Ph.D. Prospectus

Affine-Permutation Shape Invariance: Distance and Geometry

David Sepiashvili

Carnegie Mellon University
Electrical and Computer Engineering
5000 Forbes Avenue, Pittsburgh, PA
davids2@ece.cmu.edu

February 20, 2004

Contents

1 Introduction **3**

1.1 Problem statement . . . 3

1.2 Why consider 3D affine transformation for 3D shapes ? . . . 4

1.3 Why consider 2D affine transformations for 2D shapes ? . . . 4

1.4 Why do we care about permutation distortions ? . . . 5

1.5 Other shape theories . . . 5

1.6 Putting everything together . . . 6

2 Problem setup **6**

2.1 Framework . . . 6

2.2 Shape space reduction as a quotient of manifolds . . . 7

2.3 Notions of time and distance on a smooth manifold . . . 9

2.4 $\mathcal{G}r_{\text{n,p}}$ manifold and affine invariance . . . 11

2.5 $\mathcal{FS}_{\text{n,p}}$ manifold and affine-permutation invariance . . . 11

3 Geometry of the Grassmann manifold **11**

3.1 Solution for the initial value problem . . . 12

3.2 Solution for the 2-point boundary value problem . . . 13

3.3 Experimental results . . . 15

4 Distance in the residual shape space **16**

4.1 Permutation invariance as an optimization . . . 17

4.1.1 Linear programming—simplex . . . 17

4.1.2 Quadratic concave programming . . . 19

4.1.3 Derivative of the geometric cost function . . . 22

4.1.4 Experimental results . . . 23

4.2 Permutation invariance via polynomials . . . 24

5 Proposed research **26**

5.1 What will be accomplished in this research . . . 26

5.2 Other interesting questions we might pursue . . . 27

5.3 Timeline . . . 28

A Derivation of formulas for the Grassmann manifold **29**

A.1 Representing Grassmann manifold $\mathcal{G}r_{\mathrm{n,p}}$ as a quotient manifold $\mathcal{O}_{\mathrm{n}}/\mathcal{O}_{\mathrm{p}} \times \mathcal{O}_{\mathrm{n\text{-}p}}$ . . . 29

A.2 Solution for the initial value equation of the geodesic in $\mathcal{G}r_{\mathrm{n,p}}$ 32

A.3 Solution for the 2-point boundary value equation of the geodesic in $\mathcal{G}r_{\mathrm{n,p}}$ 33

1 Introduction

The problem of object representation is one of the most challenging problems in computer vision and machine learning. Learning the similarity among objects is one of its key concepts. If we do not have effective similarity measures, the important tasks of object recognition and categorization can never be performed reliably. The problem is a difficult one because we may find appearances of similar objects to be very different from each other under different viewing conditions. Moreover, different views of the same object will not have their features in correspondence with each other.

Defining an unlabeled-view-independent similarity measure is a very challenging problem by itself. Another difficulty one faces is that the similarity measure is not an Euclidean one, therefore numerical algorithms to compute distance cannot use a geometry of the Euclidean space. Thus, we need to solve an even more challenging problem of finding not just a similarity measure but a geometric manifold with unlabeled-view-independent metric properly defined on it.

1.1 Problem statement

In the view of Newtonian Mechanics, we live in a world of 3D objects. Very frequently, we capture images of the world using 2D sensors that are arbitrarily located and oriented with respect to the world. As the relative 3D position of the object is unknown and often changes in time, the view of the object captured by the sensor looks distorted in different images.

The main problem of computer vision is studying how do the captured images encode properties of the world, such as illumination and material, texture, shape, and spatial motion of the objects. In this research, we are concerned only with representation of object shapes and assume that segmentation, compensation for the illumination changes, texture, and motion recovery are already solved or need not be solved for a specific setup.

We will consider both 3D and 2D shapes of rigid objects and construct geometric shape spaces with distances defined in such a way that they achieve invariance to a class of modeled distortions. The distortions we are modeling are simultaneously applied affine transformations and permutations of unlabeled feature points.

We first motivate the use of 3D affine transformations.

1.2 Why consider 3D affine transformation for 3D shapes ?

Frequently, the imaging sensors are modeled as a pinhole perspective projection of the 3D world onto the 2D image plane. A coarser but very popular approximation is an affine projection camera model. In this model, the observed objects are assumed to undergo an affine motion before being projected ortographically onto the image plane.

Now suppose that a set of points of interest has been observed and tracked in successive images and apparent 2D motions have been estimated. An *affine structure from motion* algorithm, [1], recovers 3D shape and motion of the rigid object with respect to the camera in two steps.

The first step in this algorithm is factorization: factoring a low-rank measurement matrix of apparent 2D motions into two skinny matrices of shape and motion. The original rank-3 algorithm is explained in [2, 1], and much faster rank-1 algorithms can be found in [3]. The factorization obtained is unique up to an arbitrary 3D affine transformation. It is very simple and is based on an SVD decomposition.

On the second step, the *Euclidean metric upgrade*, [1], is performed when taking into account the rigidity constraints. These additional constraints are derived by assuming that the affine camera model is restricted to be either orthographic, weak perspective, paraperspective, or fixed-intrinsic affine. These constraints are non-linear.

As properly noted in [1], the second step need not be performed for purely affine methods. These methods do not require either camera calibration since the corresponding transformation of the image coordinates can be folded into the overall affine deformation of the object. This can be very advantageous for vision systems for which calibration parameters vary rapidly in time.

This fact motivates us defining a similarity measure for 3D shapes, which will be invariant up to an arbitrary 3D affine transformation.

1.3 Why consider 2D affine transformations for 2D shapes ?

When working with 2D shapes obtained directly from images, the 2D affine distortion model is commonly used. Under this model the 2D shapes of the same object are assumed to be related by an affine transformation. This approximation is very widely used in the literature in many image processing applications.

1.4 Why do we care about permutation distortions ?

The shapes that we consider are sets of feature points. We do not restrict how the feature points are selected. They could either be points scanned on a dense grid, or a set of points of interest with strong directional gradients in time, or points on object boundaries, or correspond to any other possible choice.

Independently of how feature points are selected, when the shape is distorted, for example, by an affine shape distortion, the feature points that comprise the shape move to different locations on the image plane. Since the image is scanned in a fixed order, such as the raster scan order, by the input device, the feature points of affine distorted shapes are not guaranteed to be read in the same order, thus adding another degree of ambiguity. We identify this type of shape irregularity as a permutation distortion. The permutation of feature points is commonly addressed as the well-known feature *correspondence* problem, see [4, 5].

1.5 Other shape theories

There are a number of different shape theories. Usually, humans define a *shape* to be the appearance of an object that is invariant under Euclidean similarity transformations. One of the first approaches seems to have been motivated by Kendall's 1977 paper [6]. Kendall assumes that the shape of an object is captured by the shape of a finite subset of its points. In his shape theory, he argues that it is not possible to measure a similarity of possible shapes of an object in standard Euclidean coordinates. Thus, he is looking for new *shape spaces* and their topological and geometric properties. Dryden and Mardia in their book on shape theory [7], following Kendall's intuitive definition of a shape, define a *shape to be "all the geometrical information that remains when location, scale, and rotational effects are filtered out from an object"*. That is, different shape spaces are obtained by factoring out a class of allowed transformations, making shapes in these spaces invariant to the modeled transformations.

Another very common shape theory is given by Grenander and his collaborators [8, 9]. These authors describe shapes in terms of closed curves that define their boundaries and outlines.

In this research we adopt the concept of a shape given by Kendall [10, 7]. However, our research is different in a number of ways. Kendall assumes a *labeled* finite subset of object's feature points. In our case, the feature points are *not* labeled, causing permutation distortions of the feature points. Also, the class of allowed transformations in Kendall's work and in our work is quite different.

Kendall restricts the class of motions to rigid motions and uniform scaling. In our research, we are modeling a class of affine motions.

This research is a continuation on our previous work given in [11], see also [12, 13, 14].

1.6 Putting everything together

The combined affine-permutation shape distortion is a frequently observed source of difficulty in many computer vision and image processing applications such as target detection, classification, pattern recognition, and registration.

We define equivalency sets of p-dim object shapes ($p \geq 2$). Each equivalency class collects the p-dim object shapes that are transformable into each other under a class of affine transformations and a class of permutations of the feature points.

In this research we propose to describe an abstract representation of these equivalency classes as points in a high-dimensional Riemannian manifold. We research efficient formulas for writing numeric algorithms in this non-Euclidean space of curvature, which allow not only to measure distance between different equivalency classes of shapes, but also to perform other geometry operations such as following geodesics.

In this research we would like to obtain formulas that are computable and take as arguments any representative shape from an equivalency class.

This geometric approach to signal processing and image processing is not common, except for few exceptions [15]. The long-term goal of this research is to bridge a gap between image processing algorithms in the Riemannian spaces and conventional algorithms embedded for simplicity in Euclidean spaces, and to take advantage of looking at these algorithms from this new perspective.

2 Problem setup

2.1 Framework

We describe the shape of an object by a set of unlabeled n feature points in $\mathcal{R}^p$ (for 2D shapes $p = 2$, for 3D shapes $p = 3$). The shape is represented by an $n \times p$ tall-skinny matrix $\mathbf{X}$ of real entries that are the Cartesian coordinates of the n feature points in $\mathcal{R}^p$. The matrix $\mathbf{X}$ is referred to as the configuration matrix. Alternatively, we will refer to $\mathbf{X}$ as the *full shape* matrix. The set of all configuration matrices $\mathcal{X}$ is defined as the configuration space, we will also refer to $\mathcal{X}$ as the *full shape space*. We note that we consider only non-singular shapes, i.e., shapes with $\text{rank}(\mathbf{X}) = p$.

We also do not allow for feature point repetitions, i.e., each feature point may appear only once in the configuration matrix.

We start by defining an equivalence relationship on $\mathcal{X}$. We say that two shapes $\mathbf{X}_1, \mathbf{X}_2 \in \mathcal{X}$ are equivalent iff they are related by:

$$\mathbf{X}_1 = P\mathbf{X}_2 A + \mathbf{1}\delta^T$$

where $A \in \mathcal{G}L_{\mathrm{p}}$ is a $p \times p$ affine distortion non-singular matrix acting on the right, and $P \in \mathcal{P}$ is an $n \times n$ permutation matrix acting on the left, which changes the order of the rows in $\mathbf{X}$, i.e., accounts for changes in the order of the feature points. The symbol δ is a $p \times 1$ translation vector that moves all feature points in $\mathcal{R}^p$ uniformly in the same direction. The vector $\mathbf{1}$ is an $n \times 1$ vector of ones. The symbol $\mathcal{G}L_{\mathrm{p}}$ represents the set of $p \times p$ invertible matrices. It is a Lie group and is called the *general linear group* [16].

Rewriting in vec$(\cdot)$ notation, defined in [17, 18], the above relationship becomes:

$$\mathrm{vec}(\mathbf{X}_1) = (A^T \otimes P)\cdot\mathrm{vec}(\mathbf{X}_2) + (\delta \otimes \mathbf{1})$$

where the symbol $\otimes$ denotes the Kronecker product, see [17, 18]. The above equation is, as shown in [12], an action of the affine-permutation group $\mathcal{G} = \left\{g : g = (A^T \otimes P, \delta \otimes \mathbf{1})\right\}$ on the configuration space.

The quotient space $\mathcal{FS}_{\mathrm{n,p}} = \mathcal{X}/G$, which we call the *residual shape space*, is the set of equivalence classes under the action of the group $\mathcal{G}$, also referred to as orbits. If we let semi-open brackets $\lceil\mathbf{X}\rfloor_{\mathcal{G}}$ to denote these equivalence classes, then we can write:

$$\lceil\mathbf{X}\rfloor_{\mathcal{G}} = \{\mathbf{X}\cdot g : g \in \mathcal{G}\} \in \mathcal{FS}_{\mathrm{n,p}}$$

We now define the *residual shape* to be the orbit $\lceil\mathbf{X}\rfloor_{\mathcal{G}} \in \mathcal{FS}_{\mathrm{n,p}}$. It is the set of all shapes that are equivalent to $\mathbf{X}$.

In this research, we would like to study the similarity of objects in the residual shape space $\mathcal{FS}_{\mathrm{n,p}}$ by defining a distance between the residual shapes. In addition to studying similarity of objects, we would like to also look at other geometric problems in $\mathcal{FS}_{\mathrm{n,p}}$, such as connecting two residual shapes by a geodesic and morphing one onto the other.

2.2 Shape space reduction as a quotient of manifolds

As we discussed above, we can go from the full shape space $\mathcal{X}$ to the residual shape space $\mathcal{FS}_{\mathrm{n,p}}$ by acting on $\mathcal{X}$ with the group $\mathcal{G}$. However, the action of the group $\mathcal{G}$ cannot be easily studied. For this reason, we break down the action of the group $\mathcal{G}$ into a set of consecutive actions of more primitive groups, as shown in Figure 1.

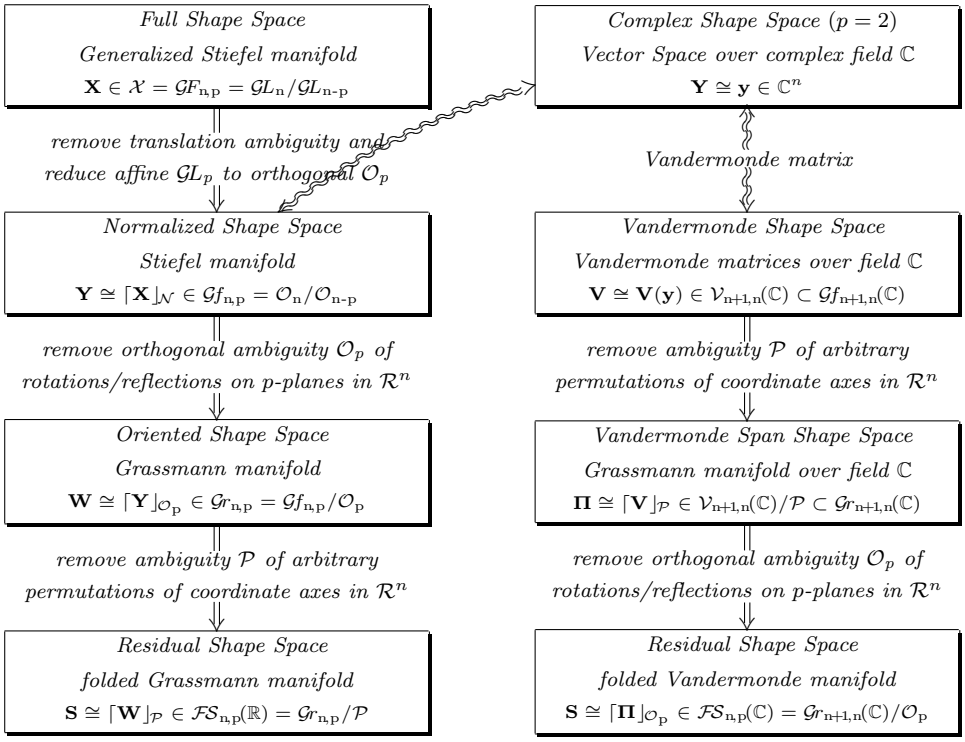

Figure 1: Shape space reduction

The left-hand side of Figure 1, going from top to bottom, shows the reduction process of the full shape space $\mathcal{X}$ into the space of residual shapes $\mathcal{FS}_{\mathrm{n,p}}$. Each factor space is obtained by acting on the space above it with some group action and forming orbits. For each orbit we could identify and keep the canonical representatives, the way it is done in [12]. But we are primarily concerned with the spaces of orbits, as they allow to work with them by picking an arbitrary member of the orbit without first finding a canonical member.

We recall that a shape is represented by an arbitrary $n \times p$ matrix $\mathbf{X} \in \mathcal{X}$ of rank p. We can view this matrix as a linearly independent p-frame in $\mathcal{R}^n$. To complete the matrix to full basis of $\mathcal{R}^n$ we can choose any orthonormal $(n-p)$-frame. Thus, we can realize $\mathcal{X}$ as a quotient of general linear groups, which are also Lie groups, [15, 16], as $\mathcal{X} \cong \mathcal{G}L_{\mathrm{n}}/\mathcal{G}L_{\mathrm{n\text{-}p}}$. We note that $\mathcal{G}F_{\mathrm{n,p}} = \mathcal{G}L_{\mathrm{n}}/\mathcal{G}L_{\mathrm{n\text{-}p}}$ obtained this way is a smooth manifold. It is frequently referred to as the *Generalized Stiefel manifold*. For background information we refer the reader to [16]. Also, we will abuse the notation and say that $\mathbf{X} \in \mathcal{G}F_{\mathrm{n,p}}$. This shape space is shown in the very top left block of Figure 1.

We construct $\mathbf{Y} \cong \lceil \mathbf{X} \rfloor_{\mathcal{N}} \in \mathcal{G}f_{\mathrm{n,p}}$ from $\mathbf{X} \in \mathcal{G}F_{\mathrm{n,p}}$ by first removing the mean, and then replacing the diagonal matrix of the singular values in its compact SVD decomposition with the identity.

Removing the mean makes the column vectors of $\mathbf{Y}$ sum to zero, while replacing the singular values with ones makes $\mathbf{Y}$ to be orthonormal ($\mathbf{Y}^T\mathbf{Y} = \mathbf{I}$). The obtained space $\mathcal{G}f_{\text{n}\natural\text{p}} = \mathcal{O}_{\text{n}}/\mathcal{O}_{\text{n-p}}$ is a space of orthonormal p-frames, and is called the *Stiefel manifold*, [16].

We note that in the literature both $\mathcal{G}F_{\text{n}\natural\text{p}}$ and $\mathcal{G}f_{\text{n}\natural\text{p}}$ are called the Stiefel manifold. We will refer to $\mathcal{G}F_{\text{n}\natural\text{p}}$ as the Generalized Stiefel manifold and to $\mathcal{G}f_{\text{n}\natural\text{p}}$ as the Stiefel manifold. Even though we could work directly with $\mathcal{G}F_{\text{n}\natural\text{p}}$, we prefer to project first $\mathcal{G}F_{\text{n}\natural\text{p}}$ onto $\mathcal{G}f_{\text{n}\natural\text{p}}$ and then work with orthonormal matrices $\mathbf{Y} \in \mathcal{G}f_{\text{n}\natural\text{p}}$. The difference between the two is that in $\mathcal{G}F_{\text{n}\natural\text{p}}$ we work with any linearly independent p-frames, while in $\mathcal{G}f_{\text{n}\natural\text{p}}$ we work with orthonormal p-frames. This is the second block of Figure 1.

In order to eliminate the orientation ambiguity, we take a Lie group $\mathcal{O}_{\text{p}}$ and consider a quotient space $\mathcal{G}f_{\text{n}\natural\text{p}}/\mathcal{O}_{\text{p}}$. Each element in the quotient space is the set of all possible rotated and reflected versions of $\mathbf{Y}$ and is given by $\lceil\mathbf{Y}\rfloor_{\mathcal{O}_{\text{p}}} = \{\mathbf{Y}V : V \in \mathcal{O}_{\text{p}}\}$. The manifold $\mathcal{G}r_{\text{n}\natural\text{p}} = \mathcal{G}f_{\text{n}\natural\text{p}}/\mathcal{O}_{\text{p}}$ obtained in this way is called the *Grassmann manifold*, [16]. It can also be defined as the space of all p-dim subspaces in $\mathcal{R}^n$. Under this definition, if $\mathbf{W}$ is a point on the Grassmann manifold $\mathcal{G}r_{\text{n}\natural\text{p}}$ then $\mathbf{W} = \text{span}(\mathbf{Y})$. Both definitions are equivalent, but we will work with the orbits $\lceil\mathbf{Y}\rfloor_{\mathcal{O}_{\text{p}}}$. We will use a shorthand symbol $\lceil\mathbf{Y}\rfloor$, omitting the group $\mathcal{O}_{\text{p}}$ as its subindex.

The final step allows arbitrary permutations of n feature points, which are row permutations of the shape matrices. This ambiguity is eliminated by taking a finite group of permutations $\mathcal{P}$ and folding the Grassmann manifold into $\mathcal{G}r_{\text{n}\natural\text{p}}/\mathcal{P}$. We will check to see if such folding is possible. So it will be part of the research to either find such folding or prove it impossible. The residual shape space $\mathcal{FS}_{\text{n}\natural\text{p}} = \mathcal{G}r_{\text{n}\natural\text{p}}/\mathcal{P}$ obtained in this way is a space of residual shapes that achieve invariance to both affine transformations and permutation distortions. This is shown in the last block on the left-hand side of Figure 1.

The right-hand side of Figure 1 shows an alternative way of going from the full shape space $\mathcal{X}$ to the residual shape space $\mathcal{FS}_{\text{n}\natural\text{p}}$. We leave the discussion of this alternative way to Sub-Section 4.2.

Now, before we continue further, we present the key notions in the theory of manifolds described in a very informal way. They lead us to the notions of time and distance on manifolds.

2.3 Notions of time and distance on a smooth manifold

We refer the reader for the necessary background on differentiable manifolds to classical textbooks on Riemannian geometry [19, 16, 20].

Smooth curves are smooth mappings $\gamma : [\tau_1, \tau_2] \to \mathcal{Y}$ from the time domain $\mathbb{R}$ to a smooth manifold $\mathcal{Y}$. They carry a way of natural differentiation in time $\gamma' = d\gamma/dt$. The derivative $\mathbf{H}_t = \gamma'(t)$ at point $\mathbf{Y}_t = \gamma(t)$ is called the *tangent vector* at $\mathbf{Y}_t$. *Tangent space* $T_p\mathbf{Y}$ at a point $\mathbf{Y} \in \mathcal{Y}$ is defined to be a set of all tangent vectors at point $\mathbf{Y}$. *Vector fields* are mappings that assign to each point on $\mathcal{Y}$ one of its tangent vectors. In Euclidean spaces we can take the directional derivative of one vector field with respect to the other vector field. We generalize this notion for manifolds by defining a *connection*. It is also called a *covariant derivative* of a vector field with respect to the other vector field. In a way, a connection is a generalization of taking second order derivatives. Also, a covariant derivative of a vector field along some curve $\gamma(t)$ is defined to be a restriction of it along the curve. A vector field along some curve is said to be a *parallel vector field* if its covariant derivative is zero along the curve. Given a curve $\gamma(t)$ and a tangent vector $\mathbf{D}_o$ at the point $\gamma(0)$, there exists a unique parallel vector field along the curve that extends the tangent vector. We say we *parallel transported* the tangent vector $\mathbf{D}_o$ along the curve to $\mathbf{D}_t$ at point $\gamma(t)$. We define a straight line, called a geodesic, to be a curve γ of zero acceleration, i.e., a curve that parallel transports its derivative tangent vector $\mathbf{H}_o = \frac{d\gamma}{dt}(0)$, see Figure 2 for the illustration. The requirement of having second order derivative to be zero comes from physics if we consider the fact that if an object moves along some curve at a constant pace, its acceleration will be zero only if this curve is a straight line.

We do not define distances directly on the manifold. Rather, we equip every tangent space $T_p\mathbf{Y}$ with an inner product, making $\mathcal{Y}$ a Riemannian manifold. There could be many connections on the manifold $\mathcal{Y}$. A connection is called *compatible* with the Riemannian metric if it preserves inner products during parallel transport. Also, if we require a connection to satisfy one more property, namely, to be torsion-free. Then there exists only one connection, called a *Levi-Civita* connection, that is compatible with a given Riemannian metric and is torsion-free. We also define an exponential map from a tangent space to a manifold $\exp : T_p\mathbf{Y} \to \mathcal{Y}$, which allows us to describe a neighborhood of $\mathbf{Y} \in \mathcal{Y}$ by a neighborhood in the flat space $T_p\mathbf{Y}$. Then a notion of distance on the manifold is defined in the following way. We pick a point $\mathbf{Y}_o \in \mathcal{Y}$ and some tangent vector $\mathbf{H}_o$ at point $\mathbf{Y}_o$. This defines a geodesic curve $\gamma(t)$ s.t. $\gamma(0) = \mathbf{Y}_o$ and $\gamma'(0) = \mathbf{H}_o$. Then the distance between points $\mathbf{Y}_o = \gamma(0)$ and $\mathbf{Y}_\tau = \gamma(\tau)$ is the length of the $\gamma(t)$ curve:

$$d\left(\mathbf{Y}_o, \mathbf{Y}_\tau\right) = \int_o^\tau \left\| \frac{d\gamma}{dt} \right\| dt = \int_o^\tau \left\| \mathbf{H}_t \right\| dt = \tau \cdot \left\| \mathbf{H}_o \right\| \tag{1}$$

where $\|\cdot\|$ denotes the Riemannian metric defined on the tangent space. Since the curve $\gamma(t)$ is a geodesic, its *derivative vector field* $\gamma'(t)$ is the parallel vector field. At the same time a Levi-Civita

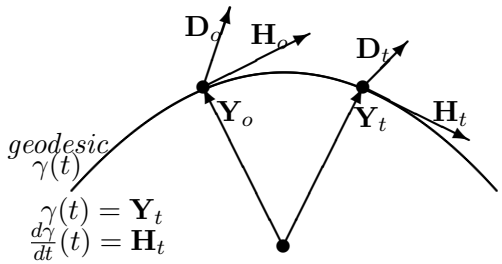

Figure 2: Connecting points $\mathbf{Y}_o$ and $\mathbf{Y}_t$ by a geodesic $\gamma(t)$

connection is compatible with the Riemannian metric and so preserves norms of parallel vector fields. Thus, the norm will be constant in the integration. We note that in general, $\mathbf{Y}_\tau$ need to be in the neighborhood of $\mathbf{Y_o}$, except for the case when a geodesic can be extended indefinitely in both directions. We define a manifold to be *geodesically complete* if every geodesic can be extended indefinitely in both directions. For example, the Grassmann manifold is geodesically complete.

2.4 $\mathcal{G}r_{\mathbf{n,p}}$ manifold and affine invariance

As discussed in Sub-Section 2.2, the third block on the left-hand side of Figure 1 is the Grassmann manifold $\mathcal{G}r_{\mathrm{n,p}} = \mathcal{G}f_{\mathrm{n,p}}/\mathcal{O}_{\mathrm{p}}$. This manifold is of great importance, as it provides us with an affine-invariant shape space. We study its geometry in Section 3.

2.5 $\mathcal{FS}_{\mathbf{n,p}}$ manifold and affine-permutation invariance

As discussed in Sub-Section 2.2, the final step on the left-hand side of Figure 1 is folding the Grassmann manifold $\mathcal{G}r_{\mathrm{n,p}}$ with a group of permutations $\mathcal{P}$. This group does not act freely, [16]. Thus, we cannot realize a geometry in $\mathcal{FS}_{\mathrm{n,p}}$ as a quotient geometry of $\mathcal{G}r_{\mathrm{n,p}}/\mathcal{P}$. One way of defining the geometry would be to require that the folded space $\mathcal{FS}_{\mathrm{n,p}}$ be *convex* in $\mathcal{G}r_{\mathrm{n,p}}$ and use the geometry of $\mathcal{G}r_{\mathrm{n,p}}$ in it. We have results on such folding for the $p = 1$ case. For $p > 1$ the folding is still under the investigation, and is an open research question.

In Section 4 we are studying a somewhat more relaxed problem than making $\mathcal{FS}_{\mathrm{n,p}}$ a geometric manifold. We look only for a distance in the residual shape space $\mathcal{FS}_{\mathrm{n,p}}$. This distance should be invariant to both affine transformations and permutation distortions.

3 Geometry of the Grassmann manifold

We note that all the formulas, given in this section, are derived in Apendix A. We remind the reader that a point on the Grassmann manifold is the orbit $\lceil \mathbf{Y} \rfloor_{\mathcal{O}_{\mathrm{p}}} \in \mathcal{G}r_{\mathrm{n,p}}$ under the action on

the Stiefel manifold $\mathcal{G}f_{\mathrm{n,p}}$ of the Lie group $\mathcal{O}_{\mathrm{p}}$. The matrix $\mathbf{Y}$ is an orthonormal $n \times p$ tall-skinny matrix representing a point on the Stiefel manifold $\mathcal{G}f_{\mathrm{n,p}}$. In this section, we study in detail $\mathcal{G}r_{\mathrm{n,p}}$, with emphasis on computational aspects. We are working with any arbitrary $\mathbf{Y} \in \lceil \mathbf{Y} \rfloor$ matrix as the defining representative of the orbit $\lceil \mathbf{Y} \rfloor \in \mathcal{G}r_{\mathrm{n,p}}$.

We consider a natural embedding of $\mathcal{G}r_{\mathrm{n,p}}$ in $\mathcal{R}^{n \times n}$ and equip the manifold with the canonical Levi-Civita connection. Then we study equations for the geodesic and its derivative vector field.

3.1 Solution for the initial value problem

Edelman, Arias, and Smith [21] have proposed a computationally practical way of defining equations for the geodesic and its derivative vector field, given an initial point and a tangent direction at the point. This is known as an *initial value problem* in the theory of differential equations. Their approach does not require an explicit formula for the connection. For completeness, we present their result below.

Let $\mathbf{Y}_o$ be a representative of the orbit $\lceil \mathbf{Y}_o \rfloor \in \mathcal{G}r_{\mathrm{n,p}}$ and let $\mathbf{H}_o$ be a valid tangent vector at $\mathbf{Y}_o$. We define the geodesic $\gamma(t)$ by the initial values $\gamma(0) = \lceil \mathbf{Y}_o \rfloor$ and $\gamma'(0) = \lceil \mathbf{H}_o \rfloor$. We perform an SVD decomposition of $\mathbf{H}_o = UEV^T$. We define matrices $C_t = \mathrm{cosm}(t \cdot E)$ and $S_t = \mathrm{sinm}(t \cdot E)$ to be the matrix-cosine and matrix-sine, which are defined well for symmetric matrices, [18]. In case of the diagonal matrix E, these will also be the diagonal matrices with diagonal elements that are the cosine and sine of the corresponding elements in E. Then the equations for the geodesic $\gamma(t) = \lceil \mathbf{Y}_t \rfloor$ and its derivative vector field along the geodesic $\gamma'(t) = \lceil \mathbf{H}_t \rfloor$ at time t are given by:

$$\begin{aligned} \mathbf{Y}_t &= \mathbf{Y}_o V C_t V^T + \mathbf{H}_o V S_t E^{-1} V^T \\ \mathbf{H}_t &= -\mathbf{Y}_o V S_t E V^T + \mathbf{H}_o V C_t V^T \end{aligned} \tag{2}$$

These equations tell us how the point $\lceil \mathbf{Y}_t \rfloor$ and its tangent $\lceil \mathbf{H}_t \rfloor$ evolve along the geodesic. This result is derived in [21].

In Figure 3 we present an example of a geodesic defined by the initial shape $\mathbf{Y_o} \in \mathcal{G}r_{500,2}$ and the tangent direction $\mathbf{H_o}$. In this example, we take our initial shape to be a contour of the pentagon. We pick a tangent direction $\mathbf{H_o}$ randomly from all tangent vectors at $\mathbf{Y_o}$. Then, we start moving on the geodesic defined by Equation (2), and compute intermediate shapes at a distance $d = \sqrt{2}\frac{\pi}{2}0.04$ from each other for a total of 25 equally spaced shapes along the same geodesic. Note that all these shapes are also defined by 500 points as the original pentagon. The shapes we obtain are not likely to be shapes that we would see in real life. This shows that there are lots of shapes in the shape

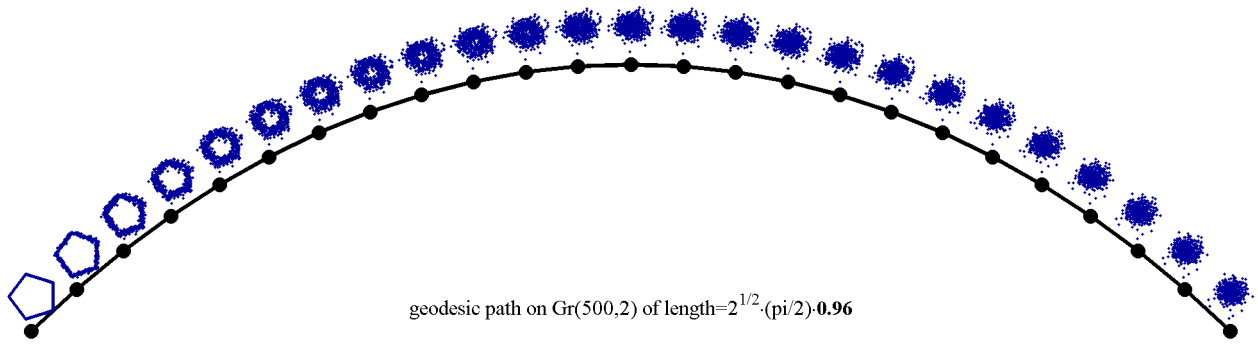

Figure 3: Extending a geodesic in a random direction

space that have a very small a-priori probability of happening in the real world. Also, this shows that the tangent vectors that are likely to be followed do have a certain structure. This motivates us for the problem presented in the following section.

3.2 Solution for the 2-point boundary value problem

The problem we would like to solve is quite different from the solution given by Equation (2). Instead of defining equations for the geodesic and its derivative vector field, given an initial point and a tangent direction at the point, it is essential for us to define these equations, given an initial point and a final point on the $\mathcal{G}r_{\text{n,p}}$, which are to be connected by a geodesic. In the theory of differential equations this corresponds to a *two-point boundary value problem*. Such a problem is commonly solved numerically, using the shooting method [22]. We present our result for an explicit expression for the solution of the two-point boundary value problem involving only standard linear algebra operations.

Let $\lceil \mathbf{Y}_o \rfloor, \lceil \mathbf{Y}_\tau \rfloor \in \mathcal{G}r_{\text{n,p}}$ be two points on the $\mathcal{G}r_{\text{n,p}}$, which are to be connected by a geodesic. We define the geodesic $\gamma(t)$ by the initial $\gamma(0) = \lceil \mathbf{Y}_o \rfloor$ and final $\gamma(\tau) = \lceil \mathbf{Y}_\tau \rfloor$ values. This is a two-point boundary value problem. Let $\mathbf{Y}_o \in \lceil \mathbf{Y}_o \rfloor$ and $\mathbf{Y}_\tau \in \lceil \mathbf{Y}_\tau \rfloor$ be some representatives for the initial and final values. We perform an SVD decomposition of $\mathbf{Y}_o^T \mathbf{Y}_\tau = V_o C_\tau V_\tau^T$, which gives two rotation matrices V_o, V_τ and a cosine matrix of the principal angles C_τ. Then we define the matrix $E_\tau = \text{acosm}(C_\tau)$, whose diagonal elements are the principal angles between the two subspaces spanned by $\lceil \mathbf{Y}_o \rfloor$ and $\lceil \mathbf{Y}_\tau \rfloor$, see [23]. The distance between the points is given by:

$$\text{d}\Big(\lceil \mathbf{Y_o} \rfloor, \lceil \mathbf{Y}_\tau \rfloor\Big) = \tau = \|E_\tau\|_F \tag{3}$$

We define the matrix $E = E_\tau/\tau$ so that $\|E\|_F = 1$. Also, we define $S_\tau = \text{sinm}(\tau \cdot E)$ as well as

cosine and sine matrices of half angles $C_m = \mathrm{cosm}(\frac{\tau}{2} \cdot E)$ and $S_m = \mathrm{sinm}(\frac{\tau}{2} \cdot E)$.

In Euclidean space, we can take two points $\mathbf{y}_o, \mathbf{y}_\tau \in \mathcal{R}^n$, connect them by a line, and then find a point $\mathbf{y}_m$ in between them, which will lie at equal distance to both points. This mid-point will be given by $\mathbf{y}_m = (\mathbf{y}_o + \mathbf{y}_\tau)/2$. We present our result for a similar construct on the Grassmann manifold $\mathcal{G}r_{\mathrm{n,p}}$. It is of great importance for studying its properties. The equations of the $\gamma(t)$ geodesic's midpoint and of the tangent vector from its derivative vector field at the midpoint can be shown to be given by:

$$\begin{aligned} \mathbf{Y}_m &= \tfrac{1}{2}\left(\mathbf{Y}_\tau V_\tau + \mathbf{Y}_o V_o\right) C_m^{-1} V_o^T \\ \mathbf{H}_m &= \tfrac{1}{2}\left(\mathbf{Y}_\tau V_\tau - \mathbf{Y}_o V_o\right) S_m^{-1} E V_o^T \end{aligned} \tag{4}$$

In order to obtain a closed form solution for the two-point boundary value problem, all we need to do is to travel back from the point $\lceil \mathbf{Y}_m \rfloor$, given by equation (4), at time $t = 0$ to the point $\lceil \mathbf{Y}_o \rfloor$ at time $t = -\tau/2$ on the geodesic defined by equation (2). First, we present a closed-form solution for $\mathbf{H}_o$ such that $\|\mathbf{H}_o\|_F = 1$:

$$\mathbf{H}_o = -\mathbf{Y}_o V_o C_\tau {S_\tau}^{-1} E V_o^T + \mathbf{Y}_\tau V_\tau {S_\tau}^{-1} E V_o^T \tag{5}$$

If we use brackets to denote stacking of matrices together to form a larger matrix, then an alternative solution can also be obtained from the equation:

$$\left[\begin{array}{c|c} \mathbf{Y}_o & \mathbf{Y}_\tau \end{array}\right]^T \cdot \left[\begin{array}{c|c} \mathbf{H}_o & \mathbf{H}_\tau \end{array}\right] = \begin{bmatrix} V_o & 0 \\ 0 & V_\tau \end{bmatrix} \cdot \begin{bmatrix} 0 & -ES_\tau \\ ES_\tau & 0 \end{bmatrix} \cdot \begin{bmatrix} V_o & 0 \\ 0 & V_\tau \end{bmatrix}^T \tag{6}$$

From this equation we simultaneously solve for $\mathbf{H}_o$ and $\mathbf{H}_\tau$ using the pseudo-inverse of $[\mathbf{Y}_o | \mathbf{Y}_\tau]^T$. These two tangents are unit norm parallel-transported versions of each other, and belong to the geodesic's derivative vector field.

We now present our result for the equations of the geodesic $\gamma(t) = \lceil \mathbf{Y}_t \rfloor$ and its derivative vector field along the geodesic $\gamma'(t) = \lceil \mathbf{H}_t \rfloor$ at time t. This is an explicit expression for the solution of the two-point boundary value problem. It is obtained by plugging Equation (5) into Equation (2). After simplification we obtain:

$$\begin{aligned} \mathbf{Y}_t = &\ \mathbf{Y}_o V_o (C_t - S_t C_\tau {S_\tau}^{-1}) V_o^T + \mathbf{Y}_\tau V_\tau S_t S_\tau^{-1} V_o^T \\ \mathbf{H}_t = &-\mathbf{Y}_o V_o (S_t + C_t C_\tau S_\tau^{-1}) E V_o^T + \mathbf{Y}_\tau V_\tau C_t S_\tau^{-1} E V_o^T \end{aligned} \tag{7}$$

The time-varying matrices in equation (7) are $C_t = \mathrm{cosm}(t \cdot E)$ and $S_t = \mathrm{sinm}(t \cdot E)$. All other matrices are obtained from an SVD decomposition of the inner product $\mathbf{Y}_o^T \mathbf{Y}_\tau$.

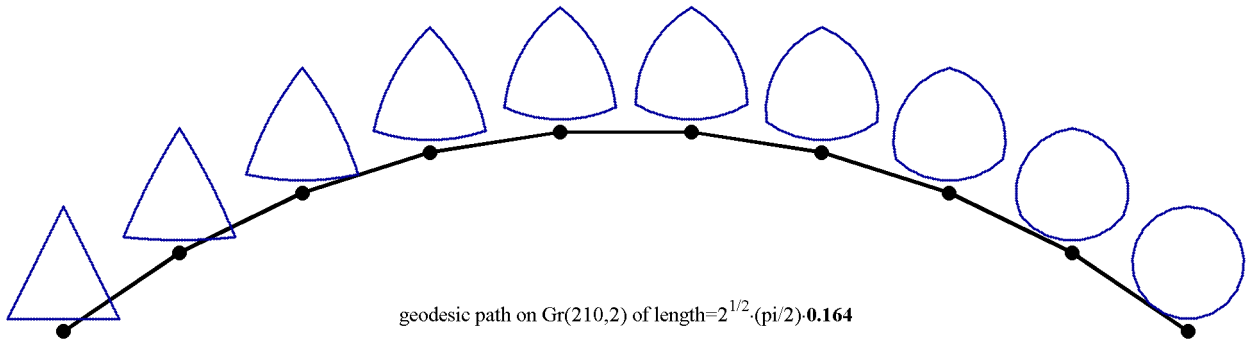

Figure 4: Morphing the contour of a triangle onto the contour of a circle

If at time $t = 0$, we start traversing the geodesic given by $\mathbf{Y}_o$ and $\mathbf{H}_o$ from the point $\lceil \mathbf{Y}_o \rfloor \in \mathcal{G}r_{\mathrm{n,p}}$, then at time $t = \tau$, we will reach a point $\lceil \mathbf{Y}_\tau \rfloor \in \mathcal{G}r_{\mathrm{n,p}}$. This is due to the fact that $\forall t \ \|\mathbf{H}_t\|_F = 1$, so from Equation (1) the distance is τ, which is also given by Equation (3).

3.3 Experimental results

Now we will present some results on computing distances and morphing shapes on the Grassmann manifold $\mathcal{G}r_{\mathrm{n,p}}$.

In Figure 4 we present a geodesic connecting two shapes that lie on the Grassmann manifold $\mathcal{G}r_{210,2}$. We pick our initial shape to be a contour of the triangle, and our final shape to be a contour of the circle. Then, using Equation (3) we determine the distance in between the two shapes. If we normalize the distance to be in the range $[0, \ldots, 1]$, then the normalized distance for this example is $d = 0.164$, as shown underneath the figure. Also, using Equation (7), we can obtain other shapes that lie on the geodesic given by these initial and final shapes. In this example, we divide the geodesic into 9 equal parts, and compute intermediate shapes along the geodesic. The distance between adjacent shapes will be $d/9$.

Another example is given in Figure 5. In this example, our initial shape is a 2D face template, and our final shape is a bicycle. These shapes lie on the $\mathcal{G}r_{2000,2}$. The normalized distance between the shapes is $d = 0.683$. Thus, these shapes lie far from each other. As before, we show some intermediate shapes as well. They show what changes a shape of the face undergoes to become the shape of a bicycle.

Yet another example is given in Figure 6. Here we are considering 3D shapes. In this example, our initial shape is a 3D fractal bowl, and our final shape is a moebius strip. These shapes lie on the $\mathcal{G}r_{4096,3}$. The normalized distance in between the shapes is $d = 0.594$.

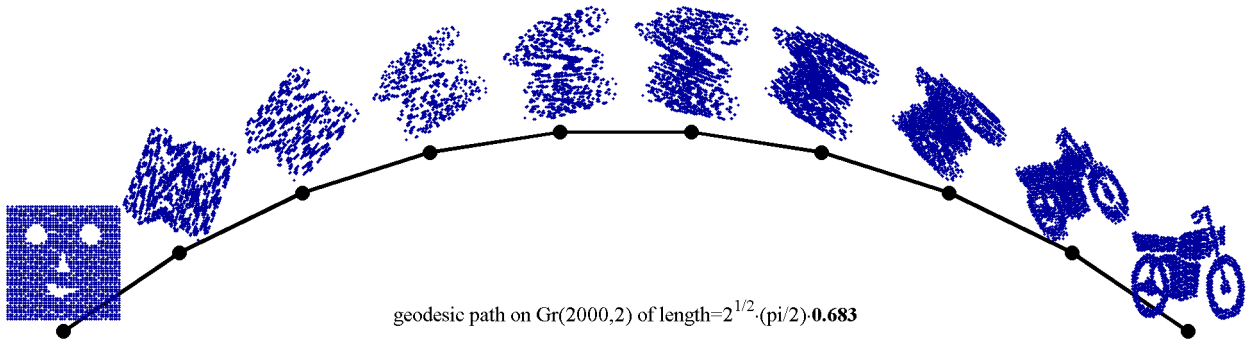

Figure 5: Morphing a face template onto a bicycle

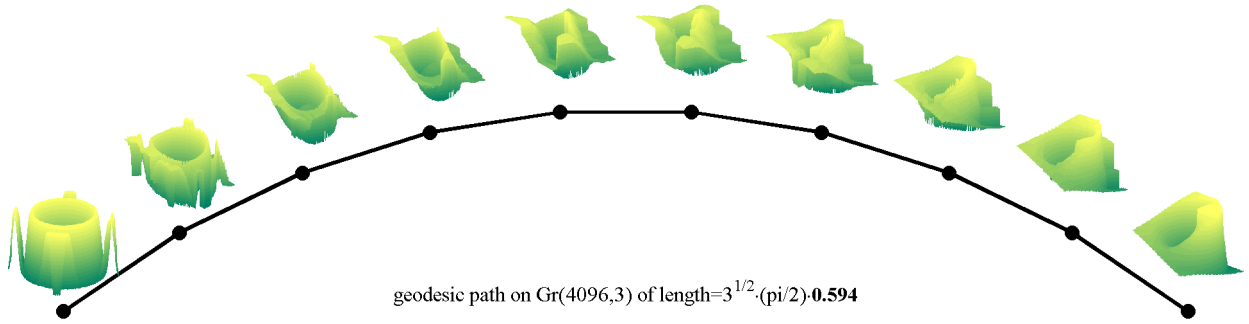

Figure 6: Morphing a fractal bowl onto a moebius strip

The last example we consider is given in Figure 7. In this example, we take the contour of a triangle, and we make two distorted shapes out of it. First, we apply two random affine distortions to obtain two different shapes. Then, we add additive gaussian noise to the location of the feature points in both shapes. In the first shape we add more noise than in the second one. Then, as before, we connect them by a geodesic, compute the distance in between them, and also compute some intermediate shapes along the geodesic. The normalized distance between them is relatively small. It is $d = 0.09$. So, we may expect that the distance on $\mathcal{G}r_{\mathrm{np}}$ should be resistant to additive noise in the location of feature points. This will be a research question we intend to pursue.

4 Distance in the residual shape space

Our goal in this research is to find a geometric manifold $\mathcal{FS}_{\mathrm{np}}$ with an unlabeled-view-independent metric properly defined on it. This is a very challenging problem. Thus, we first study a somewhat more relaxed problem. In this chapter we will look only for a distance in the residual shape space $\mathcal{FS}_{\mathrm{np}}$. This will properly define a similarity measure in between the shapes in $\mathcal{X}$. This similarity

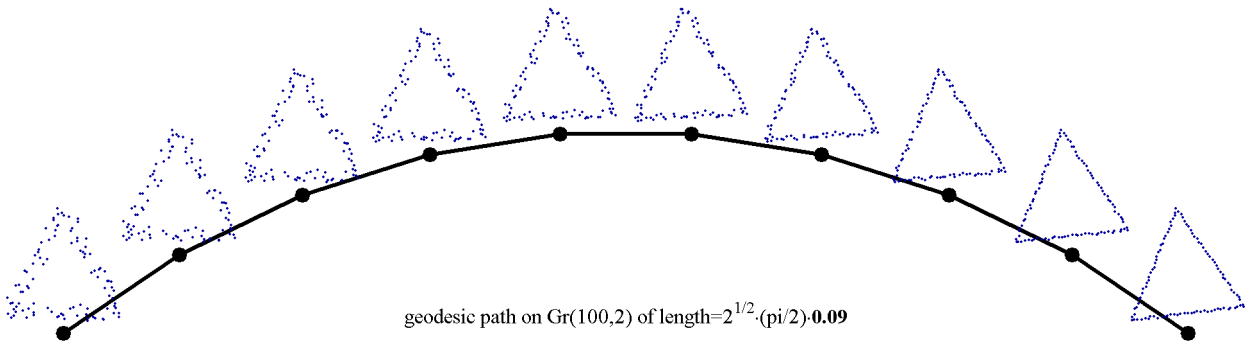

Figure 7: Morphing one distorted triangle onto the other

measure should be invariant to both affine transformations and permutations. It also should be an easily computable one.

In Figure 1, we presented two approaches of reducing the Full Shape Space $\mathcal{X}$ into the residual shape space $\mathcal{FS}_{\mathrm{n,p}}$. First, we start by reducing $\mathcal{X}$ to the normalized shape space $\mathcal{G}f_{\mathrm{n,p}}$. This reduction leaves us only with two types of distortions: permutations $\mathcal{P}$ and orthogonal transformations $\mathcal{O}_{\mathrm{p}}$. Then, we consider two approaches. In the first approach, we would like to achieve invariance to rotations by factoring out $\mathcal{O}_{\mathrm{p}}$ first, and then an invariance to permutations by factoring out $\mathcal{P}$. In the second approach, we first would like to achieve an invariance to permutations by factoring out $\mathcal{P}$, and then an invariance to rotations by factoring out $\mathcal{O}_{\mathrm{p}}$.

In this chapter we study the possibility of defining two different distances that follow the two approaches above. This leads us to some open questions to pursue further in this research.

4.1 Permutation invariance as an optimization

In this section we present a distance in the residual shape space $\mathcal{FS}_{\mathrm{n,p}}$. It is an easily computable distance that approximates a true distance, invariant to both affine transformations and permutations.

First, we present the approximate distance and an algorithm for its computation. Then we discuss how it differs from the true distance and why true distance is not easily computable. Finally, we present some experimental results that show its performance.

4.1.1 Linear programming—simplex

In the residual space $\mathcal{FS}_{\mathrm{n,p}}$ we can define a distance by minimizing the distance of the $\mathcal{G}r_{\mathrm{n,p}}$ space over a group of permutations:

$$\mathrm{d}_{\mathcal{FS}_{\mathrm{n_p}}}(\mathbf{Y}_o, \mathbf{Y}_\tau) \equiv \min_{\mathbf{P}\in\mathcal{P}} \mathrm{d}_{\mathcal{G}_{\mathrm{r_{n_p}}}}(\mathbf{Y}_o, P\mathbf{Y}_\tau) \tag{8}$$

If we approximate the distance in Equation (8) by the Euclidean distance, then the problem becomes:

$$\mathrm{d}_L(\mathbf{Y}_o, \mathbf{Y}_\tau) = \min_{\mathbf{P}\in\mathcal{P}} \|\mathbf{Y}_o - P\mathbf{Y}_\tau\|_F \tag{9}$$

This optimization problem in the Euclidean space is also known as the correspondence problem, and it is easily converted into the linear optimization framework, [4, 5]:

$$\begin{aligned} P_L &= \arg\min_{P\in\mathcal{P}} \|\mathbf{Y}_o - P\mathbf{Y}_\tau\|_F \\ &= \arg\min_{P\in\mathcal{P}} \mathrm{trace}\,(\mathbf{Y}_o - P\mathbf{Y}_\tau\ )^T\,(\mathbf{Y}_o - P\mathbf{Y}_\tau\ ) \\ &= \arg\min_{P\in\mathcal{P}} \mathrm{trace}\left(\mathbf{Y}_o^T\mathbf{Y}_o - 2\mathbf{Y}_o^T P\mathbf{Y}_\tau + \mathbf{Y}_\tau^T\mathbf{Y}_\tau\right) \\ &= \arg\max_{P\in\mathcal{P}} \mathrm{trace}\left(\mathbf{Y}_o^T P\mathbf{Y}_\tau\right) \end{aligned} \tag{10}$$

Permutation matrices are square matrices obtained by permuting rows of the identity matrix. Thus, they are doubly-stochastic and orthogonal at the same time. In fact, it can be shown that any doubly-stochastic orthogonal matrix is a permutation matrix. Since a set of permutations $\mathcal{P}$ is a subset of doubly stochastic matrices $\mathcal{D}$, we can relax the optimization problem and write:

$$\max_{P\in\mathcal{P}} \mathrm{trace}\left(\mathbf{Y}_o^T P\mathbf{Y}_\tau\right) \leq \max_{D\in\mathcal{D}} \mathrm{trace}\left(\mathbf{Y}_o^T D\mathbf{Y}_\tau\right) \tag{11}$$

Let us define the cost function to be:

$$J_L(D) = -\mathrm{trace}(\mathbf{Y}_o^T D\mathbf{Y}_\tau) = -\mathrm{trace}(D\cdot\mathbf{Y}_\tau\mathbf{Y}_o^T) \tag{12}$$

Its gradient is given by:

$$\nabla J_L(D) = -\mathbf{Y}_o\mathbf{Y}_\tau^T \tag{13}$$

The optimum of the cost function in Equation (12) over all $D \in \mathcal{D}$ is attained when D is a permutation matrix. This follows from Birkhoff's theorem [24, 25], which says that a set of doubly-stochastic matrices $\mathcal{D}$ is the convex hull over a group of permutations $\mathcal{P}$ and also from the fact that our objective function is linear, and thus achieves its optimum at the boundaries. Thus, the inequality in Equation (11) is in fact an equality and we can write:

$$P_L = \arg\min_{D\in\mathcal{D}} J_L(D) \tag{14}$$

Now we rewrite the optimization problem in a canonical form, so that it can be solved numerically using standard simplex routines. Also we note that the constraints on D to be a doubly-stochastic matrix are expressed as:

$$D \cdot \mathbf{1} = \mathbf{1} \quad , \quad \mathbf{1}^T \cdot D = \mathbf{1}^T \quad , \quad 0 \leq d_{i,j} \leq 1 \tag{15}$$

We apply vec$(\cdot)$ notation, defined in [17, 18], to convert matrices into vectors:

$$\triangledown J_L(d) = -\text{vec}(\mathbf{Y}_o \mathbf{Y}_\tau^T) \quad \text{where} \quad d = \text{vec}(D) \tag{16}$$

This leads to a canonical linear cost function with canonical linear constraints, given by:

$$\begin{aligned} J_L(d) &= -\text{vec}(\mathbf{Y}_o \mathbf{Y}_\tau^T)^T \cdot d \\ (\mathbf{1}^T \otimes I) \cdot d &= \mathbf{1} \\ (I \otimes \mathbf{1}^T) \cdot d &= \mathbf{1} \\ \mathbf{0} \leq d &\leq \mathbf{1} \end{aligned} \tag{17}$$

which is solved by standard simplex or simplex-like algorithms.

We present one example here that shows why it is important to achieve invariance to permutations. In Figure 8 we take two different shapes and connect them by a geodesic. Our initial shape $\mathbf{Y_o}$ is a filled triangle, and our final shape $\mathbf{Y}_\tau$ is a filled circle. The Grassmannian distance between them $\text{d}_{\mathcal{G}r_{\text{np}}}(\mathbf{Y_o}, \mathbf{Y}_\tau) = 0.535$. Now we perform the optimization given by Equation (9) and find a permutation P_L that minimizes the Euclidean distance between these two shapes. In Figure 9 we present the geodesic between same initial shape $\mathbf{Y_o}$ and the permuted final shape $P_L \mathbf{Y}_\tau$. The Grassmannian distance between them is $\text{d}_{\mathcal{G}r_{\text{np}}}(\mathbf{Y}_o, P_L \mathbf{Y}_\tau) = 0.166$, which is now much smaller than for the shapes in Figure 8. If we visually compare the final shapes in Figure 8 and 9, we see no difference. This is because they only differ by a permutation in the order of their feature points. However, the intermediate shapes on these geodesics are very different. The ones in Figure 8 are more distorted. They look as if feature points are getting rotated on a spiral in order to bring features of the initial and final shapes into correspondence with each other.

We remind the reader that the cost function in Equation (17) solves for the Euclidean distance given in Equation (9) instead of our true Grassmannian distance given in Equation (8). Next, we will recall how the true Grassmannian distance is computed. Also, we will look for a better approximation to the true distance.

4.1.2 Quadratic concave programming

Let $\lceil \mathbf{Y}_o \rfloor, \lceil P\mathbf{Y}_\tau \rfloor \in \mathcal{G}r_{\text{np}}$ be two points on the Grassmann manifold. Then in order to compute the distance, we first perform an SVD decomposition of $\mathbf{Y}_o^T P \mathbf{Y}_\tau = V_o C_\tau V_\tau^T$, which gives two rotation

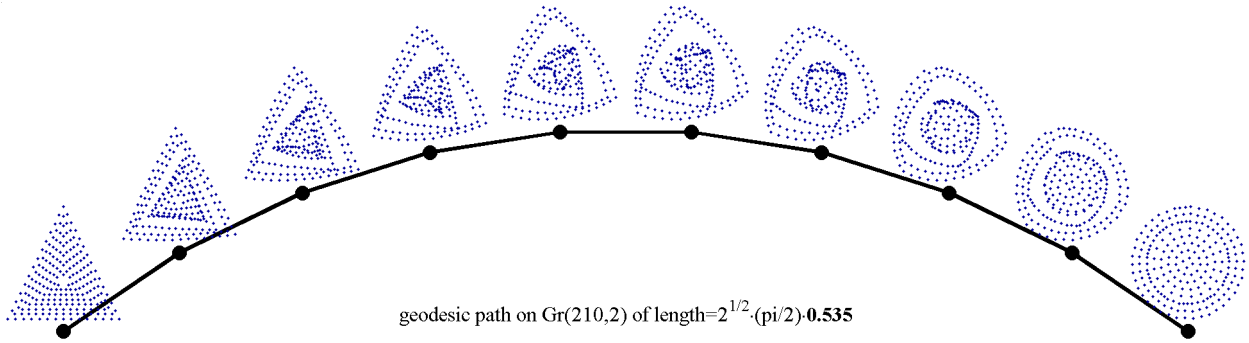

Figure 8: Morphing a filled triangle onto a filled circle

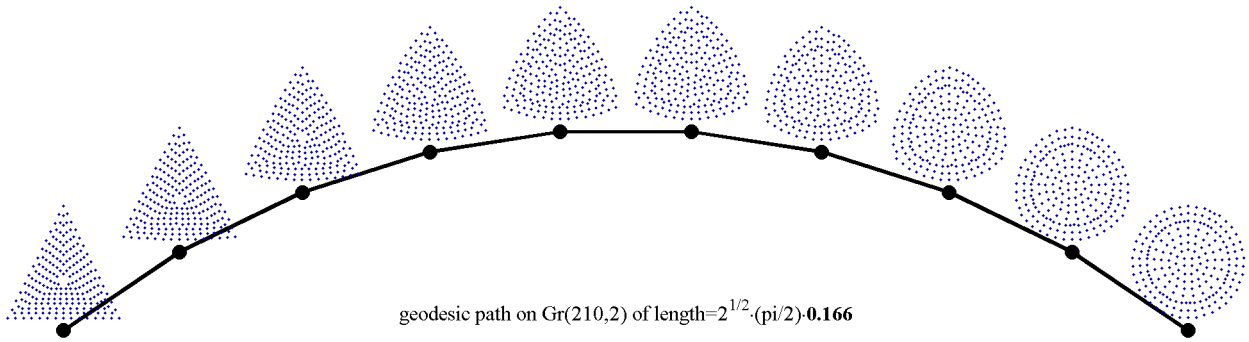

Figure 9: Morphing a filled triangle onto a filled permuted circle

matrices V_o, V_τ and a cosine diagonal matrix of the principal angles C_τ. Then we compute the matrix arc-cosine $E_\tau = \text{acosm}(C_\tau)$. The distance is given by $\text{d}_{\mathcal{G}r_{\text{n}\text{p}}}(\mathbf{Y}_o, P\mathbf{Y}_\tau) = \text{trace}(E_\tau^2)$.

Matrix sine, cosine, arc-sine, and arc-cosine are well defined for diagonal and symmetric matrices. We are going to extend the definition of these trigonometric functions to any arbitrary matrix A with the help of an SVD decomposition of A. This will help us to study the geometric distance on $\mathcal{G}r_{\text{n}\text{p}}$.

Let E be a diagonal matrix. We define $e^E = \exp(E) \equiv 1 + E/1 + E^2/2! + ...$ An interesting property of the matrix exponential is given by: $Qe^EQ^T = e^{QEQ^T}$. The matrix cosine and sine of a diagonal matrix are defined by $\text{cosm}(E) = \frac{e^{iE}+e^{-iE}}{2}$ and $\text{sinm}(E) = \frac{e^{iE}-e^{-iE}}{2i}$. They are computed by applying scalar sine and cosine to the diagonal elements, while keeping off-diagonal elements at zero. The definition is naturally extended to symmetric matrices using the property above. Now we extend the definition of matrix sine and cosine to any arbitrary square matrix A. Let $A = UEV^T$ be an SVD decomposition of A. We define $\text{cosm}(A) = U\text{cosm}(E)V^T$. Similarly we define all other matrix trigonometric functions. Now we are ready to rewrite the geometric distance on $\mathcal{G}r_{\text{n}\text{p}}$ as:

$$\text{d}_{\mathcal{G}r_{\text{n}\text{p}}}(\mathbf{Y}_o, P\mathbf{Y}_\tau) = \text{trace}\left(\text{acosm}^T(\mathbf{Y}_o^T P\mathbf{Y}_\tau) \cdot \text{acosm}(\mathbf{Y}_o^T P\mathbf{Y}_\tau)\right) \tag{18}$$

An interesting result can be obtained by using the trigonometric identity $2\text{cosm}^2(E) = I + \text{cosm}(2E)$. From it we get $4E^2 = \text{acosm}^2(2\text{cosm}^2(E) - I)$. Then, we can rewrite the optimization in Equation (8) as:

$$\begin{aligned} P_{\mathcal{FS}_{\text{n}\text{p}}} &= \arg\min_{D\in\mathcal{D}} \text{trace}(E_\tau^2) \\ &= \arg\min_{D\in\mathcal{D}} \text{trace}(4E_\tau^2) \\ &= \arg\min_{D\in\mathcal{D}} \text{trace}(\text{acosm}^2(2C_\tau^2 - I)) \end{aligned} \tag{19}$$

If we assume that minimizing the trace of the square of the matrix arc-cosine is about the same as maximizing the trace of its argument, we can write:

$$\begin{aligned} P_Q &= \arg\max_{D\in\mathcal{D}} \text{trace}(2C_\tau^2 - I) \\ &= \arg\max_{D\in\mathcal{D}} \text{trace}(C_\tau^T C_\tau) \\ &= \arg\max_{D\in\mathcal{D}} \text{trace}((V_o^T\mathbf{Y}_o^T P\mathbf{Y}_\tau V_\tau)^T(V_o^T\mathbf{Y}_o^T P\mathbf{Y}_\tau V_\tau)) \\ &= \arg\max_{D\in\mathcal{D}} \text{trace}(P^T\mathbf{Y}_o\mathbf{Y}_o^T P\mathbf{Y}_\tau\mathbf{Y}_\tau^T) \end{aligned} \tag{20}$$

In Figure 10 we show both cost functions for the $p = 2$ case. On the left-hand side, we provide a contour plot of the $f(X) = \text{trace}(\text{acosm}^2(X))$, and on the right-hand side we provide a contour

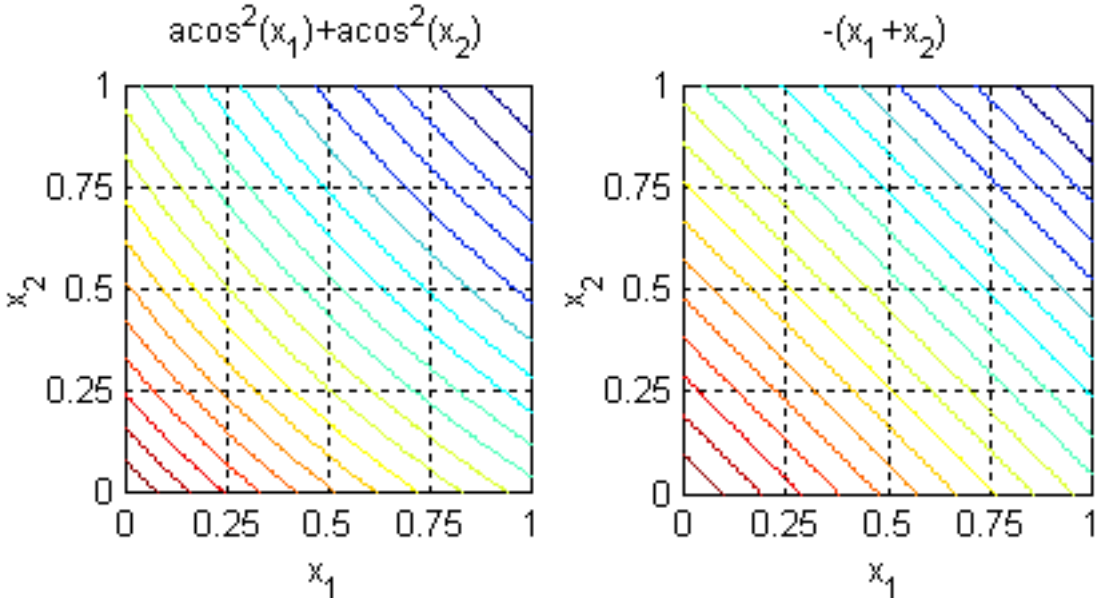

Figure 10: Approximating $\arg\min\Big(\text{trace}(\text{acosm}^2(X))\Big) \approx \arg\max\Big(\text{trace}(X)\Big)$ for diagonal X

plot of the $g(X) = -\text{trace}(X)$. Except for a small curvature, $f(X)$ and $g(X)$ behave about the same. So, the assumption we made seems to be a good approximation. Then, let us define the quadratic cost function to be:

$$J_Q(D) = -\text{trace}(D^T \mathbf{Y}_o \mathbf{Y}_o^T D \mathbf{Y}_\tau \mathbf{Y}_\tau^T) \tag{21}$$

We can easily compute its derivative as well as convert it into a canonical form, the way we did it in the case of linear programming. We do not present the computations here. We will only comment that it is much slower to minimize than the linear cost function, while providing about the same level of quality.

4.1.3 Derivative of the geometric cost function

The derivative of the true distance function Equation (18) is not easily computable. This is due to the fact that all three matrices in the SVD decomposition depend on the permutation matrix P. We were able to take such a derivative only under the assumption that a diagonal matrix E_τ commutes with any orthogonal matrix, which is only the case when both principal angles are the same. Under this assumption, the derivative is given by:

$$\nabla J_{Q,\mathcal{FS}_{\text{np}}}(D) = -2\mathbf{Y}_o \left(2E_\tau \text{sinm}^{-1}(2E_\tau)\right)^{\frac{1}{2}} \mathbf{Y}_o^T D \mathbf{Y}_\tau \left(2E_\tau \text{sinm}^{-1}(2E_\tau)\right)^{\frac{1}{2}} \mathbf{Y}_\tau^T \tag{22}$$

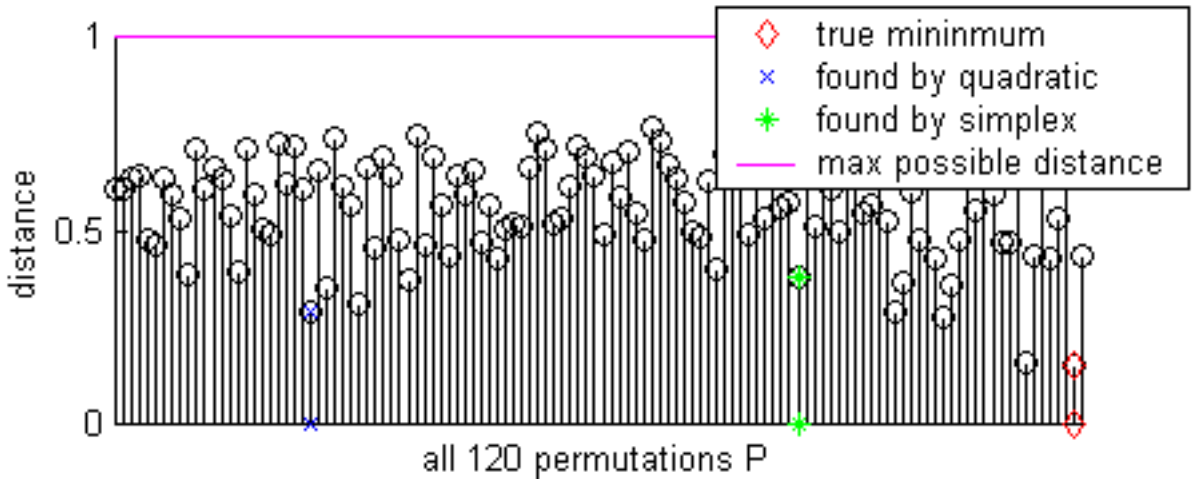

Figure 11: Exhaustive search to verify validity of linear and quadratic cost functions

We are also investigating if we can guess the derivative in the linear case by saying that it is well approximated by:

$$\nabla J_{L,\mathcal{FS}_{\mathrm{n,p}}}(D) = -\mathbf{Y}_o \left(2E_\tau \mathrm{sinm}^{-1}(2E_\tau)\right) \mathbf{Y}_\tau^T \tag{23}$$

This will be one of the questions to pursue in this research.

4.1.4 Experimental results

In Figure 11 we present an evaluation of a linear cost function from Equation (12), and a quadratic cost function from Equation (21). The evaluation is done via an exhaustive search over all possible permutations. We pick two random points $\mathbf{Y_1}$ and $\mathbf{Y_2}$ on $\mathcal{G}r_{5,2}$ manifold, but in such a way that both columns of shape $\mathbf{Y_1}$ are in ascending order, while the first column of the $\mathbf{Y_2}$ shape is in ascending order and the second column is in descending order. This is the worst case scenario for trying to bring these shapes into correspondence. Now, for every possible permutation P we compute the Grassmannian distance $\mathrm{d}_{\mathcal{G}r_{\mathrm{n,p}}}(\mathbf{Y_1}, P\mathbf{Y_2})$ and plot it in Figure 11. The maximal normalized distance we can get is 1, which is denoted by a horizontal line at level 1. For $n = 5$ there are total 120 permutations, shown on the horizontal axis. The permutation P_L found by linear programming via simplex achieves the normalized distance of $\mathrm{d}_{\mathcal{G}r_{\mathrm{n,p}}}(\mathbf{Y_1}, P_L\mathbf{Y_2}) = 0.38$. The permutation P_Q found by quadratic concave programming is different than the one found by simplex. It achieves a smaller distance of $\mathrm{d}_{\mathcal{G}r_{\mathrm{n,p}}}(\mathbf{Y_1}, P_Q\mathbf{Y_2}) = 0.29$. However, the permutation P_{True} that we find from Equation (8) achieves a much smaller distance $\mathrm{d}_{\mathcal{G}r_{\mathrm{n,p}}}(\mathbf{Y_1}, P_{\mathrm{True}}\mathbf{Y_2}) = 0.15$. This permutation was found via exhaustive search, which is not feasible for large n. However, we would like to note that

the simplex is relatively fast and does achieve adequate results for most shapes, as we could see from an example in Figures 8 and 9.

4.2 Permutation invariance via polynomials

Now, we are going to move to the second approach of defining a distance in the residual shape space $\mathcal{FS}_{\text{n,p}}$ by first achieving invariance to permutations, and secondly invariance to rotations. This approach is sketched on the right hand side of Figure 1.

The idea behind this approach is that we can achieve an invariance to permutations by going from the roots of a polynomial to the polynomial itself. It is a well known fact that an unordered set of n roots uniquely defines the n-th order polynomial, thus achieving invariance to permutations.

We remind the reader that we first start from the full shape space $\mathcal{X}$ and reduce it down to the normalized shape space $\mathcal{G}f_{\text{n,p}}$. Points $Y \in \mathcal{G}f_{\text{n,p}}$ are $n \times p$ tall-skinny orthogonal matrices, whose rows are all distinct, so $\mathcal{G}f_{\text{n,p}} \subset \mathbb{R}^{n \times p}$. In general, $\mathcal{G}f_{\text{n,p}}$ is not a vector space over some field in such a way that an action of a group of permutations is nicely preserved. For example, by going from matrices $\mathbb{R}^{n \times p}$ to vectors $\mathbb{R}^{n \cdot p}$ through the vec$(\cdot)$ notation and then allowing arbitrary permutations, we are loosing a partial correspondence of elements in the same rows of $\mathbf{Y}$. However, for the cases $p = 1$ and $p = 2$, $\mathcal{G}f_{\text{n,p}}$ can be viewed as a subset within a vector space over a field of real numbers $\mathbb{R}$ and a field of complex numbers $\mathbb{C}$ respectively: $\mathcal{G}f_{\text{n,1}} \subset \mathbb{R}^n$ and $\mathcal{G}f_{\text{n,2}} \subset \mathbb{C}^n$.

We are working on representing $\mathcal{G}f_{\text{n,p}}$ for an arbitrary p as a subset of a vector space over some field $\mathcal{G}f_{\text{n,p}} \subset \mathbb{C}_p^n$, which we call a field of generalized complex numbers and denote by $\mathbb{C}_p$. As part of this research, we will show the possibility of such representation along with all necessary proofs. But for now, we will consider only the case $p = 2$ and represent $\mathcal{G}f_{\text{n,2}}$ as a subset of $\mathbb{C}^n$. This is shown in the top-right block of Figure 1.

The map that takes us from $\mathbf{Y} \in \mathcal{G}f_{\text{n,2}}$ to $\mathbf{y} \in \mathbb{C}^n$ is defined to take each row i of the matrix $\mathbf{Y}$ and map it into a single complex number $y_i \in \mathbb{C}$. Then $\mathbf{y}$ is just an n-dim vector of y_i elements.

We define an $(n+1) \times n$ Vandermonde matrix $\mathbf{V}$ of $\mathbf{y}$ to be:

$$\mathbf{V} = \mathbf{V}(\mathbf{y}) = \begin{pmatrix} 1 & 1 & \dots & 1 \\ y_1 & y_2 & \dots & y_n \\ y_1^2 & y_2^2 & \dots & y_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ y_1^n & y_2^n & \dots & y_n^n \end{pmatrix} \tag{24}$$

This brings us down to the second block on the right-hand side of Figure 1. A permuted version $P\mathbf{Y}$ of a point $\mathbf{Y} \in \mathcal{G}f_{\mathrm{n},2}$ results in $\mathbf{y}$ getting permuted into $P\mathbf{y}$, which, in turn, leads us to the Vandermonde matrix $\mathbf{V}(\mathbf{y})P^T$ whose columns are permuted. Thus, if we consider the subspace spanned by the columns of the Vandermonde matrix, this subspace will be invariant to the permutations of the feature points. Let us denote the space of all $(n+1) \times n$ Vandermonde matrices over the complex field $\mathbb{C}$ to be $\mathcal{V}_{\mathrm{n+1,n}}(\mathbb{C})$. Since all the elements y_i of $\mathbf{y}$ are distinct, Vandermode matrices $\mathbf{V}(\mathbf{y})$ will be of rank n. Thus, we can think of $\mathcal{V}_{\mathrm{n+1,n}}(\mathbb{C}) \subset \mathcal{G}F_{\mathrm{n+1,n}}(\mathbb{C})$ to be a surface in the generalized Stiefel manifold over the field $\mathbb{C}$. If we act on the generalized Stiefel manifold by $\mathcal{G}L_{\mathrm{n}}(\mathbb{C})$, we obtain the Grassmann manifold $\mathcal{G}r_{\mathrm{n+1,n}}(\mathbb{C})$. It can be shown that the restriction of the action of the group $\mathcal{G}L_{\mathrm{n}}(\mathbb{C})$ to $\mathcal{V}_{\mathrm{n+1,n}}(\mathbb{C})$ will be a group action of $\mathcal{P}$. Then it follows that the space of Vandermonde matrices modulo permutations $\mathcal{V}_{\mathrm{n+1,n}}(\mathbb{C})/\mathcal{P} \subset \mathcal{G}r_{\mathrm{n+1,n}}(\mathbb{C})$ will be a subset of the Grassmann manifold over the field $\mathbb{C}$. This is shown in the third block of the right-hand side in Figure 1.

The points on the Grassmann manifold $\mathcal{G}r_{\mathrm{n+1,n}}(\mathbb{C})$ can be identified by $(n+1)\times(n+1)$ projection matrices of linearly independent n-frames in $\mathcal{G}F_{\mathrm{n+1,n}}(\mathbb{C})$. Thus, we can write:

$$\mathbf{\Pi} = \lceil \mathbf{V} \rfloor = \mathbf{V}(\mathbf{V}^H\mathbf{V})^{-1}\mathbf{V}^H \tag{25}$$

Now we are ready to define a distance in the space $\mathcal{G}r_{\mathrm{n+1,n}}(\mathbb{C})$. There are many possible choices [26, 27, 21]. Since we are working with projectors, we will pick the projection F-norm, which is defined to be:

$$\mathrm{d}_{\mathcal{G}r_{\mathrm{n+1,n}}(\mathbb{C})}(\mathbf{\Pi_o}, \mathbf{\Pi}_\tau) = \sqrt{2}\|\mathbf{\Pi_o} - \mathbf{\Pi}_\tau\|_F \tag{26}$$

We are claiming that the distance defined in this way will achieve invariance to permutations of the feature points in the normalized shapes, but will vary for any other affine distortions acting on the left. In other words, we are claiming that the distance will be zero if and only if two normalized shapes are related by the permutation of their rows: $\mathrm{d}_{\mathcal{G}f_{\mathrm{n},2}}(\mathbf{Y}_o, \mathbf{Y}_\tau) = \mathrm{d}_{\mathcal{G}r_{\mathrm{n+1,n}}(\mathbb{C})}\Big(\mathbf{\Pi}(\mathbf{y}_o), \mathbf{\Pi}(\mathbf{y}_\tau)\Big) = 0$ iff $\mathbf{Y}_o = P\mathbf{Y}_\tau$ where $P \in \mathcal{P}$ is a permutation.

So far we have achieved an invariance to permutations only. Now we are going to take $\mathcal{G}r_{\mathrm{n+1,n}}(\mathbb{C})$ and remove the ambiguity of rotations/reflections of normalized shapes on 2-planes in $\mathcal{R}^n$. If $\mathbf{Y} \in \mathcal{G}f_{\mathrm{n},2}$ is a normalized shape, then all its rotated/reflected shapes will be given by:

$$\lceil \mathbf{Y} \rfloor_{\mathcal{O}_2} = \{\mathbf{Y}Q : Q \in \mathcal{O}_2\} \tag{27}$$

By going to complex shape space this formula becomes:

$$\lceil \mathbf{y} \rfloor_{\mathcal{O}_2} = \{\mathbf{y}e^{j\theta}, \tilde{\mathbf{y}}e^{j\theta} : \theta \in [0 \ldots 2\pi]\} \tag{28}$$

Now we are ready to define a distance in the residual shape space $\mathcal{FS}_{\mathrm{n},2}(\mathbb{C})$ by minimizing the distance of the $\mathcal{G}r_{\mathrm{n+1,n}}(\mathbb{C})$ space over a group of rotations:

$$\mathrm{d}_{\mathcal{FS}_{\mathrm{n},2}(\mathbb{C})}\left(\mathbf{Y}_o, \mathbf{Y}_\tau\right) \equiv \min_{Q\in\mathcal{O}_2} \mathrm{d}_{\mathcal{G}f_{\mathrm{n},2}}\left(\mathbf{Y}_o, \mathbf{Y}_\tau Q\right) = \min_{\mathbf{y}\in\lceil\mathbf{y}_\tau\rfloor_{\mathcal{O}_2}} \mathrm{d}_{\mathcal{G}r_{\mathrm{n+1,n}}(\mathbb{C})}\Big(\mathbf{\Pi}(\mathbf{y}_o), \mathbf{\Pi}(y)\Big) \tag{29}$$

This would be the last block on the right-hand side of Figure 1. We are claiming that this is a valid distance, satisfying all the necessary properties. It is also a continuous distance. The last aspect we would like to study is the computation of this distance. There are two steps in computing this distance. The first step is a computation of the Vandermonde matrix and its projector. We need to look for robust numerical methods for its computation. It could be possible to compute it using the block-recursive QR decomposition. Also, we need to investigate if we effectively can calculate the $(n+1)\times 1$ orthonormal vector that is tangent to the span of the Vandermonde matrix. This would be a more compact representation of the $\mathcal{G}r_{\mathrm{n+1,n}}(\mathbb{C})$ by its complementary space $\mathcal{G}r_{\mathrm{n+1,1}}(\mathbb{C})$. Also, the Vandermonde matrix is very closely related to the coefficients of the polynomial whose n roots are the elements of the second row in the Vandermonde matrix. We could represent every point in $\mathcal{G}r_{\mathrm{n+1,n}}(\mathbb{C})$ by one such polynomial. The problem of going from the roots of a polynomial to its coefficients, and the inverse problem of finding roots of a polynomial are well studied. But there are issues with the numerical stability and additive errors that get magnified when we are dealing with a large value of n.

The second step is a minimization over the equivalence class given by Equation (28). We note that when we take numbers of the form $\mathbf{y}e^{j\theta}$ into some power, we obtain discrete Fourier exponents. As part of this research, we will study how, given the computed Vandermonde matrix, we can minimize the distance over θ by representing it as a discrete Fourier transform computation and a minimum search over the frequency coefficients.

5 Proposed research

In this section we propose the directions in which to extend and improve our approach.

5.1 What will be accomplished in this research

- **Distance in $\mathcal{FS}_{\mathbf{n,p}}(\mathbb{R})$ as an optimization over $\mathcal{P}$**

  In Sub-Section 4.1 we presented two distances that can be solved via linear programming and quadratic concave programming. They are easily computable, however, they do not

work well when a shape and its permuted versions are too far from each other in terms of their Grassmannian distance. In this research, we will look for conditions under which these approximations hold, and try to derive a better approximation to the true distance. Also, we will attempt to study the stability of the distances defined as an optimization over a group of permutations $\mathcal{P}$.

- **Distance in $\mathcal{FS}_{\mathbf{n,p}}(\mathbb{C}_p)$ as an optimization over $\mathcal{O}_{\mathbf{p}}$**

  Sub-Section 4.2 is only a sketch of an idea of going to Vandermonde matrices over a complex field in order to achieve invariance to permutations, and then searching for an optimal $Q \in \mathcal{O}_{\mathrm{p}}$ by means of a Fourier transform. This idea will need to be fully completed, as part of this research. It will include generalizing this approach to any $p > 1$, not just $p = 2$. Also, it involves looking for a robust QR decomposition algorithm that does not require a computation of the Vandermonde matrices. Once the QR decomposition is calculated, we will also try to represent minimization over $e^{j\theta}$ as the computation of a Fourier transform.

- **Geometry of the residual shape space $\mathcal{FS}_{\mathbf{n,p}}$**

  Finally, we will attempt to equip the shape space $\mathcal{FS}_{\mathrm{n,p}}$ with topological and geometric structure. One way of defining the geometry in $\mathcal{FS}_{\mathrm{n,p}}$ would be to require that the folded space $\mathcal{FS}_{\mathrm{n,p}} = \mathcal{G}r_{\mathrm{n,p}}/\mathcal{P}$ be *convex* in $\mathcal{G}r_{\mathrm{n,p}}$ and use the geometry of $\mathcal{G}r_{\mathrm{n,p}}$ on it. We have results on such folding for the $p = 1$ case. For $p > 1$ the folding is still under the investigation, and it is an open research question. Our initial impression is that such folding is not possible. We will either find a folding, or proof it is not possible. Also, we are going to look for alternative ways of defining a geometry in the space of residual shapes $\mathcal{FS}_{\mathrm{n,p}}$. This goal might not be achievable, in which case we will attempt to prove it being impossible.

5.2 Other interesting questions we might pursue

Once the geometry in the $\mathcal{FS}_{\mathrm{n,p}}$ is established, an interesting question to pursue is to bring probability distributions into the shape spaces and study robustness to noise by coming up with an analytic Cramér-Rao type of lower bounds. Another interesting question is to try to parameterize residual shapes on $\mathcal{FS}_{\mathrm{n,p}}$ manifold in such a way that a set of parametric residual shapes be represented as a sub-manifold of the $\mathcal{FS}_{\mathrm{n,p}}$ manifold.

5.3 Timeline

- Spring 2003

  1. Study equation $P = e^K$ on $\mathcal{O}_{\mathrm{n}}$ manifold in terms of K where $P \in \mathcal{P}$ is a permutation.
  2. Study properties of the mid-point in $\mathcal{G}r_{\mathrm{n,p}}$ between the shape $\mathbf{I}_{n,p}$ and its permutation $P\mathbf{I}_{n,p}$.
  3. Find out under what conditions linear and quadratic programming for minimizing distance over permutation fails.
  4. Try to address the issue of robust QR decomposition of the Vandermonde matrix for large n.

- Summer 2003

  1. Study in more detail the derivative of the geometric cost function and try to develop a good approximation.
  2. Compute the minimization over $e^{j\theta}$ as an FFT computation.
  3. Generalize the distance computation via the Vandermonde approach for any p.

- Fall 2003

  1. Try to address the issue of folding $\mathcal{G}r_{\mathrm{n,p}}/\mathcal{P}$ to make it a smooth manifold with nice behavior or prove that such folding is not possible.
  2. Study resiliance to noise in different shape spaces

- Sprint 2004

  1. Apply the theory to a real-word image processing application.
  2. Write the dissertation and defend the thesis.

- Future work

  1. Bring into the picture a-priori probabilities in the shape spaces, parametric statistics, and everything else that comes with it. Try to derive Cramér-Rao type of lower bounds.
  2. Study different shape spaces and their topological and geometric properties under different types of modeled transformations.

3. Investigate multi-resolution and shape resampling issues, i.e, changing the number of feature points in a way that does not change much its geometric properties and distances to other resampled shapes.

4. Study different distances and determine which one is the best to measure similarity of real world shapes.

A Derivation of formulas for the Grassmann manifold

A.1 Representing Grassmann manifold $\mathcal{G}r_{\mathbf{n,p}}$ as a quotient manifold $\mathcal{O}_\mathbf{n}/\mathcal{O}_\mathbf{p}\times\mathcal{O}_\mathbf{n\text{-}p}$

Here we will derive formulas for the Grassmann manifold $\mathcal{G}r_{\mathrm{n,p}}$. Some of these derivations are well known, and are given in [21, 15, 16, 19, 20]. But some of these derivations we have not seen anywhere else. For completeness, we will present all of them.

Let us consider an orthogonal group $\mathcal{O}_\mathrm{n} = \{\mathbf{Q} : \mathbf{Q}^T\mathbf{Q} = \mathbf{I}\}$. Suppose $\mathcal{O}_\mathrm{n}$ is naturally embedded in $\mathcal{R}^{n\times n}$ with a canonical Levi-Civita connection. Suppose $\gamma(t) = \mathbf{Q}_t$ is a smooth curve in $\mathcal{O}_\mathrm{n}$. By differentiating the orthogonality condition $\mathbf{Q}_t^T\mathbf{Q}_t = \mathbf{I}$ with respect to t twice we get:

$$\mathbf{Q}_t^T\dot{\mathbf{Q}}_t + \dot{\mathbf{Q}}_t^T\mathbf{Q}_t = 0 \tag{30}$$

$$\mathbf{Q}_t^T\ddot{\mathbf{Q}}_t + 2\dot{\mathbf{Q}}_t^T\dot{\mathbf{Q}}_t + \ddot{\mathbf{Q}}_t^T\mathbf{Q}_t = 0 \tag{31}$$

From Equation (30) the set of all derivatives $\mathbf{\Delta} = \dot{\mathbf{Q}}$ forms a tangent space at point $\mathbf{Q} \in \mathcal{O}_\mathrm{n}$:

$$\mathcal{T}_\mathbf{Q}(\mathcal{O}_\mathrm{n}) = \{\mathbf{\Delta} : \mathbf{Q}^T\mathbf{\Delta} + \mathbf{\Delta}^T\mathbf{Q} = 0\} \tag{32}$$

The inner product in the tangent spaces is the cannonical inner product:

$$g(\mathbf{\Delta}_1, \mathbf{\Delta}_2) = \mathrm{tr}\mathbf{\Delta}_1^T\mathbf{\Delta}_2 \tag{33}$$

We define the matrix $K = \mathbf{Q}^T\mathbf{\Delta}$. Then $\mathbf{\Delta} = \mathbf{Q}K$. By substituting $\mathbf{\Delta}$ into the Equation (32), we get $\mathbf{Q}^T(\mathbf{Q}K) + (K^T\mathbf{Q}^T)\mathbf{Q} = K + K^T = 0$, so matrix K must be skew-symmetric. Then we can write the tangent space and its orthogonal normal space as:

$$\mathcal{T}_\mathbf{Q}(\mathcal{O}_\mathrm{n}) = \{\mathbf{\Delta} : \mathbf{\Delta} = \mathbf{Q}K\ ,\ K^T + K = 0\} \tag{34}$$

$$\mathcal{N}_\mathbf{Q}(\mathcal{O}_\mathrm{n}) = \{\mathbf{\Delta} : \mathbf{\Delta} = \mathbf{Q}S\ ,\ S^T = S\} \tag{35}$$

An equation of the geodesic $\gamma(t) = \mathbf{Q}_t$ is a second order differential equation for which the second derivative lies in the normal space. If we write $\ddot{\mathbf{Q}}_t = \mathbf{Q}_tS \in \mathcal{N}_{\mathbf{Q_t}}(\mathcal{O}_\mathrm{n})$ for some symmetric matrix S and plug it into Equation (31) we obtain:

$$S + 2\dot{\mathbf{Q}}_t^T \dot{\mathbf{Q}}_t + S^T = 0 \tag{36}$$

It follows $S = -\dot{\mathbf{Q}}_t^T \dot{\mathbf{Q}}_t$ and $\ddot{\mathbf{Q}}_t = \mathbf{Q}_t S = -\mathbf{Q}_t(\dot{\mathbf{Q}}_t^T \dot{\mathbf{Q}}_t)$. Then the differential equation of the geodesic $\gamma(t)$ is given by:

$$\ddot{\mathbf{Q}}_t + \mathbf{Q}_t \left(\dot{\mathbf{Q}}_t^T \dot{\mathbf{Q}}_t \right) = 0 \tag{37}$$

Another important equation is the equation of the parallel vector field of the tangent vector $\gamma'(0)$ along the curve $\gamma(t)$. But since $\gamma(t)$ is a geodesic, the parallel vector field of the $\gamma'(0)$ will coincide with the geodesic's derivative vector field $\gamma'(t)$, and will be given by $\gamma'(t) = \mathbf{Q}_t \mathbf{Q}_o^T \dot{\mathbf{Q}}_o$. Then it is easily verified that the solution to the initial value problem in Equation (37) for the geodesic and its derivative vector field for $\mathcal{O}_{\mathrm{n}}$ manifold is given by:

$$\begin{aligned} \gamma(t) &= \mathbf{Q} \cdot e^{Kt} & \gamma(0) &= \mathbf{Q} \\ \gamma'(t) &= \mathbf{Q} \cdot e^{Kt} \cdot K & \gamma'(0) &= \mathbf{Q} \cdot K \end{aligned} \tag{38}$$

where K is an arbitrary skew-symmetric matrix. We can also obtain the dimensionality of the manifold $\mathcal{O}_{\mathrm{n}}$ from the dimensionality of $\mathcal{T}_{\mathbf{Q}}(\mathcal{O}_{\mathrm{n}})$. The dimensionality is the number of degrees of freedom of skew-symmetric matrices:

$$\dim(\mathcal{O}_{\mathrm{n}}) = \frac{n(n-1)}{2} \tag{39}$$

Let us fix $p < n$ and decompose the $n \times n$ matrix $\mathbf{Q} \in \mathcal{O}_{\mathrm{n}}$, its tangent $\mathbf{\Delta} \in \mathcal{T}_{\mathbf{Q}}(\mathcal{O}_{\mathrm{n}})$, and the skew-symmetric matrix K into block matrices as follows:

$$\begin{aligned} \underset{n\times n}{\mathbf{Q}} &= \begin{bmatrix} \underset{n\times p}{\mathbf{Y}} & \underset{n\times n-p}{\mathbf{Y}_\perp} \end{bmatrix} \\ \underset{n\times n}{\mathbf{\Delta}} &= \begin{bmatrix} \underset{n\times p}{\mathbf{H}} & \underset{n\times n-p}{\mathbf{H}_\perp} \end{bmatrix} \\ \underset{n\times n}{K} &= \begin{bmatrix} \underset{p\times p}{F} & \underset{p\times n-p}{-R^T} \\ \underset{n-p\times p}{R} & \underset{n-p\times n-p}{G} \end{bmatrix} \end{aligned}$$

where F and G are skew-symmetric and R is an arbitrary matrix. Now we can rewrite a general form of a tangent $\mathbf{\Delta} = \mathbf{Q}K$ as:

$$\mathbf{\Delta} = \left[\begin{array}{c|c} \mathbf{H} & \mathbf{H}_\perp \end{array} \right] = \left[\begin{array}{c|c} \mathbf{Y}F + \mathbf{Y}_\perp R & -\mathbf{Y}R^T + \mathbf{Y}_\perp G \end{array} \right] \tag{40}$$

We define a point $\lceil \mathbf{Q} \rfloor$ on the Grassmann manifold $\mathcal{G}r_{\mathrm{n,p}} \equiv \mathcal{O}_{\mathrm{n}}/\mathcal{O}_{\mathrm{p}} \times \mathcal{O}_{\mathrm{n\text{-}p}}$ as an orbit

$$\lceil \mathbf{Q} \rfloor = \left\{ \mathbf{Q} \cdot \begin{bmatrix} V_p & 0 \\ 0 & V_{n-p} \end{bmatrix} : \begin{array}{c} V_p \in \mathcal{O}_{\mathrm{p}} \\ V_{n-p} \in \mathcal{O}_{\mathrm{n\text{-}p}} \end{array} \right\} = \mathbf{Q} \cdot \begin{bmatrix} \mathcal{O}_{\mathrm{p}} & 0 \\ 0 & \mathcal{O}_{\mathrm{n\text{-}p}} \end{bmatrix} \tag{41}$$

The dimensionality of the manifold $\mathcal{G}r_{\mathrm{n}\mathrm{p}} = \mathcal{O}_{\mathrm{n}}/\mathcal{O}_{\mathrm{p}} \times \mathcal{O}_{\mathrm{n\text{-}p}}$ is given by:

$$\dim(\mathcal{G}r_{\mathrm{n}\mathrm{p}}) = \dim(\mathcal{O}_{\mathrm{n}}) - \dim(\mathcal{O}_{\mathrm{p}}) - \dim(\mathcal{O}_{\mathrm{n\text{-}p}}) = p(n-p) \tag{42}$$

Rewriting Equation (41) in the block-matrix form, we get:

$$\left[\begin{array}{c|c} \lceil \mathbf{Y} \rfloor & \lceil \mathbf{Y}_{\perp} \rfloor \end{array}\right] = \left\{ \left[\begin{array}{c|c} \mathbf{Y} \cdot V_p & \mathbf{Y}_{\perp} \cdot V_{n-p} \end{array}\right] : \begin{array}{c} V_p \in \mathcal{O}_{\mathrm{p}} \\ V_{n-p} \in \mathcal{O}_{\mathrm{n\text{-}p}} \end{array} \right\} = \left[\begin{array}{c|c} \mathbf{Y} \cdot \mathcal{O}_{\mathrm{p}} & \mathbf{Y}_{\perp} \cdot \mathcal{O}_{\mathrm{n\text{-}p}} \end{array}\right] \tag{43}$$

From this expression, we can see that a point on the Grassmann manifold can be viewed as a split of $\mathcal{R}^{n \times n}$ into two orthogonal and complementary to each other subspaces spanned by $\mathbf{Y}$ and $\mathbf{Y}_{\perp}$. We can easily recover a complementary subspace spanned by $\mathbf{Y}_{\perp}$ from the subspace spanned by $\mathbf{Y}$, so, abusing notation, we can write that $\lceil \mathbf{Y} \rfloor = \lceil \mathbf{Q} \rfloor \in \mathcal{G}r_{\mathrm{n}\mathrm{p}}$. Also, without loss of generality, we assume that $p \leq (n-p)$ since $\mathcal{G}r_{\mathrm{n}\mathrm{p}} \cong \mathcal{G}r_{\mathrm{n}\mathrm{n\text{-}p}}$, as it will be shown later.

The tangent space $\mathcal{T}_{\mathbf{Q}}(\mathcal{O}_{\mathrm{n}})$ at $\mathbf{Q} \in \mathcal{O}_{\mathrm{n}}$ will also get split into two components, the vertical component $\mathcal{V}_{\mathbf{Q}}(\mathcal{O}_{\mathrm{n}})$ and the horizontal component $\mathcal{H}_{\mathbf{Q}}(\mathcal{O}_{\mathrm{n}})$. The vertical component is defined such that by taking any vertical tangent and making a move along the geodesic in $\mathcal{O}_{\mathrm{n}}$ produces no move in the quotient space $\mathcal{G}r_{\mathrm{n}\mathrm{p}} = \mathcal{O}_{\mathrm{n}}/\mathcal{O}_{\mathrm{p}} \times \mathcal{O}_{\mathrm{n\text{-}p}}$. From Equation (40) it is easy to see that:

$$\boldsymbol{\Delta}_V = \left[\begin{array}{c|c} \mathbf{Y}F & \mathbf{Y}_{\perp}G \end{array}\right] = \mathbf{Q} \cdot \begin{bmatrix} F & 0 \\ 0 & G \end{bmatrix} \tag{44}$$

$$\boldsymbol{\Delta}_H = \left[\begin{array}{c|c} \mathbf{Y}_{\perp}R & -\mathbf{Y}R^T \end{array}\right] = \mathbf{Q} \cdot \begin{bmatrix} 0 & -R^T \\ R & 0 \end{bmatrix} \tag{45}$$

The fact that a move from point $\mathbf{Q} = \mathbf{Q}_o \in \lceil \mathbf{Q} \rfloor$ along the geodesic in $\mathcal{O}_{\mathrm{n}}$ in the direction of a vertical tangent will move us to point $\mathbf{Q}_\tau \in \lceil \mathbf{Q} \rfloor$, which is still the same point in the quotient space $\lceil \mathbf{Q} \rfloor \in \mathcal{G}r_{\mathrm{n}\mathrm{p}}$, can be verified by plugging $\boldsymbol{\Delta}_V$ from Equation (45) into Equation (38):

$$\mathbf{Q}_\tau = \mathbf{Q}_o \cdot \exp\left(\begin{bmatrix} F & 0 \\ 0 & G \end{bmatrix} t \right) = \mathbf{Q}_o \cdot \begin{bmatrix} e^{Ft} & 0 \\ 0 & e^{Gt} \end{bmatrix} = \mathbf{Q}_o \cdot \begin{bmatrix} V_p & 0 \\ 0 & V_{n-p} \end{bmatrix} \in \lceil \mathbf{Q} \rfloor \tag{46}$$

The tangent space of a point on the Grassmann manifold is in bijective correspondence with the horizontal component $\mathcal{T}_{\lceil \mathbf{Y} \rfloor}(\mathcal{G}r_{\mathrm{n}\mathrm{p}}) \cong \mathcal{T}_{\lceil \mathbf{Q} \rfloor}(\mathcal{G}r_{\mathrm{n}\mathrm{p}}) \cong \mathcal{H}_{\mathbf{Q}}(\mathcal{O}_{\mathrm{n}})$ under the natural projection. Thus, tangent spaces for the $\mathcal{G}r_{\mathrm{n}\mathrm{p}}$ are of the form:

$$\mathcal{T}_{\lceil \mathbf{Q} \rfloor}(\mathcal{G}r_{\mathrm{n}\mathrm{p}}) = \left\{ \lceil \boldsymbol{\Delta} \rfloor : \lceil \boldsymbol{\Delta} \rfloor = \lceil \mathbf{Q} \rfloor \cdot \begin{bmatrix} 0 & -R^T \\ R & 0 \end{bmatrix} \right\} \tag{47}$$

where R is an arbitrary $(n-p)\times p$ matrix. Equation (47) is the general form of the tangent space. In particular, when we study the Grassmann manifold by working only with the subspace spanned by the first p-columns given by $\mathbf{Y}$, we can identify the tangent space by:

$$\mathcal{T}_{\lceil \mathbf{Y} \rfloor}(\mathcal{G}r_{\mathfrak{n},\mathfrak{p}}) = \left\{ \lceil \mathbf{H} \rfloor : \lceil \mathbf{H} \rfloor = \lceil \mathbf{Y}_\perp \rfloor \cdot R \right\} \tag{48}$$

Equation (38) of the geodesic and of its derivative vector field in $\mathcal{O}_\mathrm{n}$ will also be the equation of the geodesic and of its derivative vector field in $\mathcal{G}r_{\mathfrak{n},\mathfrak{p}}$ where we take tangents of the form given by Equation (45). But these equations can be written in more computationally efficient form, when we are dealing with tangents of the form in Equation (48).

A.2 Solution for the initial value equation of the geodesic in $\mathcal{G}r_{\mathfrak{n},\mathfrak{p}}$

Let $p \leq (n-p)$ and let us decompose R using an SVD decomposition as follows:

$$\underset{n-p\times p}{R} = \begin{bmatrix} \underset{n-p\times p}{U} & \underset{n-p\times n-2p}{W} \end{bmatrix} \cdot \begin{bmatrix} \underset{p\times p}{E} \\ \underset{n-2p\times p}{0} \end{bmatrix} \cdot \underset{p\times p}{V}^T = U \cdot E \cdot V^T \tag{49}$$

Given this SVD decomposition, it is easily verified that:

$$K = \begin{bmatrix} 0 & -R^T \\ R & 0 \end{bmatrix} = \left[\begin{array}{c|cc} V & 0 & 0 \\ \hline 0 & U & W \end{array}\right] \cdot \left[\begin{array}{c|cc} 0 & -E & 0 \\ \hline E & 0 & 0 \\ 0 & 0 & 0 \end{array}\right] \cdot \left[\begin{array}{c|cc} V & 0 & 0 \\ \hline 0 & U & W \end{array}\right]^T \tag{50}$$

Using the fact that $e^{AXA^{-1}} = Ae^XA^{-1}$, we can write:

$$e^{Kt} = \left[\begin{array}{c|cc} V & 0 & 0 \\ \hline 0 & U & W \end{array}\right] \cdot \left[\begin{array}{c|cc} C_t & -S_t & 0 \\ \hline S_t & C_t & 0 \\ 0 & 0 & I \end{array}\right] \cdot \left[\begin{array}{c|cc} V & 0 & 0 \\ \hline 0 & U & W \end{array}\right]^T \tag{51}$$

where $C_t = \mathrm{cosm}(Et)$ and $S_t = \mathrm{sinm}(Et)$ are matrix-cosine and matrix-sine. Then:

$$\mathbf{Q} \cdot e^{Kt} = \left[\begin{array}{c|c} \mathbf{Y}VC_tV^T + \mathbf{Y}_\perp US_tV^T & -\mathbf{Y}VS_tU^T + \mathbf{Y}_\perp UC_tU^T + \mathbf{Y}_\perp WW^T \end{array}\right] \tag{52}$$

But $\mathbf{H} = \mathbf{Y}_\perp R = \mathbf{Y}_\perp UEV^T$, so we can write $\mathbf{Y}_\perp U = \mathbf{H}VE^{-1}$. Also, we note that $U^TW = 0$. Then Equation (52) becomes:

$$\mathbf{Q} \cdot e^{Kt} = \left[\begin{array}{c|c} \mathbf{Y}VC_tV^T + \mathbf{H}VE^{-1}S_tV^T & -\mathbf{Y}VS_tU^T + \mathbf{H}VE^{-1}C_tU^T \end{array}\right] \tag{53}$$

Post-multiplying by K and substituting $U^T R = EV^T$ we get:

$$\mathbf{Q} \cdot e^{Kt} \cdot K = \left[\; -\mathbf{Y}VS_tEV^T + \mathbf{H}VE^{-1}C_tEV^T \;\middle|\; -\mathbf{Y}VC_tV^TR^T - \mathbf{H}VE^{-1}S_tV^TR^T \; \right] \tag{54}$$

Finally, from Equations (38), (53), (54) we obtain the equations for the geodesic and its derivative vector field by constraining ourself to working just with the subspace spanned by the first p-columns given by $\mathbf{Y}$, and by going from $\mathbf{Y}$ to its orbit $\lceil \mathbf{Y} \rfloor$:

$$\begin{aligned} \gamma(t) &= \lceil \mathbf{Y} \rfloor \cdot VC_tV^T + \lceil \mathbf{H} \rfloor \cdot VS_tE^{-1}V^T \qquad \gamma(0) = \lceil \mathbf{Y} \rfloor \\ \gamma'(t) &= -\lceil \mathbf{Y} \rfloor \cdot VS_tEV^T + \lceil \mathbf{H} \rfloor \cdot VC_tV^T \qquad \gamma'(0) = \lceil \mathbf{H} \rfloor \end{aligned} \tag{55}$$

The distance on the $\mathcal{G}r_{\mathrm{n,p}}$ manifold along the geodesic $\gamma(t)$ is computed from Equation (55) as follows:

$$\mathrm{d}\Big(\gamma(t_1), \gamma(t_2)\Big) = \sqrt{\mathrm{tr}(\lceil \mathbf{H} \rfloor^T \lceil \mathbf{H} \rfloor)} \cdot |t_1 - t_2| = \sqrt{\mathrm{tr}(R^T R)} \cdot |t_1 - t_2| = \sqrt{\mathrm{tr}(E^2)} \cdot |t_1 - t_2| \tag{56}$$

A.3 Solution for the 2-point boundary value equation of the geodesic in $\mathcal{G}r_{\mathbf{n,p}}$

Equations of the geodesic, of its derivative vector field, and of the distance, given by Equations (55) and (56), are defined by the initial point $\lceil \mathbf{Y} \rfloor$ on the $\mathcal{G}r_{\mathrm{n,p}}$ manifold, and a tangent direction $\lceil \mathbf{H} \rfloor$ at the point. Now, we are going to derive the equations for the same quantities, but given an initial point $\lceil \mathbf{Y_o} \rfloor$ and a final point $\lceil \mathbf{Y}_\tau \rfloor$, which are to be connected by the geodesic.

Let $\gamma(t)$ be the geodesic given by the initial and final values $\gamma(0) = \lceil \mathbf{Y_o} \rfloor$ and $\gamma(\tau) = \lceil \mathbf{Y}_\tau \rfloor$. Let the distance between the initial and final values be $\mathrm{d}\Big(\gamma(0), \gamma(\tau)\Big) = \tau$ where the value of τ to be determined later. From the Equation (38), we can write:

$$\left[\; \mathbf{Y}_\tau \;\middle|\; \mathbf{Y}_{\tau\perp} \; \right] = \left[\; \mathbf{Y_o} \;\middle|\; \mathbf{Y_{o\perp}} \; \right] \cdot \exp\left(\begin{bmatrix} F & 0 \\ 0 & G \end{bmatrix} \tau \right) \cdot \exp\left(\begin{bmatrix} 0 & -R^T \\ R & 0 \end{bmatrix} \tau \right) \tag{57}$$

But, as it was shown in Equation (46), vertical movements only result in two independent rotations in the subspace of $\mathbf{Y}_\tau$ and in the subspace of $\mathbf{Y}_{\tau\perp}$. Thus, we can write:

$$\begin{bmatrix} V_p^T & 0 \\ 0 & V_{n-p}^T \end{bmatrix} \cdot \begin{bmatrix} \mathbf{Y_o}^T\mathbf{Y}_\tau & \mathbf{Y_o}^T\mathbf{Y}_{\tau\perp} \\ \mathbf{Y_{o\perp}}^T\mathbf{Y}_\tau & \mathbf{Y_{o\perp}}^T\mathbf{Y}_{\tau\perp} \end{bmatrix} = \exp\left(\begin{bmatrix} 0 & -R^T \\ R & 0 \end{bmatrix} \tau \right) = e^{K\tau} \tag{58}$$

Since $e^{K\tau}$ is symmetric, and its eigenvalue decomposition is given by Equation (51), we can write:

$$\left[\begin{array}{c|c} V_o^T {\mathbf{Y_o}}^T \mathbf{Y}_\tau V_\tau & \cdots \\ \hline \cdots & V_{o\perp}^T {\mathbf{Y_o}_\perp}^T \mathbf{Y}_{\tau\perp} V_{\tau\perp} \end{array}\right] = \left[\begin{array}{c|cc} C_\tau & -S_\tau & 0 \\ \hline S_\tau & C_\tau & 0 \\ 0 & 0 & I \end{array}\right] \tag{59}$$

where V_o, V_τ, $V_{o\perp}$, and $V_{\tau\perp}$ are combined orthogonal rotations from Equations (58) and (51). We recall that $C_\tau = \mathrm{cosm}(E\tau)$, so in order to recover $E_\tau = E\tau$ all we need to do is to compute an SVD decomposition of the inner product:

$$\mathbf{Y_o}^T \mathbf{Y}_\tau = V_o C_\tau V_\tau^T \tag{60}$$

then, compute $E_\tau = \mathrm{acosm}(C_\tau)$, and finally, compute $\tau = \|E_\tau^2\|_F = \sqrt{\mathrm{tr}(E_\tau^2)}$ and $E = E_\tau/\tau$. Also, we can see from Equation (59) that the distance $\mathrm{d}\Big(\lceil \mathbf{Y_o} \rfloor, \lceil \mathbf{Y}_\tau \rfloor\Big) = \mathrm{d}\Big(\lceil \mathbf{Y_o}_\perp \rfloor, \lceil \mathbf{Y}_{\tau\perp} \rfloor\Big)$ since arccosine of I will be 0. This shows that $\mathcal{G}r_{\mathrm{n,p}} \cong \mathcal{G}r_{\mathrm{n,n\text{-}p}}$. Also, it validates our assumption $p \leq (n-p)$.

Now, we will derive equations for the mid-point $\lceil \mathbf{Y_m} \rfloor = \gamma(\frac{\tau}{2})$ of the geodesic and the tangent vector $\lceil \mathbf{H_m} \rfloor = \gamma'(\frac{\tau}{2})$ at mid-point. We take the equation of the geodesic in Equation (55) and write:

$$\lceil \mathbf{Y}_\tau \rfloor = \lceil \mathbf{Y_o} \rfloor \cdot V C_\tau V^T + \lceil \mathbf{H_o} \rfloor \cdot V S_\tau E^{-1} V^T \tag{61}$$

By taking into consideration the SVD decomposition given in Equation (60), this equation becomes:

$$\Big(\mathbf{Y}_\tau V_\tau\Big) = \Big(\mathbf{Y_o} V_o\Big) \cdot C_\tau + \Big(\mathbf{H_o} V_o E^{-1}\Big) \cdot S_\tau \tag{62}$$

Now we consider the rotated sum and difference of $\mathbf{Y}_\tau$ and $\mathbf{Y_o}$ as follows:

$$\begin{array}{rcl} \Big(\mathbf{Y}_\tau V_\tau\Big) + \Big(\mathbf{Y_o} V_o\Big) & = & \Big(\mathbf{Y_o} V_o\Big) \cdot \Big(I + C_\tau\Big) + \Big(\mathbf{H_o} V_o E^{-1}\Big) \cdot S_\tau \\ \Big(\mathbf{Y}_\tau V_\tau\Big) - \Big(\mathbf{Y_o} V_o\Big) & = & -\Big(\mathbf{Y_o} V_o\Big) \cdot \Big(I - C_\tau\Big) + \Big(\mathbf{H_o} V_o E^{-1}\Big) \cdot S_\tau \end{array} \tag{63}$$

Let us define $C_m = \mathrm{cosm}(E\frac{\tau}{2})$ and $S_m = \mathrm{sinm}(E\frac{\tau}{2})$. We also have the trigonometric identities $2C_m^2 = I + C_\tau$, $2S_m^2 = I - C_\tau$, and $2C_m S_m = S_\tau$. Then we can write:

$$\begin{array}{rcl} \frac{1}{2}\Big(\mathbf{Y}_\tau V_\tau + \mathbf{Y_o} V_o\Big) & = & \Big(\mathbf{Y_o} V_o\Big) \cdot C_m C_m + \Big(\mathbf{H_o} V_o E^{-1}\Big) \cdot C_m S_m \\ \frac{1}{2}\Big(\mathbf{Y}_\tau V_\tau - \mathbf{Y_o} V_o\Big) & = & -\Big(\mathbf{Y_o} V_o\Big) \cdot S_m S_m + \Big(\mathbf{H_o} V_o E^{-1}\Big) \cdot C_m S_m \end{array} \tag{64}$$

By inspection of the original Equation (55), we can right away write equations for the mid-point:

$$\begin{array}{l} \mathbf{Y_m} = \frac{1}{2}\Big(\mathbf{Y}_\tau V_\tau + \mathbf{Y_o} V_o\Big) C_m^{-1} V_o^T \\ \mathbf{H_m} = \frac{1}{2}\Big(\mathbf{Y}_\tau V_\tau - \mathbf{Y_o} V_o\Big) S_m^{-1} E V_o^T \end{array} \tag{65}$$

In order to obtain a closed form solution for the geodesic and its derivative vector field for the two-point boundary value problem, we substitute Equation (65) into Equation (55). After simplifications, we get:

$$\begin{aligned}\mathbf{Y}_t = &\ \mathbf{Y}_o V_o (C_t - S_t C_\tau {S_\tau}^{-1}) V_o^T + \mathbf{Y}_\tau V_\tau S_t S_\tau^{-1} V_o^T \\ \mathbf{H}_t = &-\mathbf{Y}_o V_o (S_t + C_t C_\tau S_\tau^{-1}) E V_o^T + \mathbf{Y}_\tau V_\tau C_t S_\tau^{-1} E V_o^T\end{aligned} \tag{66}$$

This is an explicit solution for the two-point boundary value problem. The time-varying matrices are $C_t = \mathrm{cosm}(tE)$ and $S_t = \mathrm{sinm}(tE)$. All other matrices are obtained from an SVD decomposition of the inner product $\mathbf{Y}_o^T \mathbf{Y}_\tau$. If at time $t = 0$, we start traversing the geodesic given by $\mathbf{Y}_o$ and $\mathbf{H}_o$ from the point $\lceil \mathbf{Y}_o \rfloor \in \mathcal{G}r_{\mathrm{np}}$, then at time $t = \tau$, we will reach a point $\lceil \mathbf{Y}_\tau \rfloor \in \mathcal{G}r_{\mathrm{np}}$. This is due to the fact that $\forall t \ \|\mathbf{H}_t\|_F = 1$.

References

[1] D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*. Prentice Hall, 2003.

[2] C. T. adn T. Kanade, "Shape and motion from image streams under orthography: a factorization method," *International J.of Computer Vision*, vol. 9, no. 2, pp. 137–154, 1992.

[3] P. M. Q. Aguiar and J. M. F. Moura, "Factorization as a rank 1 problem," *IEEE Conf. on Computer Vision and Pattern Recognition*, 1999.

[4] J. Maciel and J. P. Costeira, "A global solution to sparse correspondence problems," *IEEE Transactions on pattern Analysis and Machine Intelligence*, vol. 25, no. 2, 2003.

[5] J. Maciel, *Global Matching: Optimal Solution to Correspondence Problems*. PhD thesis, Instituto Superior Tcnico,Universidade Tcnica de Lisboa, Lisbon, Portugal., August 2001.

[6] D. G. Kendall, "The diffusion of shape," *Advances in Appl. Prob.*, vol. 9, pp. 428–430, 1977.

[7] I. L. Dryden and K. V. Mardia, *Statistical Shape Analysis*. John Wiley and Sons, 1998.

[8] S. C. Joshi, M. I. Miller, and U. Grenander, "On the geometry and shape of brain submanifolds," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 11, pp. 1317–1343, 1997.

[9] U. Grenander and D. M. Keenan, "On the shape of plane images," *SIAM Journal of Applied Mathematics*, vol. 53, no. 4, pp. 1072–1094, 1993.

[10] D. G. Kendall, D. Barden, T. K. Carne, and H. Le, *Shape and Shape Theory*. John Wiley and Sons, 1999.

[11] D. A. Sepiashvili, J. M. F. Moura, and V. H. S. Ha, "Affine-permutation symmetry: Invariance and shape space," *Proceedings of the IEEE Workshop on Statistical Signal Processing*, September 2003. Special Session on Statistical Inferences on Manifolds with Apllications in Image Analysis.

[12] V. H. S. Ha, *Blind Recovery of Intrinsic Shape*. PhD thesis, Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, August 2002.

[13] V. H. S. Ha and J. M. F. Moura, "Efficient 2d shape orientation," *International Conference on IEEE Image Processing*, vol. 1, pp. 225 – 228, September 2003.

[14] V. H. S. Ha and J. M. F. Moura, "Intrinsic shape," *36th Asilomar Conference on Signals, Systems and Computers*, vol. 2, pp. 993–997, November 2002. Invited paper, Special Session on Statistical Image Processing.

[15] J. Xavier, *Blind Identification of MIMO Channels Based on 2nd Order Statistics and Colored Inputs*. PhD thesis, Department of Electrical and Computer Engineering, Instituto Superior Tcnico, Lisbon, Portugal., October 2002. The Riemannian Geometry of Algorithms, Performance Analysis and Bounds over the Quotient Manifold of Identifiable Channel Classes.

[16] W. M. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*. Academic Press, 1975.

[17] A. Graham, *Kronecker Products and Matrix Calculus with Applications*. Ellis Horwood Ltd., a Division of John Wiley and Sons, 1981.

[18] J. R. Magnus and H. Neudecher, *Matrix Differential Calculus*. John Wiley and Sons, 1988.

[19] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, vol. 1-2. John Wiley and Sons, 1963.

[20] I. Chavel, *Riemannian Geometry: A Modern Introduction*. Cambridge University Press, 1993.

[21] A. Edelman, T. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constrains," *SIAM J. Matrix Anal. Appl.*, vol. 20, pp. 303–353, 1998.

[22] R. K. Nagle, E. B. Saff, and A. D. Snider, *Fundamentals of Differential Eq. and Boundary Value Problems*. 3 ed., 1999.

[23] G. H. Golub and C. F. V. Loan, *Matrix Computations*. The John Hopkins University Press, 3 ed., 1996.

[24] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[25] S. MacLane and G. Birkoff, *Algebra*. The MacMillan Company, 1967.

[26] C. He and J. M. F. Moura, "Robust detection with the gap metric," *IEEE Transactions on Signal Processing*, vol. 45, no. 6, pp. 1591–1604, 1197.

[27] T. Kato, *Perturbation Theory for Linear Operators*. New York: Springer-Verlag, 2 ed., 1976.