/\|\dot{\beta}(t)\|$. Clearly, ϕ and θ completely specify $\dot{\beta}$, since for all t, $\dot{\beta}(t) = e^{\phi(t)}\theta(t)$. Thus, we have defined a map from the space of open curves in $\mathbb{R}^n$ to $\Phi \times \Theta$, where $\Phi = \{\phi : D \to \mathbb{R}\}$ and $\Theta = \{\theta : D \to \mathbb{S}^{n-1}\}$. This map is surjective; it is not injective, but two curves are mapped to the same pair (ϕ, θ) if and only if they are translates of each other, i.e., if they differ by an additive constant. Intuitively, ϕ tells us the (log of the) speed of traversal of the curve, while θ tells us the direction of the curve at each time t.

In order to quantify the magnitudes of perturbations of β (and enable ourselves to do geometry on the space of these curves), we wish to impose a Riemannian metric on the space of curves that is invariant under translation, and we will do this by putting a metric on $\Phi \times \Theta$. First, we note that the tangent of space of $\Phi \times \Theta$ at any point (ϕ, θ) is given by

$$T_{(\phi,\theta)}(\Phi \times \Theta) = \{(u, v) : u \in \Phi \text{ and } v : D \to \mathbb{R}^n \text{ and } v(t) \perp \theta(t), \forall t \in D\}$$

Suppose (u_1, v_1) and (u_2, v_2) are both elements of $T_{(\phi,\theta)}(\Phi \times \Theta)$. Let a and b be positive real numbers, and define an inner product by

$$\langle (u_1, v_1), (u_2, v_2)\rangle_{(\phi,\theta)} = a^2 \int_D u_1(t)u_2(t)e^{\phi(t)}\, dt + b^2 \int_D \langle v_1(t), v_2(t)\rangle e^{\phi(t)}\, dt. \tag{3}$$

[1]We make this assumption in this section only for the purpose of comparing with past work that requires this constraint. The SRV representation and the analysis presented in the rest of the paper is more general, and does not require this assumption.

(Note that $\langle \cdot, \cdot \rangle$ in the second integral denotes the standard dot product in $\mathbb{R}^n$.) This inner product, discussed in [20], has the interpretation that the first integral measures the amount of "stretching", since u_1 and u_2 are variations of the speed ϕ of the curve, while the second integral measures the amount of "bending", since v_1 and v_2 are variations of the direction ϕ of the curve. The constants a^2 and b^2 are weights, which we choose depending on how much we want to penalize these two types of deformations.

Perhaps the most important property of this Riemannian metric is that the groups $SO(n)$ and Γ both act by isometries. To elaborate on this, recall that $O \in SO(n)$ acts on a curve β by $(O, \beta)(t) = O\beta(t)$, and $\gamma \in \Gamma$ acts on β by $(\gamma, \beta)(t) = \beta(\gamma(t))$. Using our identification of the set of curves with the space $\Phi \times \Theta$ results in the following actions of these groups. $O \in SO(n)$ acts on (ϕ, θ) by $(O, (\phi, \theta)) = (\phi, O\theta)$. $\gamma \in \Gamma$ acts on (ϕ, θ) by $(\gamma, (\phi, \theta)) = (\phi \circ \gamma + \ln \circ \dot{\gamma}, \theta \circ \gamma)$.

We now need to understand the differentials of these group actions on the tangent spaces of $\Phi \times \Theta$. $SO(n)$ is easy; since each $O \in SO(n)$ acts by the restriction of a linear transformation on $\Phi \times L^2(D, \mathbb{R}^n)$, it acts in exactly the same way on the tangent spaces: $(O, (u, v)) = (u, Ov)$, where $(u, v) \in T_{(\phi,\theta)}(\Phi \times \Theta)$, and $(u, Ov) \in T_{(\phi,O\theta)}(\Phi \times \Theta)$. The action of $\gamma \in \Gamma$ given in the above formula is not linear, but affine linear, because of the additive term $\ln \circ \dot{\gamma}$. Hence, its action on the tangent space is the same, but without this additive term: $(\gamma, (u, v)) = (u \circ \gamma, \theta \circ \gamma)$, where $(u, v) \in T_{(\phi,\theta)}(\Phi \times \Theta)$, and $(u \circ \gamma, \theta \circ \gamma) \in T_{(\gamma,(\phi,\theta))}(\Phi \times \Theta)$. Combining these actions of $SO(n)$ and Γ with the above inner product on $\Phi \times \Theta$, it is an easy verification that these actions are by isometries, i.e.,

$$\langle (O, (u_1, v_1)), (O, (u_2, v_2)) \rangle_{(O,(\phi,\theta))} = \langle (u_1, v_1), (u_2, v_2) \rangle_{(\phi,\theta)}$$

and

$$\langle (\gamma, (u_1, v_1)), (\gamma, (u_2, v_2)) \rangle_{(\gamma,(\phi,\theta))} = \langle (u_1, v_1), (u_2, v_2) \rangle_{(\phi,\theta)}.$$

Since we have identified the space of curves with $\Phi \times \Theta$, we may identify the space of shapes with the quotient space $(\Phi \times \Theta)/(SO(n) \times \Gamma)$. Furthermore, since these group actions are by isometries with respect to all the metrics we introduced above, *no matter what values we assign to a and b*, we get a corresponding two-parameter family of metrics on the quotient

space $(\Phi \times \Theta)/(SO(n) \times \Gamma)$. Note that in distinguishing between the structures (for example, geodesics) associated to these metrics, only the ratio of b to a is important, since if you multiply both by the same real number you just rescale the metric, which results in the same geodesics, and all distances multiplied by the same constant. Given two shapes, we can find the geodesics between the corresponding elements of $(\Phi \times \Theta)$ with respect to any of these metrics (*i.e.* , with respect to any choice of a and b). The choice of relative weights will determine whether these deformations are biased towards allowing more stretching or more bending.

This is not the only consideration, however. While it is possible to find analytic expressions for the geodesics on the space of parameterized curves for any choice of a/b and for any n, once we restrict attention to the space of unit length curves, things are not as simple. One can ask whether there is some particular choice of weights which will be especially natural and which will result in the geodesics being easier to compute. We now show that the SRV representation is a natural answer to this question.

Recall that, given a curve $\beta : D \to \mathbb{R}^n$, its SRV representation is the function $q : D \to \mathbb{R}^n$ given by $q(t) = \frac{\dot{\beta}(t)}{\sqrt{\|\dot{\beta}(t)\|}}$. Relating this to the (ϕ, θ) representation of the curve gives $q(t) = e^{\frac{1}{2}\phi(t)}\theta(t)$. A couple of simple differentiations show that if $(u, v) \in T_{(\phi,\theta)}(\Phi \times \Theta)$, then the corresponding tangent vector to $\mathbb{L}^2(D, \mathbb{R}^n)$ at q is given by $f = \frac{1}{2}e^{\frac{1}{2}\phi}u\theta + e^{\frac{1}{2}\phi}v$. Now let (u_1, v_1) and (u_2, v_2) denote two elements of $T_{(\phi,\theta)}(\Phi \times \Theta)$, and let f_1 and f_2 denote the corresponding tangent vectors to $\mathbb{L}^2(D, \mathbb{R}^n)$ at q. Computing the $\mathbb{L}^2$ inner product of f_1 and f_2 yields

$$\langle f_1, f_2 \rangle = \int_D \left(\frac{1}{2}e^{\frac{1}{2}\phi}u_1\theta + e^{\frac{1}{2}\phi}v_1\right) \cdot \left(\frac{1}{2}e^{\frac{1}{2}\phi}u_2\theta + e^{\frac{1}{2}\phi}v_2\right) dt = \int_D \frac{1}{4}e^{\phi}u_1u_2 + e^{\phi}(v_1 \cdot v_2)\, dt. \quad (4)$$

In this computation we have used the facts that $\theta \cdot \theta = 1$, since $\theta(t)$ is an element of the unit sphere, and that $\theta(t) \cdot v_i(t) = 0$, since each $v_i(t)$ is a tangent vector to the unit sphere at $\theta(t)$. This computation shows that the $\mathbb{L}^2$ metric on the space of SRV representations (*i.e.* q-functions) corresponds precisely to the elastic metric on $\Phi \times \Theta$, with $a = \frac{1}{2}$ and $b = 1$. However, expressed in terms of the q-functions, the $\mathbb{L}^2$-metric is the "same" at every point of $\mathbb{L}^2(D, \mathbb{R}^n)$ (it is simply $\int_D f_1 \cdot f_2\, dt$, which does not depend on the q-function at which these tangent vectors are defined), and we will thus have available more efficient ways of computing

geodesics in our pre-shape and shape spaces using the SRV formulation. We emphasize again that this is true for curves in arbitrary dimension.

B. Comparisons with Prior Work

For the case $n = 1$, there is no θ component and the elastic metric in Eqn. 3 takes the form: $\langle u_1, u_2 \rangle = \int_D u_1(t) u_2(t) e^{\phi(t)} dt$. This is the *Fisher-Rao* metric and has been used for imposing a Riemannian structure on the space of probability density functions on D [1], [2], [4]. Note that $e^{\phi(t)}$ can be interpreted as a probability density function for a curve of fixed length. It is well known, at least since 1943 [2], that under the square-root representation, *i.e.* for $q(t) = e^{\frac{1}{2}\phi(t)}$, this metric reduces to the $\mathbb{L}^2$ metric; Eqn. 4 restates that point when we set $n = 1$.

To discuss $n > 1$, it is useful to use a slightly different representation. Let us replace the function ϕ by the function $\rho = e^{\phi/2} = \|\dot{\beta}(t)\|^{1/2} : D \to \mathbb{R}^+$, *i.e.* the square root of the speed. We can then re-express the elastic metric, Eqn. 3, in terms of the (ρ, θ) representation:

$$\begin{aligned} \langle (w_1, v_1), (w_2, v_2) \rangle_{(\rho,\theta)} &= 4a^2 \int_D w_1(t) w_2(t)\, dt + b^2 \int_D \rho^2(t)\, \langle v_1(t), v_2(t) \rangle\, dt \\ &= 4a^2 \int_D \Big\{ w_1(t)\, w_2(t) + c^2\, \rho^2(t)\, \langle v_1(t), v_2(t) \rangle \Big\}\, dt\ . \end{aligned} \tag{5}$$

where $c = b/2a$. Notice that when $c = 1$, the integrand is simply the Euclidean metric on $\mathbb{R}^n$ expressed in hyperspherical coordinates. As a result, the space of parameterized curves has zero Riemannian curvature when $b = 2a$, *e.g.* for $a = \frac{1}{2}$ and $b = 1$. It turns out that for $n > 2$, this is the only value of c for which the space of parameterized curves has zero curvature (although it is conformally flat for all c), and thus the only value of c for which Euclidean coordinates exist, *i.e.* for which the metric can take a simple $\mathbb{L}^2$ form. These Euclidean coordinates are the q representation, which thus occupies a special position amongst curve representations.

Let us return now to the case of $n = 2$, and focus on the integrand for a fixed value of t, abbreviating functions evaluated at this value by their function names, *e.g.* $\theta(t)$ by θ. When $c < 1$, then the circumference of a circle, ρ constant, is less than it would be on $\mathbb{R}^2$. We can

visualize this geometry as the surface of a right circular cone in $\mathbb{R}^3$, with apex angle α given by $\sin(\alpha/2) = c$. Distance from the apex along the surface is ρ, while θ is angle around the cone. Such a cone has zero Riemannian curvature except at the apex, $\rho = 0$, where there is a singularity, as we can see explicitly by eliminating c from the metric by defining $\theta_c = c\theta$. Thus unlike the cases $n > 2$, where only the metrics with $c = 1$ were flat, for $n = 2$ the metric is flat for all c (except at $\rho = 0$). The value $c = 1$ still plays a distinguished role, however, because it is only for $c = 1$ that the singularity disappears; the cone is then isometric to $\mathbb{R}^2$.

Just as in the case $c = 1$ discussed for general n, we can introduce Euclidean coordinates $q_c(t) = \rho(t)(\cos(\theta_c(t)), \sin(\theta_c(t)))$ for each t, such that the metric (5) becomes an $\mathbb{L}^2$-metric. Alternatively, we can introduce complex coordinates $z(t) = \rho(t)e^{i\theta_c(t)}$. We have to bear in mind, though, that θ is a coordinate on $\mathbb{S}^1$. Let us suppose $\theta \in [0, 2\pi)$ with 2π identified with 0. In consequence, $\theta_c \in [0, 2\pi c)$, with $2\pi c$ identified with 0. For $c > 1$, this means that the Euclidean coordinates are not useful, since as the q_c range over $\mathbb{R}^2$ or the z over $\mathbb{C}$, we do not cover the original space $\Phi \times \Theta$. For $c < 1$, however, the change from θ to θ_c corresponds to cutting open the cone along a generatrix and flattening it onto $\mathbb{R}^2$ with the apex at the origin; the coordinate θ_c is just polar angle as measured in $\mathbb{R}^2$. Now, as the q_c range over $\mathbb{R}^2$ or the z over $\mathbb{C}$, we more than cover the original space $\Phi \times \Theta$. This is a manifestation of the curvature singularity, but it is not necessarily a bad thing. For example, if $c = 1/2$, then there is a double covering. This case is treated in great detail in [36], where it is shown to possess some very nice properties. In particular, the closure condition simplifies to $\int_D z^2\, dt = 0$, which for curves of fixed length simply means that the real and imaginary parts of z must be $\mathbb{L}^2$-orthogonal. This means that the space of (parameterized) fixed length, closed planar curves is a Stiefel manifold. This makes it possible to understand and utilize the geometry of the submanifold formed by closed curves for $c = 1/2$, which until then had only been explored numerically. In particular, one can write an explicit formula for computing geodesics in the pre-shape spaces of such curves. This representation and analysis does not generalize to higher n, however, while the SRV representation does. On the other hand, explicit formulae for computing geodesics

Dimension	Representation or Space	Geodesic Computation (this paper)	Geodesic Computation ([36])
$n = 2$	Representation	$q(t) = \frac{\dot{\beta}(t)}{\|\dot{\beta}(t)\|^{\frac{1}{2}}} = \|\dot{\beta}(t)\|^{\frac{1}{2}} e^{i\theta(t)}$	$q_{\frac{1}{2}}(t) = \|\dot{\beta}(t)\|^{\frac{1}{2}} e^{i\frac{\theta(t)}{2}}$
	$\mathcal{C}^o$	Analytical	Analytical
	$\mathcal{C}^o/SO(2)$	Analytical (Procrustes)	Analytical
	$\mathcal{S}^o = \mathcal{C}^o/(\gamma \times SO(2))$	Numerical (DP or Gradient, Procrustes)	Numerical
	$\mathcal{C}^c$	Numerical (Path Straightening – PS)	Analytical
	$\mathcal{C}^c/SO(2)$	Numerical (PS, Gradient)	Analytical
	$\mathcal{S}^c = \mathcal{C}^c/(\Gamma \times SO(2))$	Numerical (PS, Gradient)	Numerical
general n	Representation	$q(t) = \frac{\dot{\beta}(t)}{\|\dot{\beta}(t)\|^{\frac{1}{2}}}$	Not Applicable
	All Spaces	Same As Above	Not Applicable

TABLE I

DIFFERENCES BETWEEN THIS PAPER AND [36]. THE MAIN STRENGTH OF THIS PAPER IS ITS APPLICABILITY TO ARBITRARY n, WHILE THE STRENGTH OF [36] IS THE AVAILABILITY OF EXPLICIT GEODESICS FOR PLANAR CLOSED CURVES.

between closed curves are no longer available and one has to rely on numerical methods. However, numerical methods were anyway necessary in [36] in order to quotient by the action of the diffeomorphism group.

Table I summarizes the differences in the computation of geodesics to which the differences in the representations used in this paper and in [36] naturally give rise. Some of the ideas mentioned there will be explained in the next section, where the computation of geodesics is studied.

IV. COMPUTATION OF GEODESICS

In this section, we focus on the task of computing geodesics between any given pair of shapes in a shape space. This task is accomplished in two steps. First, we develop tools for computing geodesics in the pre-shape spaces, $\mathcal{C}^o$ or $\mathcal{C}^c$ and, then, we remove the remaining shape-preserving transformations to obtain geodesics in the shape spaces. In the case of $\mathcal{C}^o$, the

underlying space is a sphere and the task of computing geodesic paths there is straightforward. For any two points q_0 and q_1 in $\mathcal{C}^o$, a geodesic connecting them is given by: $\alpha : [0,1] \rightarrow \mathcal{C}^o$,

$$\alpha(\tau) = \frac{1}{\sin(\theta)} \left(\sin(\theta(1-\tau))q_0 + \sin(\theta\tau)q_1\right) \ , \tag{6}$$

where $\theta = \cos^{-1}(\langle q_0, q_1 \rangle)$ is the length of the geodesic. However, the geometry of $\mathcal{C}^c$ is not so simple and we will use a path-straightening approach to compute geodesics in $\mathcal{C}^c$, as described next. Other choices of numerical methods for computing geodesics include [26].

A. Path-straightening on Pre-Shape Space

For any two closed curves, denoted by q_0 and q_1 in $\mathcal{C}^c$, we are interested in finding a geodesic path between them in $\mathcal{C}^c$. We start with an arbitrary path $\alpha(t)$ connecting q_0 and q_1, *i.e.* $\alpha : [0,1] \mapsto \mathcal{C}^c$ such that $\alpha(0) = q_0$ and $\alpha(1) = q_1$. Then, we iteratively "straighten" α until it achieves a local minimum of the energy:

$$E(\alpha) \equiv \frac{1}{2} \int_0^1 \langle \frac{d\alpha}{d\tau}(\tau), \frac{d\alpha}{d\tau}(\tau) \rangle d\tau \ , \tag{7}$$

over all paths from q_0 to q_1. It can be shown that a critical point of E is a geodesic on $\mathcal{C}^c$. However, it is possible that there are multiple geodesics between a given pair q_0 and q_1, and a local minimum of E may not correspond to the shortest of all geodesics. Therefore, this approach has the limitation that it finds a geodesic between a given pair but may not reach the shortest geodesic.

Let $\mathcal{H}$ be the set of all paths in $\mathcal{C}^c$, parameterized by $\tau \in [0,1]$, and $\mathcal{H}_0$ be the subset of $\mathcal{H}$ of paths that start at q_0 and end at q_1. The tangent spaces of $\mathcal{H}$ and $\mathcal{H}_0$ are: $T_\alpha(\mathcal{H}) = \{w |\ \forall \tau \in [0,1], w(\tau) \in T_{\alpha(\tau)}(\mathcal{C}^c)\}$, where $T_{\alpha(\tau)}(\mathcal{C}^c)$ is specified as a set orthogonal to $N_q(\mathcal{C}^c)$ in Eqn. 1. Here w is a vector field along α such that $w(\tau)$ is tangent to $\mathcal{C}^c$ at $\alpha(\tau)$. Similarly, $T_\alpha(\mathcal{H}_0) = \{w \in T_\alpha(\mathcal{H}) | w(0) = w(1) = 0\}$. To ensure that α stays at the desired end points, the allowed vector field on α has to be zero at the ends.

Our study of paths on $\mathcal{H}$ requires the use of covariant derivatives and integrals of vector fields along these paths. For a given path $\alpha \in \mathcal{H}$ and a vector field $w \in T_\alpha(\mathcal{H})$, the **covariant**

derivative of w along α is the vector field obtained by projecting $\frac{dw}{d\tau}(\tau)$ onto the tangent space $T_{\alpha(\tau)}(\mathcal{C}^c)$, for all t. It is denoted by $\frac{Dw}{dt}$. Similarly, a vector field $u \in T_\alpha(\mathcal{H})$ is called a **covariant integral** of w along α if the covariant derivative of u is w, i.e. $\frac{Du}{d\tau} = w$.

To make $\mathcal{H}$ a Riemannian manifold, an obvious metric would be $\langle w_1, w_2 \rangle = \int_0^1 \langle w_1(\tau), w_2(\tau) \rangle d\tau$, for w_1, $w_2 \in T_\alpha(\mathcal{H})$. Instead, we use the Palais metric [22], which is:

$$\langle\langle w_1, w_2 \rangle\rangle = \langle w_1(0), w_2(0) \rangle + \int_0^1 \langle \frac{Dw_1}{d\tau}(\tau), \frac{Dw_2}{d\tau}(\tau) \rangle d\tau \ ,$$

One reason for using the Palais metric is that with respect to this metric, $T_\alpha(\mathcal{H}_0)$ is a closed linear subspace of $T_\alpha(\mathcal{H})$, and $\mathcal{H}_0$ is a closed subset of $\mathcal{H}$. Therefore, any vector $w \in T_\alpha(\mathcal{H})$ can be uniquely projected into $T_\alpha(\mathcal{H}_0)$. The second reason is that the expression for the gradient of E under this metric is relatively simpler.

Our goal is to find the minimizer of E in $\mathcal{H}_0$, and we will use a gradient flow to do that. Therefore, we wish to find the gradient of E in $T_\alpha(\mathcal{H}_0)$. To do this, we first find the gradient of E in $T_\alpha(\mathcal{H})$ and then project it into $T_\alpha(\mathcal{H}_0)$.

Theorem 1: The gradient vector of E in $T_\alpha(\mathcal{H})$ is given by the unique vector field u such that $Du/d\tau = d\alpha/d\tau$ and $u(0) = 0$. In other words, u is the covariant integral of $d\alpha/d\tau$ with zero initial value at $\tau = 0$.

Proof: Define a *variation* of α to be a smooth function $h : [0,1] \times (-\epsilon, \epsilon) \to \mathcal{H}$ such that $h(\tau, 0) = \alpha(\tau)$ for all $\tau \in [0,1]$. The variational vector field corresponding to h is given by $v(\tau) = h_\tau(t, 0)$ where s denotes the second argument in h. Thinking of h as a path of curves in $\mathcal{H}$, we define $E(s)$ as the energy of the curve obtained by restricting h to $[0,1] \times \{s\}$. That is, $E(s) = \frac{1}{2} \int_0^1 \langle h_\tau(\tau, s), h_\tau(\tau, s) \rangle d\tau$. We now compute,

$$E'(0) = \int_0^1 \langle \frac{Dh_\tau}{ds}(\tau, 0), h_\tau(\tau, 0) \rangle d\tau = \int_0^1 \langle \frac{Dh_s}{d\tau}(\tau, 0), h_\tau(\tau, 0) \rangle d\tau = \int_0^1 \langle \frac{Dv}{d\tau}(\tau), \frac{d\alpha}{d\tau}(\tau) \rangle d\tau \ ,$$

since $h_\tau(\tau, 0)$ is simply $\frac{d\alpha}{d\tau}(\tau)$. Now, the gradient of E should be a vector field u along α such that $E'(0) = \langle\langle v, u \rangle\rangle$. That is, $E'(0) = \langle v(0), u(0) \rangle + \int_0^1 \langle \frac{Dv}{d\tau}, \frac{Du}{d\tau} \rangle d\tau$. From this expression it is

clear that u must satisfy the initial condition $u(0) = 0$ and the ordinary (covariant) differential equation $\frac{Du}{d\tau} = \frac{d\alpha}{d\tau}$. $\square$

We will introduce some additional properties of vector fields along α that are useful in our construction. A vector field w is called **covariantly constant** if Dw/dt is zero at all points along α. Similarly, a path α is called a **geodesic** if its velocity vector field is covariantly constant. That is, α is a geodesic if $\frac{D}{dt}(\frac{d\alpha}{dt}) = 0$ for all t. Also, a vector field w along the path α is called **covariantly linear** if Dw/dt is a covariantly constant vector field.

Lemma 3: The orthogonal complement of $T_\alpha(\mathcal{H}_0)$ in $T_\alpha(\mathcal{H})$ is the space of all covariantly linear vector fields w along α.

Proof: Suppose $v \in T_\alpha(\mathcal{H}_0)$ (i.e. $v(0) = v(1) = 0$), and $w \in T_\alpha(\mathcal{H})$ is covariantly linear. Then, using (covariant) integration by parts:

$$\langle\langle v, w\rangle\rangle = \int_0^1 \langle \frac{Dv(\tau)}{d\tau}, \frac{Dw(\tau)}{d\tau}\rangle d\tau = \langle v, \frac{Dw(\tau)}{d\tau}\rangle_0^1 - \int_0^1 \langle v(\tau), \frac{D}{d\tau}\left(\frac{Dw(\tau)}{d\tau}\right)\rangle d\tau = 0 \ .$$

Hence, $T_\alpha(\mathcal{H}_0)$ is orthogonal to the space of covariantly linear vector fields along α in $T_\alpha(\mathcal{H})$. This proves that the space of covariantly linear vector fields is contained in the orthogonal complement of $T_\alpha(\mathcal{H}_0)$. To prove that these two spaces are equal, observe first that given any choice of tangent vectors at $\alpha(0)$ and $\alpha(1)$, there is a unique covariantly linear vector field interpolating them. It follows that every vector field along α can be uniquely expressed as the sum of a covariantly linear vector field and a vector field in $T_\alpha(H_0)$. The lemma then follows. $\square$.

A vector field u is called the **forward parallel translation** of a tangent vector $w_0 \in T_{\alpha(0)}(\mathcal{C}^c)$, along α, if and only if $u(0) = w_0$ and $\frac{Du(t)}{dt} = 0$ for all $t \in [0, 1]$. Similarly, u is called the **backward parallel translation** of a tangent vector $w_1 \in T_{\alpha(1)}(\mathcal{C}^c)$, along α, when for $\tilde{\alpha}(t) \equiv \alpha(1 - t)$, u is the forward parallel translation of w_1 along $\tilde{\alpha}$. It must be noted that parallel translations, forward or backward, lead to vector fields that are covariantly constant.

According to Lemma 3, to project the gradient u into $T_\alpha(\mathcal{H}_0)$, we simply need to subtract off a covariantly linear vector field which agrees with u at $t = 0$ and $t = 1$. Clearly, the correct

covariantly linear field is simply $t\tilde{u}(t)$, where $\tilde{u}(t)$ is the covariantly constant field obtained by parallel translating $u(1)$ backwards along α. Hence, we have proved the following theorem.

Theorem 2: Let $\alpha : [0,1] \mapsto \mathcal{C}^c$ be a path, $\alpha \in \mathcal{H}_0$. Then, for u as defined in Theorem 1, the gradient of the energy function E restricted to $\mathcal{H}_0$ is $w(t) = u(t) - t\tilde{u}(t)$, where $\tilde{u}$ is the vector field obtained by parallel translating $u(1)$ backwards along α.

To finish this discussion we show that the critical points of E are geodesics.

Lemma 4: For a given pair q_0, $q_1 \in \mathcal{C}^c$, a critical point of E on $\mathcal{H}_0$ is a geodesic on $\mathcal{C}^c$ connecting q_0 and q_1.

Proof: Let α be a critical point of E in $\mathcal{H}_0$. That is, the gradient of E is zero at α. Since the gradient vector field is given by $u(t) - t\tilde{u}(t)$, we have that $u(t) = t\tilde{u}(t)$ for all t. Therefore, $\frac{d\alpha}{dt} = \frac{Du}{dt} = \frac{D(t\tilde{u})}{dt} = \tilde{u}$. Since $\tilde{u}$ is a parallel translation of $u(1)$, it is covariantly constant, and therefore, the velocity field $\frac{d\alpha}{dt}$ is covariantly constant. By definition, this implies that α is a geodesic. □ Now we describe an algorithm for computing geodesics in $\mathcal{C}^c$ using path straightening. The sub-algorithms referred to here are listed in the Appendix.

Path-Straightening Algorithm: Find a geodesic between two parameterized curves β_0 and β_1 by representing them in $\mathcal{C}^c$.

1) Compute the representations of each curve in $\mathcal{C}^c$. Denote these elements by q_0 and q_1, respectively.
2) Initialize a path α between q_0 and q_1 in $\mathcal{C}^o$ using Eqn. 6 and project it in $\mathcal{C}^c$ using Item 1 in Appendix.
3) Compute the velocity vector field $d\alpha/dt$ along the path α using Algorithm 1.
4) Compute the covariant integral of $d\alpha/dt$, denoted by u, using Algorithm 2.
5) Compute the backward parallel transport of the vector $u(1)$ along α using Algorithm 3 and denote it by $\tilde{v}$.
6) Compute the full gradient vector field of the energy E along the path α, denoted by w, using $w = u(t) - t\tilde{u}_1$ (Algorithm 4).
7) Update α along the vector field w using Algorithm 5. If $\sum_{\tau=1}^{k} \langle w(\tau), w(\tau) \rangle$ is small,

then stop. Else, return to Step 3.

B. Removing Shape-Preserving Transformations

Now we have procedures for constructing geodesics between points in a preshape space $\mathcal{C}$ ($\mathcal{C}^o$ or $\mathcal{C}^c$), and we focus on the same task for shape spaces. Towards this goal, we need to solve the joint minimization problem on (γ, O) in Eqn. 2, with the cost function being $H : \Gamma \times SO(n) \to \mathbb{R}$, $H(\gamma, O) = d_c(q_0, O(q_1 \circ \gamma)\sqrt{\dot{\gamma}})$. We will use an iterative gradient-based search to update the two group elements. Let $\gamma^{(k)} = \gamma_1 \circ \gamma_2 \circ \cdots \circ \gamma_k$ and $O^{(k)} = O_1 \cdot O_2 \cdot \cdots \cdot O_k$ be the cumulative group elements and at the kth iteration we seek the increments (γ_{k+1}, O_{k+1}) that minimize $H(\gamma^{(k+1)}, O^{(k+1)})$. Let $\tilde{q}_1$ denote the current element of the orbit $[q_1]$, *i.e.* $\tilde{q}_1 = O^{(k)}(q_1 \circ \gamma^{(k)})\sqrt{\dot{\gamma^{(k)}}}$ and let $\alpha : [0, 1] \to \mathcal{C}$ be a geodesic from q_0 to $\tilde{q}_1$. So, $\dot{\alpha}_1$ is the velocity vector at $\tilde{q}_1$ and define $v \equiv \dot{\alpha}(1)/\|\dot{\alpha}(1)\|$.

1. **Rotations**: The case of updating the orientation is relatively simple. Fix the re-parameterization $\gamma^{(k+1)}$ for this part. The tangent space to only the rotation orbit is $\{A\tilde{q}_1 | A \in \mathbb{R}^{n\times n}, A + A^T = 0\}$. Let $E_1, E_2, \ldots, E_{n(n-1)/2}$ be an orthonormal basis for the space of skew-symmetric matrices. The gradient updates for rotation are performed by projecting v in this space to obtain $A = \sum_i \langle E_i, v\rangle E_i$ and updating using $O_{k+1} = e^{\delta A}\tilde{q}_1$ for a step size δ. In the case of $\mathcal{C}^o$, this optimization can be greatly simplified using the Procrustes idea. Since $\mathcal{C}^o$ is a sphere, the geodesic length is given by an arc-length, and minimizing arc-length is same as minimizing the corresponding chord-length. Therefore, the optimal rotation is directly written as:

$$\begin{aligned}\hat{O}_{k+1} &= \operatorname*{argmin}_{O\in SO(n)} \cos^{-1}(\langle q_0, O\tilde{q}_1\rangle) = \operatorname*{argmin}_{O\in SO(n)} \|q_0 - O\tilde{q}_1\| \\ &= U \begin{bmatrix} 1 & 0 & \ldots & 0 \\ .. & .. & .. & .. \\ 0 & 0 & \ldots & \text{sign}(\det(B)) \end{bmatrix} V^T, \quad B = U\Sigma V^T, \quad B = \int_D q_0(t)\tilde{q}_1(t)^T dt \ .\end{aligned}$$

2. **Re-parameterizations**: Now we seek the incremental γ_{k+1} that minimizes H. There are two possibilities: One is to take the gradient of $H(\gamma^{(k+1)})$ directly with respect to γ_{k+1} and use it to update $\gamma^{(k+1)}$. The directional derivative of H, with respect to γ_{k+1}, in the direction $b \in T_{\gamma_{id}}(\Gamma)$

is given by $2\int_D \langle v(t), \left(\tilde{\dot{q}}_1(t)b(t) + \frac{1}{2}\tilde{q}_1(t)\dot{b}(t)\right)\rangle dt$. To implement this gradient update, one has to find a finite-dimensional subspace of $T_{\gamma_{id}}(\Gamma)$, approximate the gradient in that subspace and then update $\gamma^{(k+1)}$. The other possibility, the one we have used in this paper, is to use a square-root representation of γ that often simplifies its analysis. Define $\psi_{k+1} = \sqrt{\dot{\gamma}_{k+1}}$ and re-express γ_{k+1} as the pair $(\gamma_{k+1}(0), \psi_{k+1})$. Note that the space Ψ of all ψ-functions is a subset of a sphere (of radius $\sqrt{2\pi}$). We initialize with $\gamma_{k+1}(t) = t$, with the corresponding representation being $(0, \mathbf{1})$ and $\mathbf{1}$ being the constant function with value one. Now take the gradients of H, with respect to $\gamma_{k+1}(0)$ and ψ_{k+1}, and update these individually. The derivative with respect to $\gamma_{k+1}(0)$, evaluated at $(0, \mathbf{1})$, is $\frac{\partial H}{\partial \gamma_{k+1}(0)} = \int_D \langle v(t), \dot{\tilde{q}}_1(t)\rangle dt$. Similarly, the directional derivative of H in a direction $c \in T_{\mathbf{1}}(\Psi)$ is given by:

$$\nabla_\psi H(c) = \int_D \langle v(t), \left(2\dot{\tilde{q}}_1(t)\tilde{c}(t) + \dot{\tilde{q}}_1(t)c(t)\right)\rangle dt, \quad \tilde{c}(t) = \int_0^t c(s)ds \ .$$

Form an approximate basis for the tangent space $T_{\mathbf{1}}(\Psi) = \{f : D \to \mathbb{R} | \langle f, \mathbf{1}\rangle = \mathbf{0}\}$ using: $\{(\frac{1}{\sqrt{\pi}}\sin(nt), \frac{1}{\sqrt{\pi}}\cos(nt)) | n = 1, 2, \ldots, m/2\}$, and approximate the gradient using $c = \sum_{i=1}^m \nabla_\psi H(c_i)c_i$, where the c_is are the basis elements. Then, update the ψ component according to: $\mathbf{1} \mapsto \psi_{k+1} \equiv \cos(\|c\|)\mathbf{1} + \sin(\|c\|)\frac{c}{\|c\|}$. This ψ_{k+1} in turn gives $\gamma_{k+1}(t) = \int_0^t \psi_{k+1}(s)^2 ds$ and thus $\gamma^{(k+1)}$.

Figure 1 shows two examples of this implementation. In each case we start with a parameterized curve, denoted by q_1 shown in (a), generate a random $\gamma \in \Gamma$ (shown in (b)) and form a re-parameterized curve using $q_0 = (q_1 \circ \gamma)\sqrt{\dot{\gamma}}$ (shown in (c)). Then, we use the gradient approach described above to find an optimal re-parameterization of q_1 that best matches this q_0 by minimizing the cost function H. The evolution of the relative cost function H is shown in (d), and the final re-parameterized curve $\tilde{q}_1$ is shown in (e). In these examples, since q_0 is simply a re-parmeterization of q_1, the minimum value of H is zero. Note that in the top row, where the original γ is closer to the identity, the cost function goes to zero but in the bottom case where γ is rather drastic, the algorithm converges to a final value of H that is not close to zero. We conjecture that this is due to the algorithm being stuck in a local minimum.

To illustrate the strengths and limitations of a gradient-based approach with respect to a

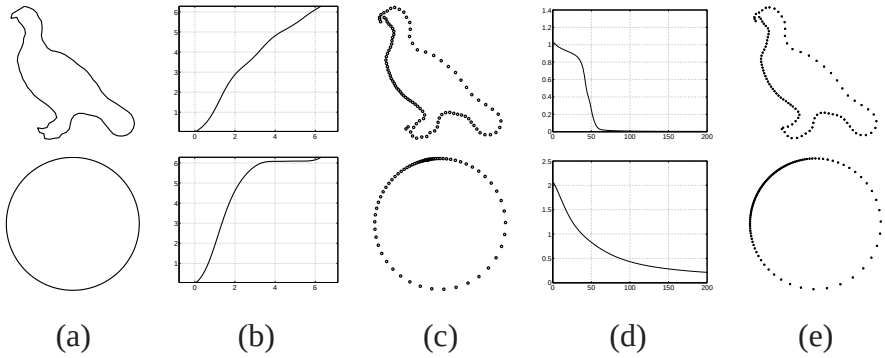

Fig. 1. (a) The original shape represented by q_1, (b) an arbitrary $\gamma \in \Gamma$, (c) the second shape formed using $q_0 = (q_1 \circ \gamma)\sqrt{\dot{\gamma}}$, (d) evolution of H in matching $\tilde{q}_1$ with q_0, (e) final curve represented by $\tilde{q}_1$.

common DP algorithm [27], we present a comparison of computational costs (using Matlab on a 2.4GHz Intel processor) and performance in Table II. The computational complexity of the gradient approach is $O(Tmk)$, where T is the number of samples on the curve, m is the number of basis functions, and k is the number of iterations, while that of DP algorithm is $O(T^2 \log(T))$. The table is generated for $T = 100$ and $k = 200$. As a measure of matching performance, we also present the relative final cost as a percentage ((H(final)/H(initial)) $\times 100$). This table shows that while the DP algorithm is very accurate in estimating the unknown γ, its computational cost is relatively high. One gets to solutions, albeit approximate, much faster when using the gradient method. An important limitation of the gradient method is that its solution is always local.

Figure 2 shows some elastic geodesics between several pairs of shapes. We have drawn ticks on these curves to show the optimal re-parametrizations of the second curves. The spacings between the ticks are uniform in the leftmost shapes (q_0) but have been adjusted for the other shapes during the minimization of H. The reader can see that the combinations of bending and stretching used in these deformations are successful in the sense that geometrical features are well preserved.

Shape	Method	DP Algorithm	Gradient Approach (m)				
			10	30	50	70	90
Circle	Time (sec)	12.00	0.88	1.72	2.55	3.39	4.22
Circle	Relative Final Cost (%)	0.06	1.19	0.40	0.28	0.24	0.21
Bird	Time (sec)	12.13	0.89	1.72	2.58	3.43	4.33
Bird	Relative Final Cost (%)	0.016	3.65	1.63	1.33	1.31	1.17

TABLE II

TIMING ANALYSIS OF GRADIENT-BASED RE-PARAMETERIZATION AND COMPARISON WITH DP ALGORITHM.

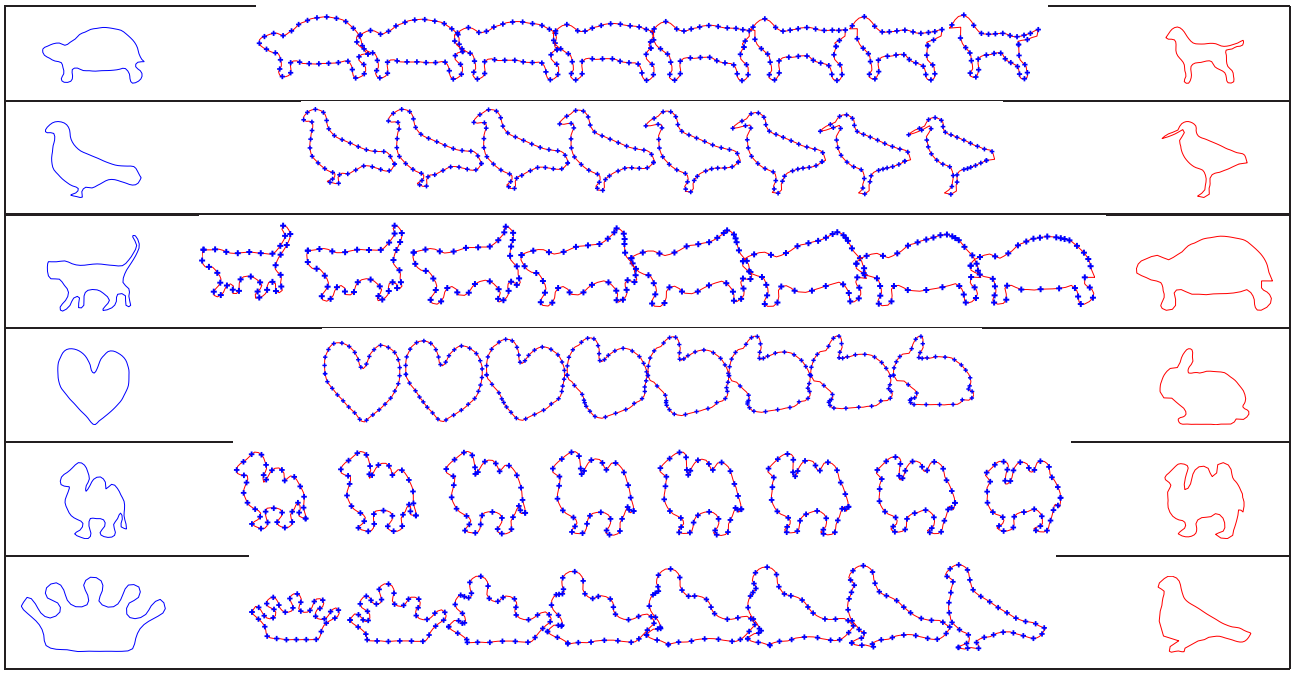

Fig. 2. Examples of elastic geodesics between pairs of 2D shapes.

V. APPLICATIONS

In this section, we illustrate this elastic shape analysis framework using some applications. Some additional applications have been presented elsewhere: shape analysis of corpus-callosum for Williams Syndrome [6]; symmetry analysis of two- and three-dimensional shapes [24];

labelling of cortical sulci [17]; shape classification of point clouds [31]; and joint gait-cadence analysis for human identification in videos [10].

A. Shapes Analysis of 3D Coils

As the first example, we will study the shapes of helices in $\mathbb{R}^3$ by matching and deforming one into another. One motivation for studying shapes of cylindrical helices comes from protein structure analysis. Proteins are linear polymers formed by concatenating amino acids. A primary structure in a protein is a linked chain of carbon, nitrogen, and oxygen atoms known as the backbone. Additionally, a protein also has side chains that are connected to the backbone and it is the combined effect of the backbone and the side chains that ultimately determines the three-dimensional structure of a protein. However, the geometry of the backbone is often a starting point in this analysis. These backbones contain certain distinct geometrical pieces and one prominent type is the so-called α-helix. In analyzing shapes of backbones it seems important to match not only their global geometries but also the local features (such as α-helices) that appear along these curves. We suggest the elastic shape analysis of curves as a framework for studying shapes of protein backbones, and we present some results to support that. However, we will start with some simulated curves since a matching of local features is hard to visualize for real protein backbones.

Shown in Figure 3 are two examples of geodesics between some cylindrical helices. In each case, the panels (a) and (b) show the two helices, and (c) is the optimal matching between them obtained using the estimated γ function shown in panel (d). The resulting geodesic paths in $\mathcal{S}^o$ between these curves are shown in the bottom row. It is easy to see the combination of bending and stretching/compression that goes into deforming one shape into another. In the left example, where the turns are quite similar and the curves differ only in the placements of these turns along the curve, a simple stretching/compression is sufficient to deform one into another. However, in the right example, where the number of turns is different, the algorithm requires both bending and stretching to reach from the first to the second shape.

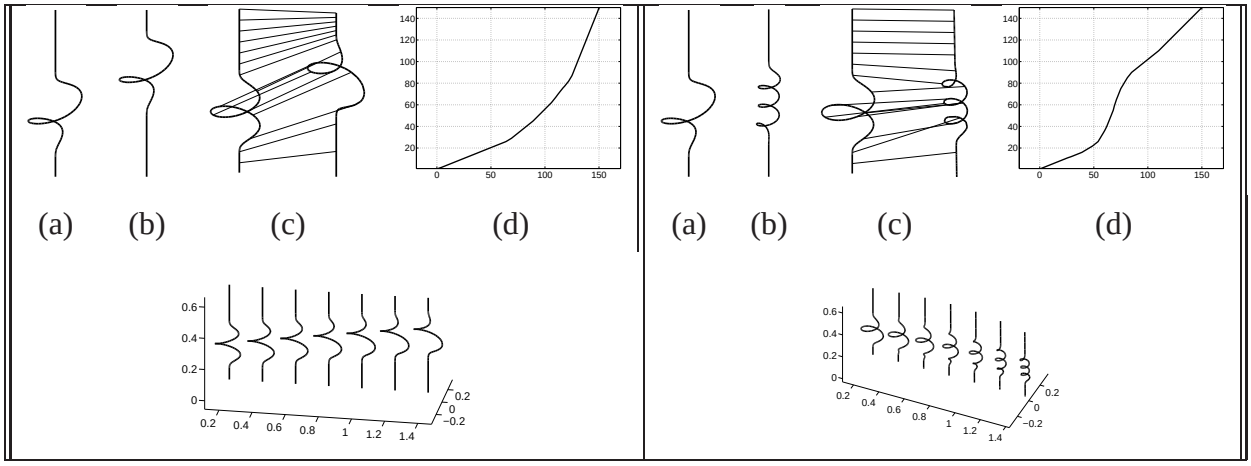

Fig. 3. Coil matching: Two example of computing geodesics between 3D curves containing spiral features. In each case (a) and (b) show the original curves, (c) shows the optimal registration between them, and (d) the optimal γ function. The lower panels show the corresponding geodesic paths between them.

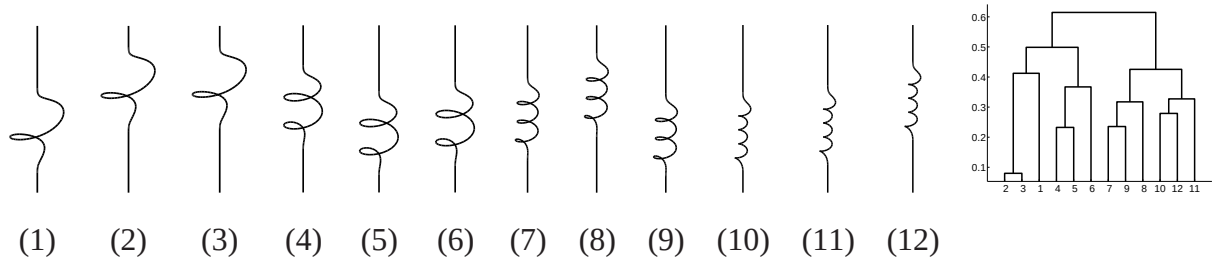

Fig. 4. A set of helices with different numbers and placements of spirals and their clustering using the elastic distance function.

Figure 4 shows an example of using the elastic distances between curves for clustering and classification. In this example, we experiment with 12 cylindrical helices that contain different number and placements of turns. The first three helices have only one turn, the next three have two turns, and so on. (We point out that the radii of these turns have some randomness that is difficult to see with the naked eye.) Using the elastic geodesic distances between them in $\mathcal{S}^o$, and the dendrogram clustering program in Matlab, we obtain the clustering shown in the right panel. This clustering demonstrates the success of the proposed elastic metric in that helices

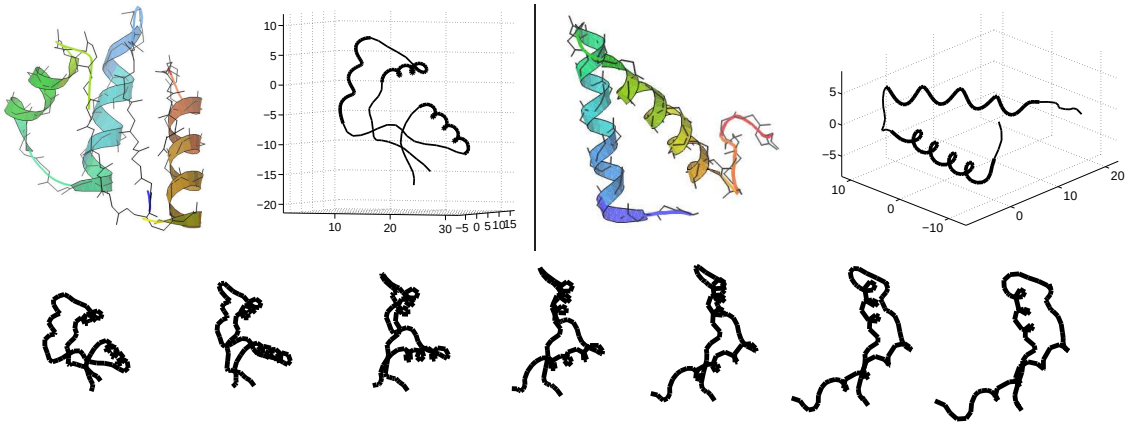

Fig. 5. Elastic deformations to compare shapes of two proteins: 1CTF and 2JVD (obtained from PDB). The top row shows the two proteins: 1CTF on the left and 2JVD on the right. The bottom row shows the elastic geodesic between them. The bottom left shows the optimal registration between the two curves and the bottom right shows the optimal γ function.

with similar numbers of turns are clustered together.

Finally, in Figure 5, we present an example of comparing 3D curves using real protein backbones. In this experiment we use two simple proteins – 1CTF and 2JVD – that contain three and two α-helices respectively. The top row of this figure shows depictions of the two backbones, while the bottom row shows the geodesic path between them in $\mathcal{S}^o$. These results on both simulated helices and real backbones suggest a role for elastic shape analysis in protein structure analysis. Further experiments are needed to ascertain this role.

B. 3D Face Recognition

Human face recognition is a problem of great interest in homeland security, client access systems, and several other areas. Since recognition performance using 2D images has been limited, there has been a push towards using shapes of facial surfaces, obtained using weak laser scanners, to recognize people. The challenge is to develop methods and metrics that succeed in classifying people despite changes in shape due to facial expressions and measurement errors. Recently Samir et al. [23], [33] have proposed an approach that: (1) computes a function on

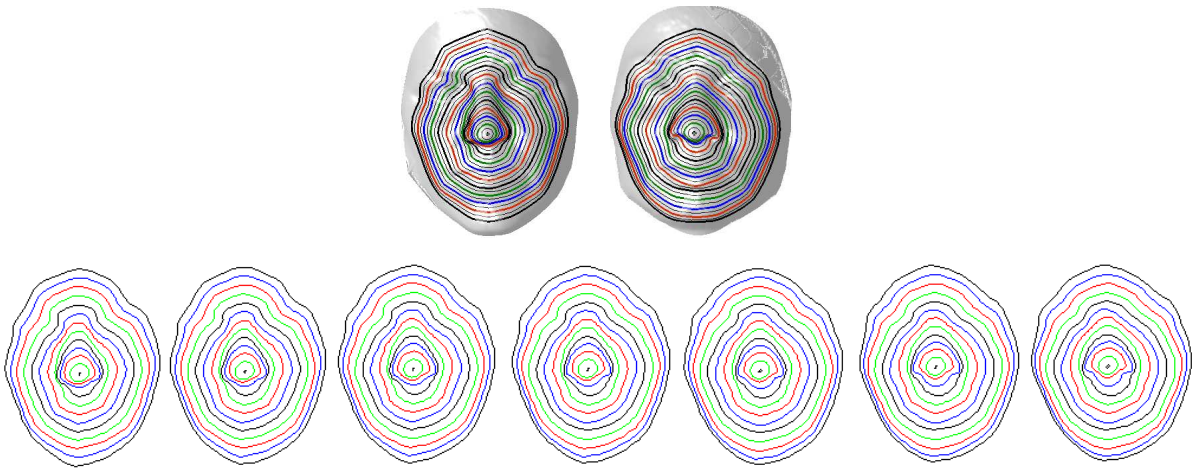

Fig. 6. The top row shows two facial surfaces represented by indexed collections of facial curves. The bottom reconstructs geodesics between the corresponding curves; these intermediate curves are rescaled and placed a the original locations to reconstruct intermediate faces.

a facial surface as the shortest-path distance from the tip of the nose (similar to [3] and [21]), (2) defines facial curves to be the level curves of that function, and (3) represents the shapes of facial surfaces using indexed collections of their facial curves. Figure 6 (top) shows two facial surfaces overlaid with facial curves. These facial curves are closed curves in $\mathbb{R}^3$ and their shapes are invariant to rigid motions of the original surface. In order to compare shapes of facial surfaces, we compare shapes of the corresponding facial curves by computing geodesics between them in $\mathcal{S}_2^c$. As an example, Figure 6 (bottom) shows geodesics in $\mathcal{S}_2^c$ between corresponding facial curves. For display, these intermediate curves have been rescaled and translated to the original values and, through reconstruction, they result in a geodesic path such that points along that path approximate full facial surfaces. Furthermore, these geodesic paths can be used to compute average faces or to define metrics for human recognition using the shapes of their faces.

C. Elastic Models for Planar Shapes

One important application of this elastic shape framework is to develop probability models for capturing the variability present in the observed shapes associated with shape classes. For example, the left panel of Figure 7 shows examples of 20 observed two-dimensional shapes of a "runner" taken from the Kimia database. Our goal is derive a probability model on the shape space $\mathcal{S}^c$, so that we can use this model in future inferences. Using ideas presented in earlier papers [5], [32], [34], we demonstrate a simple model where we: (i) first compute the sample Karcher mean [9] of the given shapes, (ii) learn a probability model on the tangent space (at the mean) by mapping the observations to that tangent space, and (iii) wrap the probability model back to $\mathcal{S}^c$ using the exponential map. In this paper, we demonstrate the model using random sampling: random samples are generated in the tangent space and mapped back to $\mathcal{S}^c$, rather than wrapping the probability model explicitly.

Let $\mu = \operatorname{argmin}_{[q]\in\mathcal{S}^c} \sum_{i=1}^{n} d_s([q],[q_i])^2$ be the Karcher mean of the given shapes $q_1, q_2, \ldots, q_n$, where d_s is the geodesic distance on $\mathcal{S}^c$. The Karcher mean of the 20 observed shapes is shown in the middle panel of Figure 7. Once we have μ, we can map $[q_i] \mapsto v_i \equiv \exp_\mu^{-1}([q_i]) \in T_\mu(\mathcal{S}^c)$. Since the tangent space is a vector space, we can perform more standard statistical analysis. The infinite-dimensionality of $T_\mu(\mathcal{S}^c)$ is not a problem here since one has only a finite number of observations. For instance, one can perform PCA on the set $\{v_i\}$ to find dominant directions and associated observed variances. One can study these dominant directions of variability as shapes by projecting vectors along these directions to the shape space. Let (σ_i, U_i)'s be the singular values and singular directions in the tangent space, then the mapping $t\sigma_i U_i \mapsto \exp_\mu(t\sigma_i U_i)$ helps visualize these principal components as shapes. The three principal components of the 20 given shapes are given in the lower three rows of Figure 7, each row displaying some shapes from $t = -1$ to $t = 1$.

In terms of probability models, there are many choices available. For the coefficients $\{z_i\}$ defined with respect to the basis $\{U_i\}$, one can use any appropriate model from multivariate statistics. In this experiment, we try a non-parametric approach where a kernel density estimator,

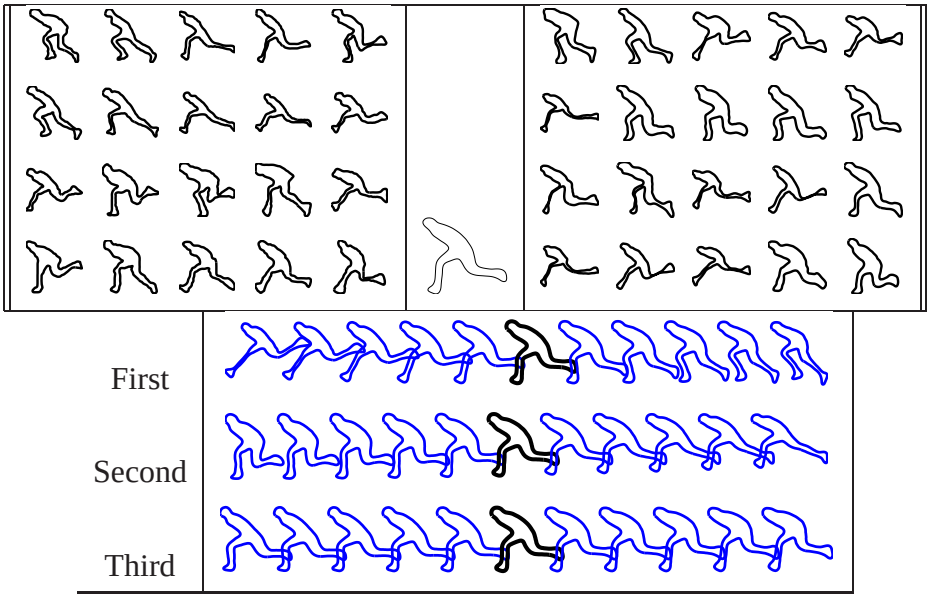

Fig. 7. The left panel shows a set of 29 observed shapes of a "runner" from the Kimia dataset. The middle panel shows their Karcher mean, and the right panel shows a random sample of 20 shapes from the learned wrapped nonparameteric model on $\mathcal{S}_2^c$. The bottom three rows show eigen variations of shapes in three dominant directions around the mean, drawn from negative to positive direction and scaled by the corresponding eigen values.

with a Gaussian kernel, is used for each coefficient z_i independently. One of the ways to evaluate this model is to generate random samples from it. Using the inverse transform method to sample z_is from their estimated kernel densities, we can form a random vector $\sum_i z_i U_i$ and then the random shape $\exp_\mu(\sum_i z_i U_i)$. The right panel of Figure 7 shows 20 such random shapes. It is easy to see the success of this wrapped model in capturing the shape variability exhibited in the original 20 shapes.

D. Transportation of Shape Deformations

One difficulty in using shapes of three-dimensional objects is that their two-dimensional appearance changes with viewing angles. Since a large majority of imaging technology is oriented towards two-dimensional images, there is a striking focus on planar shapes, their analysis and modeling, despite the viewing variability. Within this focus area, there is a problem

to predict the planar shapes of three-dimensional objects from novel viewing angles. (The problem of predicting full appearances, using pixels, has been studied by [16], [25] and others.) The idea is to study how shapes of known (training) objects deform under a change of a viewing angle, and to deform the unknown (test) shapes appropriately into shape predictions. Therefore, modeling of deformations becomes important. If we know how a known object deforms under a viewpoint change, perhaps we can apply the "same" deformation to a similar (yet novel) object and predict its deformation under the same viewpoint change. The basic technical issue is to be able to transport the required deformation from the first object to the second object, before applying that deformation. Since shape spaces are nonlinear manifolds, the deformation vector from one shape cannot simply be applied to another shape.

The mathematical statement of this problem is as follows: Let $[q_1^a]$ and $[q_1^b]$ be the shapes of an object $\mathcal{O}^1$ when viewed from two viewing angles θ_a and θ_b, respectively. The deformation in contours, in going from $[q_1^a]$ to $[q_1^b]$ depends on some physical factors: the geometry of $\mathcal{O}^1$ and the viewing angles involved. Consider another object $\mathcal{O}^2$ which is similar to $\mathcal{O}^1$ in geometry. Given its shape $[q_2^a]$ from the viewing angle θ_a, our goal is to predict its shape $[q_2^b]$ from the viewing angle θ_b. Our solution is based on taking the deformation that takes $[q_1^a]$ to $[q_1^b]$ and applying it to $[q_2^a]$ after some adjustments. Let $\alpha_1(\tau)$ be a geodesic between $[q_1^a]$ and $[q_1^b]$ in $\mathcal{S}^c$ and $v_1 \equiv \dot{\alpha}_1(0) \in T_{[q_1^a]}(\mathcal{S}^c)$ is its initial velocity. We need to *transport* v_1 to $[q_2^a]$; this is done using forward parallel translation introduced earlier in Section IV. Let $\alpha_{12}(\tau)$ be a geodesic from $[q_1^a]$ to $[q_2^a]$ in $\mathcal{S}^c$. Construct a vector field $w(t)$ such that $w(0) = v_1$ and $\frac{Dw(t)}{dt} = 0$ for all points along α_{12}. This is accomplished in practice using Algorithm 2 given in the appendix. Then, $v_2 \equiv w(1) \in T_{[q_2^a]}(\mathcal{S}^c)$ is a parallel translation of v_1. Figure 8 shows two examples of this idea, one in each row. Take the top case as an example. Here, a hexagon ($[q_1^a]$) is deformed into a square ($[q_1^b]$) using an elastic geodesic; this deformation is then transported to a circle ($[q_2^a]$) and applied to it to result in the prediction $[q_2^b]$.

Next, we consider an experiment involving shapes of tanks: the M60 as $\mathcal{O}^1$ and the T72 as $\mathcal{O}^2$ in this experiment. Give the observed shapes from different azimuths (fixed elevation) for

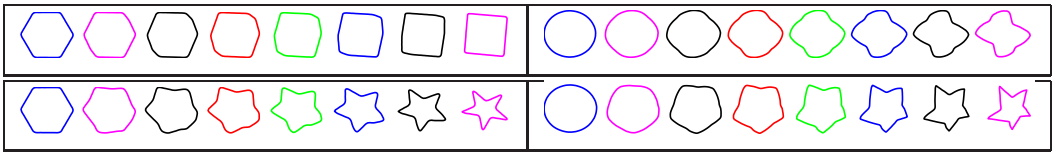

Fig. 8. For each row, the left panel shows a geodesic from the template shape (hexagon) to the training shape. The right panel shows the one-parameter flow from the test shape (circle) for the corresponding tangent vector.

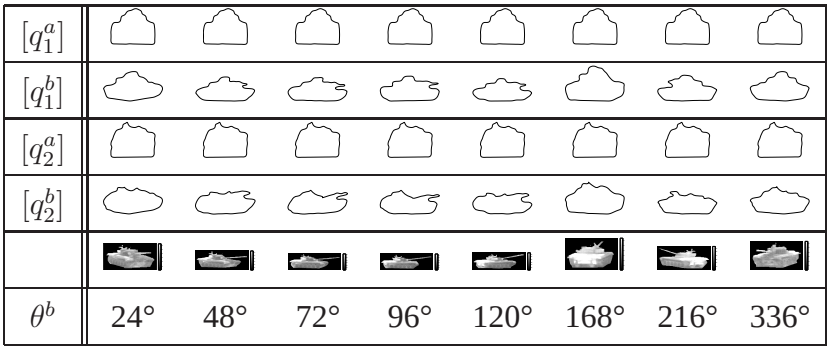

Fig. 9. Examples of shape predictions using parallel transport. In each row, the first two are given shapes of the M60 from $\theta^a = 0$ and θ^b. The deformation between these two is used to deform the T72 shape in the third row and obtain a predicted shape (fourth row). The accompanying pictures show the true shapes of the T72 at those views.

the M60 and one azimuth for the T72, we would like to predict shapes for the T72 from the other azimuthal angles. Since both the objects are tanks, they do have similar geometries but with some differences. For instance, the T72 has a longer gun than the M60. In this experiment, we select $\theta^a = 0$ and predict the shape of the T72 for several θ^b The results are shown in Figure 9. The first and the third rows show the shapes for $[q_1^a]$ and $[q_2^a]$, respectively, the shapes for the M60 and the T72 looking from head on. The second row shows $[q_1^b]$ for different θ^b given in the last column, while the fourth row shows the predicted shapes for the T72 from these same θ^b.

How can we evaluate the quality of these predictions? We perform a simply binary classification with and without the predicted shapes and compare results. Before we present the results, we describe the experimental setup. We have 62 and 59 total azimuthal views of the M60 and the T72, respectively. Of these, we randomly select 31 views of M60 and one view

	Experiment 1 ($\theta^a = 90^\circ$)		Experiment 2 ($\theta^a = 0^\circ$)	
$\downarrow$ Selected Class /True Class $\rightarrow$	M60	T72	M60	T72
M60	100% (**100 %**)	53.45% (**39.66%**)	100%(**100 %**)	93.2% (**82.8%**)
T72	0% (**0 %**)	46.55% (**60.34%**)	0% (**0 %**)	6.8% (**17.2%**)

TABLE III

CLASSIFICATION RATE WITH (BOLD FONTS) AND WITHOUT (NORMAL FONTS) USE OF PREDICTED SHAPES FOR THE T72.

of the T72 as the training data; the remaining 31 (58) views of the M60 (the T72) are used for testing. The classification results, using the nearest neighbor classifier and the elastic distance, are shown in the table below. While the classification for the M60 is perfect, as expected, the classification for the T72 is 46.55%. (Actually, this number is somewhat large – we would expect a smaller performance with only one training shape.) Now we generate an additional 31 shapes for the T72 using the prediction method described earlier. Using the 31 training shapes of the M60, we generate 31 analogous shapes of the T72 using parallel transport. The θ^a used here was 90°. The classification result after including the 31 predicted shapes is found to be 60.34%, a 15% increase in the performance when using shape predictions. We performed the same experiment for another azimuth, $\theta^a = 0^\circ$, and the results are listed under experiment 2 in the table. In this case we improve the classification performance from 6.8% to 17.2%, an increase of almost 11%, using the predicted shapes of the T72. While this experiment was performed with only one training shape, one can repeat this idea using multiple given shapes for the novel object and then perform prediction for a novel view using joint information from these views.

VI. SUMMARY

We have presented a new representation of curves that facilitates an efficient elastic analysis of their shapes and is applicable to $\mathbb{R}^n$ for all n. Its most important advantage is that the elastic metric reduces to a simple $\mathbb{L}^2$ metric, Geodesics between shapes are computed in the resulting

shape space using a path-straightening approach. This framework is illustrated using several applications: shape analysis of helical curves in $\mathbb{R}^3$ with applications in protein backbone structure analysis; shapes of 3D facial curves with applications in biometrics; wrapped probability models for capturing shape variability; and parallel transport of deformation models to predict shapes of 3D objects from novel viewpoints.

APPENDIX

Here we list some numerical procedures for computing geodesic paths between curves represented by q_0 and q_1 in $\mathcal{C}^c$. There are two basic items that are used repeatedly in these procedures – one for projecting arbitrary points in $\mathbb{L}^2(D, \mathbb{R}^n)$ into $\mathcal{C}^c$ and another for projecting arbitrary points in $\mathbb{L}^2(D, \mathbb{R}^n)$ into $T_q(\mathcal{C}^c)$ for some $q \in \mathcal{C}^c$.

Item 1: The projection from $\mathbb{L}^2(D, \mathbb{R}^n)$ to $\mathcal{C}^o$ is simple: $q \mapsto q/\|q\|$. The further projection from $\mathcal{C}^o$ to $\mathcal{C}^c$ is realized as follows. Recall the mapping $G : \mathcal{C}^o$ given by $G(q) = \int_0^{2\pi} q(t)\|q(t)\|dt \in \mathbb{R}^n$. Our idea is to iteratively update q in such a way that $G(q)$ becomes $(0, \ldots, 0)$. The update is performed in the normal space $N_q(\mathcal{C}^c)$ since changing q along the tangent space $T_q(\mathcal{C}^c)$ does not change its G value. The question is: which particular normal vector should be used in this update?

1) Calculate the Jacobian matrix, $J_{i,j} = 2\pi\delta_{ij} + 3\int_0^{2\pi} q_i(s)q_j(s)ds$, $i, j = 1, 2, \ldots, n$. Here, $\delta_{ij} = 1$ if $i = j$, else it is zero.
2) Compute the residual $r = \psi(q)$ and solve the equation $J\beta = -r$ for $\beta \in \mathbb{R}^n$.
3) Update $q = q + \sum_{i=1}^n \beta_i b_i, \quad \delta > 0$, where $\{b_i | i = 1, \ldots, n\}$ form an orthonormal basis of the normal space $N_q(\mathcal{C}^c)$ given in Eqn. 1. Rescale using $q \mapsto q/\|q\|$.
4) If $\|r(q)\| < \epsilon$, stop. Else, go to Step 1.

Item 2: For the second item, find an orthonormal basis $\{b_i\}$ of the normal space $N_q(\mathcal{C}^c)$ and project the given vector w using $w \mapsto w - \sum_{i=1}^{n+1} \langle b_i, w\rangle b_i$.

With these two items, we return to the task of straightening paths into geodesics. Let $\{\alpha(\tau/k) : \tau = 0, 1, 2, \ldots, k\}$ be a given path between q_0 and q_1 in $\mathcal{C}^c$. First, we need to compute the velocity vector $\frac{d\alpha}{dt}$ at discrete points along α.

Algorithm 1: [Compute $\frac{d\alpha}{dt}$ along α]

For all $\tau = 0, 1, \ldots, k$,

1) Compute: $c(\tau/k) = k(\alpha(\tau/k) - \alpha((\tau-1)/k))$. This difference is computed in $\mathbb{L}^2(\mathbb{S}^1, \mathbb{R}^n)$.
2) Project $c(\tau/k)$ into $T_{\alpha(\tau/k)}(\mathcal{C}^c)$ using Item 2 to get an approximation for $\frac{d\alpha}{dt}(\tau/k)$.

Next, we want to approximate the covariant integral of $\frac{d\alpha}{dt}$ along α, using partial sums, i.e. we want to add the current sum, say $u((\tau-1)/k)$, to the velocity $\frac{d\alpha}{dt}(\tau/k)$. However, these two quantities are elements of two different tangent spaces and cannot be added directly. Therefore, we project $u((\tau-1)/k)$ to the point $\alpha(\tau/k)$ first and then add it to $\frac{d\alpha}{dt}(\tau/k)$ to estimate $u(\tau/k)$.

Algorithm 2: [Compute covariant integral of $\frac{d\alpha}{dt}$ along α]

Set $u(0) = 0 \in T_{\alpha(0)}(\mathcal{C}^c)$. For all $\tau = 1, 2, \ldots, k$,

1) Project $u((\tau-1)/k)$ into the tangent space $T_{\alpha(\tau/k)}(\mathcal{S}^c)$ (Item 2) and rescale to the original length to result in $u^{\|}((\tau-1)/k)$.
2) Set $u(\tau/k) = \frac{1}{k}\frac{d\alpha}{dt}(\tau/k) + u^{\|}((\tau-1)/k)$

Next, we compute an estimate for the backwards parallel transport of $u(1)$:

Algorithm 3: [Backward parallel transport of $u(1)$]

Set $\tilde{u}(1) = u(1)$ and $l = \|u(1)\|$. For all $\tau = k-1, k-2, \ldots, 0$,

1) Project $\tilde{u}((\tau+1)/k)$ into $T_{\alpha(\tau/k)}(\mathcal{C}^c)$ using Item 2 to obtain $c(\tau/k)$.
2) Set $\tilde{u}(\tau/k) = lc(\tau/k)/\|c(\tau/k)\|$.

Now we can compute the desired gradient:

Algorithm 4: [Gradient vector field of E in $\mathcal{H}_0$]

For all $\tau = 1, 2, \ldots, k$, compute $w(\tau/k) = u(\tau/k) - (\tau/k)\tilde{u}(\tau/k)$.

By construction, this vector field, w, is zero at $\tau = 0$ and $\tau = k$. As a final step, we need to update the path α in direction opposite to the gradient of E.

Algorithm 5: [Path update]

Select a small $\epsilon > 0$ as the update step size. For all $\tau = 0, 1, \ldots, k$, perform

1) Compute the gradient update $\alpha'(\tau/k) = \alpha(\tau/k) - \epsilon w(\tau/k)$. This update is performed in the ambient space $\mathbb{L}^2(\mathbb{S}^1, \mathbb{R}^n)$.

2) Project $\alpha'(\tau/k)$ to $\mathcal{C}^c$ using Item 1 to obtain the updated $\alpha(\tau/k)$.

REFERENCES

[1] S. Amari. *Differential Geometric Methods in Statistics*. Lecture Notes in Statistics, Vol. 28. Springer, 1985.

[2] A. Bhattacharya. On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calcutta Math. Soc.*, 35:99–109, 1943.

[3] A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Three-dimensional face recognition. *International Journal of Computer Vision*, 64(1):5–30, 2005.

[4] N. N. Čencov. *Statistical Decision Rules and Optimal Inferences*, volume 53 of *Translations of Mathematical Monographs*. AMS, Providence, USA, 1982.

[5] I. L Dryden and K.V. Mardia. *Statistical Shape Analysis*. John Wiley & Son, 1998.

[6] Z. Han et al. Comparison of group average and individual differences in brain morphometry in williams syndrome. In *International Society for Magnetic Resonance in Medicine*, April 18-24, 2009.

[7] S. H. Joshi, E. Klassen, A. Srivastava, and I. Jermyn. Removing shape-preserving transformations in square-root elastic (SRE) framework for shape analysis of curves. In *Proc. of* 6^{th} *Intl. Conf. on Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR), Hubei, China*, pages 387–398, 2007.

[8] S. H. Joshi, E. Klassen, A. Srivastava, and I. H. Jermyn. A novel representation for riemannian analysis of elastic curves in $\mathbb{R}^n$. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–7, 2007.

[9] H. Karcher. Riemannian center of mass and mollifier smoothing. *Comm. Pure and Applied Mathematics*, 30(5):509–541, 1977.

[10] D. Kaziska and A. Srivastava. Joint gait-cadence analysis for human identification using an elastic shape framework. *Communications in Statistics Theory and Methods*, accepted for publication, 2009.

[11] D. G. Kendall. Shape manifolds, procrustean metrics and complex projective spaces. *Bulletin of London Mathematical Society*, 16:81–121, 1984.

[12] N. Khaneja, M.I. Miller, and U. Grenander. Dynamic programming generation of curves on brain surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1260–1264, Nov 1998.

[13] M. Kilian, N. J. Mitra, and H. Pottmann. Geometric modeling in shape space. In *Proceedings of SIGGRAPH*, 2006.

[14] E. Klassen, A. Srivastava, W. Mio, and S. H. Joshi. Analysis of planar shapes using geodesic paths on shape spaces. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 26(3):372–383, 2004.

[15] S. Lang. *Fundamentals of Differential Geometry*. Springer, 1999.

[16] F.-F. Li, R. Fergus, and P. Perona. One-shot learning of object categories. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(4):594–611, 2006.

[17] M. Mani, A. Srivastava, and C. Barillot. The labeling of cortical sulci using multidimensional scaling. In *In Proceedings of MICCAI Workshop on Manifolds in Medical Imaging*, September 2008.

[18] P. W. Michor and D. Mumford. Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc.*, 8:1–48, 2006.

[19] J. W. Milnor. *Topology from the Differentiable Viewpoint*. Princeton University Press, 1997.

[20] W. Mio, A. Srivastava, and S. H. Joshi. On shape of plane elastic curves. *Intl. Journal of Computer Vision*, 73(3):307–324, 2007.

[21] I. Mpiperis, S. Malassiotis, and M. G. Strintzis. 3-D face recognition with the geodesic polar representation. *IEEE Transactions on Information Forensics and Security*, 2(3):537 – 547, 2007.

[22] R. S. Palais. Morse theory on Hilbert manifolds. *Topology*, 2:299–349, 1963.

[23] C. Samir, A. Srivastava, and M. Daoudi. Three-dimensional face recognition using shapes of facial curves. *IEEE Trans. Pattern Anal. Mach. Intell.*, 28(11):1858–1863, 2006.

[24] C. Samir, A. Srivastava, M. Daoudi, and S. Kurtek. On analyzing symmetry of objects using elastic deformations. In *In Proceedings of International Conference on Computer Vision Theory and Applications (VISAPP)*, February 2009.

[25] S. Savarese and F.-F. Li. View synthesis for recognizing unseen poses of object classes. In *ECCV*, 2008.

[26] F. R. Schmidt, M. Clausen, and D. Cremers. Shape matching by variational computation of geodesics on a manifold. In *Pattern Recognition (Proc. DAGM)*, volume 4174 of *LNCS*, pages 142–151, Berlin, Germany, September 2006. Springer.

[27] T. B. Sebastian, P. N. Klein, and B. B. Kimia. On aligning curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(1):116–125, 2003.

[28] J. Shah. H^0-type Riemannian metrics on spaces of planar curves. *Preprint*, 2005. http://arxiv.org/abs/math.DG/0510192.

[29] J. Shah. An H^2 type riemannian metric on the space of planar curves. In *Workshop on the Mathematical Foundations of Computational Anatomy, MICCAI*, 2006.

[30] C. G. Small. *The Statistical Theory of Shape*. Springer, 1996.

[31] A. Srivastava and I. H. Jermyn. Looking for shapes in two-dimensional, cluttered point clouds. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, to appear, doi: 10.1109/TPAMI.2008.223, 2009.

[32] A. Srivastava, S. H. Joshi, W. Mio, and X. Liu. Statistical shape anlaysis: Clustering, learning and testing. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 27(4):590–602, 2005.

[33] A. Srivastava, C. Samir, S. H. Joshi, and M. Daoudi. Elastic shape models for face analysis using curvilinear coordinates. *Journal of Mathematical Imaging and Vision*, 33(2):253–265, February 2009.

[34] M. Vaillant, M.I. Miller, L. Younes, and A. Trouve. Statistics on diffeomorphisms via tangent space representations. *NeuroImage*, 23:161–169, 2004.

[35] L. Younes. Computable elastic distance between shapes. *SIAM Journal of Applied Mathematics*, 58(2):565–586, 1998.

[36] L. Younes, P. W. Michor, J. Shah, D. Mumford, and R. Lincei. A metric on shape space with explicit geodesics. *Matematica E Applicazioni*, 19(1):25–57, 2008.

[37] L. Younes, A. Qiu, R. L. Winslow, and M. I. Miller. Transport of relational structures in groups of diffeomorphisms. *J. Math. Imaging and Vision*, 32(1):41–56, 2008.