2D Affine and Projective Shape Analysis
Darshan Bryner, Eric Klassen, Huiling Le, and Anuj Srivastava

Abstract—Current techniques for shape analysis tend to seek invariance to similarity transformations (rotation, translation and scale), but certain imaging situations require invariance to larger groups, such as affine or projective groups. Here we present a general Riemannian framework for shape analysis of planar objects where metrics and related quantities are invariant to affine and projective groups. Highlighting two possibilities for representing object boundaries – ordered points (or landmarks) and parameterized curves – we study different combinations of these representations (points and curves) and transformations (affine and projective). Specifically, we provide solutions to three out of four situations and develop algorithms for computing geodesics and intrinsic sample statistics, leading up to Gaussian-type statistical models, and classifying test shapes using such models learned from training data. In the case of parameterized curves, we also achieve the desired goal of invariance to re-parameterizations. The geodesics are constructed by particularizing the path-straightening algorithm to geometries of current manifolds and are used, in turn, to compute shape statistics and Gaussian-type shape models. We demonstrate these ideas using a number of examples from shape and activity recognition.

Index Terms—Affine shape analysis, projective shape analysis, path straightening method, geodesic computation, shape models

1 INTRODUCTION

Shape analysis plays an important role in computer vision and image analysis with important applications in medical diagnostics, target recognition, activity recognition, and many other branches of science. Although there are many ideas in the literature for shape analysis, a common theme in shape analysis is to develop quantities – metrics, averages, modes of variations – that are invariant to certain shape-preserving transformations. Most of the past methods focus primarily on being invariant to similarity transformations (translation, rotation, and global scaling); however, in a variety of practical situations, especially those arising in imaging contexts, the observations become transformed in a more complicated manner than what can be modeled by the similarity group alone. These types of transformations occur, for example, when the image plane of a camera is not parallel to the plane containing the defining part of the shape or when a camera images the same scene from different viewing angles. Here shapes become transformed by perspective effects, and one often needs to go beyond the similarity group to define shape equivalences. The affine and projective groups are both larger than (and contain) the similarity group, and are commonly used to model such shape deformations brought about by perspective skew. In this context, the focus is now on developing shape analysis methods that are invariant to affine and projective transformations.

For a point \( x \in \mathbb{R}^2 \), these transformations and their degrees of freedom (D) are given as follows:

<table>
<thead>
<tr>
<th>Group</th>
<th>Action</th>
<th>Space</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Similarity</td>
<td>( x \mapsto aOx + y )</td>
<td>( O \in SO(2), a &gt; 0, y \in \mathbb{R}^2 )</td>
<td>4</td>
</tr>
<tr>
<td>Affine</td>
<td>( x \mapsto Ax + y )</td>
<td>( A \in GL_+(2), y \in \mathbb{R}^2 )</td>
<td>6</td>
</tr>
<tr>
<td>Projective</td>
<td>( x \mapsto \frac{(Bx + y)^T}{|Bx + y|} )</td>
<td>( B \in GL_+(3) )</td>
<td>8</td>
</tr>
</tbody>
</table>

Here \( GL_+(k) \) is the subgroup of \( GL(k) \) containing only orientation-preserving transformations. (One can replace \( SO(k) \) and \( GL_+(k) \) by \( O(k) \) and \( GL(k) \), respectively, and obtain a similar theory but including reflections.) Compared to the similarity group, which only models rotation, translation and scale, the affine and projective groups can better capture the distortion introduced in imaged contours when objects are imaged from different angles. Although the affine group allows for skew effects, parallel lines still remain parallel after transformation. The extra degrees of freedom present in the projective group can model perspective distortions where parallel lines are skewed to converge to vanishing points located at infinity. The affine group is used to approximate the perspective effects seen in situations such as when the camera is far enough away from the imaging plane or when there is only a slight change in viewing angle [8].

In terms of mathematical representations of planar contours, the past techniques have relied on many different ideas: ordered points or landmarks, level sets, medial axes, parameterized curves, and so on. We will focus on two of these – ordered landmarks and parameterized curves – and develop a framework that incorporates the desired invariance to affine and projective transformations under these representations. Landmark based shape analysis, pioneered by Kendall in [10] and [6], considers a finite set of ordered points, or landmarks, along an object’s boundary. Since the choice of landmarks is difficult and assumes pre-registration of points, there has been a large interest in shape analysis methods of continuous, parameterized curves. These methods use
functions to represent curves and impose Riemannian metrics that are invariant not only to similarity transformations but also to re-parameterizations of curves. This setup has two advantages: the points across curves are optimally registered during shape comparison, and the resulting quantities are independent of the given parametrization of curves. This framework, termed elastic shape analysis, effectively uses a combination of bending and stretching of parts in order to better match features across curves. The recent papers [13], [18], [22] develop the theory and techniques needed for similarity invariant, elastic shape analysis of parametrized curves. In the following, we will use both representations—landmarks and parameterized curves—of objects while incorporating invariances to affine and projective groups.

1.1 Past Work in Affine & Projective Shape Analysis

We summarize past research in these two topic areas individually, starting with the affine shape analysis.

There have been several papers dedicated to handling the affine variability of shapes. Sparr [17] develops affine shape theory in the context of landmark-based shape analysis using subspace computations. Begelfor and Werman in [1] extend this idea and compute distances between shapes as points on a Grassmannian manifold, on which it is possible to compute geodesics and perform statistical analyses. Berthilsson and Åström in [2] further extend the theory in [17] to more general point sets and continuous curves but without handling the parameterization variability of curves. Another common idea is to transform the original shape into a canonical or standardized form. This transformation rescales and centers the shape, and removes the effect of skewing by de-correlating the shape’s $x$ and $y$ coordinates. If the original shapes are within affine transformations, then the resulting shapes will be within rotations of each other (more on this later). This technique is demonstrated in [9] in the context of Independent Component Analysis (ICA) on landmark-based shapes and further demonstrated in [23] and [15] in the context of region-based shapes. Even though region-based methods avoid the need for handling the parameterization variability of curves, they lack the ability to incorporate optimal registration of points across curves.

Projective-invariant shape analysis is not as widely explored as its similarity or affine invariant counterparts due to the complex nature of projective transformations. Although there are many papers in computer vision utilizing projective geometry, they mostly focus on shape recovery, or so called “shape from X” methods. We emphasize that this body of work is distinctly different from our focus, which is projective shape analysis. The majority of past work in projective shape analysis is based on one of two standardizations: (1) the selection and standardization of a projective frame, and (2) standardizations of global properties of shapes. The papers [5], [7], [14], [16] focus on the first technique, but they fail to develop a shape space of projective standardized configurations, or to compute geodesics and shape statistics. Also another major issue with this method is how to select a frame on a shape in a consistent or optimal way. The second, and more appealing, type of projective standardization method focuses on standardization of global properties of the shape. Introduced by Kent and Mardia in [11], this type of projective standardization is in the same vein as affine standardization. Using estimation techniques from [19], they perform “Tyler-standardization” in a landmark-based setting, but do not study the projective shape space nor the resulting geodesics.

1.2 Our Goals and Contributions

While several papers provide procedures to eliminate the variability from an affine or projective transformation before comparing two shapes, i.e. standardization methods, they do not define a formal shape space of standardized shapes, provide a way to compute geodesics on such a space, or even mention statistics of standardized shapes. Table 1 organizes progress in shape analysis according to representation space and level of invariance. The references in bold denote papers that provide all the above-mentioned elements of statistical analysis. Note that in the continuous case, the method must also achieve invariance to the re-parameterization group.

<table>
<thead>
<tr>
<th>Similarity</th>
<th>Landmark</th>
<th>Continuous Curves</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affine</td>
<td>[9], [17], [23], [1], etc.</td>
<td>[2], [4], etc.</td>
</tr>
<tr>
<td>Projective</td>
<td>[7], [11], [14]</td>
<td>[15]</td>
</tr>
</tbody>
</table>

TABLE 1

Past work in each shape analysis category.

Our main goal is to formalize the aforementioned standardization method into a complete statistical shape analysis framework, thus providing a general method to perform shape analysis invariant to affine and projective group actions. Towards this goal we (1) define a space of standardized shapes, (2) choose a Riemannian metric on this space, (3) provide the numerical tools for calculating geodesics between shapes, and (4) compute intrinsic means and covariances and build probability models on this space. Items (3) and (4) depend heavily on the differential geometry of the space defined in (1).

The main contributions of the paper are centered around computations of geodesics, based on particularizing the general path-straightening algorithm to three spaces of standardized shapes: (i) landmark-affine, (ii) curve-affine, and (iii) landmark-projective. The computation of geodesics leads to statistical summaries and Gaussian-type models and is an important tool for intrinsic statistical shape analysis. We point out that while the standardization tools have existed previously for all three cases, the tools for geodesic computations and statistical analyses are first developed in this paper.
Also, the standardization in the curve-affine case, mentioned earlier in [4], has been developed further here, including a novel theoretical result about the existence and uniqueness of a standardization. We also point out that geodesics and statistics have been formalized on space (i) in [1], but in a completely different setting (the Grassmann manifold) than what we present here. Our approach is based on defining sets of standardized shapes, or sections, as representative spaces of shapes by establishing one-to-one maps between them and the desired quotient spaces formed to remove pertinent transformations. This aspect of our construction is novel.

The rest of this paper is as follows. Section 2 introduces the notion of shape as an equivalence class and describes the set of standardized shapes as sections of appropriate orbits. The three sections introduced are: (1) landmark-affine section, (2) curve-affine section, and (3) landmark-projective section. Section 3 outlines the path-straightening technique [18], a numerical procedure for calculating geodesics on a general manifold. Sections 4, 5, and 6 provide the details necessary to implement path-straightening on the three respective sections introduced in Section 2. Various examples of geodesic paths and statistical models computed on each of the three sections are presented. Section 7 shows a variety of experimental results on simulated and real data, and Section 8 the concludes the paper.

2 STANDARDIZED SHAPE REPRESENTATIONS

2.1 Shapes As Equivalence Classes

Shape analysis methods in general are built on a representation space, a manifold $V$ of all possible configurations, and are designed to be invariant under certain shape-preserving transformations. The main mathematical tools to reach these invariances come from group theory. Here one considers the action of a group of transformations $G$ on $V$—any transformation $g \in G$ transforms an element $p \in V$ into $p * g$ (or in the case of a left action, into $g * p$). Under this action we can define an equivalence class, or orbit, of a point $p \in V$ as $[p] = \{p \ast g \mid g \in G\}$. Shapes are now represented by orbits on $V$, and the process of shape analysis refers to comparing these orbits rather than just individual points in $V$. Given a certain metric $d_v(\cdot, \cdot)$ on the space $V$, a group acts by isometries if $d_v(p_1 \ast g, p_2 \ast g) = d_v(p_1, p_2)$ for any $g \in G$. Under this condition, we can define a distance between orbits as $d_v([p_1], [p_2]) = \inf_{g \in G} d_v(p_1 \ast g, p_2) = \inf_{g \in G} d_v(p_1, p_2 \ast g)$. The important property of a group action by isometries allows one to define shape space as the set of orbits or, equivalently, the quotient space $V/G = \{[p] \mid p \in V\}$, and to compute distances in the quotient space. Past works on similarity shape analysis of landmarks and curves are mostly based on this idea.

2.2 Shape Spaces as Sections

However, the pertinent group of shape transformations may not always act by isometries under the commonly used $\mathbb{L}^2$ metric. This is the case for actions of affine and projective groups, and thus we have to develop a new approach to define distances between orbits in the absence of this isometry. It may still be possible to select a representative, or standard, member of the orbit $[p]$, say $p_0 = p \ast g_0$, that uniquely satisfies some specific mathematical criteria. We term the process of applying $g_0$ to $p$ to obtain $p_0$ as standardization, and similarly, we define the process of applying $(g_0)^{-1}$ to $p_0$ to obtain the original point $p$ as de-standardization.

Consider the space $M \subset V$ of all such standard elements. If $M$ intersects each orbit exactly once, then we call $M$ a section of $V$ under the group action of $G$. The usefulness of $M$ comes from the fact that the quotient space $V/G$ can be identified with $M$ through a bijection, and a statistical analysis on $M$ represents, implicitly, an analysis on $V/G$. A Riemannian structure on $M$ allows for computing geodesics and distances for performing statistical analysis on $M$.

Fig. 1 illustrates this setup using a simple example. Consider the standard action of the scaling group $\mathbb{R}_+ \times \mathbb{R}^3\setminus\{0\}$. The orbits under this action are given by rays emanating from (but not including) the origin; the figure shows two such orbits $[p^{(1)}]$ and $[p^{(2)}]$. Since the scaling action is not by isometries under the Euclidean norm, we take a section $M = S^2$ where each orbit intersects $M$ at only one point. Now, the distance between the two orbits is defined as the length of the geodesic $\alpha$ between standard elements $[p_0^{(1)}]$ and $[p_0^{(2)}]$ on $M$.

![Fig. 1. Sphere as section under scaling group.](image)

The framework used in the following cases differs slightly from this classical notion of section, though. Instead of a bijection directly from $M$ to $V/G$, we find a subgroup $G_0$ of $G$ for which there exists a bijection from $M/G_0$ to $V/G$. In other words for each case, we define an $M$ that intersects each orbit under $G$ in exactly one orbit under a subgroup $G_0$. Therefore, we study the differential geometry and perform statistical analysis on $M/G_0$ in place of $V/G$. With a slight abuse of terminology, we will still call the quotient space $M/G_0$ a section of $V$ under the action of $G$.

Case 1: Landmark-Affine Section: Consider $\mathbb{R}^{n \times 2}$, the space of all configurations of the form $X = [x_1, \ldots, x_n]^T$, where each $x_i \in \mathbb{R}^2$. The action of the affine group
\[ G_a = GL_a(2) \times \mathbb{R}^2 \] defines the orbits \([X] = \{XA + B \mid A \in GL_a(2), B = 1 \text{diag}(b)\}\), where \(1\) is the \(n \times 2\) matrix consisting of all ones, and \(\text{diag}(b)\) is the \(2 \times 2\) diagonal matrix with entries \(b_1, b_2 \in \mathbb{R}\) on the diagonal. The centroid, or center of mass, of \(X\) is defined as \(C_X = \frac{1}{n} \sum_{k=1}^{n} x_k \in \mathbb{R}^2\), and the covariance of \(X\) is defined as \(\Sigma_X = (X - 1 \text{diag}(C_X))^T (X - 1 \text{diag}(C_X)) \in \mathbb{R}^{2 \times 2}\). It is a fact that for an \(X\) of full rank, there exists a point \(X_0 \in [X]\) that satisfies the properties

1. \(C_{X_0} = 0\) and 2. \(\Sigma_{X_0} = I\).

Furthermore, any two such selections for the same configuration, say \(X_0^{(1)}\) and \(X_0^{(2)}\), are related by \(X_0^{(2)} = X_0^{(1)} O\), where \(O \in SO(2)\). This existence/uniqueness fact is easily verifiable through linear algebra techniques, and thus we omit its proof. We call this representative \(X_0\) of the affine orbit \([X]\) an affine standardization of \(X\).

Denote \(M^I_a \subset \mathbb{R}^{n \times 2}\) as the space of all affine-standardized landmarks, where

\[ M^I_a = \{X \in \mathbb{R}^{n \times 2} \mid C_X = 0, \Sigma_X = I\} \]

Note that \(SO(2)\) is a subgroup of the affine group \(G_a\), and it acts on \(M^I_a\) by isometries with respect to the standard Euclidean metric. Due to the existence/uniqueness fact stated previously, there exists a bijection from \(M^I_a/O(2)\) to \(\mathbb{R}^{n \times 2}/G_a\). Therefore, we study the geometry of and perform shape analysis on \(M^I_a/\Sigma\) instead of \(\mathbb{R}^{n \times 2}/G_a\). Denote the section \(S^I_a = M^I_a/\Sigma\) as landmark-affine shape space.

We note that, except for the condition \(C_X = 0\), \(M^I_a\) is actually the Stiefel manifold \(\mathcal{N}_{n, 2}\) and, after removal of \(SO(2)\), becomes a Grassmann manifold, where it is possible to write geodesic equation explicitly under the standard norm. A path \(\alpha : [0, 1] \rightarrow S^I_a\) is a geodesic if and only if \(\dot{\alpha}(\tau)^T \ddot{\alpha}(\tau) = \dot{\alpha}(\tau)^T \dot{\alpha}(\tau)\) and \(\dot{\alpha}(\tau)^T \dot{\alpha}(\tau) = 0\) for all \(\tau \in [0, 1]\) (for derivation, see Appendix A). Since these differential equations are difficult to solve analytically, the numerical recipe of path-straightening to find geodesics, presented in the next section, becomes an important tool. One can then check that the output of this algorithm satisfies these equations.

**Case 2: Curve-Affine Section:** Next we derive a section of the affine group for shapes represented by parameterized closed curves. Let \(B\) be the space of all absolutely continuous curves of the form \(\beta : \mathbb{S}^1 \rightarrow \mathbb{R}^2\). The action of the affine group \(G_a\) defines the orbits \([\beta] = \{A\beta + b \mid A \in GL_a(2), b \in \mathbb{R}^2\}\). Let \(L_\beta = \int_0^1 |\dot{\beta}(t)| dt\) be the length of the curve \(\beta\), where \(|\cdot|\) is Euclidean 2-norm. (Please contrast it from \(|\cdot|_2\) which is used to denote the \(L^2\)-norm of a curve or function.) The centroid of \(\beta\) is defined as \(C_\beta = \frac{1}{L_\beta} \int_0^1 \dot{\beta}(t) |\dot{\beta}(t)| dt \in \mathbb{R}^2\). The covariance of \(\beta\) is defined as \(\Sigma_\beta = \frac{1}{L_\beta} \int_0^1 (\dot{\beta}(t) - C_\beta) (\dot{\beta}(t) - C_\beta)^T |\dot{\beta}(t)| dt \in \mathbb{R}^{2 \times 2}\). It is a fact that for any non-degenerate \(\beta \in B\) there exists a standard element \(\beta_0 \in [\beta]\) that satisfies the following three conditions:

1. \(L_{\beta_0} = 1\), (2) \(C_{\beta_0} = 0\), and (3) \(\Sigma_{\beta_0} \propto I\).

Furthermore, for any two curves \(\beta_1\) and \(\beta_2\) within an affine transformation of each other, the corresponding standardized elements, \(\beta_0^{(1)}\) and \(\beta_0^{(2)}\), are related by 

\[ \beta_0^{(2)} = O(\beta_0^{(1)} \circ \gamma) \]

where \(O \in SO(2)\) and \(\gamma \in \Gamma\). Here \(\Gamma\) is the re-parameterization group, the group of all orientation-preserving homeomorphisms \(\gamma : \mathbb{S}^1 \rightarrow \mathbb{S}^1\), such that \(\gamma\) and \(\gamma^{-1}\) are absolutely continuous. We provide a mathematically rigorous proof of this existence/uniqueness theorem in Appendix D.

Fig. 2 provides an illustration of this fact. The images on the left show four rows of landmark configurations in \(\mathbb{R}^{100 \times 2}\) (top) and four rows of curves in \(B\) (bottom), where objects in each row are within the same affine orbit. The images on the right show the standardized versions of each respective configuration in the left image, i.e. with elements in \(M^I_a\) (top) and elements in \(M^I_a\) (bottom). Within each row, the configurations are all the same modulo rotation and, in the curve case, re-parameterization.

Denote \(M^I_a \subset B\) as the space of all affine standardized closed curves, where

\[ M^I_a = \{\beta \in B \mid L_\beta = 1, C_\beta = 0, \Sigma_\beta = \lambda_\beta I\}\]

where \(\lambda_\beta\) is a real scalar. Due to the existence/uniqueness fact stated previously, there exists a bijection from \(M^I_a/\Sigma \times \mathbb{S}^1\) to \(B/G_a\). However, unlike the landmark-affine case, the subgroup \(SO(2) \times \mathbb{S}^1\) does not act on \(M^I_a\) by isometries with respect to the \(L^2\) metric. We overcome this shortcoming via the following transformation given by Srivastava et al in [18]. Here, the authors define the square root velocity function \((\text{SRVF})\) of a curve \(\beta\) as \(\varphi(t) = \sqrt{\dot{\beta}(t)^2}\), and show that \(SO(2) \times \mathbb{S}^1\) acts on the space of \(\text{SRVF}'s\) of unit length closed curves by isometries with respect to the \(L^2\) metric. Denote \(\mathcal{C}\) as the space of \(\text{SRVF}'s\) of elements of \(M^I_a\). Since \(\mathcal{C}\) is contained in the space of \(\text{SRVF}'s\) of unit length closed curves, \(SO(2) \times \mathbb{S}^1\) acts by isometries on \(\mathcal{C}\). Furthermore, since there exists a
bijection between $C$ and $M_p$, we study the geometry of and perform shape analysis on $C/(SO(2) \times \Gamma)$ instead of on $M_p^i/(SO(2) \times \Gamma)$ or equivalently on $B/(G_a \times \Gamma)$. Denote the section $S^i_p = C/(SO(2) \times \Gamma)$ as curve-affine shape space. The derivation of geodesic equations for $S^i_p$ is left as future work, but the algorithms for computing geodesics are presented later in the paper.

Case 3: Landmark-Projective Section: A projective transformation is easier explained through the projective space $\mathbb{P}^2$ which is defined as follows. A point $(x, y)^T \in \mathbb{R}^2$ is identified to a point $p \in \mathbb{P}^2$ through a homogeneous vector representation, which is the following equivalence class of vectors: $p = \{(\lambda x, \lambda y, \lambda) \mid \lambda \in \mathbb{R} \setminus \{0\}\}$. This equivalence class represents a line through (but not including) the origin and the point $(x, y, 1)^T \in \mathbb{R}^3$, and any member of this equivalence class is considered a representation of the point $(x, y)^T$ in homogeneous coordinates. Therefore, any point in $\mathbb{P}^2$ is identified with a line through but not including the origin in $\mathbb{R}^3$. An arbitrary point in homogeneous coordinates $(x_1, x_2, x_3)^T$ is equivalent to $(x_1/x_3, x_2/x_3, 1)^T$ and therefore represents the point $(x_1/x_3, x_2/x_3)^T$ in $\mathbb{R}^2$. Henceforth, when visualizing a point $p \in \mathbb{P}^2$ in $\mathbb{R}^2$, we select the representative of the equivalence class $p$ that lies on the plane $z = 1$, and when visualizing a point $p \in \mathbb{R}^3$ in $\mathbb{R}^2$, we select the representative that lies on the northern hemisphere of the unit sphere.

Now consider a configuration of $n$ points $X_0 = [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^{n \times 2}$. To avoid trivial cases, assume $n \geq 5$. Let $X \in \mathbb{R}^{n \times 3}$ be the homogeneous representation of $X_0$ by appending a column of all 1s to $X_0$. The action of the projective group $G_p = GL_n(3)/(\mathbb{R}\{0\})$ defines an orbit $[X] = \{DXB \mid B \in GL_n(3), D \text{ an } n \times n \text{ diagonal matrix}\}$. The multiplication by $B$ represents equivalence under linear transformation, and the multiplication by $D$ represents the equivalence under arbitrary non-zero scaling of each homogeneous point that occurs in $\mathbb{P}^2$. Assume that $X$ has the regularity property that for any vector subspace $V \subset \mathbb{R}^{n \times 3}$ such that $1 \leq \dim(V) < 3$, the number of points of $X$ lying in $V$ is less than $(n/3)\dim(V)$. According to Kent and Mardia [11], for any such $X$ it is possible to choose $B$ and $D$ such that $X_0 \in [X]$ has the following properties:

1. $|x_{0,k}| = 1$, for $k = 1, \ldots, n$, and
2. $X^T X_0 = n^3 I_3$.

Furthermore, any two such selections, say $X_0^{(1)}$ and $X_0^{(2)}$, are related by $X_0^{(2)} = S X_0^{(1)} O$, where $O \in SO(3)$ and $S$ is an $n \times n$ diagonal matrix with entries consisting of $\pm 1$. (This matrix $S$ is identity in practice due to our representation of points on the northern hemisphere of $\mathbb{S}^2$.) We call this representative $X_0$ of the projective orbit $[X]$ a projective standardization of $X$. In [11], this $X_0$ is alternatively referred to as a “Tyler standardization.”

Fig. 3 is analogous to Fig. 2. The left image shows four rows of six landmark configurations, where each row consists of elements within a projective orbit as visualized in $\mathbb{R}^2$. The right image shows the standardized versions in $M_p^i$ of each respective configuration in the left image, as visualized on the unit sphere in $\mathbb{R}^3$. Within each of these rows, the configurations are all the same modulo rotation.

![Fig. 3. Projective standardization of landmarks.](image)

Denote $M_p^i \subset \mathbb{R}^{n \times 3}$ as the space of projective standardized landmarks, where

$$M_p^i = \{ X \in \mathbb{R}^{n \times 3} \mid X^T X = \frac{n}{3} I_3, |x_{k}| = 1 \text{ for } \forall k \}.$$  

Note that $SO(3)$ is a subgroup of the projective group $G_p$, and it acts on $M_p^i$ by isometries with respect to the standard Euclidean metric. Similar to the previous cases, due to the existence/uniqueness fact stated above, there exists a bijection from $M_p^i/\text{SO}(3)$ to $\mathbb{R}^{n \times 3}/G_p$. Therefore, we study the geometry of and perform shape analysis on $M_p^i/\text{SO}(3)$ instead of on $\mathbb{R}^{n \times 3}/G_p$. Denote the section $S_p^i = M_p^i/\text{SO}(3)$ as landmark-projective shape space.

As in the landmark-affine case, we can derive the geodesic equations for $S_p^i$ since the manifold $M_p^i$ is a variation of the Stiefel manifold $\mathcal{S}_{n,3}$. A path $\alpha(\tau) \in S_p^i$ is a geodesic if and only if it satisfies the following:

For all $\tau$, (1) the diagonal elements of $\dot{\alpha}(\tau)\alpha(\tau)^T$ are all 0, (2) $\dot{\alpha}(\tau)^T\alpha(\tau) = 0$, and (3) $\ddot{\alpha}(\tau) = -(3/n)\alpha(\tau)\dot{\alpha}(\tau)^T\dot{\alpha}(\tau)$. Since these differential equations are difficult to solve, path-straightening becomes a valuable tool.

3 GEODESICS ON SECTIONS

In this section we outline the path-straightening method for geodesic calculation on a Riemannian manifold $M$ embedded in an ambient Hilbert space $V$ and with metric inherited from $V$. The original ideas are presented in [12], [18] (and also given in Appendix B), but are presented here briefly for convenience. The main idea is to initialize a path $\alpha$ between two given points on $M$ and iteratively straighten it according to the gradient of energy $\dot{E}$ until it cannot be straightened any further. The energy of a path $\alpha : [0, 1] \rightarrow M$ is given by $E = \int_0^1 \left(\dot{\alpha}(\tau), \dot{\alpha}(\tau)\right) d\tau$, where inner-product in the integrand is that of $V$. The resulting path $\alpha$ can be shown to be a geodesic. The energy gradient can be written analytically, and thus the computational cost is relatively low. For its implementation one needs the following three subroutines:

1. **Subroutine 1**: Projection of an arbitrary point in $V$ to a point $p \in M$, 

![image1](image1)

![image2](image2)
2) **Subroutine 2**: Projection of an arbitrary vector \( v \in V \) onto the tangent space \( T_p(M) \),

3) **Subroutine 3**: Parallel translation of a vector \( v \) from \( T_{p_1}(M) \) to \( T_{p_2}(M) \), where \( p_1 \) and \( p_2 \) are close enough to use a straight line in \( V \) as a geodesic.

These subroutines, in turn, require an orthonormal basis for \( N_p(M) \), the normal space of \( M \) at any point \( p \). To implement a path-straightening approach on a computer, we discretize the path \( \alpha \) on a uniform partition \( \{ \tau_i = i/T, i = 0, 1, \ldots, T \} \) of the interval \([0,1]\).

The first procedure is to initialize a path in \( M \) between two given points \( p_1 \) and \( p_2 \).

**Algorithm 1 – Path Initialization**: Given \( p_1, p_2 \in M \), and for \( i = 0, 1, \ldots, T \)

1) Compute straight line geodesic \( \alpha \) in \( V \) between \( p_1 \) and \( p_2 \) via \( \alpha(\tau_i) = (1 - \tau_i)p_1 + (\tau_i)p_2 \).

2) Project \( \alpha(\tau_i) \) to \( M \) using Subroutine 1.

The next procedure is to compute the velocity vector \( \frac{d\alpha}{d\tau} \) using finite differences. For numerical stability, we suggest using a forward difference for \( i = 0 \), a backward difference for \( i = T \), and a centered difference for all other intermediary values of \( i \).

**Algorithm 2 – Compute \( \frac{d\alpha}{d\tau} \) along \( \alpha \)**: For all \( i = 0, 1, \ldots, T \)

1) If \( i = 0 \), compute \( c(\tau_i) = (\alpha(\tau_{i+1}) - \alpha(\tau_i))/\Delta \). If \( i = T \) compute \( c(\tau_i) = (\alpha(\tau_1) - \alpha(\tau_{i-1}))/\Delta \). Else compute \( c(\tau_i) = (\alpha(\tau_{i+1}) - \alpha(\tau_{i-1}))/2\Delta, \Delta = 1/T \).

2) Project \( c(\tau_i) \) onto \( T_{\alpha(\tau_i)}(M) \) using Subroutine 2 to get an approximation of \( \frac{d\alpha}{d\tau}(\tau_i) \).

Next we present the computation of the covariant integral of \( \frac{d\alpha}{d\tau} \) along \( \alpha \) using partial sums. As we move from \( \alpha(0) \) to \( \alpha(1) \), we wish to add the current sum, say \( u(\tau_{i-1}) \), to the velocity \( \frac{d\alpha}{d\tau}(\tau_i) \) to estimate \( u(\tau_i) \).

However, these two quantities are elements of two different spaces and cannot be added together directly; therefore, we first parallel translate \( u(\tau_{i-1}) \) to the point \( \alpha(\tau_i) \), then add it to \( \frac{d\alpha}{d\tau}(\tau_i) \) to estimate \( u(\tau_i) \).

**Algorithm 3 – Compute Covariant Integral of \( \frac{d\alpha}{d\tau} \) along \( \alpha \)**: Set \( u(0) = 0 \in T_{\alpha(0)}(M) \). For all \( i = 1, 2, \ldots, T \)

1) Using Subroutine 3, compute \( u^\tau(\tau_{i-1}) \) as the parallel translation of \( u(\tau_{i-1}) \) in \( T_{\alpha(\tau_{i-1})}(M) \) to \( T_{\alpha(\tau_i)}(M) \).

2) Set \( u(\tau_i) = \Delta \frac{d\alpha}{d\tau}(\tau_i) + u^\tau(\tau_{i-1}) \).

This covariant integration results in a vector field \( u \) that is the unconstrained gradient of \( E \). To obtain the gradient in the space restricted to all paths that begin and end at the correct endpoints \( p_1 \) and \( p_2 \), we need to subtract off the appropriate covariantly-linear component as follows. First we compute an estimate of the backwards parallel translation of \( u(1) \).

**Algorithm 4 – Backward Parallel Translation of \( u(1) \)**: Set \( \tilde{u}(1) = u(1) \). For all \( i = T - 1, T - 2, \ldots, 0 \)

1) Parallel translate \( \tilde{u}(\tau_{i+1}) \) onto \( T_{\alpha(\tau_i)}(M) \) using Subroutine 3 to obtain \( \tilde{u}(\tau_i) \).

Now we can compute the desired gradient.

**Algorithm 5 – Gradient vector field of \( E \)**: For all \( i = 1, 2, \ldots, T \), compute: \( w(\tau_i) = u(\tau_i) - \tau_i \tilde{u}(\tau_i) \).

By construction, this vector field \( w \) is zero at \( i = 0 \) and \( i = T \). As a final step, we update the path \( \alpha \) in the direction of the negative gradient of \( E \).

**Algorithm 6 – Path Update**: Select a small \( \epsilon > 0 \) as the update step size. For all \( i = 0, 1, \ldots, T \), perform

1) Compute the gradient update \( \tilde{\alpha}(\tau_i) = \alpha(\tau_i) - \epsilon w(\tau_i) \).

2) Using Subroutine 1, project \( \tilde{\alpha}(\tau_i) \) to \( M \) to obtain \( \alpha(\tau_i) \).

This completes a numerical recipe for computing geodesics on \( M \subset V \) via path-straightening.

4) **Case 1: Landmark-Affine Shapes**

We start with the details of the three subroutines necessary to perform path-straightening for geodesic computation on the submanifold \( S^6_{n,2} \subset R^{n \times 2} \). Subroutine 1 is straightforward, and Fig. 2 in Section 2.2 shows its results on four different affine orbits.

**Subroutine 1 – Landmark-Affine Standardization**: Given \( X \in R^{n \times 2} \),

1) Let \( A_0 = (\Sigma_X)^{-1/2} \) and \( B_0 = I \text{diag}(C_X) \).

2) Compute \( X_0 = (X - B_0)A_0 \).

The next two subroutines require a basis for the normal space \( N_X(M^1_a) \), and for that we must first understand the geometry of \( M^1_a \). Notice that \( M^1_a \) is the intersection of two sub-manifolds of \( R^{n \times 2} \): \( D \), the set of \( n \) points in \( R^2 \) with centroid at the origin, and \( S_{n,2} \), the Stiefel manifold consisting of \( n \times 2 \) orthonormal frames. For any point in \( D \), the normal direction is defined by any translation in \( R^2 \), and thus \( N_X(D) \) is the 2-dimensional space spanned by the basis \( B_1 = \{ I(\text{diag}(e_1)) \} \), where \( e_1 \) is the canonical basis of \( R^2 \). Furthermore, given \( X \in S_{n,2} \), the normal space is defined as \( N_X(S_{n,2}) = \{ XS \mid S \in R^{2 \times 2} \text{ symmetric} \} \). This means that \( N_X(S_{n,2}) \) is a three-dimensional space spanned by the basis \( B_2 = XB \), where \( B = \{ B_1, B_2, B_3 \} \) is the standard basis of the space of \( 2 \times 2 \) symmetric matrices. Thus the space spanned by the combined five basis vectors \( \{ B_1 \cup B_2 \} \) defines \( N_X(M^1_a) \). It can be shown that all 5 basis vectors are linearly independent, and hence \( N_X(M^1_a) \) is a 5-dimensional space. We can now write down the remaining two subroutines.

**Subroutine 2 – Projection onto \( T_X(M^1_a) \)**: Given any vector \( v \in R^{n \times 2} \) and \( X \in M^1_a \)

1) Compute \( B_1 \) a basis for \( N_X(D) \), and compute \( B_2 \) a basis for \( N_X(S_{n,2}) \).

2) Compute \( \{ b_i, i = 1, \ldots, 5 \} \) a Gram-Schmidt orthonormalization of \( \{ B_1 \cup B_2 \} \).

3) Let \( \bar{v} = v - \sum_{i=1}^{5}(v, b_i)b_i \).

**Subroutine 3 – Parallel Translation**: Given \( X_1, X_2 \in M^1_a \) and \( v \in T_{X_1}(M^1_a) \)

1) Compute \( l = |v| \).
2) Project \( v \) to \( TX_2(M^l(t)) \) via Subroutine 2 to obtain \( \hat{v} \).

3) Rescale to obtain \( \hat{v} = \hat{v}l/|\hat{v}| \).

We now have all the tools necessary to compute geodesics via path-straightening on \( S^l_a \), and the procedure is as follows. Given \( X^{(1)}, X^{(2)} \in \mathbb{R}^{n \times 2} \), we first use Subroutine 1 to obtain the standardizations \( X^{(1)}_a, X^{(2)}_a \in M^l_a \) and hence the unique orbits \( \hat{X}^{(1)}_0 \) and \( \hat{X}^{(2)}_0 \) under the group action of \( SO(2) \). Since \( SO(2) \) acts by isometries, we define the distance on \( S^l_a \) as

\[
d_{S^l_a}([X^{(1)}_0], [X^{(2)}_0]) = \min_{O \in SO(2)} d_{M^l}([X^{(1)}_0], X^{(2)}_0 O). \tag{1}
\]

The optimal group element that achieves this distance minimization, say \( \hat{O} \in SO(2) \), is solved by Procrustes rigid body alignment, which is a standard procedure. The distance \( d_{M^l}([X^{(1)}_0], X^{(2)}_0 \hat{O}) \) is then computed with path-straightening. Our computational time for geodesic calculation on \( S^l_a \) with a 2.8 GHz processor using Matlab is on average about 0.8 seconds. It can be verified that the results do satisfy the geodesic equations for \( S^l_a \) presented in Section 2.2.

In Figs. 4 and 5 we provide visualizations of geodesics in \( S^l_a \). Fig. 4 shows an example of path-straightening iterations that converge to a geodesic path in \( S^l_a \) between two slightly different “T” shapes, say \( X^{(1)}_0 \) and \( X^{(2)}_0 \). Recall that the output of each iteration in the path-straightening procedure is a path of shapes \( \alpha(\tau) \) in the shape space, where the endpoints \( \alpha(0) = X^{(1)}_0 \) and \( \alpha(1) = X^{(2)}_0 \) are fixed for each iteration. Each row of Fig. 4 represents an iteration of path-straightening, where the top row is a randomly initialized path, and the bottom row is the final converged path, a geodesic between \( X^{(1)}_0 \) and \( X^{(2)}_0 \). In this particular implementation of path-straightening, we use the partition \( \{\tau_i = i/6, i = 0, 1, \ldots, 6\} \), and hence each row shows a discretized path consisting of seven shapes. Note that in order to save space, we do not show all iterations of the procedure; we show the first four iterations followed by the final iteration. The plot of the energy evolution is shown on the right, confirming that the algorithm converges to a local minimum, which corresponds to a geodesic.

Fig. 4. Path-straightening on \( S^l_a \): Iterations of the path-straightening algorithm (rows) and energy evolution.

Fig. 5 shows four pairs of shape paths, with each pair separated by a horizontal line. Each pair of paths is created by the following procedure. Select two random shapes \( X^{(1)} \) and \( X^{(2)} \). Standardize each shape to obtain \( X^{(1)}_0 \) and \( X^{(2)}_0 \) in \( S^l_a \), saving the matrices \( A^{(1)}_0, A^{(2)}_0 \in GL_+(2) \) that achieve each respective standardization. Form the top path by obtaining a geodesic \( \alpha(\tau) \) via path-straightening between \( X^{(1)}_0 \) and a rotationally aligned \( X^{(2)}_0 \hat{O} \). Form the bottom path by de-standardizing the geodesic to create a path \( \alpha_d(\tau) \) in the ambient space between original shapes \( X^{(1)} \) and \( X^{(2)} \).

The de-standardized path is created in the following manner. First, form a geodesic path \( MA(t) \) in \( GL_+(2) \) between \( (A^{(1)}_0)^{-1} \) and \( (A^{(2)}_0)^{-1} \). Here, \( MA(t) = M_1 \exp(\tau \log((M_1^{-1}M_2))) \), where \( M_1 = (A^{(1)}_0)^{-1} \) and \( M_2 = (A^{(2)}_0)^{-1} \). Then, form a geodesic path \( ML_d(t) \) in \( SO(2) \) by using the previous formula with \( M_1 = I \) and \( M_2 = \hat{O}^{-1} \). Finally, form the de-standardized path \( \alpha_d(\tau) = \alpha(\tau)M\hat{O}(\tau)M_A(\tau) \), which is such that \( \alpha_d(0) = X^{(1)} \) and \( \alpha_d(1) = X^{(2)} \). Note that we use the de-standardized path only for visualization and not for any statistical analysis.

Fig. 5. Geodesics in \( S^l_a \) and their de-standardizations.

Fig. 6 shows a clear contrast between the statistical models created using similarity and affine invariances. Here, we estimate a Gaussian model from 20 training shapes from the MPEG-7 shape class “crown” under random affine transformations. The top left-most box in Fig. 6 shows the training data, the top center shows the Karcher mean of the same shapes. The bottom row shows the improved shape model using \( S^l_a \) for the same training data. The bottom left-most box shows the respective standardizations in \( M^l_a \) of the same training shapes, the bottom middle shows the Karcher mean in \( S^l_a \), and the bottom right box first shows random samples from the estimated Gaussian distribution on \( S^l_a \) and then the respective de-standardizations of these samples. (Note that the term Karcher mean is reserved for an intrinsic mean that satisfies a local minimum of a cost function. The global minimizer is also called a Fréchet or Riemannian mean.) We perform each of these de-standardizations by sampling from an estimated Gaussian model of the transformations in \( GL_+(2) \) that are used to standardize the original data, and then we apply the inverse of that matrix sample. Algorithms to compute the mean, covariance, and random samples

\[
\begin{align*}
\end{align*}
\]
on $S_0^t$ all make use of the three subroutines and basis for the normal space described in this section. It is important to note that by separating the two sources of variability – shape variability in $S_0^t$ and affine variability in $GL_+(2)$ – we capture the true variability of the training shapes much more accurately than by modeling in just similarity shape space alone.

The space $P(2)$ is a three-dimensional vector space with standard basis $\mathcal{B} = \{B_1, B_2, B_3\}$, and therefore any matrix in $P(2)$ can be expressed as an element of $\mathbb{R}^3$ consisting of coefficients from a linear combination of these basis vectors. Let $dF_I = J\beta$, where $J$ is the $3 \times 3$ Jacobian matrix of coefficients. Fig. 2 shows the result of applying this subroutine to four different affine orbits.

**Subroutine 1a – Curve-Affine Standardization:** Given $\beta$ with $L_\beta = 1$ and $C_\beta = 0$, initialize $P_0 = P^{(0)} = I$, $\beta^{(0)} = \beta$ and select a step size $\delta$. Let $m = 0$.

1) Calculate the residual $r^{(m)} = F(I; \beta^{(m)}) - I$. Rewire $r^{(m)}$ as a vector of coefficients with respect to the basis vectors $B_1, B_2, B_3$. If $|r^{(m)}| < \epsilon$ go to (7), else

2) Calculate $dF_I(B; \beta^{(m)})$ for $i = 1, 2, 3$ using Eqn. 3, and form the Jacobian $J$.

3) Compute $x = J^{-1}r^{(m)}$ and let $D = \sum_{i=1}^3 x_i B_i$.

4) Let $P_{(m+1)} = I - \delta D$.

5) Update $\beta^{(m+1)} = P_{(m+1)}\beta^{(m)} - C_{P_{(m+1)}}\beta^{(m)}$ and $p^{(m+1)} = P_{(m+1)}p^{(m)}$.

6) Return to (1) and let $m = m + 1$.

7) Let $P_0 = P^{(m+1)}/L_{\beta^{(m+1)}}$, and let $\beta_0 = \beta^{(m+1)}/L_{\beta^{(m+1)}}$.

Now we proceed to study the geometry of the manifold $C$ as defined in Section 2.2. This task is an extension of the work presented in [18] on elastic shape analysis of closed curves. For a parameterized curve $\beta : [0, 1] \rightarrow \mathbb{R}^2$, its SRVF is given by the function $q(t)$ as defined previously. Note that since $q(t)|q(t)| = \beta(t)$, the curve $\beta(t)$ can be recovered from the $q$-function up to a translation. Let $x(q(t) = \int_0^t q(u)|q(u)|du$ be our original curve $\beta(t)$ but with seed located at the origin, i.e. $\beta(0) = 0$. Also note that since $|q(t)|^2 = |\beta(t)|$, the set of all unit length curves $\{q \in L^2([0, 1], \mathbb{R}^2)|\int_0^1 |q(t)|^2dt = 1\}$ is the unit hypersphere $S^\infty$ in $L^2$ space. The centroid and covariance of a curve $x(q, t)$ can be stated in terms of the $q$-function as follows: The centroid in $\mathbb{R}^2$ is given by $C_q = \int_0^1 x(q(t))|q(t)|^2dt$, and the covariance in $\mathbb{R}^2 \times \mathbb{R}^2$ is given by $\Sigma_q = \int_0^1 (a + x(q(t)))(a + x(q(t)))|q(t)|^2dt$, where $a = -C_q$. In order to impose the condition that the curve $\beta$ be closed, we set $x(q; t) = 0$, i.e. the endpoint of the curve $x(q; t)$ is equal to the initial point, the origin. Define a mapping $\Psi : S^\infty \rightarrow \mathbb{R}^4$ as:

$$\Psi_1(q) = \int_0^1 ((a_1 + x_1(q(t)))^2 - (a_2 + x_2(q(t)))^2)|q(t)|^2dt$$

$$\Psi_2(q) = \int_0^1 ((a_1 + x_1(q(t)))(a_2 + x_2(q(t)))|q(t)|^2dt$$

$$\Psi_3(q) = \int_0^1 q_1(t)|q(t)|dt, \quad \Psi_4(q) = \int_0^1 q_2(t)|q(t)|dt,$$

where a subscript indicates the $i$th coordinate in Euclidean space, that is, $\Psi_i(q) \in \mathbb{R}$ for each $i$. $\Psi_i(q) = 0$ implies that the difference of diagonal entries in the covariance matrix is 0. $\Psi_2(q) = 0$ implies that the off-diagonal entry in the covariance matrix is 0. Together, they imply that $\Sigma_\beta = \lambda_3 I$. The constraint $\Psi_3(q) = \Psi_4(q) = 0$ implies the closure condition. Since SRVF’s are translation invariant, we don’t need any explicit condition on the centroid. The space of all affine-standardized, unit length, closed curves is...
therefore the level set \( C = \Psi^{-1}((0, 0, 0, 0)) \subset \mathbb{S}^\infty \).

**Defining the Normal Space of \( C \):** An integral step in each of the three subroutines necessary for path-straightening on \( C \) is obtaining a basis for \( N_q(C) \). The normal space is a 4-dimensional space, and here we show the analytical formulas for the functions \( \{ h^j(t), j = 1, \ldots, 4 \} \) that serve as a basis for this space. These functions arise from the calculation of the directional derivative of \( \Psi \) in Eqn. 3, and we provide the full derivation in Appendix C. Let \( f^i(t) = |q(t)| e_1 + \frac{q(t)}{|q(t)|} q(t) \) for \( i = 1, 2 \). Let \( Q(t) = \int_0^t |q(\tau)|^2 d\tau \). Then \( G_i(t) = \int_0^t |q(\tau)|^2 x_i(q; \tau) d\tau \) for \( i = 1, 2 \). Now

\[
\begin{align*}
    h^1(t) &= 4q(t)a_1 x_1(q; t) - 4q(t)a_2 x_2(q; t) + 2a_1 f^1(t)(1 - Q(t)) - 2a_2 f^2(t)(1 - Q(t)) + 2q(t)x_1(q; t^2) - 2q(t)x_2(q; t^2) + 2f^1(t)(a_1 - G_1(t)) - 2f^2(t)(a_2 - G_2(t)), \\
    h^2(t) &= 2a_2 f^1(t)x_1(q; t) + 2a_1 f^2(t)x_2(q; t) + a_2 f^1(t)(1 - Q(t)) + a_1 f^2(t)(1 - Q(t)) + 2q(t)x_1(q; t)x_2(q; t) + f^1(t)(a_2 - G_2(t)) + f^2(t)(a_1 - G_1(t)), \\
    h^3(t) &= f^1(t), \quad \text{and} \quad h^4(t) = f^2(t).
\end{align*}
\]

The restriction of these functions inside \( T_q(\mathbb{S}^\infty) \) is obtained by removing the projection along the function \( q: b^i(t) = h^i(t) - q(t)(q(t), h^i(t)) \) and \( \{ b^i \} \) span the normal space \( N_q(C) \) inside \( T_q(\mathbb{S}^\infty) \).

Next we describe the three subroutines needed to implement the path-straightening algorithm on \( C \). For any point \( q \in \mathbb{S}^\infty \), we need a tool to project \( q \) to the nearest point in \( C \). The procedure presented below is different from Subroutine 1a in that, although they accomplish a similar task, standardization is restricted to the same orbit while the projection is not. To clarify further, Subroutine 1a is used on coordinate functions for curve standardization, and Subroutine 1b is used on SRVF’s within the path-straightening algorithm. Subroutines 2 and 3 are straightforward and similar to the landmark case.

**Subroutine 1b – Projection onto \( C \):** Let \( \epsilon > 0 \).

1. **Compute the residual vector** \( r = \Psi(q) \). If \( |r| < \epsilon \), stop. Otherwise continue to step 2.
2. **Calculate the basis vectors** \( \{ b^j, j = 1, \ldots, 4 \} \), for \( N_q(C) \), and form the Jacobian matrix \( J \) via \( J_{ij} = \langle b^i, b^j \rangle \).
3. Solve \( J x = -r \) for \( x \).
4. Define \( dq = \sum_{j=1}^4 |x_j| b^j \). Update \( q \rightarrow \cos(||dq||)q + \sin(||dq||) \frac{dq}{||dq||} \). Go to step 1.

**Subroutine 2 – Projection onto \( T_q(C) \):** Given any function \( w \in L^2([0, 1], \mathbb{R}^2) \),

1. If \( w \notin T_q(\mathbb{S}^\infty) \), then project \( w \) to \( T_q(\mathbb{S}^\infty) \) via \( w \rightarrow w - \langle w, q \rangle q \).
2. Compute a basis \( \{ b^j, j = 1, 2, 3, 4 \} \), for \( N_q(C) \), and then obtain a Gram-Schmidt orthonormalization, \( \{ b^j_0 \} \).
3. Project \( w \) into \( T_q(C) \) via \( w \rightarrow w - \sum_{j=1}^4 \langle w, b^j_0 \rangle b^j_0 \).

Note that we can skip the first step even if \( w \notin T_q(\mathbb{S}^\infty) \), but it will compromise numerical stability.

**Subroutine 3 – Parallel Translation:** Given two points \( q_1, q_2 \in C \) and \( w \in T_q(C) \).

1. Compute the analytic expression for parallel translation on \( \mathbb{S}^\infty \) as \( w \rightarrow \bar{w} = \frac{q_1 + q_2}{|q_1 + q_2|} (q_1 + q_2) \).
2. Let \( l = |\bar{w}| \). Project \( w \) onto \( T_{q_2}(C) \) using Subroutine 2 to obtain \( \bar{w} \).
3. Rescale \( w \) via \( \bar{w} \rightarrow \bar{w} / |\bar{w}| \).

We now have all the tools necessary to compute geodesics via path-straightening on \( \mathcal{C}_0 \) and the general procedure of affine-invariant, elastic shape analysis is as follows. We are given two curves \( \beta^{(1)} \) and \( \beta^{(2)} \). First, use Subroutine 1a to obtain \( \beta^{(1)}_0 \) and \( \beta^{(2)}_0 \). Then convert to SRVF representation to obtain \( q^{(1)}, q^{(2)} \in C \) and hence the unique orbits \( [q^{(1)}] \) and \( [q^{(2)}] \) under the group action of \( SO(2) \times \Gamma \). Since this group action is by isometries on \( C \), we define the distance on \( \mathcal{C}_0 \) as

\[
d_{\mathcal{C}_0}([q^{(1)}], [q^{(2)}]) = \min_{O \in SO(2) \times \Gamma} d_C(q^{(1)}, O(q^{(2)} \ast \gamma)),
\]

where the group action by \( \Gamma \) is given by \( q \ast \gamma = (q \circ \gamma) \sqrt{7} \). The optimal group elements that achieve this distance minimization, say \( \bar{O} \) and \( \bar{\gamma} \), are solved respectively by Procrustes rigid body alignment and by either dynamic programming or gradient based optimization (see [18]). The distance \( d_C(q^{(1)}, O(q^{(2)} \ast \gamma)) \) is thus computed with path-straightening. Our computational time for geodesic calculation on \( \mathcal{C}_0 \) with a 2.8 GHz processor using Matlab is on average about 1.25 seconds.

Various examples of geodesic paths in \( \mathcal{C}_0 \) can be seen in Figs. 7, 8, and 9. As in Fig. 4, Fig. 7 shows an example of path-straightening iterations that lead to a geodesic path in \( \mathcal{C}_0 \). Likewise, analogous to Fig. 5, Fig. 8 shows four examples of geodesic paths in continuous affine shape space along with their corresponding de-standardized paths. The de-standardized paths are created with the same procedure as described earlier in the landmark-affine case.
two shapes in various shape spaces, whereby the two shapes differ in placement of bumps and an affine transformation. We compute the geodesics in the following shape spaces: (a) closed curve shape space with a non-elastic, bending only metric [13], (b) closed curve, similarity-invariant shape space with the elastic metric, denoted \( S^e_c \), (c) \( S^a_c \), and for display purposes (d) the de-standardization of (c). We can see that the deformation in geodesic path (b) is more natural and smaller than that of path (a) due to the addition of elasticity to appropriately stretch and match features. Furthermore, path (c) shows an even smaller deformation than that of path (b) due to the addition of affine invariance. Note that any affine transformation of either the beginning or ending shape will not change the geodesic path shown in (c) due to our affine invariant shape analysis framework.

**Fig. 9.** Geodesic paths between two shapes in different spaces. (a) Similarity invariant with bending-only metric, \( S^e_c \); (b) \( S^a_c \); (c) \( S^a_c \); (d) De-standardized version of path (c).

Analogous to Fig. 6, Fig. 10 demonstrates the advantage of using statistical models in \( S^a_c \) over the similarity shape space \( S^e_c \). Here, we use the MPEG-7 shape class “hat” under random affine transformations. The curve-similarity model (top row) in this case captures the underlying shape variability much better than the landmark-similarity case in Fig. 6 due to the addition of elastic curve matching. However, in the curve-affine model (bottom row) we see that further improvements to modeling these affine shapes are gained by separating the variability into the spaces \( S^a_c \) and \( GL_+(2) \), as we showed in the landmark-affine case.

**Fig. 10.** Statistics on \( S^e_c \) (top row) and \( S^a_c \) (bottom row).

### 6 CASE 3: LANDMARK-PROJECTIVE SHAPES

We start with Subroutine 1 that finds the standardization \( X_0 \in [X] \) for an arbitrary \( X \):

**Subroutine 1 – Landmark-Projective Standardization:**

Given \( X \in \mathbb{R}^{n \times 3} \),

1. Normalize each row of \( X \) via \( x_k \mapsto x_k/|x_k| \) for all \( k \).
2. Set \( A^{(0)} = I \). Iterate until convergence

\[
A^{(m+1)} = \sum_{k=1}^{n} \frac{x_k^T x_k}{x_k(A^{(m)})^{-1} x_k^T} / \sum_{k=1}^{n} \frac{1}{x_k(A^{(m)})^{-1} x_k^T}
\]

3. Define \( d_k = 1/\sqrt{x_k A^{-1} x_k^T} \) and let \( D_0 = \text{diag}(\{d_k\}) \).
4. Define \( B_0 \) to be a square root of \( A^{-1} \) such that \( B_0 B_0^T = A^{-1} \). Let \( X_0 = D_0 X B_0 \).

Fig. 3 shows examples of applying Subroutine 1 to shapes taken from four projective orbits.

Next, we need a basis for the normal space \( N_X(M^p_u) \), and in order to find such a basis, we must describe the geometry of \( M^p_u \). Note that \( M^p_u \) is in fact the intersection of two well known manifolds: \( (\mathbb{S}^2)^n \), the product of \( n \) unit spheres in \( \mathbb{R}^3 \), and \( \mathcal{J}_{n,3} \), the Stiefel manifold consisting of \( n \times 3 \) orthonormal matrices (scaled by \( \sqrt{n/3} \)). Given \( x \in \mathbb{S}^2 \), the normal space of \( \mathbb{S}^2 \) at \( x \) is the ray defined by \( x \) itself. Therefore, it is easy to see that a basis for the \( n \)-dimensional space \( N_X(M^p_u) \) is given by \( B_1 = \{e_k e_k^T X, k = 1, ..., n\} \), where \( \{e_k\} \) is the canonical basis for \( \mathbb{R}^n \). Furthermore, given \( X \in \mathcal{J}_{n,3} \), the normal space is defined as \( N_X(\mathcal{J}_{n,3}) = \{X S | S \text{ is } 3 \times 3 \text{ symmetric}\} \). This means that \( N_X(\mathcal{J}_{n,3}) \) is a 6-dimensional space spanned by the basis \( B_2 = \{Y(e_i e_i^T + e_j e_j^T) | 1 \leq i \leq j \leq 3\} \). Thus the space spanned by the combined \( n + 6 \) basis vectors \( \{B_1 \cup B_2\} \) defines \( N_X(M^p_u) \). It can be shown that these two normal spaces share one common direction, that is, one vector in \( B_2 \) is a linear combination of the vectors in \( B_1 \), and hence the dimension of \( N_X(M^p_u) \) is \( n + 5 \). We can now write down the remaining two subroutines.
Subroutine 2 – Projection onto $T_X(M^l_p)$: Given any vector $v \in \mathbb{R}^{n \times 3}$ and $X \in M^l_p$.

1) Compute $B_1$ a basis for $N_X([S^3]^n)$, and compute $B_2$ a basis for $N_X(S_{n,3})$.
2) Compute $\{b_i, i = 1, ..., n+5\}$ a Gram-Schmidt orthonormalization of $\{B_1 \cup B_2\}$.
3) Let $\tilde{v} = v - \sum_{i=1}^{n+5} \langle v, b_i \rangle b_i$.

Subroutine 3 – Parallel Translation: Given $X_1, X_2 \in M^l_p$ and $v \in T_{X_1}(M^l_p)$,

1) Compute $l = |v|$.
2) Project $v$ to $T_{X_2}(M^l_p)$ via Subroutine 2 to obtain $\tilde{v}$.
3) Rescale to obtain $\bar{v} = \tilde{v}/|\tilde{v}|$.

7 Experiments and Results

In this section, we describe some experimental results obtained using the algorithmic tools derived thus far.

Retrieval of Affine Curves: The MPEG-7 shape database [3] is commonly used to test the performance of shape descriptors. We select a subset by taking 10 different shapes and applying 9 random affine transformations to each shape. We then re-parameterize each curve to have uniform-speed parameterization. This gives us 10 samples from 10 affine orbits of continuous closed curves, totaling 100 curves. Then, we compute a $100 \times 100$ distance matrix for each of the metrics in these five shape analysis frameworks: (1) $S^l_s$, (2) $S^c_s$, (3) $S^l_a$, (4) $S^c_a$, and (5) $S^l_p$. To quantify the performance of each metric, we compute a precision-recall curve to visualize classification rate with respect to different sized retrieval sets (see [4] for a detailed explanation of a precision-recall curve). Results are shown in Fig. 14 for two of the five distance matrices as well as all five precision-recall curves. The knee-point values of the P-R curves are as follows: (1) 0.337, (2) 0.513, (3) 0.721, (4) 1.000, and (5) 0.772. Therefore, affine invariant, elastic shape analysis method classifies this dataset perfectly. Even though the projective group is larger than the affine group, the landmark projective approach does not handle re-parameterization of curves and consequently does not perform as well in this case.
Database (MCD): The MCD [23] has been constructed from the MPEG-7 database. Here, 40 shapes were selected from the MPEG-7 database and printed on white paper as binary images. Seven variations of each curve were recorded by photographing the binary images under seven different camera angles and extracting the resultant boundary curves. Each shape in the database is re-parameterized to have uniform speed and resampled to have 100 sample points. The MCD is similar to the dataset we used in the previous experiment on retrieval of affine curves, but it differs in one important aspect. Previously we used data consisting of MPEG-7 shapes to which we applied random affine transformations; whereas, the MCD consists of MPEG-7 shapes that have been extracted from real images that are affected by natural perspective skew.

Since affine or projective standardizations only approximately model the true perspective skew resulting from real imaging scenarios, the performance here will not be perfect. We use a statistical modeling of shapes in each class to perform classification and cross-validation as follows. First, we select a shape space in which to work. Then, we designate 5 shapes at random from each class of 7 as training, designate the remaining 2 shapes as test shapes, and then we build “Gaussian” probability distribution functions (using means and covariances) for each of the 40 shape classes using the training shapes. For each of the 80 test shapes we compute the likelihood that it belongs to each of the 40 distributions, and we assign to it the label of the class that yields the highest likelihood. Since we know the ground truth class labels of each test shape, we compute the percentage of correctly labeled test shapes. We carry out this entire process a total of three times and average the results as per a cross-validation experimental design.

Classification results are as follows: (1) 43.33%, (2) 90.00%, (3) 95.42%, (4) 98.75%, (5) 95.00%. Clearly, the shape models formed on landmark-similarity space are insufficient for classifying continuous curves under natural perspective distortions. Since the MCD consists of uniformly parameterized curves, elasticity improves the classification results tremendously in a similarity-invariant setting. In each of the affine and projective shape spaces, classification rates are excellent, as most of the within-class shape variability is removed after standardization. The curve-affine case performs the best at 98.75% due to its invariance to parameterization.

Pose-Invariant Activity Classification: An important application of affine invariant shape analysis is in the field of human activity, or human motion analysis where a major need here is to be invariant under differing pose or camera angles. In the case of moderate pose changes (< 45°) without occlusion, the proposed affine-invariant algorithms can be used for a pose-invariant activity classification. In this experiment we use the UMD activity dataset [20], which consists of 100 sequences of 80 shapes each, where each sequence represents a frame-by-frame outline of a person performing a task. The dataset is divided into 10 classes, or “activities,” of 10 sequences each. Our goal is to classify a test sequence under an arbitrary viewing angle. To simulate a different viewing angle on a shape sequence, we compute a stochastic process on $GL(2)$ with an appropriate mean and apply each point of the process to the corresponding shape in the sequence. Since these data were captured using broadside imaging, we simulate test sequences for different views: original, narrowside, top view, and top-left (see Fig. 15 for an example). Then, we classify this test sequence using the nearest neighbor classifier under different metrics. (The distance for classifying a sequence is the sum of the distances for individual shapes.)

<table>
<thead>
<tr>
<th>Shape Space</th>
<th>Broadside Level</th>
<th>Narrowside Level</th>
<th>Broadside Elevated</th>
<th>Narrowside Elevated</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{aff}$</td>
<td>98%</td>
<td>97%</td>
<td>51%</td>
<td>87%</td>
</tr>
<tr>
<td>$S_{proj}$</td>
<td>98%</td>
<td>98%</td>
<td>98%</td>
<td>98%</td>
</tr>
</tbody>
</table>

TABLE 2
LOO classification rates for human activity sequence dataset under various simulated camera angles.

Note that the classification rate of the affine standardized test sequences matches that of the original, un-transformed sequences. Since classification rate decreases under different camera angles without standardization, we conclude that an affine-invariant metric helps in classifying sequences of human activity under various camera angles. From the drastic improvement in classification rate from 51% to 98% in the top camera angle, we can see that standardization would be especially useful when applied to sequences from mounted surveillance cameras.
Retrieval of Projective Landmarks: We derive a new database of shapes similar to the affine curve database by taking 10 shapes from the MPEG-7 shape database and applying 9 random projective transformations to each shape. This gives a database that samples from 10 projective orbits, each 10 times (9 plus the original), for a total of 100 shapes. We then compute a $100 \times 100$ distance matrix for each of the 5 metrics used previously. Results are shown in Fig. 16 for two of the five distance matrices as well as all five precision-recall curves. The knee-point values of the P-R curves are as follows: (1) 0.319, (2) 0.437, (3) 0.721, (4) 0.798, and (5) 1.000, and thus landmark-projective shape analysis is perfect. Fig. 17 shows the geodesic paths between two queries within the same projective shape class in this classification experiment. In each instance the figure shows the three geodesic paths corresponding to the landmark shape spaces (1), (3), and (5) respectively from top to bottom. Distance values for the heart are (1) 0.4336, (3) 0.3689, and (5) $O(10^{-6})$, and distance values for the star are (1) 0.3671, (3) 0.2286, and (5) $O(10^{-6})$. That is, when comparing members of the same projective orbit, affine invariant shape analysis still exhibits a non-trivial amount of deformation; whereas, projective shape analysis yields deformations that are close to the numerical precision.

Street Sign Identification: A projective transformation models the perspective skew we observe on a planar object that is imaged by a pinhole camera at different viewing angles. We perform the same classification experiment on a small dataset of point sets taken from real images of street signs at different camera angles. We form the dataset from 24 images, 12 of a stop sign and 12 of a “do not enter” sign, whereby we select 8 landmarks on each sign by hand in a consistent fashion. Our dataset is therefore a collection of 24 point sets in $\mathbb{R}^{8 \times 2}$ belonging to 2 shape classes that, according to theory, each consist of 12 samples along a projective orbit. Fig. 18 shows a selection of images of stop signs and do not enter signs from this database with the 8 landmarks marked on each. Fig. 19 shows three pairwise distance matrices corresponding to the landmark based metrics (1), (3), and (5). From these plots it is easy to see that a projective transformation models perspective skew of planar objects very well because projective invariant shape analysis leads to the smallest within-class shape distances compared to other methods. For the stop signs, the average within-class shape distances are (1) 0.1716, (3) 0.0641, and (5) 0.0061, and for the do not enter signs, they are (1) 0.2186, (3) 0.0799, and (5) 0.0121.

Fig. 16. Retrieval of projective landmarks. Left: Distance matrix using metric (4) $S_c^a$. Center: Using metric (5) $S_l^p$. Right: Precision-recall curves for metrics (1) blue, (2) green, (3) red, (4) cyan, (5) magenta.

Fig. 17. Geodesics between two shapes in the same landmark-projective orbit. Top: $S_l^i$, Middle: $S_l^i$, Bottom: $S_p^i$.

Fig. 18. Pictures of street signs under different camera angles. Landmarks are selected points in black.

Fig. 19. Distance matrices computed from metrics (1), (3), and (5) respectively.

8 Conclusion and Future Work
It is important to develop shape analysis frameworks that are robust to the perspective skews inherent in imaging a 3D scene. While many previous methods have incorporated transformations that model perspective skew – the affine and projective transformation – they seldom develop a complete statistical shape analysis framework. Here, we provide a general foundation based in Riemannian geometry that allows us to compute distances, geodesics, and sample statistics of shapes invariant to complex transformations. Specifically, we form three shape spaces – landmark-affine, curve-affine, and landmark-projective – and show examples of analysis in each of these spaces. Experimental results confirm their respective invariances. Future efforts will study shape space of projective invariant continuous curves.
ACKNOWLEDGEMENTS

This work was supported in part by the grants NSF IIS-1217515 and NSF DMS-1208959 to AS, and by NSWC PCD ILIR program to DB. The authors acknowledge the support of Dr. Kerry Commander and Dr. Quyen Huynh.

REFERENCES


[6] Eric Klassen earned a PhD in mathematics from Cornell University in 1987, spent a 3-year postdoc at CalTech, and is now a Professor in the Florida State University Department of Mathematics. His interests include topology, geometry, shape theory, and computer image analysis.


Eric Klassen. During this time his research has focused on statistical shape analysis and spatial point processes.

Huling Le received the PhD degree in mathematics from the University of Cambridge, UK in 1989. Currently, she is at the University of Nottingham, UK. Her research interests include the statistical analysis of shape.

Anuj Srivastava is a Professor of Statistics in Florida State University. He obtained his PhD degree in Electrical Engineering from Washington University in St. Louis in 1996 and was a visiting research associate at Division of Applied Mathematics at Brown University during 1996-1997. He joined the Department of Statistics at the Florida State University in 1997 as an Assistant Professor. He was promoted to the Associate Professor position in 2003 and to the full Professor position in 2007. He has been a visiting Professor to INRIA, France and the University of Lille, France. His areas of research include statistics on nonlinear manifolds, statistical image understanding, functional analysis, and statistical shape theory. He has published more than 175 papers in refereed journals and proceedings of refereed international conferences. He has been the associate editor for the Journal of Statistical Planning and Inference, and the IEEE Transactions on Signal Processing and the IEEE Transactions on Pattern Analysis and Machine Intelligence.