Shape Analysis of Elastic Curves in Euclidean Spaces

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

Abstract—This paper introduces a square-root velocity (SRV) representation for analyzing shapes of curves in Euclidean spaces under an elastic metric. Due to this SRV representation the 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 unit sphere. The shape space of closed curves is quotient space of (a submanifold of) the unit sphere, modulo rotation and re-parameterization groups, and we find geodesics in that space using a path-straightening approach. These geodesics and geodesic distances provide a framework for optimally matching, deforming and comparing shapes. These ideas are demonstrated using: (i) Shape analysis of cylindrical helices for studying protein backbones, (ii) Shape analysis of facial curves for recognizing faces, (iii) A wrapped probability distribution for capturing shapes of planar closed curves, and (iv) Parallel transport of deformations for predicting shapes from novel poses.

Index Terms—Elastic curves, Riemannian shape analysis, elastic metric, Fisher-Rao metric, square-root representations, path-straightening method, elastic geodesics, parallel transport, shape models.

1 Introduction

Shape is an important feature for characterizing objects in several branches of science, including computer vision, medical diagnostics, 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 came from Kendall and colleagues [6], [12]. While this formulation took major strides in shape analysis, 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 [33] who first defined shape spaces of *planar* curves and imposed Riemannian metrics on them. In particular, he computed geodesic paths between curves under these metrics as open curves and "closed" the curves along those geodesics to obtain deformations between closed curves. 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 do so directly on the space of closed curves and in a manner that is invariant to re-parameterization. Among other things, they applied this framework to statistical modeling and analysis using large databases of shapes [30]. Michor and Mumford [18] and Mennucci [17], [32] have exhaustively studied several choices of Riemannian metrics on spaces of planar curves for the purpose of comparing their shapes. Mio et al. [20] presented a family of elastic metrics that quantified the relative amounts of bending and stretching needed to deform shapes into each other. Similarly, Shah [27] derived geodesic equations for planar closed curves under different elastic metrics and different representations of curves. In all 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 belong to the same orbits 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 can inherit this Riemannian structure and become a quotient manifold or an 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 [2], [28]. A similar idea was introduced by Younes [33] and later used in Younes et al. [34] for studying shapes of *planar* curves

• *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 Department of Mathematical Sciences, Durham University Science Laboratories, Durham, UK.*

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 [8], [9]. 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 reparameterization 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 been applied to shape representations using deformable templates [35] and for studying shapes of 3D triangulated meshes [13].

The proposed representation spaces for curves are infinite-dimensional manifolds, or rather their quotient spaces under the actions of infinite-dimensional groups. The infinite-dimensionality of such representations is an important challenge. At a conceptual level, however, it may help a reader to understand the proposed solutions on finite-dimensional manifolds at first and consider the issue of infinite-dimensionality later. Also, we clarify the use of word geodesic in this paper. We refer to a path with a (covariantly) constant velocity (defined later in Section IV) as a *geodesic* and the shortest geodesic between any two points as a *minimizing geodesic*.

The paper is organized as follows. Section 2 introduces the proposed elastic shape framework, while Section 3 discusses its merits relative to existing literature. Section 4 describes a path-straightening approach for finding geodesics and a gradient-based approach for elastic curve registration. Section 5 presents four applications of this framework. The paper ends with a short summary in Section 6.

2 Shape Representation

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

2.1 SRV Representation and Pre-Shape Space

Let β be a parameterized curve ($\beta : D \to \mathbb{R}^n$), where D is a certain domain for the parameterization. We are going to restrict to those β that are absolutely continuous on D. In general D will be $[0,1]$, but for closed curves it will be more natural to have $D = \mathbb{S}^1$. We define a mapping: $F : \mathbb{R}^n \to \mathbb{R}^n$ according to $F(v) \equiv v/\sqrt{\|v\|}$ if $\|v\| \neq 0$ and 0 otherwise. Here, $\|\cdot\|$ is the Euclidean 2-norm in $\mathbb{R}^n$; 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 $q : D \to \mathbb{R}^n$, 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 presented in the Section 3.1.

To remove the scaling variability, we rescale all curves to be of unit length. This restriction to an orthogonal section of the full space of curves is identical to Kendall's [12] approach for removing the scale variability. The remaining transformations (rotation, translation, and reparameterization) will be dealt with differently. This is due to the differences in the actions of scaling and other groups on the representation space of curves, as described later. The restriction that β is of unit length translates to the condition that $\int_D \|q(t)\|^2 dt = \int_D \|\dot{\beta}\| dt = 1$. Therefore, the SRV functions associated with these curves are elements of a unit hypersphere in the Hilbert manifold $\mathbb{L}^2(D, \mathbb{R}^n)$; we will use $\mathcal{C}^o$ 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 the curve starts and ends at the same point. In view of this condition, it is natural to have the domain D be the unit circle $\mathbb{S}^1$ for closed curves. For a certain placement of the origin on $\mathbb{S}^1$, it can be identified with $[0,1]$ using the function $t \mapsto (\cos(2\pi t), \sin(2\pi t))$. We will use either one 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:

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

The superscript c implies the closure condition. With the earlier identification of $[0,1]$ 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 $\mathcal{C}^c$? In the Appendix, we sketch a proof that $\mathcal{C}^c$ is a codimension-n submanifold of $\mathcal{C}^o$.

Now we have two submanifolds – $\mathcal{C}^o$ and $\mathcal{C}^c$ – containing all curves and only 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 just

to emphasize that 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 $\mathbb{L}^2([0,1],\mathbb{R}^n)$, its tangent space at a point q is given by: $T_q(\mathcal{C}^o) = \{v \in \mathbb{L}^2([0,1],\mathbb{R}^n) | \langle v, q\rangle = 0\}$. Here $\langle v, q\rangle$ denotes the inner product in $\mathbb{L}^2([0,1],\mathbb{R}^n)$: $\langle v, q\rangle = \int_0^1 \langle v(t), q(t)\rangle dt$.

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 $q \mapsto G(q) = \int_{\mathbb{S}^1} q(t)\|q(t)\|dt$. 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 $\mathbb{L}^2(\mathbb{S}^1,\mathbb{R}^n)$ that are perpendicular to $T_q(\mathcal{C}^c)$. This normal space is found using the directional derivatives of G and is given by:

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

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 the two manifolds $\mathcal{C}^o$ and $\mathcal{C}^c$ to form Riemannian structures on them. 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: $L[\alpha] = \int_0^1 \langle \dot\alpha(\tau), \dot\alpha(\tau)\rangle^{1/2} d\tau$, and α is a said to be a *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 a numerical methods described in Section 4.

2.2 Shape Space as Quotient Space

By representing a parameterized curve β by its SRV function q, and imposing the constraint $\int_D \langle q(t), q(t)\rangle dt = 1$, 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 (as shown in Fig. 1); 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 us the *right* action

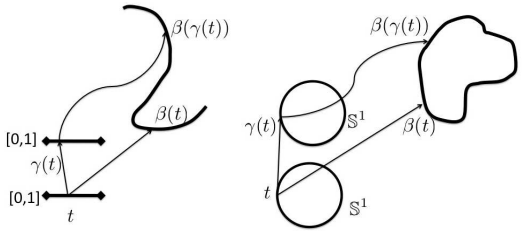

Fig. 1. Re-parameterizations of open and closed curves using orientation-preserving diffeomorphisms.

$\mathcal{C} \times \Gamma \to \mathcal{C}, \quad (q, \gamma) = (q \circ \gamma)\sqrt{\dot\gamma}$. In order for our shape comparison 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: It follows from the definition.

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

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

$$\int_D \langle \tilde u(t), \tilde v(t)\rangle dt = \int_D \langle u(\gamma(t))\sqrt{\dot\gamma(t)}, v(\gamma(t))\sqrt{\dot\gamma(t)}\rangle dt$$
$$= \int_D \langle u(\tau), v(\tau)\rangle d\tau, \quad \text{with} \quad \tau = \gamma(t)\ .$$

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. □

Since the action of the product group is by isometries, we can form a quotient space of $\mathcal{C}$ modulo $\Gamma \times SO(n)$ and try to inherit the Riemannian metric from $\mathcal{C}$ to that quotient space. The orbit of a function $q \in \mathcal{C}$ is given by:

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

An orbit is associated with a shape uniquely 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. We would like to use the basic fact that if a compact Lie group H acts freely on a Riemannian manifold M (*i.e.* , no elements of M are fixed by $h \in H$ unless h is the identity) by isometries, and if the orbits are closed, then the quotient M/H is a manifold, and inherits a Riemannian metric from M. The trouble is that while we have our group $\Gamma \times SO(n)$ acting by isometries, the orbits are not closed. The reason for this is that the space of diffeomorphisms is not closed with respect to either the $\mathbb{L}^2$ or the Palais metric, since a

sequence of diffeomorphisms might approach a map which is not a diffeomorphism under either of these two metrics. To resolve this theoretical difficulty, we propose that instead of modding out by the orbits, we mod out by the closures of these orbits. Thus, if we there is a sequence q_i in the orbit $[q]$, and this sequence converges to a function $\tilde{q}$ in $\mathcal{C}^o$ (with respect to the $\mathbb{L}^2$-metric), then we identify q with $\tilde{q}$ in this quotient construction. As evidence that this idea has merit, one can prove that in this situation, if we let β and $\tilde{\beta}$ be the curves corresponding to q and $\tilde{q}$, both β and $\tilde{\beta}$ contain exactly the same points. (This is assuming that we set $\beta(0) = \tilde{\beta}(0)$.) With a slight abuse of notation, we will use $[q]$ to denote the closure of the orbit of q. 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 $\Gamma \times SO(n)$ acts 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:

$$d_s([q_0],[q_1]) = \inf_{(\gamma,O)\in\Gamma\times SO(n)} d_c(q_0, O(q_1 \circ \gamma)\sqrt{\dot{\gamma}}) \ . \quad (2)$$

We state without proof that if q_0 and q_1 lie in two different orbits which are not in each other's closure, then this distance is strictly positive.

3 Motivation & Comparisons

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

3.1 Motivation for the SRV Representation

Let $\beta : D \to \mathbb{R}^n$ be a curve in $\mathbb{R}^n$. Assume that for all $t \in D$, $\dot{\beta}(t) \neq 0$ (this is only for comparing with past works, our method does not require it). 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 Φ and Θ are sets of smooth maps. 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. In physical terms, ϕ is the (log of the) speed of traversal of the curve, while θ is the direction of the curve at each t.

The tangent space of $\Phi \times \Theta$ at any point (ϕ, θ) is given by $T_{(\phi,\theta)}(\Phi \times \Theta) = \Phi \times \{v \in \mathbb{L}^2(D, \mathbb{R}^n) | v(t) \perp \theta(t), \forall t \in D\}$. We now define a Riemannian metric on $\Phi \times \Theta$.

Definition 1 (Elastic Metric): Let a and b be positive real numbers. For $(u_1, v_1),\ (u_2, v_2) \in T_{(\phi,\theta)}(\Phi \times \Theta)$, define an inner product:

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

Note that $\langle \cdot, \cdot \rangle$ in the second integral on the right denotes the standard dot product in $\mathbb{R}^n$. This elastic metric, introduced in [20], has the interpretation that the first integral measures the amount of "stretching", since u_1 and u_2 are variations of the log 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 that 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 $(O, (\phi, \theta)) = (\phi, O\theta)$ and $\gamma \in \Gamma$ 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.* ,

$$\begin{aligned}\langle (O, (u_1, v_1)), (O, (u_2, v_2)) \rangle_{(O,(\phi,\theta))} &= \langle (u_1, v_1), (u_2, v_2) \rangle_{(\phi,\theta)} \\ \langle (\gamma, (u_1, v_1)), (\gamma, (u_2, v_2)) \rangle_{(\gamma,(\phi,\theta))} &= \langle (u_1, v_1), (u_2, v_2) \rangle_{(\phi,\theta)}.\end{aligned}$$

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 a to b is important, since if we multiply both by the same real number we just rescale the metric, which results in the same geodesics.

This is not the only consideration, however. The issue of computing geodesics between curves for different choices

of $c = b/2a$ remains, especially once we restrict attention to the space of unit length curves. One can ask: Is there some particular choice of weights which will be especially natural and which will result in the geodesics becoming easier to compute? We now show that the SRV representation provides an answer to this question.

In terms of (ϕ, θ) SRV is given by $q(t) = e^{\frac{1}{2}\phi(t)}\theta(t)$. A simple derivation shows 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

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

In this computation we have used the fact that $\langle\theta(t), \theta(t)\rangle = 1$, since $\theta(t)$ is an element of the unit sphere, and that $\langle\theta(t), v_i(t)\rangle = 0$, since each $v_i(t)$ is a tangent vector to the unit sphere at $\theta(t)$. This expression, when compared with Eqn. 3, shows that the $\mathbb{L}^2$ metric on the space of SRV representations corresponds precisely to the elastic metric on $\Phi \times \Theta$, with $a = 1/2$ and $b = 1$. However, expressed in terms of the SRV functions, the $\mathbb{L}^2$-metric is the "same" at every point of $\mathbb{L}^2(D, \mathbb{R}^n)$ (it is simply $\langle f_1, f_2\rangle = \int_D \langle f_1(t), f_2(t)\rangle\, dt$, which does not depend on the point at which these tangent vectors are defined), and we will thus have access to 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.

3.2 Comparison with Prior Work

The previous subsection showed that the SRV representation provides Euclidean coordinates for the space of parameterized curves in $\mathbb{R}^n$ equipped with the elastic metric. In this subsection, we compare the SRV representation to previous work, and provide evidence that this is the only case for which Euclidean coordinates can be found.

When $n = 1$, 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 called 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, given by Eqn. 4 with $n = 1$.

To discuss $n > 1$, it is useful to use a slightly different representation. Let us define $q_c = \dot{\beta}(t)/\|\dot{\beta}(t)\|^{1-\frac{1}{2c}}$. For v_c, w_c in the tangent space at q_c, the elastic metric becomes:

$$\langle v_c, w_c\rangle_{q_c} = b^2 \int_D \|q_c(t)\|^{(2c-2)}\, \langle v_c(t), w_c(t)\rangle\, dt\ . \quad (5)$$

Notice that when $c = 1$, the integrand is the Euclidean metric on $\mathbb{R}^n$, otherwise it is not. If we use a discrete representation of curves, say using N points sampled on each curve, one can calculate the curvature of the resulting finite-dimensional representation space (details are omitted). This calculation shows that:

- when $c \neq 1$: for $n = 2$, the representation space of curves is flat except at $q_c = 0$, where it is singular; for $n > 2$, the curvature is again singular at $q_c = 0$, otherwise it is non-flat (the curvature is not zero).
- when $c = 1$: the curvature is *identically zero for all* n; the space of curves is flat.

Euclidean coordinates thus exist for all n *only* when $c = 1$: these coordinates are the *SRV representation*. We conjecture that this situation continues to hold in the infinite-dimensional case. This would mean that the SRV representation occupies a *unique* position amongst curve representations. We are unaware of any previous work that discusses an SRV-type representation for $n > 2$; the method described in Younes et al. [34] is for $n = 2$.

4 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: $\alpha : [0, 1] \to \mathcal{C}^o$,

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

where $\theta = \cos^{-1}(\langle q_0, q_1\rangle)$ is the length of the geodesic. However, we will use a path-straightening approach to compute geodesics in $\mathcal{C}^c$.

Notationally, we are using τ to parameterize paths on spaces of curves and t to parameterize individual curves.

4.1 Path-Straightening Method: Theory

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 $\alpha(\tau)$ 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:

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

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 a minimizing geodesic. Therefore, this approach has the

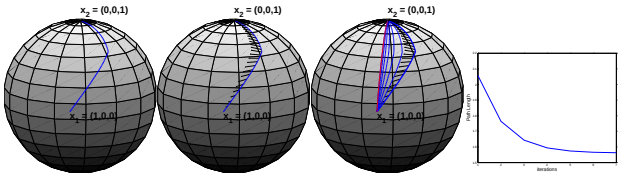

Fig. 2. An example of path-straightening method for computing geodesics between two points on $\mathbb{S}^2$. The right panel shows the decrease in the path length.

limitation that it finds a geodesic between a given pair but may not reach the minimizing geodesic, if it exists.

Let $\mathcal{H}$ be the set of all paths in $\mathcal{C}^c$ 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 $N_q(\mathcal{C}^c)$ (defined in Eqn. 1). A tangent w is actually a tangent 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 τ, and is denoted by $\frac{Dw}{d\tau}(\tau)$. 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:

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

where $\langle \cdot, \cdot \rangle$ is the chosen metric on $\mathcal{C}^c$. The 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 $T_\alpha(\mathcal{H}_0)$. This enables us to derive the gradient of E as a vector field on α.

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: Please refer to the appendix.

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 $Dw/d\tau$ 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 $\frac{D}{d\tau}(\frac{d\alpha}{d\tau}) = 0$ for all τ. Also, a vector field w along the path α is called **covariantly linear** if $Dw/d\tau$ 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: Please refer to the appendix.

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 $\frac{Du(\tau)}{d\tau} = 0$ for all $\tau \in [0,1]$. 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 $\tilde{\alpha}(\tau) \equiv \alpha(1-\tau)$, 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 $\tau = 0$ and $\tau = 1$ (recall that $u(0) = 0$). Clearly, the correct covariantly linear field is simply $\tau\tilde{u}(\tau)$, where $\tilde{u}(\tau)$ 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 $w(\tau) = u(\tau) - \tau\tilde{u}(\tau)$, 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 $u(\tau) - \tau\tilde{u}(\tau)$, we have that $u(\tau) = \tau\tilde{u}(\tau)$ for all τ. Therefore, $\frac{d\alpha}{d\tau} = \frac{Du}{d\tau} = \frac{D(\tau\tilde{u})}{d\tau} = \tilde{u}$. Since $\tilde{u}$ is a parallel translation of $u(1)$, it is covariantly constant, and therefore, the velocity field $\frac{d\alpha}{d\tau}$ is covariantly constant. By definition, this implies that α is a geodesic. □

4.2 Path-Straightening Method: Implementation

We present 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: 1. For projecting arbitrary points in $\mathbb{L}^2(\mathbb{S}^1, \mathbb{R}^n)$ into $\mathcal{C}^c$, and 2. For projecting arbitrary points in $\mathbb{L}^2(\mathbb{S}^1, \mathbb{R}^n)$ 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, $J_{i,j} = \delta_{ij} + 3\int_{\mathbb{S}^1} q_i(s)q_j(s)ds$, $i, j = 1, 2, \ldots, n$. 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, take the 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 can address 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 $\frac{d\alpha}{d\tau}$ at discrete points along α.

Algorithm 1: [Compute $\frac{d\alpha}{d\tau}$ 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 $\frac{d\alpha}{d\tau}(\tau/k)$.

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 $\frac{d\alpha}{d\tau}(\tau/k)$. However, these two quantities are elements of two different tangent spaces and cannot be added directly. Therefore, we project $u((\tau - 1)/k)$ into the tangent space at 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 $\frac{d\alpha}{d\tau}$ 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 $u^{\parallel}((\tau - 1)/k)$.
2) Set $u(\tau/k) = \frac{1}{k}\frac{d\alpha}{d\tau}(\tau/k) + u^{\parallel}((\tau - 1)/k)$

Next, we compute an estimate for the backward 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 $\alpha(\tau/k)$.

4.3 Path-Straightening Algorithm

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 previous section.

Path-Straightening Algorithm: To find a geodesic between two curves β_0 and β_1 in $\mathcal{C}^c$.

1) *Compute their representations q_0 and q_1 in $\mathcal{C}^c$.*
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.*
3) *Compute the velocity vector field $d\alpha/d\tau$ along the path α using Algorithm 1.*
4) *Compute the covariant integral of $d\alpha/d\tau$, denoted by u, using Algorithm 2.*
5) *Compute the backward parallel transport of the vector $u(1)$ along α using Algorithm 3; denote it by $\tilde{v}$.*
6) *Compute the full gradient vector field of the energy E along the path α, denoted by w, using $w(\tau) = u(\tau) - \tau\tilde{u}_1(\tau)$ (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.*

In these implementations, each curve is represented by its coordinates at some sampled points and the algorithm smoothly interpolates between them when needed. The derivatives are approximated using symmetric finite differences and integrals are approximated using summations.

4.4 Removing Shape-Preserving Transformations

Now that we have procedures for constructing geodesics between points in a preshape space $\mathcal{C}$ ($\mathcal{C}^o$ or $\mathcal{C}^c$), we focus on the same task for shape spaces. Towards this goal, we need to solve the joint minimization problem on (γ, O) stated in Eqn. 2, with the cost function being $H : \Gamma \times SO(n) \to \mathbb{R}$, $H(\gamma, O) = d_c(q_0, O(q_1 \circ \gamma)\sqrt{\dot{\gamma}})$. This optimization problem is depicted using a cartoon diagram in Fig. 3 (left). Our strategy is to fix one variable and iteratively optimize over the other. In case of $\mathcal{C}^o$, this procedure is simple since the solutions to individual optimizations are well known. For a fixed γ, the optimization of $H_\gamma = H(\gamma, \cdot)$ over $SO(n)$ is obtained using the SVD while, for a fixed O, the optimization of $H_O = H(\cdot, O)$ over Γ is performed using the dynammic programming (DP) algorithm.

In case of $\mathcal{C}^c$, these direct solutions do not apply and we resort to a gradient-based approach. Let $\gamma^{(m)} = \gamma_1 \circ \gamma_2 \circ \cdots \circ \gamma_m$ and $O^{(m)} = O_1 \cdot O_2 \cdot \cdots \cdot O_m$ be the cumulative group elements and at the k^{th} iteration we seek the increments (γ_{m+1}, O_{m+1}) that minimize $H(\gamma^{(m+1)}, O^{(m+1)})$. Let $\tilde{q}_1$ denote the current element of the orbit $[q_1]$, *i.e.* $\tilde{q}_1 = O^{(m)}(q_1 \circ \gamma^{(m)})\sqrt{\dot{\gamma}^{(m)}}$ 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 $v \equiv \dot{\alpha}(1)/\|\dot{\alpha}(1)\|$. This v is precisely the gradient of $d_c(q_0, \tilde{q}_1)$ with respect to $\tilde{q}_1$.

1) **Rotations**: In the case of $\mathcal{C}^o$, 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

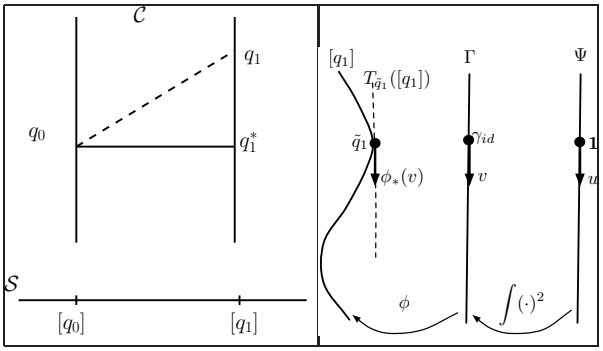

Fig. 3. Left: Computing geodesics in the quotient space $\mathcal{C}/(\Gamma \times SO(n))$. Right: The mapping from $u \in T_{\mathbf{1}}(\Psi)$ to the tangent vector in $T_{\tilde{q}_1}([q_1])$ in two steps.

corresponding chord-length. Therefore, the optimal rotation is directly written as:

$$\hat{O}_{m+1} = \underset{O \in SO(n)}{\operatorname{argmin}} \|q_0 - O\tilde{q}_1\| = UV^T \ , \quad (8)$$

where $U\Sigma V^T = svd(B)$ and $B = \int_D q_0(t)\tilde{q}_1(t)^T dt$. If the $\det(B) < 0$, then the last column of V^T changes sign before multiplication.

In the case of $\mathcal{C}^c$, the update uses the gradient of H_γ. The tangent space to 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 $n \times n$ skew-symmetric matrices. The gradient updates for rotation are performed by projecting v in this space to obtain $A = \sum_i \langle E_i\tilde{q}_1, v\rangle E_i$ and updating using $O_{k+1} = e^{\delta_o A}\tilde{q}_1$ for a step size $\delta_o > 0$.

2) **Re-parameterizations**: In case of $\mathcal{C}^o$, the optimization over H_O can be performed using the DP algorithm but for $\mathcal{C}^c$ we develop the following gradient iteration. We seek the incremental γ_{m+1} that minimizes H_O. There are two possibilities: One is to take the gradient of $H_O(\gamma^{(m+1)})$ directly with respect to γ_{m+1} and use it to update $\gamma^{(m+1)}$. The other possibility, the one that we have used in this paper, is to use a square-root representation of $\dot{\gamma}$ that often simplifies its analysis. Define $\psi_{m+1} = \sqrt{\dot{\gamma}_{m+1}}$ and re-express γ_{m+1} as the pair $(\gamma_{m+1}(0), \psi_{m+1})$. With a slight abuse of notation, let H_O be a function of $(\gamma_{m+1}(0), \psi_{m+1})$. Note that the space Ψ of all ψ-functions is the unit hypersphere in $\mathbb{L}^2(D, \mathbb{R})$ (of radius one). We initialize with $\gamma_0(t) = t$, with the corresponding representation being $(0, \mathbf{1})$ and $\mathbf{1}$ being the constant function with value one. At the iteration m, we take the gradients of H_O, with respect to $\gamma_{m+1}(0)$ and ψ_{m+1}, and update these individually. The derivative with respect to $\gamma_{m+1}(0)$, evaluated at $(0, \mathbf{1})$, is $\frac{\partial H_O}{\partial \gamma_{m+1}(0)} = \int_D \langle v(t), \frac{d\tilde{q}_1(t)}{dt}\rangle dt$. To obtain the derivative with respect to ψ_{m+1}, consider the sequence of maps $\psi \overset{\int_0^t \psi(s)^2 ds}{\longmapsto} \gamma \overset{\phi}{\longmapsto} r$, where $r \equiv \phi(\gamma) = (\tilde{q}_1 \circ \gamma)\sqrt{\dot{\gamma}}$, as shown in Fig. 3 (right). For the constant function $\mathbf{1} \in \Psi$ and a tangent $u \in T_{\mathbf{1}}(\Psi)$, the differential of the first mapping at $\mathbf{1}$ is $u(t) \mapsto 2\bar{u}(t) = 2\int_0^t u(s)ds$ and for a tangent $w \in T_{\gamma_{id}}(\Gamma)$, the differential of the second mapping at γ_{id} is: $w(t) \mapsto \phi_*(w) \equiv \frac{d\tilde{q}_1}{dt}w + \frac{1}{2}\tilde{q}_1\dot{w}$. Concatenating these two linear maps, we obtain the directional partial derivative of H_O in a direction $u \in T_{\mathbf{1}}(\Psi)$ as:

$$\nabla_\psi H_O(u) = \int_D \langle v(t), \left(2\frac{d\tilde{q}_1(t)}{dt}\bar{u}(t) + \tilde{q}_1(t)u(t)\right)\rangle dt \ .$$

Since $T_{\mathbf{1}}(\Psi)$ is an infinite-dimensional space, we can approximate the gradient of H_O, with respect to the ψ-component, by considering a finite-dimensional subspace of $T_{\mathbf{1}}(\Psi)$, as follows. Form a subspace of $T_{\mathbf{1}}(\Psi) = \{f : D \to \mathbb{R} | \langle f, \mathbf{1}\rangle = \mathbf{0}\}$ using: $\{(\frac{1}{\sqrt{\pi}}\sin(2\pi nt), \frac{1}{\sqrt{\pi}}\cos(2\pi nt)) | n = 1, 2, \ldots, m/2\}$. Then, approximate the partial derivative of H with respect to ψ using $c = \sum_{i=1}^m \nabla_\psi H_O(c_i)c_i$, where the c_is are the basis elements of that subspace. Then, update the ψ component according to: $\mathbf{1} \mapsto \psi_{k+1} \equiv \cos(\delta_g\|c\|)\mathbf{1} + \sin((\delta_g\|c\|)\frac{c}{\|c\|}$, for a step size $\delta_g > 0$. Since Ψ is a hypersphere, this update is simply the exponential map on that sphere, at the point $\mathbf{1}$ and applied to the tangent vector c. This ψ_{m+1} in turn gives $\gamma_{m+1}(t) = \gamma_{m+1}(0) + \int_0^t \psi_{m+1}(s)^2 ds$ and thus $\gamma^{(m+1)}$.

We can now state the algorithm for computing geodesics on shape spaces.

Shape Geodesic Algorithm: Find a geodesic between shapes of two parameterized curves β_0 and β_1 in $\mathcal{S}$ ($\mathcal{S}^o$ or $\mathcal{S}^c$). Compute the representations of each curve in $\mathcal{C}$; denote them by q_0 and q_1, respectively. Set $\tilde{q}_1 = q_1$.

1) *Compute the geodesic α between q_0 and $\tilde{q}_1$ in the preshape space. For $\mathcal{C}^o$, use the analytical expression, while for $\mathcal{C}^c$ use the path-straightening algorithm given in the previous section.*
2) *Removal of nuisance variables:*
   a) **Rotation***: For $\mathcal{C}^o$, use the SVD-based solution (Eqn. 8). For $\mathcal{C}^c$ compute A, the derivative of H_γ with respect to $SO(n)$. and form the rotation update O_{m+1}.*
   b) **Re-parameterization***: For $\mathcal{C}^o$ one can use the DP algorithm. More generally, compute the derivatives of H_O with respect to ψ_{m+1} and $\gamma_{m+1}(0)$, and for the re-parameterization update γ_{m+1}.*
3) *Update $\tilde{q}_1 \mapsto O_{m+1}(\tilde{q}_1 \circ \gamma_{m+1})\sqrt{\dot{\gamma}_{m+1}}$.*
4) *If the norms of the increments are small, then stop. Else return to step 1.*

The two rows in Fig. 4 shows two examples of optimization over Γ. In each case we start with a parameterized curve, shown in (a) and represented by q_1, 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_O. The evolution of the cost

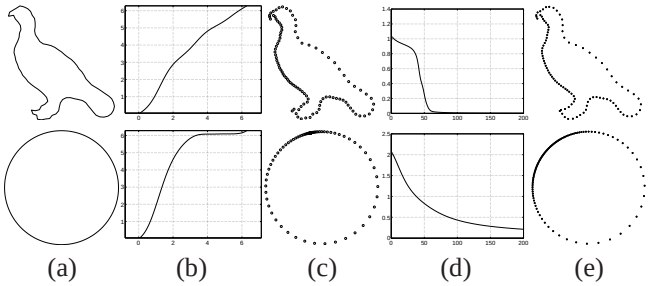

Fig. 4. (a) The original shape represented by q_0, (b) an arbitrary $\gamma \in \Gamma$, (c) the second shape formed using $q_0 = (q_1 \circ \gamma)\sqrt{\dot{\gamma}}$, (d) evolution of H in matching $\tilde{q}_1$ with q_0, (e) final curve represented by $\tilde{q}_1$.

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 1
Timing analysis of gradient-based re-parameterization and comparison with DP algorithm.

function H_O 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_O should be 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 can be mitigated by a improved numerical implementation of the basic procedure.

To illustrate the strengths and limitations of a gradient-based approach with respect to a common DP algorithm [7], [26], we present a comparison of computational costs (using Matlab on a 2.4GHz Intel processor) and performance in Table 1. In this experiment we consider the shape space $\mathcal{S}^o$ since DP is not applicable for optimization in the case of closed curves. 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 $O(T^2)$. 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_O(final)/H_O(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 5 shows some elastic geodesics between several pairs of shapes. We have drawn ticks on these curves to show the optimal re-parametrizations. 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.

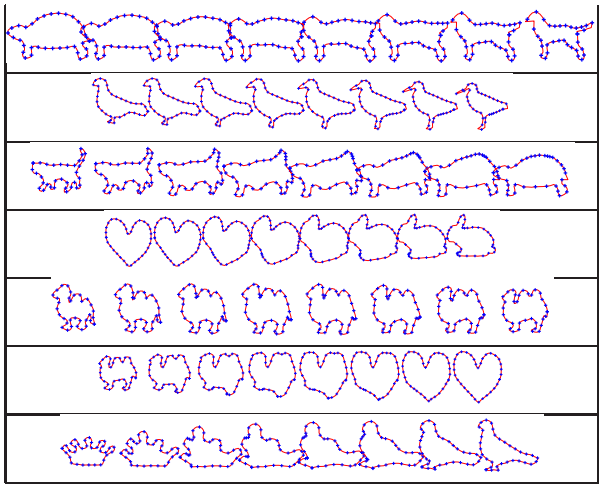

Fig. 5. Examples of planar elastic geodesics.

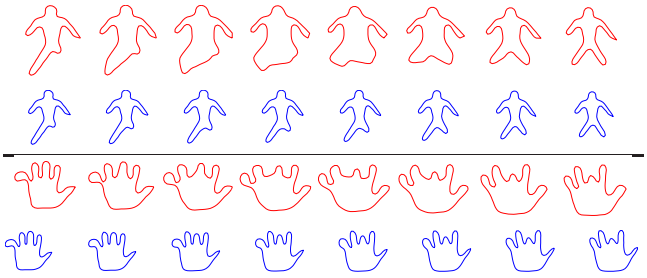

Fig. 6. In each case the top row shows a non-elastic geodesic ([14]) while the bottom rows the elastic geodesic between the same shapes.

Figure 6 compares the elastic geodesics in $\mathcal{S}^c$ with the non-elastic method of Klassen et al. [14] where the representation is restricted to arc-length parameterizations. The resulting deformation is purely bending and no stretching is allowed. We observe that the elastic shape analysis results in a better matching of features across shapes and a more natural deformation along the geodesic path.

5 Applications

In this section we illustrate the proposed elastic shape analysis using some applications. Some additional applications have been presented elsewhere: symmetry analysis of two- and three-dimensional shapes [24]; shape classification of point clouds [29]; and joint gait-cadence analysis for human identification in videos [11].

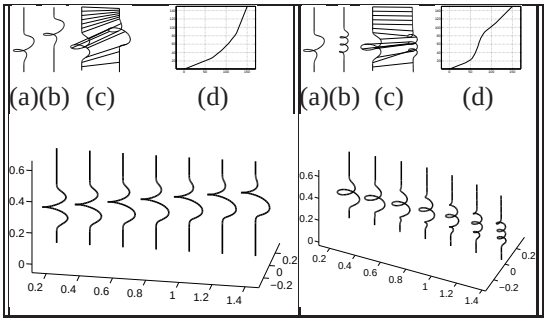

Fig. 7. (a), (b): original curves, (c) optimal registration between them, and (d) optimal γ^*. Bottom: corresponding geodesic paths.

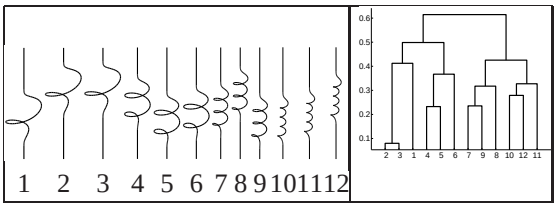

Fig. 8. A set of helices with different numbers and placements of spirals and their clustering using the elastic distance function.

5.1 Shapes Analysis of 3D Helices

As the first example we will study shapes of helical in $\mathbb{R}^3$ by matching and deforming one into another. One motivation for studying shapes of cylindrical helices comes from protein structure analysis. A primary structure in a protein is a linked chain of carbon, nitrogen, and oxygen atoms known as the backbone, and the geometry of the backbone is often a starting point in structural analysis of proteins. 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 use of elastic shape analysis of curves as a framework for studying shapes of protein backbones and present some results involving both synthetic and real data.

Shown in Fig. 7 are two examples of geodesics between some cylindrical helices. In each case, the panels (a) and (b) show 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.

Figure 8 shows an example of using the elastic distances between curves for clustering and classification. In this example, we study 12 cylindrical helices that contain different number, radii, and placements of turns. The first three helices have only one turn, the next three have two turns, and so on. 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 with similar numbers of turns are clustered together.

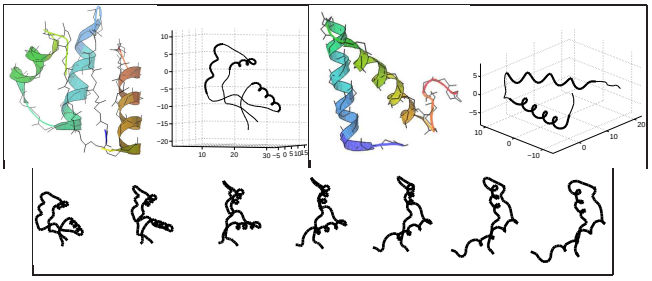

Fig. 9. Two proteins: 1CTF (left) and 2JVD (right) and the elastic geodesic between their shapes.

Finally, in Fig. 9, we present an example of comparing 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 suggest a role for elastic shape analysis in protein structure analysis. Additional details and experiments are presented in [16].

5.2 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 shapes due to facial expressions and measurement errors. Samir et al. [23], [31] have proposed an approach that: (1) computes a function on a facial surface as the shortest-path distance from the tip of the nose (similar to [3], [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 10 (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. We compare shapes of facial surfaces by comparing shapes of the corresponding facial curves, using geodesics between them in $\mathcal{S}^c$. As an example, Fig. 10 (bottom) shows geodesics in $\mathcal{S}^c$ between the two sets of facial curves. For display, these intermediate curves have been rescaled and translated to the original values and, through reconstruction,

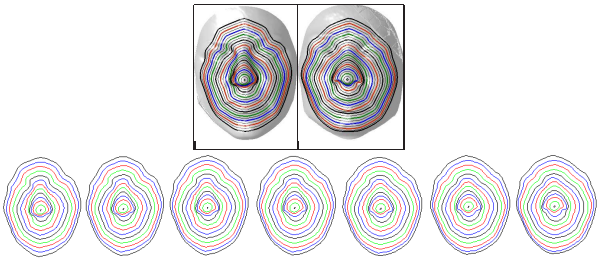

Fig. 10. Top: Two facial surfaces represented by indexed collections of facial curves. Bottom: Geodesics between shapes of corresponding curves.

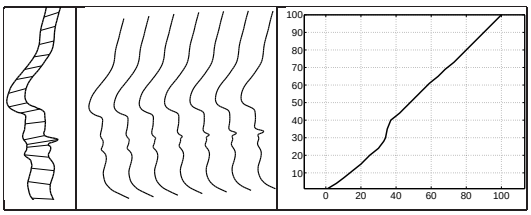

Fig. 11. Elastic geodesics between facial profiles.

they result in a geodesic path such that points along that path approximate full facial surfaces. These geodesic paths can be used to compute average faces or facial parts, or to define metrics for human recognition [5].

Another example of elastic shape analysis of faces, this time using facial profiles is shown in Fig. 11.

5.3 Elastic Models for Planar Shapes

An important application of this elastic shape framework is in developing probability models for capturing the variability present in the observed shapes. For example, the left panel of Fig. 12 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 [6], [30], we demonstrate a simple model where we: (i) first compute the sample Karcher mean [10] 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 $\mathcal{S}^c$.

Let $\mu = \text{argmin}_{[q]\in\mathcal{S}^c} \sum_{i=1}^{n} d_s([q],[q_i])^2$ 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 Fig. 12. Once we have μ, we can map $[q_i]$ into $T_\mu(\mathcal{S}^c)$ using the inverse exponential map: $[q_i] \mapsto v_i \equiv \exp_\mu^{-1}([q_i])$. 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 since we usually have only a finite number of observations. For instance, one can perform PCA on the set $\{v_i\}$ to find dominant directions

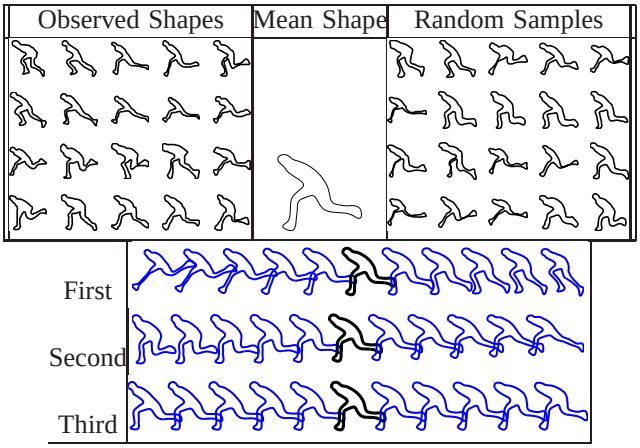

Fig. 12. The left panel shows a set of 20 observed shapes of a "runner" from the Kimia dataset. The middle panel shows their Karcher mean, and the right panel shows a random sample of 20 shapes from the learned wrapped nonparameteric model on $\mathcal{S}^c$. The bottom three rows show eigen variations of shapes in three dominant directions around the mean, drawn from negative to positive direction and scaled by the corresponding eigen values.

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 $\tau\sigma_i U_i \mapsto \exp_\mu(\tau\sigma_i U_i)$ helps visualize these principal modes as shapes. The three principal components of the 20 given shapes are given in the lower three rows of Fig. 12, each row displaying some shapes from $\tau = -1$ to $\tau = 1$.

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, 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 Fig. 12 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.

5.4 Transportation of Shape Deformations

One difficulty in using shapes for recognizing 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 an interesting problem of predicting

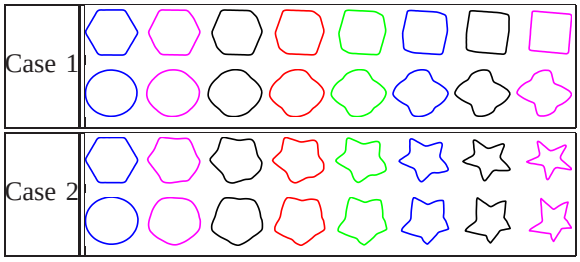

Fig. 13. In each case: a geodesic from the template shape (hexagon) to the training shape (top) and deformation of the test shape (circle) with the transported deformation (bottom).

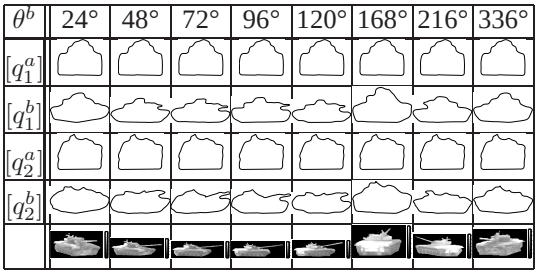

Fig. 14. Shape predictions for novel pose. In each column, the first two are given shapes of the M60 from $\theta^a = 0$ and θ^b. The deformation between these two is used to deform the T72 shape in the third row and obtain a predicted shape in the fourth row. The accompanying pictures show the true shapes of the T72 at those views.

shapes of three-dimensional objects from novel viewing angles. (The problem of predicting full appearances, using pixels, has been studied by [25] and others.) Our solution to the problem of shape prediction is the following. 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 deformations of one shape cannot simply be applied to another.

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 but not identical 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 deforms $[q_1^a]$ to $[q_1^b]$ and applying it to $[q_2^a]$ after some adjustments.

1) Let $\alpha_1(\tau)$ be a geodesic between $[q_1^a]$ and $[q_1^b]$ in $\mathcal{S}^c$ and $v_1 \equiv \dot{\alpha_1}(0) \in T_{[q_1^a]}(\mathcal{S}^c)$ be its initial velocity.
2) We need to *transport* v_1 to $[q_2^a]$; this is done using forward parallel translation. 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 $\frac{Dw}{d\tau} = 0$ for all points along α_{12}. This is accomplished in practice using Algorithm 2 in Section 4.2. Then, $v_2 \equiv w(1) \in T_{[q_2^a]}(\mathcal{S}^c)$ is a parallel translation of v_1.
3) Construct a geodesic starting from $[q_2^a]$ in the direction of v_2.

Figure 13 shows two examples of this idea. In the top case, 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 $[q_2^b]$. A similar transport is carried out in the bottom example.

Next, we consider an experiment involving the M60 tank as $\mathcal{O}^1$ and the T72 as $\mathcal{O}^2$. Given shapes for different azimuthal pose (fixed elevation) of 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 have similar but not identical geometries. For instance, both have mounted guns but 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 Fig. 14. 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 those 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. Here is 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 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 d_s (Eqn. 2), 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 higher than expected – we would expect a smaller performance with only one training shape.) Now we generate additional 31 shapes for the T72 using the prediction method described earlier. Using the 31 training shapes of the M60, we generate 31 corresponding 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, $\theta^a = 0°$, 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

	Experiment 1 ($\theta^a = 90°$)		Experiment 2 ($\theta^a = 0°$)	
Est. /True	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 2
Classification rate with (bold fonts) and without (normal fonts) use of predicted shapes for the T72.

these views.

6 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 of closed curves are computed 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.

Acknowledgments

This work was partially supported by AFOSR FA9550-06-1-0324, ONR N00014-09-1-0664 and NSF DMS-0915003, and by INRIA/FSU Associated Team "SHAPES".

Appendix

Proof that $\mathcal{C}^c$ is a 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. For the i^{th} component $G_i(q) = \int_{\mathbb{S}^1} q_i(t)\|q(t)\|dt$, $i = 1, 2, \ldots, n$, its directional derivative in a direction $w \in \mathbb{L}^2(\mathbb{S}^1, \mathbb{R}^n)$ is given by:

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

where $\mathbf{e}^i$ is a unit vector in $\mathbb{R}^n$ along the i^{th} coordinate axis. To show that G is surjective, we need to show the functions $\{\frac{q_i(t)}{\|q(t)\|}q(t) + \|q(t)\|\mathbf{e}_i; i = 1, 2, \ldots, n\}$ are linearly independent. Suppose not. This implies that there exists a constant vector $b = (b_1, b_2, \ldots, b_n)$ such that, for all t, $\sum_i b_i(\frac{q_i(t)}{\|q(t)\|}q(t) + \|q(t)\|\mathbf{e}_i) = 0$. This, in turn, implies that for all t, $q(t)$ is in the same direction as a constant vector $\sum_{i=1}^n b_i\mathbf{e}_i$. This proves that for any q function that does not lie in a single one-dimensional subspace, the mapping G is surjective. So the space $\mathcal{C}^c$ is a manifold except at those points. 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" 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, $\mathbb{L}^2(\mathbb{S}^1, \mathbb{R}^n)$. □

Proof of Theorem 1: Define a *variation* of α to be a smooth function, $h(\tau, s)$ with $h : [0, 1] \times (-\epsilon, \epsilon) \to \mathcal{H}$ such that $h(\tau, 0) = \alpha(\tau)$ for all $\tau \in [0, 1]$. The variational vector field corresponding to h is given by $v(\tau) = h_s(\tau, 0)$ where s denotes the second argument in h. Thinking of h as a path of curves in $\mathcal{H}$, indexed by s, 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,

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

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 $\dot{E}(0) = \langle\langle v, u\rangle\rangle$. That is, $\dot{E}(0) = \langle v(0), u(0)\rangle + \int_0^1 \langle \frac{Dv}{d\tau}, \frac{Du}{d\tau}\rangle d\tau$. From this expression it is clear that u must satisfy the initial condition $u(0) = 0$ and the ordinary (covariant) differential equation $\frac{Du}{d\tau} = \frac{d\alpha}{d\tau}$. □

Proof of Lemma 3: 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:

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

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 $T_\alpha(\mathcal{H}_0)$. The lemma follows. □.

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] H. Drira, B. Ben Amor, A. Srivastava, and M. Daoudi. A riemannian analysis of 3d nose shapes for partial human biometrics. In *Proceedings of ICCV*, 2009.

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

[7] M. Frenkel and R. Basri. Curve matching using fast marching method. In *Proc. of EMMCVPR)*, pages 35–51, 2003.

[8] 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 EMMCVPR, Hubei, China*, pages 387–398, 2007.

[9] 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 *Proceedings of IEEE CVPR*, pages 1–7, 2007.

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

[11] D. Kaziska and A. Srivastava. Joint gait-cadence analysis for human identification using an elastic shape framework. *Communications in Statistics Theory and Methods*, 39(10):1817–1831, 2010.

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

[13] M. Kilian, N. J. Mitra, and H. Pottmann. Geometric modeling in shape space. In *Proceedings of SIGGRAPH*, 2007.

[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] W. Liu, A. Srivastava, and J. Zhang. Protein structure alignment using elastic shape analysis. In *ACM Conference on Bioinformatics and Computational Biology*, August 2010.

[17] A. C. G. Mennuci. *Metrics of curves in shape optimization and analysis*. 2009.

[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 *Proceedings of VISAPP*, February 2009.

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

[26] 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.

[27] 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.

[28] A. Srivastava, I. Jermyn, and S. H. Joshi. Riemannian analysis of probability density functions with applications in vision. *IEEE Conference on Computer Vision and Pattern Recognition*, 0:1–8, 2007.

[29] A. Srivastava and I. H. Jermyn. Looking for shapes in two-dimensional, cluttered point clouds. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 31(9):1616–1629, 2009.

[30] 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.

[31] 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.

[32] G. Sundaramoorthi, A. C. G. Mennucci, S. Soatto, and A. Yezzi. A new geometric metric in the space of curves, and applications to tracking deforming objects by prediction and filtering. 2010.

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

[34] 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.

[35] 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.

Anuj Srivastava Anuj Srivastava is a Professor of Statistics at the Florida State University in Tallahassee, FL. He obtained his MS and PhD degrees in Electrical Engineering from the Washington University in St. Louis in 1993 and 1996, respectively. After spending the year 1996-97 at the Brown University as a visiting researcher, he joined FSU as an Assistant Professor in 1997. He has received the Developing Scholar and the Graduate Faculty Mentor Awards at FSU. His research is focused on pattern theoretic approaches to problems in image analysis, computer vision, and signal processing. In particular, he has developed computational tools for performing statistical inferences on certain nonlinear manifolds and has published over 130 journal and conference articles in these areas.

Eric Klassen Eric Klassen received the PhD degree at Cornell University in 1987 in the field of low-dimensional topology. After postdocs at Caltech and the University of California at San Diego, he joined the Department of Mathematics at Florida State University in 1991, where he is currently a professor. He has worked in topology, geometry, gauge theory, and Riemann surfaces, as well as applications to computer vision and pattern recognition.

Shantanu Joshi Shantanu H. Joshi reveived his BE degree in Electronics and Telecommunication from the University of Pune, India in 1998, and MS, and PhD degrees in Electrical Engineering from the Florida State University. He is currently a postdoctoral research fellow in the Laboratory of Neuro Imaging, Department of Neurology, University of California Los Angeles.

Ian H. Jermyn Ian Jermyn received the BA Honours degree (First Class) in Physics from Oxford University in 1986, and the PhD degree in Theoretical Physics from the University of Manchester, UK, in 1991. After working for a total of three years at the International Centre for Theoretical Physics in Trieste, Italy, he began study for a PhD in Computer Vision in the Computer Science department of the Courant Institute of Mathematical Sciences at New York University, receiving the PhD in July 2000. He joined the Ariana research group at INRIA Sophia Antipolis as a postdoctoral researcher in August 2000. From September 2001 to August 2010, he was a Senior Research Scientist in the Ariana group. Since September 2010, he has been Reader in Statistics in the Department of Mathematical Sciences at the University of Durham in the UK. His main research interests include the statistical modelling of shape and texture, and information geometry as applied to inference.