2015

Keeping Pace with the Times: Quantifying Variation of Newly Emerging Biological Shape Data

Qiuping Xu

FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

KEEPING PACE WITH THE TIMES: QUANTIFYING VARIATION OF NEWLY EMERGING BIOLOGICAL SHAPE DATA

By

QIUPING XU

A Dissertation submitted to the
Department of Mathematics
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Degree Awarded:
Spring Semester, 2015

Qiuping Xu defended this dissertation on March 19, 2015.
The members of the supervisory committee were:

Washington Mio

Professor Directing Dissertation

Piyush Kumar

University Representative

Richard Bertram

Committee Member

Xiuwen Liu

Committee Member

Jack Quine

Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the dissertation has been approved in accordance with university requirements.

I dedicate this dissertation to to my parents and my family for their constant support and unconditional love. You always suspected I'd end up here. I love you all dearly.

ACKNOWLEDGMENTS

I am profoundly grateful to my supervisor, Dr. Washington Mio. Much of this work would have not been possible without his ideas, encouragement, interest, dedication, great vision and permanent support. I am also thankful to Dr. Richard Bertram, Dr. Piyush Kumar, Dr. Xiuwen Liu and Dr. Jack Quine for their comments, support and reviews. I would like to thank all faculty and staff members at various departments at The Florida State University who have taught me or helped me in the past several years.

I also want to thank our collaborators from the Hallgrimsson lab at University of Calgary and the Marcucio lab at University of California, San Francisco for sharing the biological shape data used in this dissertation.

I would like to thank my classmates and friends for their friendship and help. Thanks for the valuable lessons you have taught me about life and for helping me survive this great adventure.

Finally and most importantly, I owe my deepest gratitude to my family. Thank for your unconditional love, dedication, support and patience through all my life. I know I always have my family to count on when times are rough.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Shape represents a complex and rich source of biological information that is fundamentally linked to underlying mechanisms and functions. Many fields of biology employ mathematical tools for the statistical analysis of shape variation. However, difficulties in reliably quantifying biological shape, especially for newly emerging shape data, still present an obstacle for researchers to understand how shape variation relates to biological functions and development processes. To overcome these difficulties, it is desirable to build efficient ways to quantify shapes. Having a quantitative tool in hand, we can further design methods to correlate shape with biological information. The integration of these models with machine learning and statistical inference methods will allow biologists to explore how morphological variation correlates to biological variates and to help advance various areas of research.

One goal of this dissertation is to construct new type of shape representation to quantify gene expression data. Advances in microscopy and techniques such as Optical Projection Tomography (OPT) allow researchers to visualize and to study 3D morphological patterns of gene expression domains. Quantitative analysis of gene expression domains and investigation of relationships between gene expression and developmental and phenotypic outcomes are central to advancing our understanding of the genotype-phenotype map. However, quantification of shape variation in gene expression domains poses particularly challenging problems, as these domains typically have no clearly defined forms, often appearing seemingly amorphous. Those properties of the gene expression domains make it difficult to analyze shape variation with the tools of landmark-based geometric morphometrics. In addition, 3D image acquisition and processing introduce many artifacts that further exacerbate the problem. To overcome these difficulties, we present a method that combines OPT scanning, a shape regularization technique and a landmark-free approach to quantify variation in the morphology of sonic hedgehog expression domains in the frontonasal ectodermal zone (FEZ) of avians and investigate relationships with embryonic craniofacial shape. The landmark-free approach quantifies variation in shape of amorphous gene expression domains, enhancing their most salient morphological characteristics and being robust to uninformative local shape variation and irregularities associated with image acquisition. The correlation analysis reveals axes in FEZ and embryonic-head morphospaces along which variation exhibits a sharp linear relationship at

high statistical significance. Combined with qualitative findings, these results have the potential to benefit biologists in exploring the gene expression pathway and in understanding the underlying expression mechanisms. The techniques we used to deal with FEZ meshes should be applicable to analyses of other 3D surface-like biological structures that have ill-defined shape and are relevant to understanding developmental processes and phenotypic variation.

Existing biological shape models, such as those based on landmarks, rely on sparse landmarks on the shapes to model shape variations. However, on soft-tissue surfaces as the face there are few such landmarks. Across the cheek and forehead, for instance, there are no points that have exact biological correspondence and yet aspects of their shape contains useful biological information. The analysis based on the sparse landmarks will compromise the deep and comprehensive morphological information collected by advanced image processing technologies. Thus, instead of using only the limited number of landmarks, we propose to use the spline method to construct dense surface model which covers the entire shape. This brings another goal of this dissertation - to develop such a spline method to build a dense correspondence across all shapes. Although, spline is an active area in shape analysis and also in many other disciplines for interpolation, approximation and regression. Most results have been focused on Euclidean domain. However, data living on manifold occurs often, especially when dealing with shape surfaces, so constructing spline with manifold domain and providing effective computation method for such spline are desirable in real-life problems. To fulfill this goal, we present a general theoretical framework of spline in which the Euclidean domain can be extended to manifold domain. Additionally, we provide computationally effective algorithm to compute such spline function based on bounded rectangular domain. We demonstrate the advantages of this framework by using examples on closed and open manifold domains and by comparing performance with other spline methods. The computation framework shows comparable result with the spline directly constructed on the manifold and also shows clear improvement respect to the thin plate spline method. This manifold spline method has been applied to construct dense surface models of avian embryos shapes. Those dense surface models can establish a correspondence of thousands of points across each 3D image and provide dramatic visualization of shape variation.

CHAPTER 1

INTRODUCTION

This dissertation is concerned with developing shape analysis methods for newly emerging data. Those methods include the construction of shape representation, shape space and shape metrics. The notion of shape we refer to are geometric objects that are invariant of certain transformations, e.g., the invariance of translation, scaling or rotation. The fundamental question in shape analysis is to quantify shape similarities and dissimilarities in a reliable, effective and computationally efficient way, and this is the main concern of this dissertation as well.

Shape analysis applications to problems such as shape registration, statistical shape analysis, modeling of shape, shape classification, and shape recognition are widely addressed in the fields of biological morphology, biomedical imaging, computer vision and computer graphics. Specifically, with the evolution of scanning technologies, shape analysis is playing an increasingly important role in evolutionary study. For example, the use of shape analysis in quantifying the morphological features of gene expression domain has provided a new direction to study the relationships between genotype and phenotype.

1.1 Remarks on Shape Analysis

A traditional approach in shape analysis, carried out mainly by biologists, is known as multivariate morphometrics or "traditional" morphometrics. A variety of measurements of shapes, such as lengths, angles, masses, areas and ratios, are collected and subjected to multivariate analysis. However, these kinds of measurement only provide simple summaries about the shape. One main drawback is that the geometry is very difficult to reconstruct after the analysis, so that it is impossible to visualize the change of the shape. A more geometrical approach is to employ the entire configuration of the shape. These kinds of studies can be traced back to as early as D'Arcy Thompson (1917)[62], who began to consider the direct transformation between complete geometrical objects. Thompson helped to advance the notion that perhaps nature could be studied in a more quantitative way, instead of just by the empirical methods that were the normal at the time.

More importantly, he introduced the idea that the variation in similar anatomical shape might be described by mathematical transformations, and his work is often credited with inspiring the field of morphometrics. For illustration, Fig. 1.1 shows Thompson's image of the transformation of Argyropelecus olfersi into Sternoptyx diaphana by applying a 70^o shear mapping. The deformation illustrated by this case of Argyropelecus is precisely analogous to the simplest and commonest kind of deformation to which fossils are subject as the result of shearing-stresses in the solid rock.

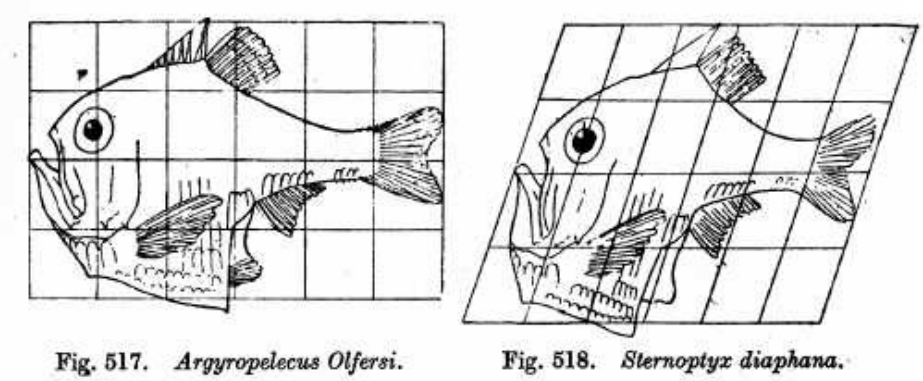


Figure 1.1: The transformation of Argyropelecus olfersi into Sternoptyx diaphana by applying a 70^o shear mapping.

Shape analysis developed rapidly around late 1970s and early 1980s, based on contribution from D.G.Kendall (1984, 1999) [29, 30] and F.L. Bookstein (1986) [6], where the former brought shape analysis to a more theoretical level. Kendall's model is based on a collection of ordered landmark points. Landmarks begin to more completely represent the geometry of the anatomy and can be more easily interpreted in the context of statistical analysis of shapes. An important feature of such geometric models is that shapes are naturally morphable as geodesics yield natural shape interpolators. An example of landmark representation of shapes is shown in Fig. 5.1A. The Kendall's model has solid theoretical foundation and is computationally easy to handle. A general framework for landmark-based shape analysis has been developed which is still in widespread use today. Many of the current geometric morphometric analysis of shape in evolutionary biology is still based on these theories. However, this model can not apply to shapes those do not have natural landmarks.

In recent years, theories of many new shape models have been proposed which involve different shape representations and metrics, such as those presented in Younes (1998, 1999) [72, 73]; Miller and Younes (2001) [46]; Younes et al. (2007) [74]; Klassen et al. (2004) [32]; Mio et al. (2007, 2007, 2009) [47, 49, 48]; Joshi et al. (2007) [28]; Michor and Mumford (2006, 2007) [43, 44]; Yezzi and Mennucci (2005) [71]; Michor et al. (2008) [45]; and Fuchs et al. (2009) [16]. In practical applications, shapes in 3D space are represented either by their outer contour surfaces or by the whole volumes. However, the development of applicable methods of handling noisy and complex shape data is still incipient.

1.2 Contributions of this Dissertation

To accommodate newly emerging shape data, which differs from traditional data in size and type, the existing models have to be modified and new methods have to be developed. We are exploring these ideas in two projects.

The first project involves quantifying variations in morphology of gene expression domains. However, quantification of shape variations in gene expression domains poses particularly challenging problems, as these domains typically have no clearly defined forms, often appearing seemingly amorphous. Those properties make it difficult to analyze shape variation with the tools of landmark-based geometric morphometrics. Another serious difficulty encountered in all methods for quantifying expression patterns is that the extracted domains are often corrupted with different types of artifacts. Therefore, to analyze the morphological features, it will be necessary to propose a method to noise removal, pattern enhancement and shape preservation. Since the loss of anatomical landmarks, thus, in order to build model for expression patterns, it will be desirable to build new "correspondence" that is biological meaningful and captures the fundamental properties of the pattern.

This idea is initiated from studying the relationship between the spatial information of the sonic hedgehog (*Shh*) expression domain and the facial features during embryonic stage of two avian species (chickens and ducks). *Shh* is a protein whose signaling plays an essential role in the epithelial mesenchymal interactions that control proximodistal extension and dorsoventral polarity of the vertebrate upper jaw [25, 39, 23, 75]. In amniotes, including both mice and avians, *Shh* is first expressed in the forebrain prior to outgrowth of the facial prominences. As neural crest cells migrate

into the midface, *Shh* is activated in the adjacent epithelium, and this frontonasal ectodermal zone (FEZ) acts as a signaling center that controls growth [39]. The spatial organization of this signaling center differs among avians, and these differences correspond to sonic hedgehog (*Shh*) expression in the basal forebrain and embryonic facial shape. Hu and Marcucio [23] show spatial organization of the FEZ regulates morphological variation in the developing upper jaw.

In this dissertation, we present a computational technique that overcome those difficulties. Our method involves two key components, the first is shape regularization to remove the artifacts. We consider each dissected gene expression pattern as the graph of a function f of x and y. The bounded domain of this function is estimated by principle component analysis and kernel density estimation. We have also adopted the thin plate spine algorithm which has only been used extensively in morphometric analysis [7], to estimate the smooth function f from pre-triangulated domain. The second component is the representation and registration method. We estimated the sagittal plane for each embryo and sectioned each correspondent gene expression domain with multiple parallel translates of the plane. It is a continuum of sections. Sweeping from left to right, we measured the length of each of these intersected curves that registers these expression patterns. The canonical correlation analysis reveals axes in FEZ and embryonic-head morphospaces along which variation exhibits a sharp linear relationship at high statistical significance. This relationship has been validated by the experimental data of two transplanted groups. Combined with qualitative findings, these results have the potential to benefit biologists in exploring the gene expression pathway and in understanding the underlying expression mechanisms. The techniques we used to deal with FEZ meshes should be applicable to analyses of other 3D surface-like biological structures that have ill-defined shape and are relevant to understanding developmental processes and phenotypic variation.

The analysis based on the sparse landmarks will compromise the deep and comprehensive morphological information collected by advanced image processing technologies. Thus, instead of using only the limited number of landmarks, in the second project, we present a general theoretical framework of spline to construct dense correspondence. In this spline method, the Euclidean domain can be extended to manifold domain to better accommodate the nature of surface shape data. Additionally, to apply this spline method onto modern data set, we provide computationally effective algorithm to compute such spline function based on bounded rectangular domain. The advantages

of this framework is demonstrated by using examples on closed and open domain and by comparing performance with other spline methods. The computation framework shows comparable result with the spline directly constructed on the manifold and shows clear improvement respect to the thin plate spline method [67, 7, 68].

Beside providing the theoretical derivation and computational algorithm, we also apply this spline method to the construction of dense surface model of avian embryos. Those dense surface models can establish a correspondence of thousands of points across each 3D image, so that it can take full advantages of geometrical information. The dense surface models not only provide dramatic visualization of 3D embryo-shape variation but also enable modular analysis which is impossible under sparse landmark representation. The dense facial models carry the potential for precisely identifying the facial features and the syndrome effects that can benefit a serious of following up studies. Beside applications in shape modeling, the spline method can be easily extended to general maps from manifold domain. e.g in order to compute dense point correspondences between two objects.

The reminder of this dissertation is organized as follows: In Chapter 2, we cover the mathematical preliminaries, including Riemannian manifolds and the Laplace-Beltrami operator on Riemannian manifold. The discretization of the Laplacian on triangular meshes and point clouds is discussed in Chapter 3. Several multivariate statistic models used in this dissertation are brief discussed in Chapter 4 which includes principal component analysis [53], canonical correlation analysis [21] and thin spline spline method [69]. In Chapter 5, statistical shape analysis including algorithms to calculate mean shapes and to perform tangent principal component analysis is sketched. The algorithm of handling noisy gene expression domains and the results of quantifying variations in morphology of FEZ shapes are presented in Chapter 6. The correlation analysis between FEZ and embryonic-head morphospaces is also presented in this chapter. The spline model on manifold domain and the comparison result on simulated surfaces are shown in Chapter 7. A summary and discussion is in Chapter 8.

CHAPTER 2

MATHEMATICAL PRELIMINARIES

In this chapter, we will review some basic facts about Riemannian manifold [8] and the Laplace-Beltrami operator on a connected oriented Riemannian manifold.

2.1 Riemannian Manifolds

As a preliminary to the definition of a differentiable manifold, we recall the definition of a topological manifold M of dimension n: it is a Hausdorff space with a countable basis of open sets and with the further property that each point has a neighborhood homeomorphic to an open subset of $\mathbb{R}^n$. Each pair (U, φ), where U is an open set of M and φ is a homeomorphism of U to an open subset of $\mathbb{R}^n$, is called a *coordinate neighborhood*.

If q also lies in the second coordinate neighborhood (V, ψ), since φ, ψ are both homeomorphisms, this further defines a homeomorphism as

$$\psi \circ \varphi^{-1} : \varphi(U \cap V) \to \psi(U \cap V) \tag{2.1}$$

The domain and range being the two open subsets of $\mathbb{R}^n$ which correspond to the points $U \cap V$ by the two coordinate maps φ, ψ, respectively. If $\psi \circ \varphi^{-1}$ is a diffeomorphism, then these two coordinate neighborhoods U, φ and V, ψ are C^∞ compatible.

The basic idea that leads to differentiable manifolds is to try to select a family or sub collection of neighborhoods so that the change of coordinates is always given by differentiable functions.

Definition of differentiable manifold

A *n-dimensional differentiable* or C^∞ *structure* on a topological manifold M is a family $\mu = \{U_\alpha, \varphi_\alpha\}$ of coordinate neighborhoods such that:

1. The U_α cover M
2. For any α, $(U_\alpha, \varphi_\alpha)$ is coordinate neighborhood.
3. Each pair of coordinate neighborhoods is C^∞ compatible.

A C^∞ manifold is a topological manifold together with a C^∞-differentiable structure. Each φ_α is called *chart* [35].

Given a differentiable manifold, there is no natural way to define the Laplacian of a function $r : M \to \mathbb{R}$, without additional "geometry". This "geometry" is prescribed in the form of a Riemannian metric.

Let M be a differentiable manifold. A real-valued function $f : M \to \mathbb{R}$ belongs to $C^\infty(M)$, if f^{-1} is infinitely differentiable for every chart φ_α. The definition of $C^\infty(M)$ is stated here in order to motivate the general definition of the tangent spaces of a differentiable manifold. The tangent space T_pM to M at $p \in M$ is defined to be all the linear maps $D_p : C^\infty(M) \to \mathbb{R}$ which has the property for all $f, g \in C^\infty(M)$:

$$D_p(fg) = D_p(f) \cdot g(p) + f(p) \cdot D_p(g) \tag{2.2}$$

The elements of the tangent space T_pM are called *tangent vectors* at p. All the tangent spaces have the same dimension, and this dimension equals to the dimension of the manifold. For example, the 2-sphere S^2 is a 2-manifold, one can picture the tangent space at a point p as the plane which touches the sphere at that point and is perpendicular to p.

A *Riemannian manifold* or *Riemannian space* (M, Φ) is a differentiable manifold M in which each tangent space is equipped with an inner product Φ, a *Riemannian metric*, which varies smoothly from point to point.

More preciously, let M be a differentiable manifold of dimension n. A Riemannian metric Φ on M is a smooth family of inner products Φ_p such that,

$$\Phi_p : T_pM \times T_pM \to \mathbb{R}, \ p \in M \tag{2.3}$$

Given a Riemannian metric, the length of a curve $\gamma : [0, 1] \to M$ is defined as

$$l(\gamma) = \int_0^1 \Phi_{\gamma(t)} \left(\gamma'(t), \gamma'(t)\right)^{\frac{1}{2}} dt \tag{2.4}$$

For any point p and q on the manifold, $d_M(p, q)$ which is the geodesic distance between p and q is defined to be the infimum of the lengths of curves joining p and q. Two Riemannian manifolds (M, g) and (N, h) are called *isometric* if there exists a smooth diffeomorphism $f : M \to N$ such that $d_M(p, q) = d_N(f(p), f(q))$ for any p and q on M.

2.2 The Laplace-Beltrami Operator on a Riemannian Manifold

The Laplacian Δ of smooth function on M is often referred to as the Laplace-Beltrami operator. It is defined as $\Delta = -\text{ div grad}$,where div and grad are the divergence and gradient on a Riemannian manifold, respectively. The minus sign is just a convention to make eigenvalues of Δ nonnegative. The standard Laplacian on Euclidean domain $\Delta_{\mathbb{R}^n} = -\sum_{i=1}^{n}(\partial^2/\partial_{x_i^2})$ agrees with this definition. We will define the volume element of a Riemannian metric, followed by a definition of L^2 functions on a connected oriented Riemannian manifold. We will also review the basic properties of eigenvalues and eigenfunctions of Δ [56].

Let (M, Φ) be a connected oriented Riemannian manifold of dimension n, if $\partial_{x_1}, ..., \partial_{x_n}$ is a positively oriented basis of T_pM, then the volume element $dV(p)$ is defined in local coordinates as $dV(x) = \sqrt{g(x)}dx_1 \wedge ... \wedge dx_n$, where $g_{ij} = \Phi(\partial_{x_i}, \partial_{x_j})$, $\sqrt{g(x)} = \sqrt{\det(g(x))}$. Then the volume of (M, Φ) is defined to be $V_m = \int_M dV(p)$. Let $L^2(M, \Phi)$ be the Hilbert space of real-valued square-integrable functions on (M, Φ) with the inner product.

$$< f, g >= \int_M f(p)g(p)dV(p) \tag{2.5}$$

Let g^{ij}=$(g_{ij})^{-1}$, and $\hat{f} = f \circ \varphi^{-1}$. The Laplacian Δ on differentiable functions of (M, Φ) is the linear differential operator given in local coordinates by

$$\Delta f = -\frac{1}{\sqrt{g}} \sum_{ij} \partial_{x_j}(g^{ij}\sqrt{g}\partial_{x_i}\hat{f}) \tag{2.6}$$

An Example: The Laplacian on S^2

There is a coordinate neighborhood (U, φ) of S^2 s.t $\varphi : S^2 \to \mathbb{R}^2$

$$\varphi^{-1}(u, v) = (\cos u \sin v, \sin u \sin v, \cos v) \tag{2.7}$$

Simple calculation shows that

$$\partial_u = (-\sin u \sin v, \cos u \sin v, 0); \partial_v = (\cos u \cos v, -\sin u \cos v, -\sin v) \tag{2.8}$$

The Euclidean metric on $\mathbb{R}^3$ induces a Riemannian structure on tangent space of S^2. The metric is expressed by

$$(g_{ij}) = \begin{bmatrix} \sin^2 v & 0 \\ 0 & 1 \end{bmatrix} \tag{2.9}$$

and so

$$(g^{ij}) = \begin{bmatrix} \frac{1}{\sin^2 v} & 0 \\ 0 & 1 \end{bmatrix} \tag{2.10}$$

For any $f \in C^\infty(M)$, define $\hat{f} = f \circ \varphi^{-1}$. Then Eqs. 2.6, 2.9 and 2.10 imply that

$$\Delta f = -\frac{1}{\sin^2 v}\hat{f}_{uu} - \frac{1}{\sin v}\partial_v(\sin v \hat{f}_v) \tag{2.11}$$

Theorem 2.2.1 *(Hodge Theorem for Functions) Let (M, Φ) be a compact connected oriented Riemannian manifold. There exists a complete orthonormal set of $L^2(M, \Phi)$ consisting of eigenfunctions of the Laplacian. All the eigenvalues are positive, except that zero is an eigenvalue with multiplicity one. Each eigenvalue has finite multiplicity, and the eigenvalues accumulate only at infinity.*

The positive eigenvalues of Δ on a compact connected Riemannian manifold M may have multiplicity more than one, this phenomenon is mostly introduced by the symmetrical characteristic of a shape. In practice, due to noise, shape irregularity and computation error, all the eigenvalues are distinct [63, 64]. Hence, if we represent eigenfunctions and eigenvalues of Δ by $\phi_i, \lambda_i, i = 0, 1, \ldots,$ respectively, where they satisfy $\Delta\phi_i = \lambda_i\phi_i$. We can assume all eigenvalues form a strictly monotone increasing set and can be ordered as $\lambda_{i+1} > \lambda_i$. The eigenfunctions ϕ_i making up the complete orthonormal set of $L^2(M, \Phi)$ only carry a sign ambiguity. According to the Hodge Theorem for Functions in theorem 2.2.1, connected manifold M implies $\lambda_0 = 0$ and $\lambda_i > 0$ for $i \geq 1$. The set $\{\lambda_i\}$ is called the *spectrum* of Δ. In shape analysis, this set is also known as *shape DNA*. And this shape DNA can be used as shape representation in problems such as shape classification and shape retrieval. More detail of application of shape DNA can be found in [55].

2.3 A Definition of Shape and Shape Metric

Two Riemannian manifolds M and N are isometric if there is a diffeomorphism $f : M \to N$ such that $d_M(p, q) = d_N(f(p), f(q))$ for all $p, q \in M$. This f is then called an *isometry* between M and N. Isometry gives a natural equivalence relation $\sim$ on Riemannian manifolds. We consider the collection of closed, connected Riemannian manifolds M. M is partitioned into isometric equivalence classes $M/\sim$, which are sometimes called geometric structures. The intrinsic geometry of (M, g) refers to all properties of its equivalence class in $M/\sim$.

Definition 2.3.1 *(Shape) A shape is an equivalence class in $S = M/\sim$. Note that every element of S has a fixed underlying differentiable manifold.*

Ideas from morphometry have motivated this definition of shape, and it should not be taken to be universal. For instance, a shape and its mirror image are equivalent from the point of view of intrinsic geometry. However, other studies may need to consider chirality, for example, and in such instances isometry is too weak a relation.

A shape metric

We ultimately compare differences between shapes, such as healthy and diseased structures, native and mutated forms. The shape metric we consider is the Gromov-Hausdorff distance d_{GH},

Definition 2.3.2 *(Hausdorff Distance) Let (Z, d_Z) be a metric space and M, N compact subsets of Z. The Hausdorff distance between M and N is defined as*

$$d_H^Z(M,N) = \max\{\sup_{p\in M} d_Z(p,N), \sup_{q\in N} d_Z(M,q)\} \tag{2.12}$$

Where $d_z(p,N) = \inf_{q\in N} d_Z(p,q)$. Intuitively, d_H measures the overhang of M and N. The full expression simply imposes symmetry for the definition of distance.

The Gromov-Hausdorff distance between any two compact metric spaces (M, d_M) and (N, d_N) is then defined as

$$d_{GH}(M,N) = \inf_{Z, i_M, i_N} d_H^Z(i_M(M), i_N(N)) \tag{2.13}$$

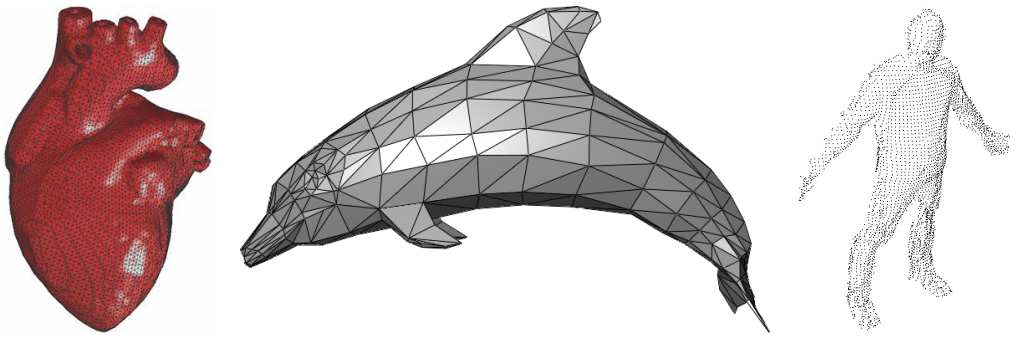
Where $i_M : M \to Z$, $i_N : N \to Z$ are any isometric embedding of M, N into any common metric space (Z, d_Z).

CHAPTER 3

DISCRETE LAPLACIAN

In order to achieve computer implementation, a shape need to be represented and stored as certain data type. Hence we need introduce numerical computable discrete analogy of Δ operator. In this chapter, the case where a 3D shape is represented as a triangular mesh is mainly discussed. For triangle mesh in the computer data structure, we will have the coordinate for each 3D point, the index for each point and the information of the triangles in term of the indexes of the three vertices of the triangle and triangle itself index. Point cloud is another major type of data structure which contains only the coordinate of points. In this section, we will discuss the discretization of Δ on triangular mesh and point cloud.

Before diving into the discussion, examples of mesh representation and point cloud representation are shown in Fig. 3.1.



(A) (B) (C)

Figure 3.1: Examples of shape data. (A) Triangular surface of a heart shape. (B) Triangular surface of a dolphin. (C) Point cloud of a human body shape.

3.1 Mesh Laplacian

Let $V = [v_1, ..., v_n]$ be the vertex set of a triangular mesh. The order of vertices is arbitrary. The $R(i)$ is the 1-*ring* of v_i, that is, $R(i)$ is the index set of vertices adjacent to v_i. Thus, if

$j \in R(i)$, then v_i and v_j are connected by an edge e_{ij} (denotes by blue line in Fig. 3.2). We denote the length of the edge e_{ij} by l_{ij}. We denote the length of the 1-*cell* $\tilde{e}_{ij}$, which is the green line in Fig. 3.2, as $\tilde{l}_{ij}$. The 1-cell $\tilde{e}_{ij}$ is the pair of line segments that connects the centroids of the two triangles with e_{ij} in common to the midpoint of e_{ij}. The area of the 2-*cell* dual (area inside the red and green curve in Fig. 3.2) to v_i as A_i. Note that the l, $\tilde{l}$ are both symmetric operator, which means $l(ij) = l(ji)$ and $\tilde{l}(ij) = \tilde{l}(ji)$.

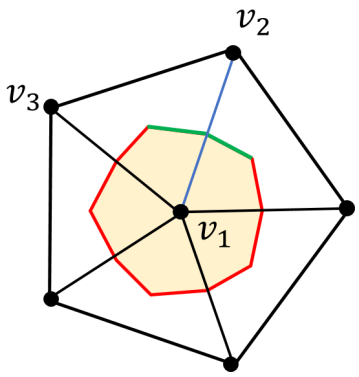


Figure 3.2: Example of a dual cell of triangular mesh

A function $f : V \to \mathbb{R}$ is represented by the vector $f = [f_1 \dots f_n]^T$, where $f_i = f(v_i)$. Follow the general definition of Laplace-Beltrami operator $\Delta = -\text{div}(\text{grad})$, div and grad are the Riemannian divergence and gradient operator, respectively, we can define the Laplacian operator on mesh structures.

The Laplacian of f at v_i is defined such that the discrete divergence theorem is satisfied

$$\int_{A_i} \Delta f dV = -\int_{\partial A_i} \frac{\partial f}{\partial n} ds \tag{3.1}$$

The right part of Eq. 3.1 measures all the out flow from v_i. Discretization of this part becomes

$$\int_{\partial A_i} \frac{\partial f}{\partial n} ds = \sum_{j \in R(i)} \frac{(f_j - f_i)}{l_{ij}} \tilde{l}_{ij} \tag{3.2}$$

The discrete counterpart of left part of Eq. 3.1 becomes

$$\int_{A_i} \Delta f dV = \Delta f_i A_i \tag{3.3}$$

Then Eqs. 3.1, 3.2 and 3.3 imply that

$$\Delta f_i A_i = -\sum_{j \in R(i)} \frac{(f_j - f_i)}{l_{ij}} \tilde{l}_{ij} \tag{3.4}$$

The summation estimates the total outward flux of the gradient filed ∇f across the boundary of the 2-*cell*, and the full expression represents the minus of the flux density over the 2-*cell*. More details of this discrete version of Δ are shown in [20].

Eq. 3.3 introduces the discrete $L^2(M)$ inner product over the mesh

$$< f, g >= \sum_{i=1}^{n} (f_i g_i) A_i \tag{3.5}$$

It can be seen from Eq. 3.4 that the discrete Laplacian is a linear operator. Hence, It can be written in matrix form. If we define

$$w_{ij} = \begin{cases} \tilde{l}_{ij}/l_{ij} & \text{if } j \in R(i) \\ 0 & \text{otherwise} \end{cases} \tag{3.6}$$

and $D_i = \sum_j w_{ij}$, then Eq. 3.4 becomes

$$\Delta f_i A_i = -\sum_j w_{ij}(f_j - f_i) \tag{3.7}$$

We can further define matrices $A = \text{diag}(A_1, ..., A_n)$, $D = \text{diag}(D_1, ..., D_n)$ and $L = D - W$, then Eq. 3.4 can be written in the following form,

$$A(\Delta f) = -(Wf - Df) = Lf \tag{3.8}$$

Note that A, D, W and L are all symmetric matrices. Next, I will use the definition to show that L is positive semi-definite matrix.

$$\begin{aligned} & 2f^T L f \\ = \; & 2f^T D f - 2f^T W f \\ = \; & 2\sum_i D_i f_i^2 - 2\sum_{ij} w_{ij} f_i f_j \\ = \; & \sum_i D_i f_i^2 + \sum_j D_j f_j^2 - 2\sum_{ij} w_{ij} f_i f_j \\ = \; & \sum_{ij} w_{ij} f_i^2 + \sum_{ij} w_{ij} f_j^2 - 2\sum_{ij} w_{ij} f_i f_j \end{aligned}$$

$$
\begin{aligned}
&= \sum_{ij} w_{ij}(f_i^2 + f_j^2 - 2f_i f_j) \\
&= \sum_{ij} w_{ij}(f_i - f_j)^2
\end{aligned} \tag{3.9}
$$

Since all $w_{ij} \geq 0$, based on Eq. 3.9 we know that $f^T L f$ is always nonnegative. We should also notice that if f is a non zero constant vector, $f^T L f = 0$. Therefore, L is positive semi-definite matrix. Assume there are solutions λ, f to the eigenvalue problem $\Delta f = \lambda f$, Eq. 3.8 becomes $Lf = \lambda A f$, the solutions to this generalized eigenvalue problem can be found and they satisfy $\Phi^T L \Phi = \Lambda, \Phi^T A \Phi = I$. Here Φ is the eigenvector matrix defined as $[\phi_0, \phi_1, \ldots, \phi_n]^T$, $\Lambda = \text{diag}(\lambda_0, \lambda_1, \ldots, \lambda_n)$, where $0 = \lambda_0 < \lambda_1 <, \ldots, < \lambda_n$. The eigenvectors can be scaled to form an orthonormal basis with respect to the inner product defined in Eq. 3.5. The eigenvalues $(\lambda_0, \ldots, \lambda_n)$ and a corresponding set of orthonormal eigenvectors $\tilde{\phi}_0, \ldots, \tilde{\phi}_n$ within normal Euclidean inner product can be computed as

$$
\tilde{\phi}_i = [\phi_{i1}\sqrt{A_1}, \ldots, \phi_{in}\sqrt{A_n}]^T \tag{3.10}
$$

All the eigenvalues can be put in the ascending order, henceforth, for convenience, I will use the smaller eigenfunction to denote the eigenfunction with the smaller eigenvalue.

In above setup, both matrices A and D are diagonal matrices, L is a large symmetric sparse matrix, $l_{ij} \neq 0$ if and only if the vertices i and j are connected by an edge. It has efficient algorithm to compute small number of eigenvalues and eigenvectors by krylov subspace method.

To close up this part of discussion, we show how these eigenvalues and eigenvectors are invariant under the effects of rotation, translation and scaling. Because they are isometry invariant, they only depend on the gradient and divergence which in turn just depend on the structure and topology of the Riemannian manifold. The effects of translation and rotation have been removed naturally. For scaling, if the manifold is scaled by an factor $a > 0$ in every dimension, the eigenvalues is scaled by $1/a^2$. In order to keep orthonormality, the eigenvectors need to be scaled by $1/a$. Particularly, if we set $a = \sqrt{\lambda_1}$, the normalized eigenvalues are defined by λ_i/λ_1 and the normalized eigenvectors are given by $\phi_i/\sqrt{\lambda_1}$. These normalized eigenvalues and eigenvectors are insensitive to scale. After normalization, λ_0 is always 0, and λ_1 is always 1. Hence, if scaling has no effect on shape, we will use normalized eigenvalues and eigenfunctions of Δ in the further application.

3.2 Point Cloud Laplacian

A Laplacian on point clouds is discussed in [3]. The algorithm of constructing point cloud Laplacian is briefly listed below,

Given k points $x_1, \ldots, x_k$ in $\mathbb{R}^l$, we construct a weighted graph with k nodes, one for each point, and a set of edges connecting neighboring points.

1. Constructing the adjacency graph
   - $\epsilon - neighborhood$. Nodes i and j are connected if $||x_i - x_j||^2 < \epsilon$.
   - n nearest neighbors. Nodes i and j are connected if i is among n nearest neighbors of j or the other way around. This condition is required to ensure the symmetric relation.
2. Weighting the edge
   - Heat kernel. i and j are connected then $W_{ij} = e^{-\frac{||x_i - x_j||^2}{t}}$, otherwise $W_{ij} = 0$.
   - Simple-minded. $W_{ij} = 1$ if and i and j are connected, $W_{ij} = 0$ if they are not.
3. Constructing the generalized eigenvalue problem

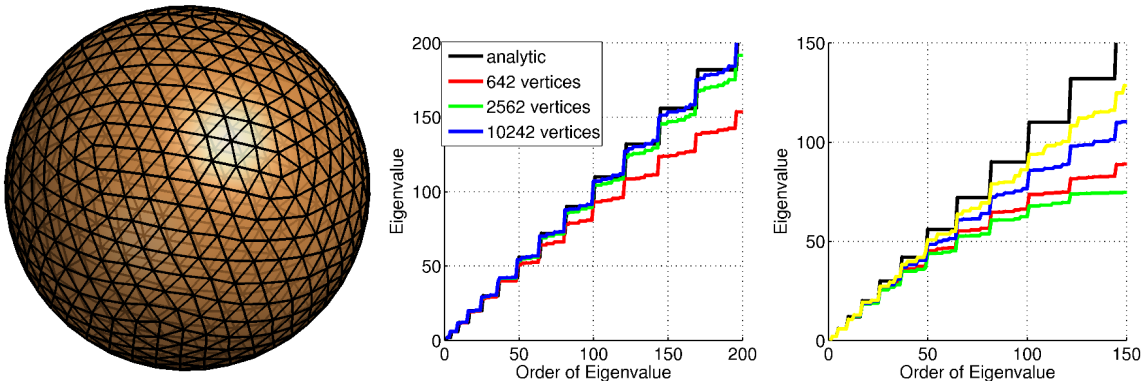
$$Lf = \lambda Df \tag{3.11}$$

   where the i_{th} element D_{ii} in the diagonal matrix D is defined as $\sum_j W_{ji}$ and $L = D - W$.

The approach considered here uses a graph which is connected with exponentially decay weights. The computation of eigenvalues and eigenfunctions is an analog to mesh Laplacian after constructing large sparse matrix L and the diagonal matrix D. The normalized version can be computed as the same fashion in mesh Laplacian. But it is not easy to choose the decay parameter t and the number of neighbors n. "How to choose the appropriate parameters" is still an open problem. Other definitions of discrete Laplacian operator are also discussed in [51, 4, 5, 55].

Like many other discretization methods, the accuracy of discrete Laplacian are depending on the quality of the shape data and the choice of the parameters. The following Fig. 3.3 shows the comparison of the eigenvalues of the Laplacian operator on the sphere with several shape discretization and different choice of parameters. From those pictures, we can see although the sphere is generated as almost equal area triangular mesh (Fig. 3.3A), there is still difference between analytic eigenvalues and the computed eigenvalues. This phenomenon is more obvious as the order of the eigenvalue increases. The computed eigenvalues not only depend on the number

of vertices on the spherical mesh (Fig. 3.3B) but also depend on the choice of parameters (Fig. 3.3C).



(A) (B) (C)

Figure 3.3: The comparison of eigenvalues of Laplacian operator on sphere. (A) Triangular surface of a sphere with 642 vertices. (B) The comparison of analytical eigenvalues (black curve) on sphere with several sphere triangulation with different number of vertices (colored curves). (C) The comparison of analytical eigenvalues (black curve) on sphere with point cloud of sphere with 2562 vertices with difference combination of ϵ and t in point cloud laplacian (colored curves).

CHAPTER 4

MULTIVARIATE STATISTICAL MODELS

The term "multivariate statistics" is appropriately used to include all statistics where there are more than two variables simultaneously analyzed. Multivariate statistics concerns understanding the different aims and background of each of the different forms of multivariate analysis, and how they relate to each other. The practical implementation of multivariate statistics to a particular problem may involve several types of univariate and multivariate analyses in order to understand the relationships between variables and their relevance to the actual problem being studied. In this section, we will cover several multivariate statistical models that have been used in this dissertation.

4.1 Principal Component Analysis

Principal component analysis (PCA) is a mathematical algorithm that reduces the dimensionality of the data while retaining most of the variation in the data set. It accomplishes this reduction by identifying directions, called principal components, along which the variation in the data is maximal. By using a few components, each sample can be represented by relatively few numbers instead of by values for thousands of variables. Samples can then be plotted, making it possible to visually assess similarities and differences between samples and determine whether samples can be grouped/clustered.

4.1.1 Interpretation of PCA

PCA is mathematically defined as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by projecting the data into the new coordinate system comes to lie on the first coordinate (first principal component).

For a given set of points $P_1, P_2, ..., P_n \in \mathbb{R}^m$, the data matrix can be defined as $X = [P_1, P_2, \ldots, P_n]^T$. This X matrix is $n \times m$ matrix with each row as an observation with m features. We can assume X is a matrix with column-wise zero empirical mean. Even though most time this is not the case, zero empirical mean can be simply achieved by subtracting mean of each column.

First Component. The first loading vector $w_1 = [w_1^1, w_1^2, ..., w_1^{m-1}, w_1^m]^T$ thus has to satisfy

$$w_1 = \underset{||w||=1}{\text{argmax}}\{||Xw||^2\} = \underset{||w||=1}{\text{argmax}}\{w^T X^T X w\} \tag{4.1}$$

Since w has been defined to be a unit vector, it also equivalently satisfies

$$w_1 = \text{argmax}\{\frac{w^T X^T X w}{w^T w}\} \tag{4.2}$$

The quantity to be maximized can be recognized as a Rayleigh quotient. A standard result for a symmetric matrix, such as $X^T X$, is that the quotient's maximum possible value is the largest eigenvalue λ_1 of the matrix $X^T X$, which occurs when w_1 is the corresponding eigenvector. The projection (scores) of the data onto the first principal component can be computed as Xw_1.

Following Components. Suppose that column vectors $w_1, w_2, ...w_{k-1}$ are the first $k-1$ principal components respectively, then the k^{th} principal component satisfies

$$w_k = \text{argmax}\{\frac{w^T X^T X w}{w^T w}, 0 =< w_k, w_1 >=< w_k, w_2 >= ... =< w_k, w_{k-1} >\} \tag{4.3}$$

It turns out that this gives the remaining eigenvectors of $X^T X$, with the maximum values for the quantity in brackets given by their corresponding eigenvalues.

In summary, the k eigenvectors $W = [w_1, w_2, w_3, ..., w_{k-1}, w_k]$ associated with the first k largest eigenvalues $[\lambda_1, \lambda_2, \lambda_3, ..., \lambda_{k-1}, \lambda_k]$ of $X^T X$ are the first k principal components. The scores in the transformed coordinates are given as XW.

Selection of k in Dimensionality Reduction. The XW maps the data from the original space of m dimension into $k << m$ dimension which are *uncorrelated* over the dataset. Such dimensionality reduction can be a very useful step for visualizing and processing high-dimensional datasets, while still retaining as much of the variance in the dataset as possible.

Dimensionality reduction may also be appropriate when the variables in a dataset are noisy. If each column of the dataset contains independent identically distributed Gaussian noise, then the columns of scores will also contain similarly identically distributed Gaussian noise. However, with more of the total variance concentrated in the first few principal components compared to the same noise variance, the proportionate effect of the noise is less - the first few components achieve a higher signal-to-noise ratio. PCA thus can have the effect of concentrating much of the signal into the first few principal components, which can usefully be captured by dimensionality reduction;

while the later principal components may be dominated by noise, and so disposed of without great loss.

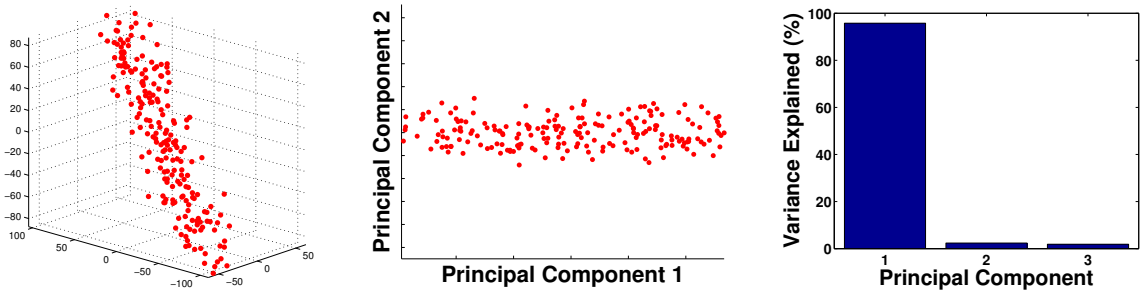
The appropriate k is normally selected by the proportion of the variance accounted for by first k dimensional scores. This proportion can be computed as the quotient between $\lambda_1 + \lambda_2 + ... + \lambda_k$ and the summation of all the eigenvalues.

Alternative Way of Computation. Normally, we compute covariance matrix $X^T X$ and do singular value decomposition (SVD) of the the covariance matrix. However, for data with small sample size n and high dimensional feature m (this happens a lot in shape/image and bioinformatics analysis), the covariance matrix will be a large matrix, the dimension of this matrix is the number of feature $m \times m$. In this case, applying SVD directly on covariance matrix is time-consuming and totally unnecessary. An better way is to decompose matrix XX^T instead, because the matrices XX^T and $X^T X$ share the same nonzero eigenvalues, and if w_i is the ith largest eigenvector of $X^T X$, v_i is the ith largest eigenvector of XX^T, and they both correspondent to the eigenvalue λ_i. One can check these eigenvectors are related in identity $w_i = X^T v_i$.

An Example: PCA Analysis on 3D Data with Linear Structure and 3D Random Data. The following example in Fig. 4.1 shows an illustration of the PCA process. The left panel shows the original data. The middle panel shows the centered data projected onto the first two principal components. We can see from this picture, the spread is mainly along the first PC direction. The right panel shows the percent variability explained by each principal component. The first PC contains more than 90% variance of the three dimensional data. If there is a dimensional reduction problem, it is safe to reduce the original data to one dimension.

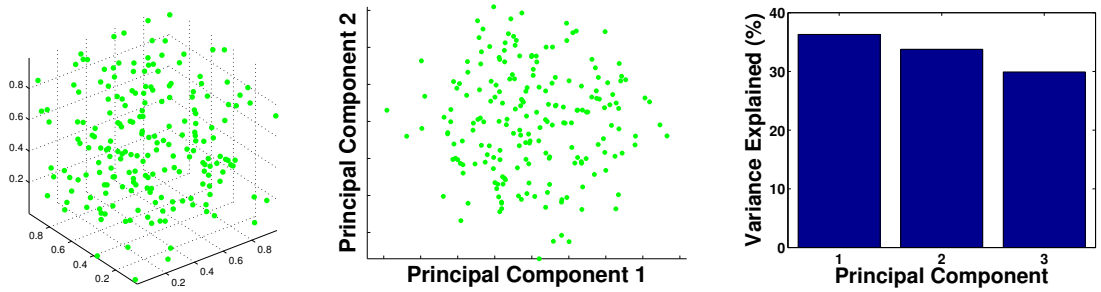
We also look at an example of 3D random data in Fig. 4.2. The random generated data is shown in left panel. The randomness is also inherited in the projection plot in the middle panel. When we look at the variability plot in the right panel, we can see that the three principal components contain roughly the same amount of variance.

PCA on Big Data. Now, principal component analysis (PCA) is a widely-used tool in genomic and statistical genetics, as well as in big industrial data. However, traditional approaches to compute the PCA are computationally expensive. For example, PCA based on the singular value decomposition (SVD) scales as $\mathcal{O}(\min(n^2 m, nm^2))$, where n and m are the number of samples and dimension of the features, respectively. This makes it time-consuming to perform PCA on large data



(A) Original Data (B) PC1 VS PC2 (C) Proportion of Variance Explained

Figure 4.1: PCA example of 3D data with linear structure. (A) Original data. (B) Projection on the first two principal components. (C) Percent variability explained by each principal component.



(A) Original Data (B) PC1 VS PC2 (C) Proportion of Variance Explained

Figure 4.2: PCA example of 3D random data. (A) Original data. (B) Projection on the first two principal component. (C) Percent variability explained by each principal component.

sets such as those routinely being analyzed in genetics and customer behavior studies, involving millions of features and tens of thousands of individuals, with this difficulty only likely to increase in the future with the availability of even larger studies.

In recent years, research into randomized matrix algorithms has yielded alternative approaches for performing PCA and producing these top principal components (PCs), while being far more computationally tractable and maintaining high accuracy relative to the traditional "exact" algorithms. These algorithms are especially useful when we are interested in finding only the first few PCs of the data, as is often the case in PCA analysis. A recently published method, flashpca[1], provide one efficient way to handle PCA on big dataset.

4.2 Canonical Correlation Analysis

Principal component analysis involves only one data set. Canonical correlation analysis (CCA) is a standard statistical technique for finding linear projections of two data sets that are maximally correlated.

Proposed by Hotelling in 1936 [21], CCA can be seen as the problem of finding basis vectors for two sets of variables such that the correlation between the projections of the variables onto these basis vectors is mutually maximized. Correlation analysis is dependent on the coordinate system in which the variables are described, so even if there is a very strong linear relationship between two sets of multidimensional variables, depending on the coordinate system used, this relationship might not be visible as a correlation. CCA was designed to overcome this problem. CCA seeks a pair of linear transformations, one for each of the sets of variables, such that when the set of variables is transformed, the corresponding coordinates are maximally correlated.

Let $(X_1, X_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ denote random vectors with covariance $(\Sigma_{11}, \Sigma_{22})$ and cross-covariance Σ_{12}. CCA finds pairs of linear projections of two vectors, $(w_1^T X_1, w_2^T X_2)$ that are maximally correlated:

$$\begin{aligned}(w_1^*, w_2^*) &= \underset{w_1, w_2}{\operatorname{argmax}} \operatorname{corr}(w_1^T X_1, w_2^T X_2), \\ &= \underset{w_1, w_2}{\operatorname{argmax}} \frac{w_1^T \Sigma_{12} w_2}{\sqrt{w_1^T \Sigma_{11} w_1 w_2^T \Sigma_{22} w_2}}\end{aligned}$$

Since the objective is invariant to scaling of w_1 and w_2, the projections are constrained to have unit variance:

$$(w_1^*, w_2^*) = \underset{w_1^T \Sigma_{11} w_1 = w_2^T \Sigma_{22} w_2 = 1}{\operatorname{argmax}} w_1^T \Sigma_{12} w_2 \tag{4.4}$$

When finding multiple pairs of vectors (w_1^i, w_2^i), subsequent projections are also constrained to be uncorrelated with previous ones, that is $w_1^i \Sigma_{11} w_1^j = w_2^i \Sigma_{22} w_2^j = 0$ for $i < j$. Assembling the top k projection vector w_1^i into the columns of a $n_1 \times k$ matrix A_1, and similarly placing w_2^i into $n_2 \times k$ matrix A_2, we obtain the following formulation to identify the top $k \leq \min(n_1, n_2)$ projections:

$$\begin{aligned}&\text{maximize:} \quad \operatorname{tr}(A_1^T \Sigma_{12} A_2) \\ &\text{subject to:} \quad A_1^T \Sigma_{11} A_1 = A_2^T \Sigma_{22} A_2 = I\end{aligned}$$

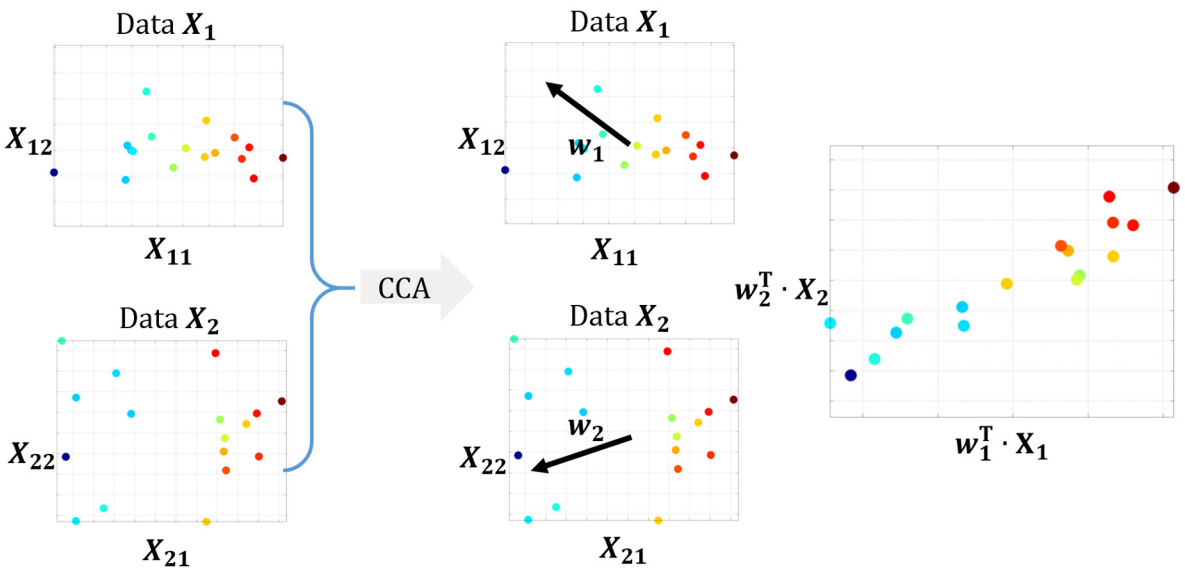


Figure 4.3: The illustration of the CCA analysis. For two multivariate dataset X_1 and X_2, CCA detects two direction w_1 and w_2 within each dataset, such that $w_1^T X_1$ and $w_2^T X_2$ are maximally correlated.

There are several ways to express the solution to this objective. We follow the one shows in [40]. Define T as $\Sigma_{11}^{-\frac{1}{2}}\Sigma_{12}\Sigma_{22}^{-\frac{1}{2}}$, and let U_k and V_k be the matrices of the first k left and right singular vectors of T. Then the optimal objective value is the sum of the top k singular value of T and the optimum is attained at $(A_1^*, A_2^*) = (\Sigma_{11}^{-\frac{1}{2}}U_k, \Sigma_{22}^{-\frac{1}{2}}V_k)$. Note that this solution assumes that the covariance matrices Σ_{11} and Σ_{22} are nonsingular, which is satisfied in practice because they are estimated from data with regularization: given centered data matrices $\bar{H}_1 \in \mathbb{R}^{n_1 \times m}$, $\bar{H}_2 \in \mathbb{R}^{n_2 \times m}$, one can estimate, e.g.

$$\hat{\Sigma}_{11} = \frac{1}{m-1}\bar{H}_1\bar{H}_1{}^T + r_1 I \tag{4.5}$$

where $r_1 > 0$ is a regularization parameter. Estimating the covariance matrices with regularization also reduces the detection of spurious correlation in the training data.

4.3 Thin Plate Spline on Euclidean Domain

Thin-plate spline interpolation can be stated as a multivariate interpolation problem: given a number n of corresponding point landmarks $\{X_i, y_i\}$, $i = 1, \cdots, n$ with $X_i = (x_{1i}, \cdots, x_{di}) \in \mathbb{R}^d$

and $y_i \in \mathbb{R}$, find a continuous transformation $f : \mathbb{R}^d \to \mathbb{R}$ within a suitable Hilbert space of admissible functions that minimizes the following energy function:

$$E_{tps} = \sum_{i=1}^{n} \frac{1}{n} |y_i - f(X_i)|^2 + w \sum_{\alpha_1 + \cdots + \alpha_d = m} \frac{m!}{\alpha_i! \cdots \alpha_d!} \int_{\mathbb{R}^d} (\frac{\partial^m f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}})^2 dX \tag{4.6}$$

Where α_k being non-negative integers and w is the parameter that controls the smoothness of the function f. Note that for the special case of $d = m = 2$ we obtain the widely used energy function for interpolating two dimensional data

$$E_{tps} = \sum_{i=1}^{n} \frac{1}{n} |y_i - f(X_i)|^2 + w \iint [(\frac{\partial^2 f}{\partial x_1^2})^2 + 2(\frac{\partial^2 f}{\partial x_1 \partial x_2})^2 + (\frac{\partial^2 f}{\partial x_2^2})^2] dx_1 dx_2 \tag{4.7}$$

Let a set of functions ϕ_j span the space of all polynomials on $\mathbb{R}^d$ up to order $m-1$, which is the null space of the functional $\sum_{\alpha_1 + \cdots + \alpha_d = m} \frac{m!}{\alpha_i! \cdots \alpha_d!} \int_{\mathbb{R}^d} (\frac{\partial^m f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}})^2 dX$. The dimension of this space is $M = \frac{(d+m-1)!}{d!(m-1)!}$ and must be lower than n. This condition determines the minimum number of landmarks, e.g., for $d = m = 2$ the number of landmarks must be larger than three, and the null space is spanned by $\phi_1(X) = 1, \phi_2(X) = x$, and $\phi_3(X) = y$. The solution of of minimizing the functional in Eq. 4.6 can now be written in analytic form:

$$f(X) = \sum_{j=1}^{M} d_j \phi_j(X) + \sum_{i=1}^{n} c_i U_m(X, X_i) \tag{4.8}$$

Where U_m is a Green's function for the m-iterated Laplacian ($\Delta^m U_m(., X_i) = \delta_{X_i}$, where δ_{X_i} is the Dirac delta function). Choosing the space of functions on $\mathbb{R}^d$ for which all partial derivatives of total order m are square integrable, i.e., are in $L_2(\mathbb{R}^d)$, this results in the basis functions is showing in Eq. 4.9 as defined in Wahba [69].

$$U_m(X, X_i) = \begin{cases} \theta_{m,d} |r|^{2m-d} \ln |r| & \text{if } 2m - d \text{ is an even integer,} \\ \theta_{m,d} |r|^{2m-d} & \text{otherwise,} \end{cases} \tag{4.9}$$

Where $|r|$ is the Euclidean distance between X and X_i. And the coefficient $\theta_{m,d}$ is defined as

$$\theta_{m,d} = \begin{cases} \frac{(-1)^{\frac{d}{2}+1+m}}{2^{2m-1} \pi^{\frac{d}{2}} (m-1)! (m-\frac{d}{2})!} & \text{if } 2m - d \text{ is an even integer,} \\ \frac{\Gamma(\frac{d}{2} - m)}{2^{2m} \pi^{\frac{d}{2}} (m-1)!} & \text{otherwise,} \end{cases} \tag{4.10}$$

For $m = d = 2$ we have the well-known function $U_2(X, X_i) = \frac{1}{8\pi} ||X - X_i||^2 \ln ||X - X_i||$. To compute the coefficients $D = (d_1, \cdots, d_M)^T$ and $C = (c_1, \cdots, c_n)^T$ of the analytic solution to Eq.

4.8 we have to solve the following system of linear equations:

$$(K + nwI)C + PD = Y \tag{4.11}$$

$$P^T C = \mathbf{0} \tag{4.12}$$

Where $K_{ij} = U_m(X_i, X_j)$, the ith row of $P_{ij} = \phi_j(X_i)$, and Y is the column vector of y_i. $\mathbf{0}$ is a $M \times 1$ column vector of zeros.

The parameter $w > 0$ controls the smoothness of the resulting transformation. If w is small, we obtain a solution with good adaption to the local structure of the deformations and if w is large, we obtain a very smooth transformation with little adaption to the deformations. There are two limiting cases: For $w \to 0$ we put more emphasis on the goodness of fit, and for $w \to \infty$ we have a global polynomial of order up to $m - 1$, which has no bending energy at all.

We do a QR decomposition of P, we can have $P = \left(Q_{1_{[n\times M]}} | Q_{2_{[n\times(n-M)]}}\right) \begin{pmatrix} R_{[M\times M]} \\ 0_{[(n-M)\times M]} \end{pmatrix}$. The solution can be found by solving the system

$$Q_2^T Y = Q_2^T (K + nwI)C \tag{4.13}$$

$$RD = Q_1^T (Y - (K + nwI)C) \tag{4.14}$$

4.3.1 Smoothing Parameter Estimation

We use the generalized cross validation [2] to choose the smooth parameter w. The idea is find $\{f_k, k = 1, 2, \cdots, n\}$ based on all samples except $\{X_k, y_k\}$ for given w. Then among all the w_s, choose the one minimizes Eq. 4.15

$$\sum_{k=1}^{n} |y_k - f_k(X_k)|^2 \tag{4.15}$$

Based on Q_2, Y and K, we can find L, W_2, U, D_1, Z and V form the following identities.

$$Q_2^T K Q_2 = L^T L \tag{4.16}$$

$$W_2 = Q_2^T y \tag{4.17}$$

$$L^T = U D_1 V^T \tag{4.18}$$

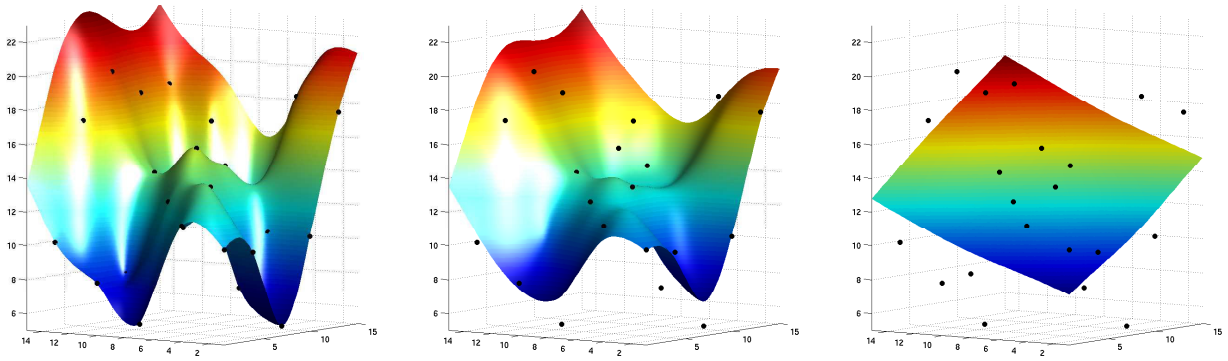
$$Z = U^T W_2 \tag{4.19}$$

D_1 is diagonal matrix with diagonal element $d_i > d_{i+1}$.

The optimal w can be found by minimize

$$V(w) = \frac{n \sum_{j=1}^{n-1} \left[\frac{n\lambda}{d_j^2+n\lambda}\right]^2 z_j^2}{\sum_{j=1}^{n-1} \left[\frac{n\lambda}{d_j+n\lambda}\right]^2} \tag{4.20}$$

The effect of the smoothing parameter w is illustrated in Fig. 4.4 with $d = m = 2$. For 2d $\{X_k\}$, the surfaces in Fig. 4.4 show the fitted scalar functions. The black dots are $\{X_k, y_k\}$.



(A) Small $w = 0.01 \times \tilde{w}$ (B) Optimal $\tilde{w}$ (C) Large $w = 100 \times \tilde{w}$

Figure 4.4: The effect of smoothing parameter w. (A) For small smoothing parameter w, the fitted surface tends to go through each point of the data set without much of the smoothness. (B) The optimal parameter w gives a transformation which approximates the distance between the landmark sets and is sufficiently smooth. (C) For large smoothing parameter w, we tend to have a global polynomial fitting up to the order of $m-1$. Because $m = 2$ in this case, so the fitted surface looks like a plane.

CHAPTER 5

KENDALL'S MODEL AND STATISTICAL ANALYSIS OF SHAPE

5.1 Procrustes Alignment

The first mathematical model of shape was developed by Kendall [29] and is based on Procrustean alignment. The model uses a geodesic metric to quantify shape divergence. This model is based on landmark representations and on a metric based on the position of landmark points. An example of the landmark representation of two shapes is shown in Fig. 5.1A. Most commonly, landmark points are selected manually according to biological correspondences, however, with the evolution of scanning and computation techniques, algorithms are being developed to automate this landmarking process.

In Kendall's model, two shapes in $\mathbb{R}^k$ are represented by collections of n homologous landmark point sets $p_1,\ p_2,\ \cdots,\ p_n \in \mathbb{R}^k$ and $q_1,\ q_2,\ \cdots,\ q_n \in \mathbb{R}^k$. Formally, shapes are encoded in a $k \times n$ matrices $P = [p_1, p_2, ..., p_n]$ and $Q = [q_1, q_2, ..., q_n]$. As a reminder, the notion of shape we refer to are geometric objects that are invariant of certain transformations, e.g., the invariance of translation, scaling or rotation. Thus, in Procrustes analysis, two shapes being compared need to be centered, scaled and orthogonal aligned first in order to be invariant under those transformations. To factor out indeterminacies in the representation due to translation and scale, one places the centroid of the points at the origin and scales the matrix to have unit Frobenius norm. To deal with orientation, one also wants to find an orthogonal matrix U, such that $||P - UQ||$ is minimized for centered and scaled P and Q.

Definition 5.1.1 $||.||_F$ *is the* **Frobenius norm** *associated with the inner product* $< P, Q >= \sum_{i=1}^{n} p_j q_j$.

- Step 1: Centering (Translation)

  The mean (centroid) $\bar{p}$ of $p_1, \cdots, p_n \in \mathbb{R}^k$ is computed as

$$\bar{p} = \sum_i p_i / n \tag{5.1}$$

Centering a matrix P is the operation that corresponds to placing the centroid of $\bar{p}$ of $p_1, \cdots, p_n$ at the origin

$$P \rightarrow [p_1 - \bar{p}, ..., p_n - \bar{p}] \tag{5.2}$$

From this point, we assume that all shapes are centered.

- Step 2: Normalizing (Scale)

  The Frobenius norm of P is

  $$||P||_F = \sqrt{\sum_i ||p_i||_F^2} \tag{5.3}$$

  Here, we assume that the cases where all landmarks are the same are excluded. This guarantees that the $||P||_F \neq 0$.

  The following operation

  $$P \rightarrow \frac{P}{||P||_F} \tag{5.4}$$

  will make the matrix into unit Frobenius norm.

  Note 5.1.1 *The collection of all centered and scaled $k \times n$ matrix forms a pre-shape space $\Omega(k, n)$, a hypersphere of a unit radius. Each pre-shape can be viewed as a point in $\Omega(k, n)$.*

- Step 3: Invariant under rotation

  In the following, we will discuss the details of how to solve this minimization problem to find the optimal orthogonal matrix U. Suppose P and Q are two pre-shapes. We want to find the $U \in O(k)$, such that $U^T U = UU^T = I$, we define distance between P and Q, *dist*, as the following

  $$dist = \min_{A \in U(k)} ||P - UQ||_F \tag{5.5}$$

  Where $O(k)$ is the orthogonal group formed by $k \times k$ orthogonal matrices.

  It is the same as find the minimizer for

  $$\min_{U \in O(k)} ||P - UQ||_F^2 \tag{5.6}$$

  Rewrite this into inner product form

  $$\begin{aligned} & \min_{U \in O(k)} ||P - UQ||_F^2 & (5.7) \\ = & \min_{U \in O(k)} < P - UQ, P - UQ > & (5.8) \\ = & \min_{U \in O(k)} < P, P > -2 < P, UQ > + < Q, Q > & (5.9) \end{aligned}$$

For giving P and Q, $< P, P >$ and $< Q, Q >$ are constant, so the minimization problem can be computed from maximizing $< P, UQ >$.

According to the properties of inner product, $< P, UQ >=< PQ^T, U >$ and then PQ^T can be decomposed into

$$PQ^T = U_1 \Sigma V_1^T \tag{5.10}$$

Where $U_1, V_1 \in O(k)$, and Σ is a diagonal matrix, we denote the diagonal elements by $\lambda_1, ..., \lambda_k$.

$$< P, UQ >=< PQ^T, U >=< U_1 \Sigma V_1^T, U >=< \Sigma, U_1^T U V_1 > \tag{5.11}$$

If we define $U_1^T U V_1 = W$, Eq. 5.11 can be write explicitly as

$$\lambda_1 w_{11} + \lambda_2 w_{22} + ... + \lambda_k w_{kk} \leq \sum_i \lambda_i = trace(\Sigma) \tag{5.12}$$

Since all U_1, U, and V_1 all belong to $O(k)$, all $w_{ii} < 1$. To achieve the equal sign, we need to have $U_1^T U V_1 = I$, so $U = U_1 V_1^T$.

- Step 4: Shape distance

  Since all the sample data has been normalized, so the sample data can be considered as a data point in a spherical surface of the shape space, if we define $\tilde{Q} = UQ$, the geodesic distance between two shapes

$$\tilde{d}([P], [Q]) = \arccos < P, \tilde{Q} >= \arccos(trace(\Sigma)) \tag{5.13}$$

  Here $[P], [Q]$ are the representation of shapes which are invariant under scaling, translation and rotation. The $\tilde{d}$ is the geodesic distance defined in the pre-shape space. It is also the length of the shortest path among all the paths connect these two shapes within pre-shape space.

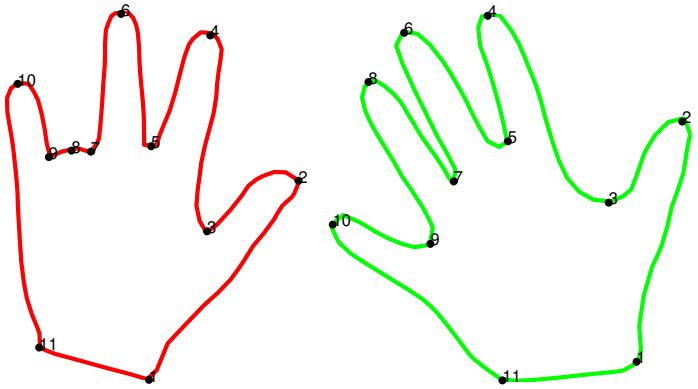
- Step 5: Geodesic interpolation

  If $P \neq \tilde{Q}$, we can interpolate intermediate shapes between P and $\tilde{Q}$ by using the following equation

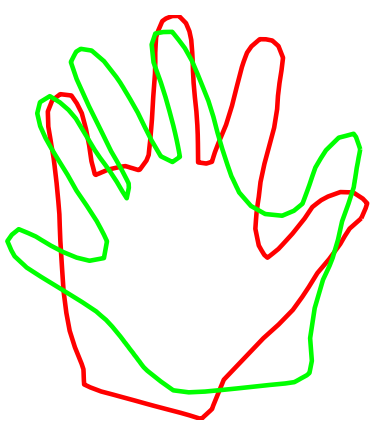
$$PQ(t) = \cos(wt)P + \sin(wt)\frac{\tilde{Q} - (tr\Sigma)P}{||\tilde{Q} - (tr\Sigma)P||_F} \tag{5.14}$$

  where $w = \arccos(tr(\Sigma))$ and $0 \leq t \leq 1$.

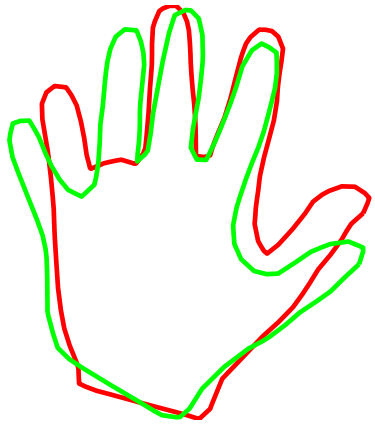
  A toy example about Procrustes alignment of two hand shapes and shape geodesic interpolation between those two hands is shown in Fig. 5.1. Each of the hand is represented by 200 landmark points in $\mathbb{R}^2$. In Fig. 5.1A, we only show 11 landmarks on those hands.



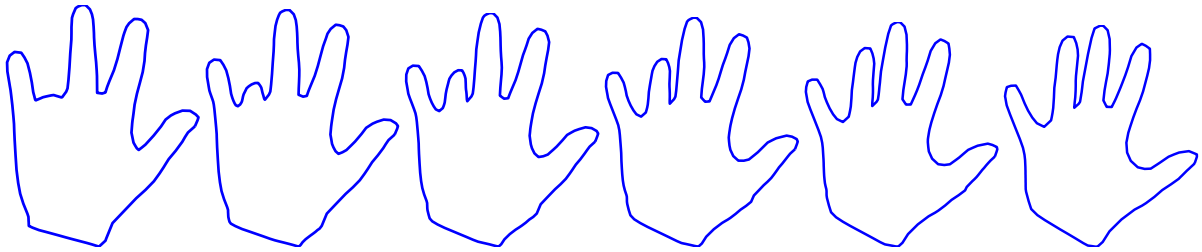
(A) Original Shapes with Landmarks



(B) Pre Shapes



(C) Final Alignment



(D) $t = 0$ (E) $t = 0.2$ (F) $t = 0.4$ (G) $t = 0.6$ (H) $t = 0.8$ (I) $t = 1$

Figure 5.1: A geodesic interpolation between two shapes in pre-shape space.

5.2 Statistical Models of Shape

Summary statistics are used to summarize a set of observations, in order to communicate the largest amount of information as simply as possible. Because of this property, mean shape and its variation of sample shapes have always been an important topic of shape analysis research.

In this section, we propose to build statistical models to capture characteristic variation within a given population of shapes. To achieve this, definition and calculation of the mean shape become necessary. Based on mean shape, we can define the variation as the spread around the mean shape. In this section, we will give the definition of the Frechet mean shapes and propose algorithm to compute it. The numerical method we used here is based on an attracting fixed point method, which is extended from Huckemann and Ziezold's method [27]. More details can be found in Liu et al.[38]. As in Dryden and Mardias work [10], to realize PCA on the non-linear pre-shape space, we approximate the problem by using its tangent space. The idea is: first project each pre-shape to a tangent plane at the mean shape; then perform PCA on this tangent plane; and finally project the principal components back to the pre-shape space. Since the biological shape data supposed to be analyzed tends to be concentrated around the mean, which suggests that the tangent plane is a good approximation of the local geometry on the pre-shape space. The tangent space PCA provides an effective way to analyze the main modes of variation within a population of shapes.

5.2.1 Mean Shape

Let us consider the normal meaning of mean in Euclidean space. Let $x_1, \ldots, x_n$ be n data samples in $\mathbb{R}^m$, and their arithmetic mean is given by

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{5.15}$$

The mean can also be interpreted as the minimizer of the scatter function $V : \mathbb{R}^m \to \mathbb{R}$, where $V(x)$ is defined as

$$V(x) = \sum_{i=1}^{n} ||x - x_i||^2 \tag{5.16}$$

In the following, we show in details that μ is the minimizer of $V(x)$:

$$\begin{aligned}
V(x) &= \sum_{i=1}^{n} ||x - \mu - (x_i - \mu)||^2 \\
&= \sum_{i=1}^{n} < x-\mu, x-\mu > -2\sum_{i=1}^{n} < x-\mu, x_i-\mu > + \sum_{i=1}^{n} < x_i-\mu, x_i-\mu > \\
&= n||x-\mu||^2 - 2 < x-\mu, \sum_{i=1}^{n}(x_i-\mu) > + \sum_{i=1}^{n} ||x_i-\mu||^2 \qquad (5.17) \\
&= n||x-\mu||^2 + \sum_{i=1}^{n} ||x_i-\mu||^2 \qquad (5.18)
\end{aligned}$$

The reason Eq. 5.18 and Eq. 5.18 are equal is that, from the definition of μ showing in Eq. 5.15, $\sum_{i=1}^{n}(x_i-\mu)$ is the zero vector in $\mathbb{R}^m$. Since both terms in Eq. 5.18 are nonnegative and $\sum_{i=1}^{n}||x_i-\mu||^2$ is a fixed number for given sample. Thus, V is minimal when $n||x-\mu||^2=0$. This requires $x=\mu$.

We show in the following how the scatter function can be extended to shape space. Let $p_1,\ldots,p_n$ be shapes in pre-shape space $\Omega(k,n)$. Given a pre-shape p, let $U_i(p)$ be the orthogonal transformation that optimally aligns p_i to p. The mean shape p of the family of shapes is a shape that minimizes the total scatter function.

$$V(p)=\frac{1}{2}\sum_{i=1}^{n}\tilde{d}^2(p,p_i) \tag{5.19}$$

$\tilde{d}$ is the shape distance as defined in Eq. 5.13, and the 1/2 is added just for computational convenience. We plug $\tilde{d}$ into Eq. 5.13, $V(s)$ can be written as

$$V(s)=\frac{1}{2}\sum_{i=1}^{n}\arccos^2<p,U_ip_i> \tag{5.20}$$

Then the unconstrained gradient of V is given by

$$\nabla V(p)=-\sum_{i=1}^{n}\frac{\arccos(\zeta_i(p))}{\sqrt{1-\zeta_i^2(p)}}U_ip_i \tag{5.21}$$

where $\zeta_i =< p,U_ip_i >$. At a minimum of V restricted to the pre-shape space, we must have $\nabla V(p)=\lambda p$, the λ here is also the projection of $\nabla V(p)$ onto p, we get

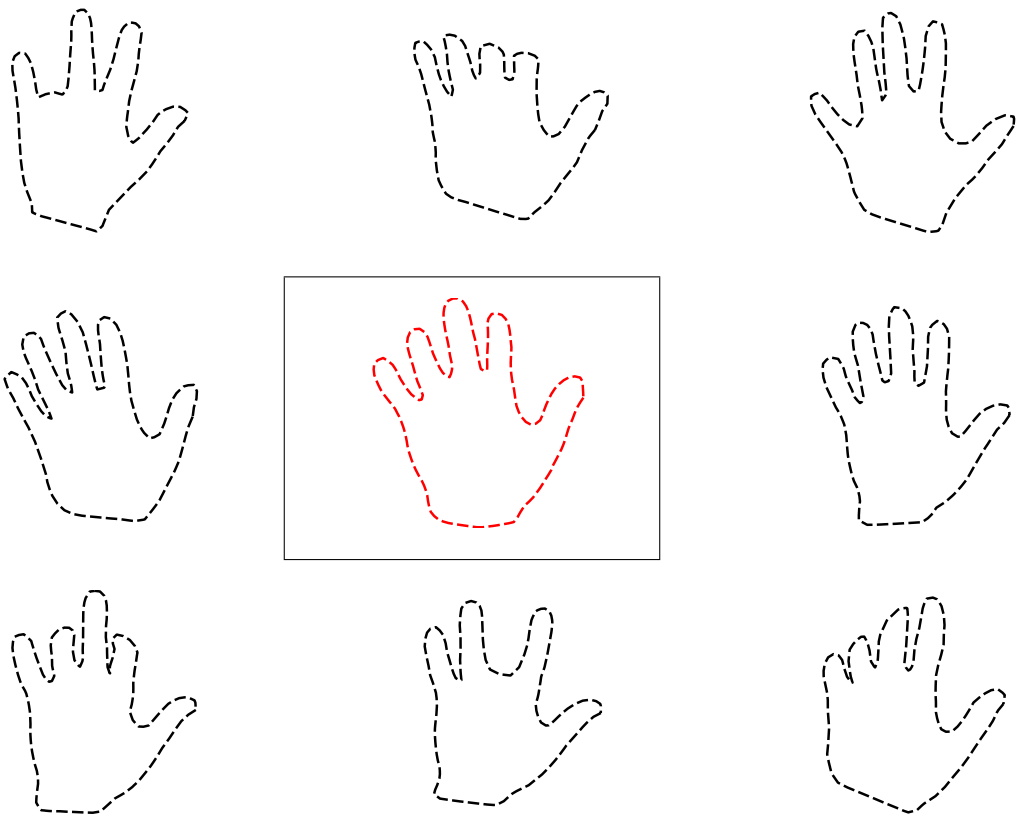
$$\lambda =<\nabla V(p),p>=-\sum_{i=1}^{n}\frac{\zeta_i(p)\arccos(\zeta_i(p))}{\sqrt{1-\zeta_i^2(p)}} \tag{5.22}$$

At a minimum, it requires

$$p=\text{sign}(\lambda)\frac{\nabla V(p)}{||\nabla V(p)||_F} \tag{5.23}$$

Thus, if a function $T:\Omega\to\Omega$ defined by $T(p)=\text{sign}(\lambda)\frac{\nabla V(p)}{||\nabla V(p)||_F}$ as they lead to the stable minimal of V. The strategy of moving towards the fixed point by an iteration of T, initializing the procedure with a pre-shape p, say, one of the pre-shape in the given family. For a given threshold value $\epsilon>0$, one calculates $T(p),\ldots,T^n(p)$ iterating until $||T^n(p)-T^{n-1}(p)||_F<\epsilon$.

Table 5.1: Eight hands and their mean (in frame).



5.2.2 Tangent Space PCA

The algorithm of PCA in tangent space are list below.

1. Given n shapes, compute pre-shapes respectively.
2. Suppose we have n samples in pre-shape space, we can calculate the mean shape p.
3. Calculate the W_i for $i = 1, \ldots, n$

$$W_i = \frac{U_i p_i - < p, U_i p_i > p}{||U_i p_i - < p, U_i p_i > p||_F} \tilde{d}(p, p_i) \tag{5.24}$$

From a geometric point of view, W_i is the tangent vector at p on the great circle, as shown in Fig. 5.2B. $< p, U_i p_i > p$ is the projection of $U_i p_i$ onto mean shape p; W_i is in the direction of $U_i p_i - < p, U_i p_i > p$, with the length $\tilde{d}(p, p_i)$. Thus, we can see the W_i as the velocity vector between p and p_i.

Table 5.2: Nine registered horses and their mean (in frame).

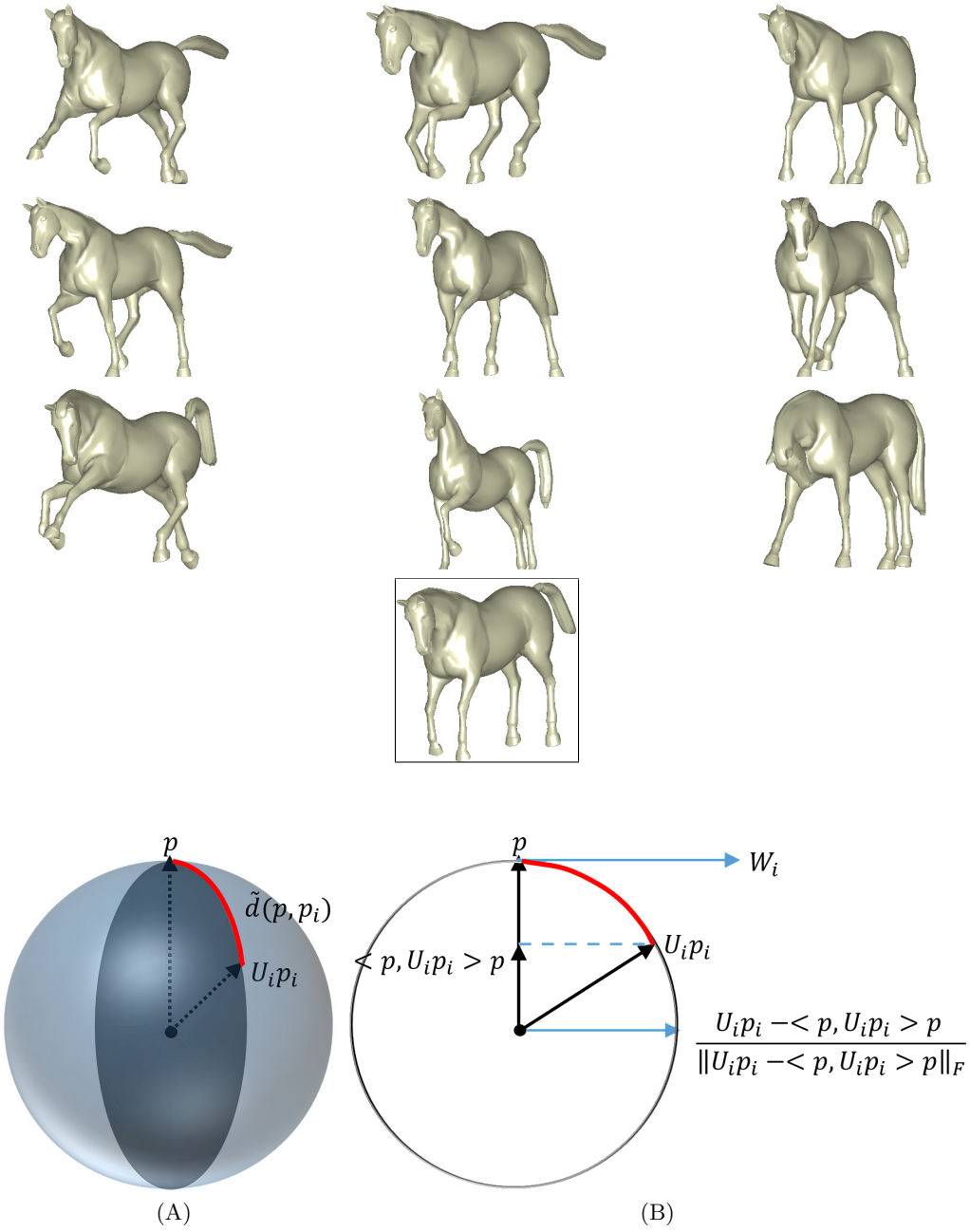


Figure 5.2: Tangent space projection. (A) p_i is optimally aligned to the mean shape p, the distance between p and p_i is the length of the shortest path (red in Fig. 5.2A) among all the paths connect these two shapes within pre-shape space. (B) The great circle through mean shape p and $U_i p_i$.

4. Let $X = [W_1, \ldots, W_n]$, if we compute XX^T, then this will be a large matrix, the dimension is the number of corresponding points. An alternative way is to compute X^TX, because the matrices XX^T and X^TX share the same nonzero eigenvalues, and if u_i is the eigenvector of XX^T, v_i is the eigenvector of X^TX, and they both correspondent to the eigenvalue σ_i^2. One can check these eigenvectors are related in identity $u_i = Xv_i$.

   So we compute $K = X^TX$, the $K_{ij} =< W_i, W_j >$.

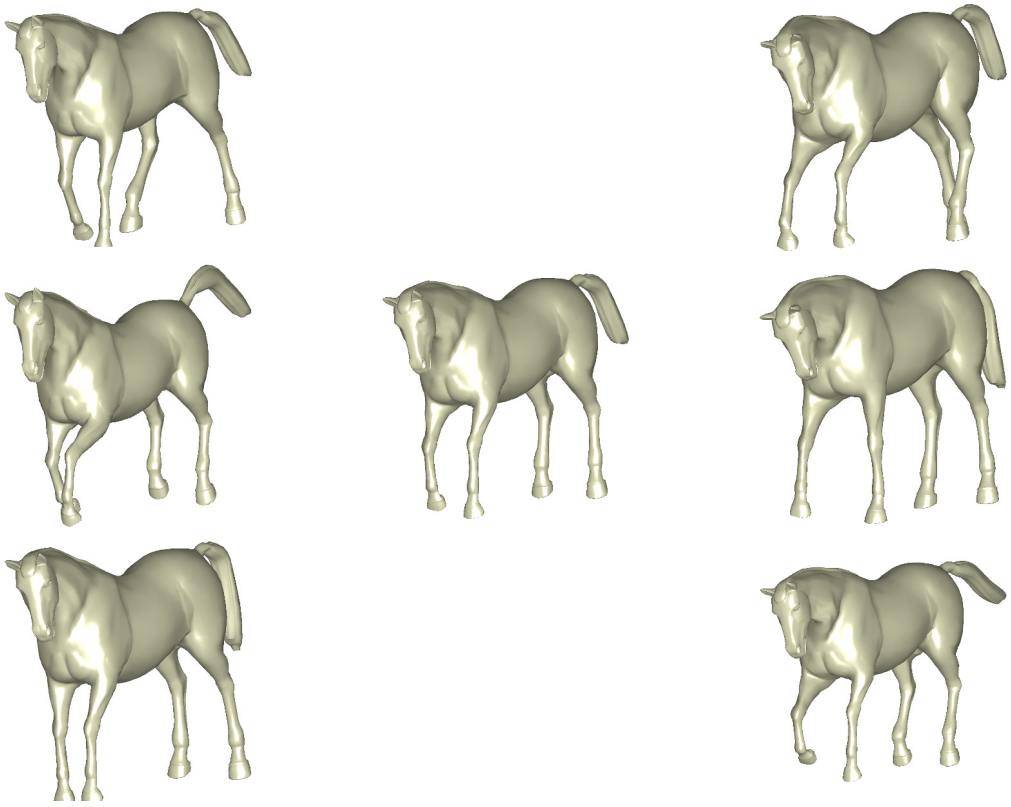
5. Diagonalize X^TX, here $X^TX = V\Sigma V^T$, the Σ is a diagonal matrix with a decreasing order σ_i^2 on the main diagonal. So the ith principle component PC_i can be computed by XV_i, V_i is the ith column of V. Where σ_i^2 is the variation along the direction of PC_i, and σ_i is the standard deviation.

6. Project the PCs back to the pre-shape space by

$$p_{\alpha\sigma_i} = \cos(\alpha\sigma_i)p + \sin(\alpha\sigma_i)\frac{PC_i}{||PC_i||_F} \tag{5.25}$$

   So the distance $\tilde{d}(p_{\alpha\sigma_i}, p) = \alpha\sigma_i$.

Table 5.3: Horse shape variation along the first three principal directions.



	Mean-1SD	Mean	Mean+1SD
PC1			
PC2			
PC3			

CHAPTER 6

CORRELATIONS BETWEEN THE MORPHOLOGY OF SONIC HEDGEHOG EXPRESSION DOMAINS AND EMBRYONIC CRANIOFACIAL SHAPE

Quantitative analysis of gene expression domains and investigation of relationships between gene expression and developmental and phenotypic outcomes are central for advancing our understanding of the genotype-phenotype map. Gene expression domains typically have smooth but irregular shapes lacking homologous landmarks, making it difficult to analyze shape variation with the tools of landmark-based geometric morphometrics. In addition, 3D image acquisition and processing introduce many artifacts that further exacerbate the problem. To overcome these difficulties, we present a method that combines optical projection tomography scanning, a shape regularization technique and a landmark-free approach to quantify variation in the morphology of Sonic hedgehog expression domains in the frontonasal ectodermal zone (FEZ) of avians and investigate relationships with embryonic craniofacial shape. The model reveals axes in FEZ and embryonic-head morphospaces along which variation exhibits a sharp linear relationship at high statistical significance. The technique should be applicable to analyses of other 3D surface-like biological structures that have ill-defined shape and are relevant to understanding developmental processes and phenotypic variation.

6.1 Biological Background and Introduction

A current line of thought is that the process of embryonic development acts to structure and modulate genetic and phenotypic variation in ways that influence how natural selection can act on that variation to produce morphological change. However, it is unclear how various developmental processes generate and structure variation [66, 19, 22, 18]. Developmental processes can be identified and characterized by specific gene expression patterns, and are often visualized through mRNA or protein localization. To uncover relationships between developmental processes and phenotypic variation, spatial patterns of protein and mRNA expression are being systematically recorded over

a range of spatial and temporal scales in several animal systems [34, 50, 65, 37, 61, 70]. However, quantitative studies have focused primarily on building atlases for studies of variation of gene expression level and relationships between different genes [37, 15, 13]. We focuses on methodology for investigation of a different facet of this problem, how variation in the morphology of gene expression domains relates to developmental and phenotypic outcomes. Here, we develop a morphometric method to quantify 3D shape variation in Sonic hedgehog (*Shh*) mRNA expression domains in avians (chickens and ducks) and explore relationships with embryonic facial shape. The focus here is on methods and a more thorough discussion of the experimental model and its biological context appears elsewhere [26].

Signaling by Sonic hedgehog plays an essential role in the development of the vertebrate upper jaw [25, 39, 23, 75]. In amniotes, including mice and avians, *Shh* is first expressed in the forebrain prior to outgrowth of the facial prominences. As neural crest cells migrate into the midface, *Shh* expression is activated in the frontonasal ectodermal zone (FEZ), which acts as a signaling center that controls growth of the upper jaw [39]. Hu and Marcucio [23] demonstrate empirically that spatial organization of the FEZ regulates morphological variation in the developing upper jaw. The methods of this paper let us uncover quantitative relationships between the morphology of *Shh* mRNA expression in the FEZ and embryonic facial shape.

Quantification of shape variation in gene expression domains poses particularly challenging problems, as these domains typically have no clearly defined forms, often appearing seemingly amorphous, as illustrated in Figs. 6.1C and 6.1D. In particular, 3D morphometrics based on landmarks [29, 30] is not easily applicable to this problem, severely limiting the effectiveness of some existing methods of statistical shape analysis [36, 10]. Another layer of difficulty is related to image acquisition. The geometric 3D meshes representing gene expression domains tend to be noisy and contain multiple local topological and geometrical defects such as holes and irregularities that are not really present in the tissues. For these reasons, our approach has two key components:

(i) *A Shape Regularization Technique* The FEZ is a thin, surface-like structure. A difficulty in fitting a smooth surface model to a 3D FEZ image is that FEZes among organisms lack homologous landmarks. If landmarks were available, we could construct a smooth template and morph the template to fit the image using the landmarks as guides. Dense template morphing can be done with such techniques as thin-plate spline (TPS) interpolation [11, 41]. Section 6.3 describes a method that combines TPS interpolation with probability density estimation to bypass landmarks and obtain smooth FEZ models that remove noise and enhance shape.

(ii) *FEZ Topography Vectors* Lack of well-defined form makes it difficult to develop statistical FEZ shape models using standard techniques. To obtain an informative model of shape variation, we use the relative position of the FEZ in the embryonic face. We introduce FEZ topography vectors that let us construct effective shape summaries that retain the most salient morphological features and filter out confounding details. A topography vector essentially describes how the FEZ height varies across its extension.

Using optical projection tomography scans of 17 specimens (7 chickens and 10 ducks), the shape regularization method and FEZ topography vectors, we show that there is a strong linear association between particular characteristics of FEZ morphology and embryonic craniofacial shape. Variation in craniofacial shape is quantified using geometric morphometrics based on 67 landmarks, covering the face, mouth, eyes, and forebrain. The landmarks are depicted in Fig. 6.1A and 6.1B. Analysis of *Shh* expression in the FEZ guided the development of the landmark-free morphometric technique, but the method should be useful in many other settings, particularly in analysis of irregular, surface-like shapes that lack landmarks.

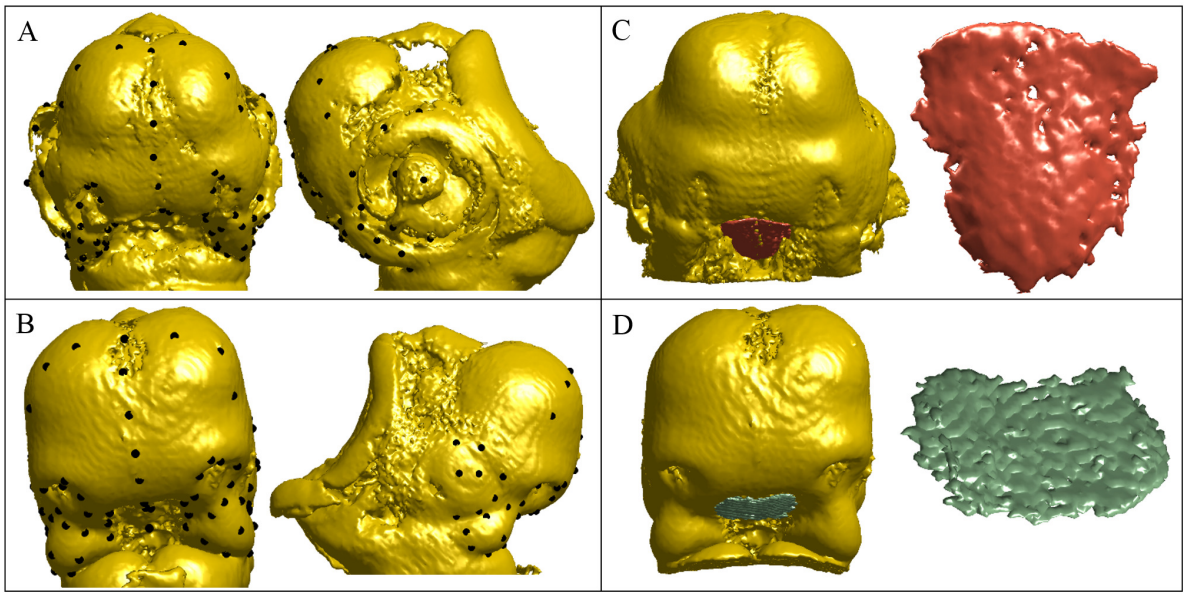


Figure 6.1: The visualization of landmarks on the embryonic head and the visualization of FEZ shape. (A, B) Frontal and side views of landmarks on embryonic head of a chicken and a duck. (C, D) Embryonic heads of avians and close-up view of *Shh* expression domains in the FEZ.

6.2 Experimental Procedures

6.2.1 Embryo Preparation

To determine the extent to which FEZ spatial organization differs among the two bird species, fertilized eggs from white Leghorn chicken (Gallus gallus) and White Pekin duck (Anas platyrhynchos) were incubated at 37^{o}C in a humidified chamber until stage-matched at HH [1] 22, at which time *Shh* expression in the surface ectoderm is robust, and then performed whole mount in situ hybridization to examine *Shh* expression in the FEZ.

6.2.2 Forebrain Transplantation

Our goal was to test the intrinsic ability of the forebrain to direct unique expression patterns of *Shh* in the FEZ to control divergent facial morphology. To achieve this goal, fertilized eggs were incubated at 37^{o}C in a humidified chamber until stage-matched at HH 7/8 prior to neural crest cell emigration. Eggs were prepared for surgery as described in [24]. To develop this method, the entire basal portion of the prosencephalon from donor duck embryos were removed at HH 7 and transplanted orthotopically into stage-matched chicken hosts. Tissue grafts of the basal forebrain were harvested from stage 7/8 embryos using sharpened tungsten needles. The grafts, measuring 0.3 mm in height by 0.2 mm in width, were transferred into DMEM containing Neutral Red (23^{o}C, 2 minutes), which was added to facilitate visualization when transferred to the host. Care was taken to avoid excessive disruption of underlying endoderm. The donor grafts were positioned to replace the extirpated tissue. For the experiments duck embryos were used as donors and chicken embryos were used as hosts. To examine the effects of the surgery, as controls, chick-chick chimeras were also created. Both of these species have more similar rates of brain growth. Chimeric embryos were incubated for 48 and 72 hours, and to day 6 or 10 for molecular, cellular, histological, and morphological analysis.

6.2.3 In Situ Hybridization and Optical Projection Tomography (OPT) Imaging

Shh expression in the avian embryos was detected by standard in situ hybridization resulting in the domain of interest being stained dark blue [26, 9]. 3D data of both this domain, as well

[1] Indevelopmental biology, theHamburgerHamilton stages(HH) are a series of 46 chronological stages inchickdevelopment, starting from laying of theeggand ending with a newly hatched chick. It is named for its creators, Viktor Hamburger and Howard L. Hamilton.

as exterior surface of the avian embryo were acquired on Bioptonics 3000 OPT system (Sky Scan, Germany). OPT has the advantage over more conventional imaging methods that it can generate 3D surface data of 1-3 cm^3 objects as well as detect colorimetric or fluorescent signals. It therefore, allows the correlation of multiple data types [54, 58]. Embryos were imaged on the OPT system as previously published [54]. Briefly, embryos were embedded in 1% low melt agarose (Invitrogen), which was then cut into a hexagonal block, mounted onto a magnetic chuck dehydrated for 48 hours in methanol and cleared in BABB (2 parts Benzyl Benzoate: 1 part Benzyl Alcohol). Embryos were imaged in the UV range on the GFP channel (480 nm) and in the visual light range. The NRecon software package (SkyScan, Germany) was used to align the stacks and reconstruct the images. Image stacks were imported into Amira (Version 5.0, FEI, Hillsboro OR, USA) for landmark placement and segmenting of the *Shh* FEZ domain. Sixty-seven landmarks were registered on each embryo, as shown in Figs. 6.1A and 6.1B.

6.3 Quantitative Methods

In this section we develop (i) a shape regularization technique that is used to construct smooth surface models of *Shh* expression domains in the FEZ and (ii) FEZ topography vectors that capture the most salient morphological characteristics of *Shh* expression and yet are robust to uninformative details. Henceforth, we refer to an expression domain in the FEZ simply as FEZ.

6.3.1 Shape Regularization

As illustrated in Figs. 6.1C and 6.1D, FEZ meshes acquired through 3D imaging are very irregular with numerous artifacts, whereas the FEZ itself resembles a smooth surface. To remove these irregularities and make the meshes more tractable, we build smooth surface models from imaging data. We approximate a FEZ mesh K by the graph of a smooth function f defined over a plane region D, as indicated in Fig. 6.2A. We begin with the construction of a plane P that contains D and then proceed to the estimation of D and construction of f.

Viewing the vertices $v_1, \cdots, v_n$ of the mesh K as n data points in 3D space, we use principal component analysis (PCA) to construct a plane P parallel to the first two principal directions, as indicated in Fig. 6.2A. The usual practice is to choose P containing the mean $\bar{v} = (v_1 + \cdots + v_n)/n$, but for the present purposes the plane may be translated off the mean. This is done primarily to

facilitate visualization. Orthogonal projection of $v_1, \cdots, v_n$ onto P gives a point cloud $p_1, \cdots, p_n$ that delineates a region D in the plane P, as shown in Fig. 6.2A. We now describe a procedure to estimate the domain D from these points.

We assume that the vertices $v_1, \cdots, v_n$ produced by the imaging and segmentation processes are n independent and identically distributed (i.i.d.) samples from a distribution on the FEZ surface. Thus, $p_1, \cdots, p_n$ may be viewed as i.i.d. samples from the corresponding distribution on the plane P induced by projection. We estimate this distribution using the Gaussian density estimator [52, 56].

$$\phi(p) = \frac{1}{n2\pi\sigma^2} \sum_{i=1}^{n} \exp\left(-\frac{\|p - p_i\|^2}{2\sigma^2}\right) \tag{6.1}$$

ϕ is the uniform mixture of isotropic Gaussians of width σ centered at the points p_i. For a discussion of selection of the bandwidth parameter σ, one may consult, for example, [59]. A key observation is that ϕ is large within the domain D relative to the values it attains outside D because only the interior of D is well populated by (projected) data points. Thus, the contour ∂D of the region D comprises points where a transition occurs. This suggests that we estimate ∂D as an isocontour of ϕ, that is, $\partial D = \{p \in P \colon \phi(p) = \epsilon\}$, where $\epsilon > 0$ is a fixed small value learned from data. In the unlikely event that the isocontour consists of multiple curves, we take the component that encloses the most points. Fig. 6.2B shows an example of a contour obtained with this technique.

The final step in the mesh regularization process is the construction of a smooth function f on the domain D whose graph interpolates the points $v_1, \cdots, v_n$. We introduce a coordinate system, where P becomes the $x-y$ plane and the $z-$ axis is orthogonal to P, as indicated in Fig. 6.2A. Let $v_i = (a_i, b_i, c_i)$ be the coordinates of the vertex v_i, so that the $(x, y)-$ coordinates of its projection p_i are (a_i, b_i). The goal is to construct a smooth function $f(x, y)$ such that $f(a_i, b_i) \approx c_i$, which ensures that the graph of f smoothly interpolates the data points. We use thin-plate spline (TPS) interpolation to construct such a function. TPS interpolation is a technique widely used in data analysis, including geometric morphometrics [7]. In simple terms, a TPS interpolation balances out minimization of the average residual $\sum_i |f(a_i, b_i) - c_i|^2/n$ and the smoothness of f. One may consult [69] for further details. The TPS interpolant f is defined over the entire plane P, but the part of the graph over the domain D gives the desired smooth FEZ model. To discretize the model, we use the (restricted) Delaunay triangulation [17, 14] τ of the region D associated with the points

$p_1, \cdots, p_n$. We lift the vertices of τ to 3D space via the TPS interpolant f and construct a mesh with the same connectivity as τ. Fig. 6.2C shows a FEZ model constructed with this method.

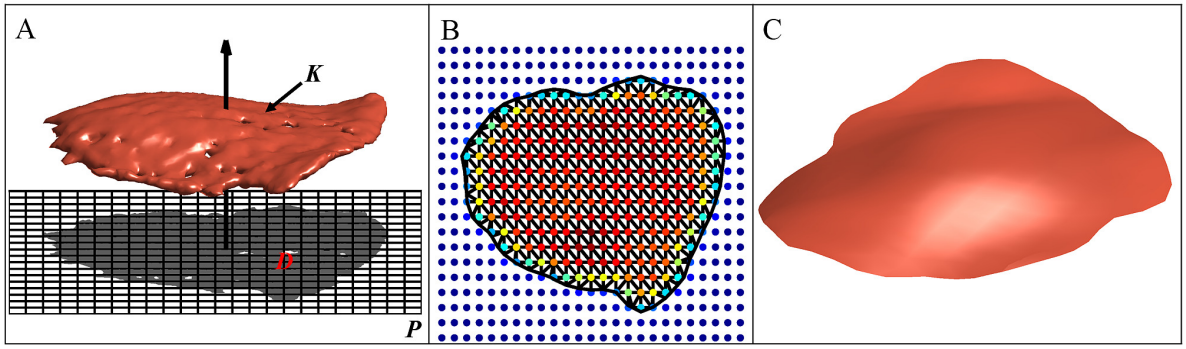


Figure 6.2: FEZ regularization process: (A) Original FEZ mesh K and projection D onto the plane P spanned by first two PCs. (B) Estimation of interpolation domain with colormap of the density value. The wire shows the triangulation of the estimated domain. (C) Regularized FEZ.

6.3.2 FEZ Topography Vectors

We develop a quantitative representation, termed FEZ topography vector, which summarizes the most salient morphological properties of the FEZ and to a large extent is blind to confounding details. This yields a representation that is robust to the large variability observed in local and regional FEZ morphology. Although the lack of homologous landmarks makes it difficult to find point correspondences between FEZ meshes for different specimens, a relaxed notion of shape correspondence that exploits the position of the FEZ in the embryonic head is implicit in topography vectors.

Using 67 landmarks, depicted in Figs. 6.1A and 6.1B, covering the face, mouth, eyes, and forebrain, we normalize centroid size and employ Procrustes superimposition to standardize position and spatial orientation of an embryonic head by aligning it to a template. In particular, this fixes a scale and orientation for the FEZ. We construct a sagittal plane, as shown in Fig. 6.3A, and use parallel translates of this plane to slice up the FEZ along a series of curves. These curves are used in the construction of the FEZ topography vector. To estimate the sagittal plane, we exploit the following facts: (i) the head landmarks are nearly symmetrical about the sagittal section and (ii) the dominant spatial spread of the head landmarks occurs in a direction perpendicular to the sagittal

plane. Thus, PCA on the landmarks gives a simple way of estimating the sagittal plane as the plane through the centroid of the landmarks that is parallel to the second and third principal directions since we expect the first principal direction to be orthogonal to the sagittal plane. Sweeping the FEZ from left to right with parallel translates of the sagittal plane yields a continuous family of sections of the FEZ by spatial curves, as shown in Figs. 6.3B and 6.3C. As the shape of these spatial curves is still sensitive to the high variability of the FEZ across specimens, we only use their lengths to describe FEZ morphology. In this manner, we obtain FEZ shape descriptors that encode how the "height" of the FEZ varies as we sweep it from left to right. As this variation is gradual, in practice, it suffices to consider a discrete, sparse family comprising k (equally spaced) sections across the width of the FEZ. Denoting the length of the ith section by h_i, FEZ shape is summarized in a FEZ topography vector h, the $k-$ dimensional vector whose coordinates are h_i, $1 \leq i \leq k$.

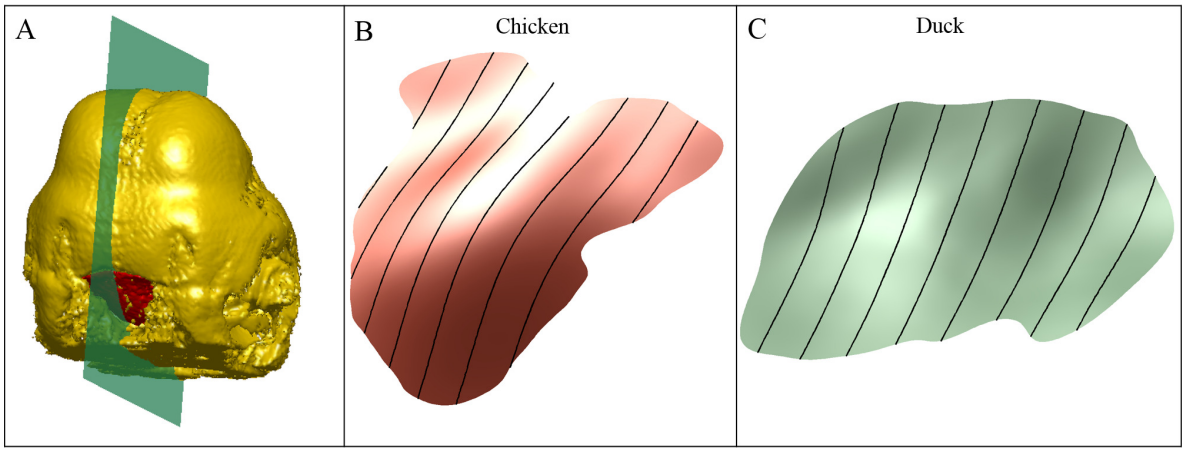


Figure 6.3: Sectioning the FEZ: (A) Sagittal plane of an embryo head (green plane). (B) FEZ sections by translates of the sagittal plane of a Chicken FEZ. (C) FEZ sections by translates of the sagittal plane of a duck FEZ.

6.4 Correlations between FEZ Morphology and Craniofacial Shape

After eliminated individuals obviously damaged during preparation for OPT, the sample for 3D morphometric analysis consisted of 10 duck embryos, 7 chicken embryos, 11 duck-chick transplants

and 14 chick-chick transplants.

We employed FEZ regularization and topography vectors to model variation in FEZ morphology and to study correlations with embryonic craniofacial shape. The analysis was based on optical projection tomography scans of the embryonic heads of 7 chickens and 10 ducks. FEZ meshes were smoothed with the regularization technique of Section 6.3.1 and FEZ shape variation was quantified using topography vectors. Head shape variation was modeled with standard techniques of geometric morphometrics using sixty-seven manually placed landmarks. The landmarks cover the face, mouth, eyes, and forebrain, and are depicted in Fig. 6.1.

6.4.1 Modeling Native Groups

We employed topography vectors of dimension $k = 8$ to model variation in FEZ morphology. We adopted a sparse representation with only eight FEZ sections because topographical variation across the FEZ is gradual and experiments indicated that little additional information relevant to this analysis is obtained with denser samplings. Figs. 6.3C and 6.3D show examples of FEZ sections used in the construction of topography vectors for a chicken and a duck.

Principal component analysis showed that PC1 and PC2 explain 92% of FEZ topography variation. Fig. 6.4A shows a plot of the PC scores. The figure also indicates that the PC1 and PC2 scores discriminate chickens from ducks sharply. An examination of the PC loadings revealed that PC1 is primarily about FEZ height at its center and PC2 about height gradient from the center to the left and right ends. To quantify variation in head shape, we standardized centroid size and used Procrustes superimposition to spatially align all head meshes to the mean head of the entire group. The mean was calculated with the fast converging fixed-point algorithm developed by Liu et al.[38]. PC1-PC6 explained approximately 80% of the variation. Fig. 6.4B shows a plot of the first two PC scores. PC1 sharply discriminates chickens from ducks and reflects the fact that embryonic duck heads are longer and narrower than chicken heads.

To explore relationships between FEZ morphology and embryonic facial shape, we used canonical correlation analysis (CCA) [21, 57, 33] on the FEZ morphospace determined by the first two PC scores and head shape space determined by the first six PC scores. CCA uncovers a pair of directions and in the FEZ and head morphospaces, respectively, such that variation along these axes has maximal correlation coefficient. CCA produced a pair of axes along which the correlation coefficient is $\rho = 0.96$ at high significance ($p = 0.0012$). Fig. 6.4C shows a plot of the scores

along these axes, the regression line of head shape over FEZ topography, as well as illustrations of shape variation along these directions. The first canonical direction in the FEZ morphospace nearly coincides with the anti-diagonal direction in the PC1-PC2 plane, whereas the axis in the head morphospace captures the fact that the heads of ducks are narrower, deeper and longer than the heads of chickens. The scores along these directions yielded FEZ and head shape signatures that sharply discriminate chickens and ducks. For the second pair of canonical directions, the correlation coefficient dropped to $\rho = 0.5$, revealing no additional statistically significant correlations between FEZ topography and head shape.

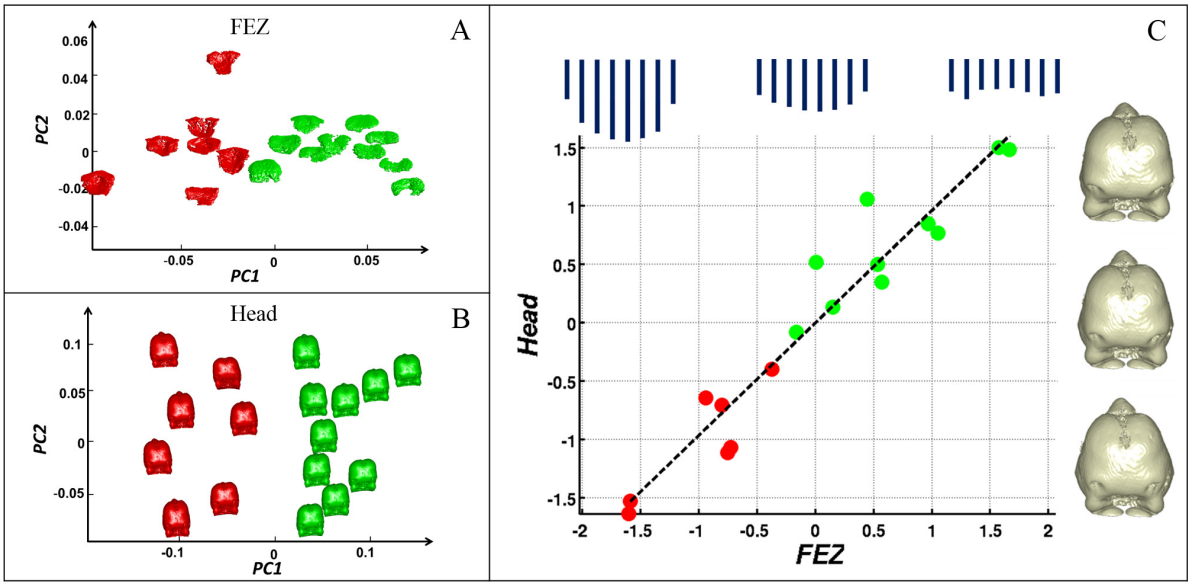


Figure 6.4: 3D Morphometric analysis of FEZ and craniofacial shape based on optical projection tomography imaging of native groups (chicken in red and duck in green). (A) PC scores for FEZ morphology. PC1 captures 81% of the variation, PC2 captures 11% of the variation. (B) PC scores for head shape. PC1 captures 35% of the variation, PC2 captures 17% of the variation. (C) First pair of canonical directions in FEZ and head morphospaces. The first canonical direction in the FEZ morphospace nearly coincides with the anti-diagonal direction in the PC1-PC2 plane, whereas the axis in the head morphospace captures the fact that the heads of ducks are narrower, deeper and longer than the heads of chickens.

6.4.2 Mapping the Chimeras

We mapped the chimeras onto the FEZ V. Head CCA plot as Fig.6.4C. In Fig.6.5A, the axes are the first pair of the canonical directions whose construction only involves the native chick and duck

group. The native chick specimens are indicated in red, the native duck specimens are indicated in green, the duck-chick chimeras are indicated in purple, and the chick-chick chimeras are indicated in yellow.

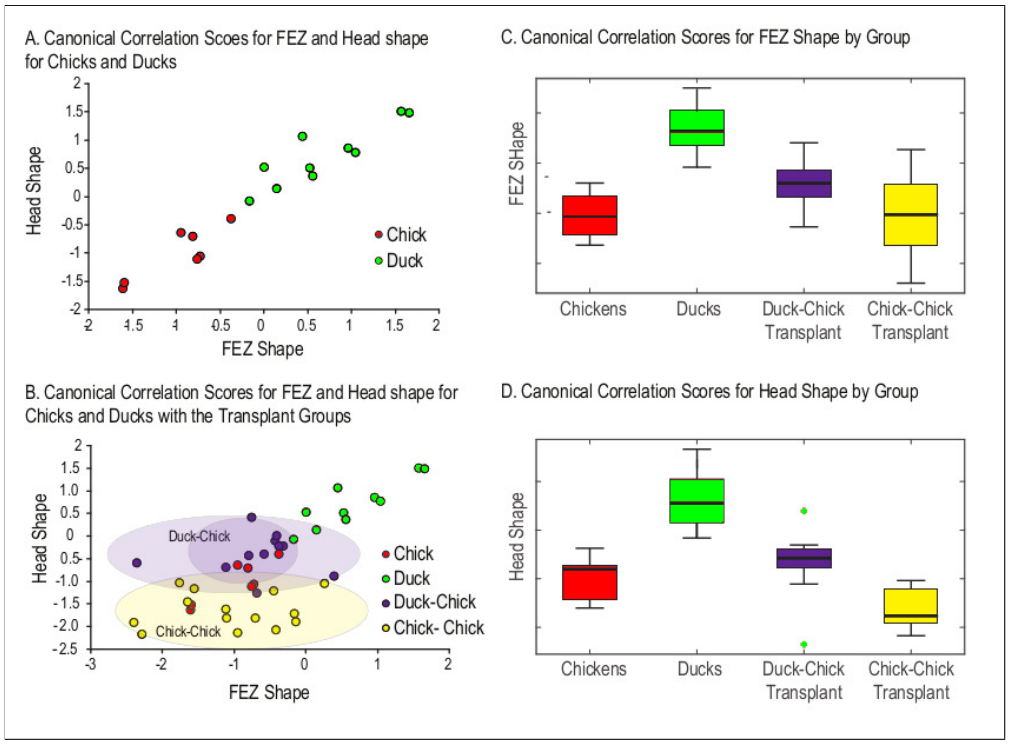


Figure 6.5: 3D Morphometric analysis of FEZ and craniofacial shape based on optical projection tomography imaging on native specimens and chimeras (chicken in red, duck in green, duck-chick chimera in purple and chick-chick chimera in yellow). (A) Canonical correlation scores for chick and duck embryos for FEZ and head shape. This plot show the clear separation of both FEZ and head shape in duck and chick embryos as well as the correlation between *Shh* expression in the FEZ and head shape. (B) Here the two hemi-forebrain transplant groups (duck-chick and chick-chick) are added to the data shown in graph A. Morphometric analysis for the transplants is performed only on the transplant side. Here, the clear separation of facial shape among these groups is shown with the duck-chick transplant group shifted significantly towards the duck group. (C) and (D) show the medians and dispersions of the canonical correlation scores for FEZ and head shape.

This revealed clear separation of facial shape for the two transplant groups (t-test, df=24, p=0.008). Procrustes permutation tests in MorphoJ confirmed this result with head shape differing significantly between the two transplant groups (p <0.001). The duck-chick transplant group is moved in the direction of the duck group although it was significantly different form both the duck

and chick group. The chick-chick is closer to the chick group (Procrustes distance = 0.12) than to the duck group (Procrustes distance = 0.21). Both transplant groups, though, appear shifted towards the lower end of the head shape range in plot 6.5A such that both overlap the range of variation in the chick group. The chick-chick group likely differs from the chick group due to the perturbing effect of the transplant surgery. While FEZ shape was not significantly different between the transplant groups, the shape of the FEZ in the duck-chick chimeras was shifted in the direction of the duck shape. The 3D data suggest that transplanting duck forebrain into a chick embryo moves facial morphology significantly in the direction of the duck.

CHAPTER 7

SPLINE MODEL ON MANIFOLD DOMAIN

Splines are widely used in shape analysis and also in many other disciplines for interpolation, approximation and regression, but most experimental results have focused on Euclidean domain [68]. Although the theory of spline methods has been previously generalized to the flat torus and the standard sphere [11, 41, 67], as well as to closed compact Riemannian manifold [31], on arbitrary domain, these authors have not presented numerical methods for computing such splines. Since manifold-valued data occur in numerous problems, constructing splines with manifold domains and providing effective computational methods are desirable.

In this chapter, we present the spline method, which naturally extend the concept of the popular thin plate spline (TPS) [68, 7] to compact Riemannian manifold domains. The key approach is to use mathematical framework of reproducing kernel Hilbert space, along with integrating spectral geometry associated with compact Riemannian manifolds. Incorporate with Neumann boundary condition, we propose an computational scheme based on a bounding box. The efficiency of our spline method is proved by comparing the interpolation results with TPS on closed and open surfaces.

This spline method has also been applied to construct dense surface model of avian embryos. More details of this application can be found in section 7.6. Those dense surface models can establish a correspondence of thousands of points across each 3D embryo image, so that it can take full advantages of the high resolution geometrical information. This models also carry the potential for precisely identifying the local shape features and the syndrome effects that can benefit a series of following up studies.

7.1 Model Derivation

7.1.1 Preliminaries on Reproducing Kernel Hilbert Space (RKHS)

To setup the problem, let M be a compact Riemannian manifold. We assume that we are given n observations $\{x_i, y_i\}, i = 1, \cdots, n$, with $x_i \in M$ and $y_i \in \mathbb{R}$. The y_i are function values at location

x_i. Our goal is to infer "best" transformation f between $\{x_i\}$ and $\{y_i\}$. In order to define "best", we introduce the energy function

$$E(f) = \frac{1}{n}\sum_{i=1}^{n} |y_i - f(x_i)|^2 + w\int_M (\Delta^p f)^2 dM \tag{7.1}$$

Here Δ is the Laplace-Beltrami operator taking functions on M to functions and defined as $\Delta f = -\text{div grad}(f)$. p is an integer with value ≥ 1.

Definition 7.1.1 *A symmetric, real-valued function $K(s,t)$ of two variables $s,t \in M$ is said to be positive definite if, for any real $a_1, \cdots, a_n$, and $t_1, \cdots, t_n \in M$*

$$\sum_{i,j=1}^{n} a_i a_j K(t_i, t_j) \geq 0$$

that is, if every square matrix $[K(t_i,t_j)]_{i,j=1}^n$ is positive definite matrix.

The minimizer of the E defined in Eq. 7.1 gives the "best" transformation f between $\{x_i\}$ and $\{y_i\}$. To find this minimizer we first need to introduce a search space for f. Because continuous linear functional of point-wise evaluation is necessary in interpolation problems, it is natural to consider reproducing kernel Hilbert space (RKHS) in which point-wise evaluation is a continuous linear functional. Based on the Riesz representation theorem, for a given a positive definite function $K(.,.) : M \times M \to \mathbb{R}$, we can associate it with a unique RKHS H, which is a Hilbert space of real valued function on M with the property that for each $t \in M$, the evaluation functional $L_t : H \to \mathbb{R}$ defined as $L_t(f) = f(t)$ is continuous.

The construction of the space H can be described as follows. We denote $K(t,.)$ as $K_t(.)$ for $t \in M$, we can see $K_t(.)$ is a function defined from $M \to \mathbb{R}$. We first construct a vector space H by taking all finite linear combination of the form in Eq. 7.2

$$\sum_{i=1}^{n} a_i K_{t_i} \tag{7.2}$$

for all choices of n, a_i, t_i.

Inner product in H is defined as

$$\left\langle \sum_{i=1}^{n} a_i K_{t_i}, \sum_{j=1}^{m} b_j K_{s_j} \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j K(t_i, s_j) \tag{7.3}$$

It is not hard to check this is a well defined inner product because of the positive definiteness of $K(.,.)$, and we can also check the validity of $f(x) =< f, K_x >$. To make this space complete, we can add in the limits of all Cauchy sequences over H into this space. To better show the dependency between kernel K and space H, henceforth, we will use H_k to represent the RKHS associate with kernel K.

Let M be a compact Riemannian manifold. Consider the integral operator $T : L^2(M) \to L^2(M)$ defined by

$$Tf(x) = \int_M K(x,t)f(t)dV(t) \tag{7.4}$$

If kernel K satisfies

$$\iint_M K^2(s,t)dV(s)dV(t) < \infty \tag{7.5}$$

Then there exists eigenvalues $0 \leq \lambda_0^k \leq \lambda_1^k \leq \dots$. By the fundamental theorem of self-adjoint compact operator [63, 64], the corresponding L^2 normalized eigenfunctions $\left\{\phi_i^K\right\}_{i=1}^{\infty}$ of operator T form an orthonormal basis for $L^2(M)$. For $\forall f \in L^2(M)$, f can be expressed in the form $f = \sum_{i=0}^{\infty} a_i\phi_i^k$, where $a_i^k = \int_M f(t)\phi_i^K(t)dV(t)$. Mercer's theorem [42] states that $K(s,t) = \sum_{i=0}^{\infty} \lambda_i^k\phi_i^k(s)\phi_i^k(t)$.

Then we can find the condition for f belongs to H_k and the relation between $||.||_{L^2}$ and $||.||_{H_k}$ by lemma 7.1.1.

Lemma 7.1.1 $f \in H_k \Leftrightarrow \sum_{i=0}^{\infty} \frac{\left(a_i^k\right)^2}{\lambda_i^K} < \infty$ *and* $||f||_{H_k}^2 = \sum_{i=0}^{\infty} \frac{\left(a_i^k\right)^2}{\lambda_i^k}$.

More detail of this proof can be found in [68].

7.1.2 Kernel for Spline Energy Function

From Hodge theorem for functions, if M is compact connected oriented Riemannian manifold, there exists a complete orthonormal set of $L^2(M, \Phi)$ consisting of eigenfunctions of the Δ. Suppose $\{\phi_i, \lambda_i, i = 0, \cdots, \infty\}$ are the eigenfunctions and eigenvalues associated with operator Δ, $\lambda_i \leq \lambda_{i+1}$, for $\forall i \geq 0$. All the eigenvalues are positive, except $\lambda_0 = 0$. Each eigenvalue has finite multiplicity, and the eigenvalues accumulate only at infinity. Thus we can express f in $L^2(M)$ as $f = \sum_{i=0}^{\infty} a_i\phi_i$, where $a_i = \int_M f(t)\phi_i(t)dV(t)$. From the fact that $\{\phi_i, \lambda_i, i = 0, \cdots, \infty\}$ are the normalized eigenfunctions and eigenvalues associated with operator Δ, we have $\int_M (\Delta^p f)^2 dM = \sum_{i=0}^{\infty} a_i^2\lambda_i^{2p}$.

Motivated by Lemma 7.1.1, the kernel defined as Eq. 7.6 satisfies $\int_M (\Delta^p f)^2 dM = ||f||^2_{H_k}$.

$$K(s,t) = \sum_{i=1}^{\infty} \frac{\phi_i(s)\phi_i(t)}{\lambda_i^{2p}} \tag{7.6}$$

Suppose $\iint_M K^2(s,t)dV(s)dV(t) = \sum_{i=1}^{\infty} \frac{1}{\lambda_i^{4p}} < \infty$ and $\sum_{i=1}^{\infty} a_i^2 \lambda_i^{2p} < \infty$, we can define a RKHS associate with this kernel K in Eq. 7.6 with norm $||f||^2_{H_k} = \sum_{i=1}^{\infty} a_i^2 \lambda_i^{2p}$. Since the null space of H_k is spanned by ϕ_0, for $\forall f \in L^2(M)$ satisfies $\sum_{i=1}^{\infty} a_i^2 \lambda_i^{2p} < \infty$, we can decompose this function into the form shows in Eq. 7.7

$$f(t) = \sum_{i=0}^{\infty} a_i \phi_i(t) = a_0\phi_0 + < K_t, f > \tag{7.7}$$

Using the inner product to rewrite Eq. 7.1, we have this simple linear system

$$E(f) = E(c_0, f) = \frac{1}{n}\sum_{i=1}^{n}(y_i - c_0 - < K_{x_i}, f >)^2 + w < f, f > \tag{7.8}$$

To find the minimizer f, we could take directional derivative for Eq. 7.8

$$\begin{aligned} \partial_f E(h) &= \frac{2}{n}\sum_{i=1}^{n}(y_i - c_0 - < K_{x_i}, f >) < -K_{x_i}, h > + 2w < f, h > \\ &= < h, \frac{2}{n}\sum_{i=1}^{n}(y_i - c_0 - < K_{x_i}, f >)(-K_{x_i}) + 2wf > \end{aligned} \tag{7.9}$$

In order to have $\partial_f E(h) = 0$ for $\forall h$, it is necessary to have $wf = \frac{1}{n}\sum_{i=1}^{n}(y_i - c_0 - < K_{x_i}, f >)K_{x_i}$. This implies f must be some linear combination of K_{x_i}. As we mentioned before, the constant function is in the null space of H_k and can not be represented by $\{K_{x_i}\}$. Put all those together, we have $f = c_0 + \sum_{i=1}^{n} c_i K_{x_i}$.

To determine f, the only parameters we need to find are c_0 and $\mathbf{c} = [c_i, c_2, \cdots, c_n]^T$. By substituting in the expression f into Eq. 7.8.

$$\begin{aligned} E(c_0, c) &= \frac{1}{n}\sum_{i=1}^{n}(y_i - c_0 - \sum_{j=1}^{n} c_j K(x_i, x_j))^2 + w\mathbf{c}^T\Sigma\mathbf{c} \\ &= \frac{1}{n}||Y - Tc_0 - \mathbf{c}\Sigma||^2 + w\mathbf{c}^T\Sigma\mathbf{c} \end{aligned} \tag{7.10}$$

Where $Y = [y_1, y_2, \cdots, y_n]^T$, Σ is a $n \times n$ matrix with (i, j) element $K(x_i, x_j)$, and T is a $n \times 1$ vector with 1_s.

A straightforward calculation for taking derivatives respect to c_0 and $\mathbf{c}$ shows that the minimizer c_0 and $\mathbf{c}$ will satisfy the following equations.

$$\begin{aligned} Y - Tc_0 - \Sigma\mathbf{c} &= nw\mathbf{c} \\ \mathbf{c}^T T &= 0 \end{aligned} \tag{7.11}$$

We can have the solution by solving the linear system Eq. 7.11. The optimal smoothing parameter w is chosen by generalized cross validation method introduced in [2].

7.2 Extension from Single Variate to Multivariate Problem

In shape analysis, the spline interpolation can usually be stated as a multivariate problem: Given n observations$\{x_i, \mathbf{y_i}\}$, with $x_i \in M$ and $\mathbf{y_i} \in \mathbb{R}^d$, finding a transformation $\mathbf{f} : M \to \mathbb{R}^d$ which satisfies the smoothing interpolation conditions can be decomposed in d interpolation problems each one of them only depends on only one component $\{\mathbf{y_i^k}\}$ of $\{\mathbf{y_i}\}$. We use f^k to denote the best transformation for the k_{th} component. Those d problems can be considered as the minimizer of the following energy function.

$$E(\mathbf{f}) = \frac{1}{n}\sum_{i=1}^{n} ||\mathbf{y_i} - \mathbf{f}(x_i)||^2 + w\int_M (\Delta^p \mathbf{f})^2 dM \tag{7.12}$$

Where $\displaystyle\int_M (\Delta^p \mathbf{f})^2 dM = \sum_{k=1}^{d}\int_M (\Delta^p f^k)^2 dM$.

The minimizer of Eq. 7.12 can be written in analytical form of the same basis functions as before, the $n \times d$ coefficients matrix $\mathbf{c}$ and $1 \times d$ constant matrix $\mathbf{c_0}$ of the analytic solution satisfy the following equation system

$$\mathbf{Y} - T\mathbf{c_0} - \Sigma\mathbf{c} = nw\mathbf{c} \tag{7.13}$$

$$\mathbf{c}^T T = \mathbf{0} \tag{7.14}$$

Where $\mathbf{Y} = [\mathbf{y_1}, \mathbf{y_2}, \cdots, \mathbf{y_n}]^T$.

7.3 Extension from Interpolation to Approximation

7.3.1 Multivariate Problem with Isotropic Errors

The interpolation describe previously assume that the correspondence matched exactly, however, feature points in shape analysis are most learnt by automatic landmarking methods, there are unavoidable corresponding variance or even mismatched landmarks. Even for human experts, there is still unavoidable inter-observer and intra-observer errors. In order to take account of the landmark variances or errors, we propose to extend this by involve a symmetric positive definite matrix W into the cost function. It is same as we redefine our energy function as

$$E(\mathbf{f}) = \frac{1}{n}||W(\mathbf{Y} - \mathbf{f}(X))||^2 + w\int_M (\Delta^p \mathbf{f})^2 dM \tag{7.15}$$

Where $\mathbf{f}(X) = [\mathbf{f}(x_1), \mathbf{f}(x_2), \cdots, \mathbf{f}(x_n)]^T$.

The first term in Eq. 7.15 measures the sum of the quadratic Euclidean distance between the transformed feature points and the target feature points plus possible interactions between them. The second term still measures the smoothness of the transformation.

The computation is nearly the same and the solution to the approximation problem in Eq. 7.15 can also be stated analytically by solving the following linear system,

$$W^2(\mathbf{Y} - T\mathbf{c_0} - \Sigma\mathbf{c}) = nw\mathbf{c} \tag{7.16}$$

$$\mathbf{c}^T T = \mathbf{0} \tag{7.17}$$

If W is the identity matrix, this is equivalent to the previous model. A diagonal matrix W means only assigning different weights to the landmarks. A special case is that we take a diagonal matrix $\left(W^2\right)^{-1}$ and define diagonal elements as the variances σ_i^2 of each landmark

$$\left(W^2\right)^{-1} = \begin{pmatrix} \sigma_1^2 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \sigma_2^2 & \ddots & \cdots & \cdots & \vdots \\ \vdots & \ddots & 0 & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \sigma_{n-1}^2 & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & \sigma_n^2 \end{pmatrix} \tag{7.18}$$

This is same as define the cost function as

$$E(\mathbf{f}) = \frac{1}{n}\sum_{i=1}^{n} \frac{||\mathbf{y_i} - \mathbf{f}(x_i)||^2}{\sigma_i^2} + w\int_M (\Delta^p \mathbf{f})^2 dM \tag{7.19}$$

To interpret Eq. 7.19, we can think that each quadratic Euclidean distance between the transformed feature points and the target feature points is weighted by the variances σ_i^2 representing landmark localization errors. If, for example, the variance is high, i.e., landmark localization is uncertain, the the influence on the overall approximation error is weighted low. Those σ_i^2 can be learned from empirical data, training replica experiments or learning algorithms for specific purpose, such as the learning metric learned for classification purpose [12].

7.3.2 Multivariate Problem with Anisotropic Errors

In the above case of isotropic errors, we assume that the error occurs only within the landmark level. But generally, this assumption will not hold, since the errors are different in different directions and thus are anisotropic. In order to accommodate this, we can further extend the model by involving a $nd \times nd$ symmetric positive definite matrix WL. Each element represents the weights for each coordinate of every landmark.

$$E(\mathbf{f}) = \frac{1}{n}||WL(\mathbf{YL} - \mathbf{fL}(X))||^2 + w\int_M (\Delta^p \mathbf{f})^2 dM \tag{7.20}$$

where $\mathbf{YL}$ and $\mathbf{fL}$(X) are $nd \times 1$ vectors reshaped point wisely from $\mathbf{Y}$ and $\mathbf{f}$(X) respectively.

Indeed, the computation scheme has the same structure as before and the solution can be stated in analytical form with the same basis functions as well.

$$WL^2(\mathbf{YL} - (T \otimes I_d)\mathbf{c_0} - (\Sigma \otimes I_d)\mathbf{C}) = nw\mathbf{C} \tag{7.21}$$

$$\mathbf{C}^T(T \otimes I_d) = \mathbf{0} \tag{7.22}$$

Where $\otimes$ means Kronecker product. $\mathbf{C}$ is a $nd \times 1$ coefficient vector which is the row-wisely reshaped version of the $n \times d$ coefficients matrix $\mathbf{c}$.

If we take the a diagonal block matrix $\left(WL^2\right)^{-1}$ and diagonal blocks are the covariances matrices Σ_i of i_{th} landmark. For instance, for 3 dimensional shape, $\{\Sigma_i\}$ are 3×3 matrices.

$$\left(WL^2\right)^{-1} = \begin{pmatrix} \Sigma_1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \Sigma_2 & \ddots & \cdots & \cdots & \vdots \\ \vdots & \ddots & 0 & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \Sigma_{n-1} & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & \Sigma_n \end{pmatrix} \tag{7.23}$$

This is same as define the cost function as

$$E(\mathbf{f}) = \frac{1}{n}\sum_{i=1}^{n} < (\mathbf{y_i} - \mathbf{f}(x_i)), \Sigma_i^{-1}(\mathbf{y_i} - \mathbf{f}(x_i)) > + w \int_M (\Delta^p \mathbf{f})^2 dM \tag{7.24}$$

With the extended scheme it is possible to include different types of fearure points, e.g.,"normal" point landmarks as well as other feature points. Normal anatomical/mathematical landmarks have unique positions and low localization uncertainties in all directions. An example of other feature points this algorithm can be extremely beneficial for are arbitrary edge points. Such points are not uniquely definable in all directions, and they are used, for example, in the reference system of Talaitach [60] to define the 3-D bounding box of the human brain. The incorporation of such landmarks is important when normal point landmarks are hard to define, for example, at the outer parts of the human head; or when the analysing process requires a large number of points.

7.4 Box Spline: Algorithm

A broad range of applications in shape modeling and analysis is concerned with processing two-dimensional surface. Those surfaces are typically represented by point clouds or meshes. Thus, interpolation on these objects requires the availability of discrete Δ operator to numerically compute the eigenfunctions and eigenvalues. The discretization of the Laplacian on triangular meshes and point clouds is discussed in Chapter 3. We can constructed the kernel in the manner of Eq. 7.6. With the discrete kernel in hand, in principle, we have already resolved the interpolation problem. However, several drawbacks of this discretization should be noticed. First, computation efficiency. The naive way of computing the kernel in Eq. 7.6 requires discretization of Laplace-Beltrami operator and explicit computation of the entire set or large number of the eigenvectors and eigenvalues of Laplacian matrix L. This approach performs poorly on modern data set. Second, accuracy. The approximation of L is highly depending on the point cloud or mesh structure and choice of parameters. One example is shown in Fig.3.3. The convergence to the Δ of underlying surface is only guaranteed as the mesh becomes finer [4, 5]. Thus, for a given shape, convergence is not guaranteed. As a result, the computed eigenvalues and eigenfunctions are not accuracy. To solve those problems, we compromise ourselves to the spline domain as the convex 3D box $[x_1, x_2] \times [y_1, y_2] \times [z_1, z_2]$ of the mesh. The kernel function within in the box is constructed by

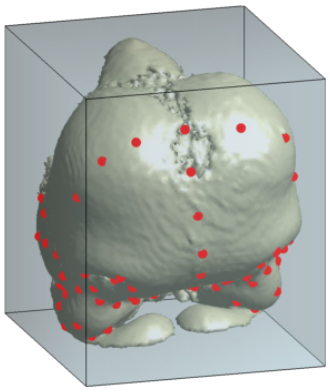


Figure 7.1: Duck Embryo shape with convex 3D box spline domain (light blue).

the analytical eigenfunctions and eigenvalues of this rectangular domain. An example of such as convex box is shown in Fig. 7.1. The light blue box shows the spline domain.

The eigenfunctions of Δ with Neumann boundary conditions are explicit and they are given as

$$\phi_{mnl}(x,y,z) = \cos(\frac{m\pi(x-x_1)}{x_2-x_1})\cos(\frac{n\pi(y-y_1)}{y_2-y_1})\cos(\frac{l\pi(z-z_1)}{z_2-z_1}) \tag{7.25}$$

And the corresponding eigenvalues are given as

$$\lambda_{mnl} = (\frac{m\pi}{x_2-x_1})^2 + (\frac{n\pi}{y_2-y_1})^2 + (\frac{l\pi}{z_2-z_1})^2 \tag{7.26}$$

Where m, n, l are positive integers. Then the kernel of can be computed by truncated version of Eq. 7.6.

For 2D domain, we can construct kernel in a similar manner. Suppose we have fixed 2d convex square domain as $[x_1, x_2]\times[y_1, y_2]$, the eigenfunctions and eigenvalues of Δ with Neumann boundary conditions are given as

$$\begin{aligned} \phi_{mn}(x,y) &= \cos(\frac{m\pi(x-x_1)}{x_2-x_1})\cos(\frac{n\pi(y-y_1)}{y_2-y_1}) \\ \lambda_{mn} &= (\frac{m\pi}{x_2-x_1})^2 + (\frac{n\pi}{y_2-y_1})^2 \end{aligned} \tag{7.27}$$

Where m, n are positive integers.

Similarly, for 1D domain $[x_1, x_2]$, The eigenfunctions and eigenvalues of Δ with Neumann boundary conditions are given as

$$\begin{aligned} \phi_m(x,y) &= \cos(\frac{m\pi(x-x_1)}{x_2-x_1}) \\ \lambda_m &= (\frac{m\pi}{x_2-x_1})^2 \end{aligned} \tag{7.28}$$

Table 7.1: Rank of the three spline methods on spherical domain.

Method \ Rank	1st	2nd	3rd
TPS	0%	3%	97%
Spherical	62%	38%	0%
Box	38%	59%	3%

Where m is positive integer.

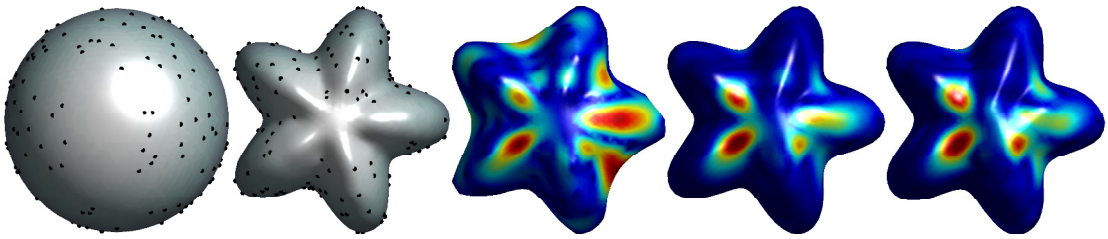
Because we are using the solution of the Laplacian equation with Neumann boundary condition. To better deal with the interpolation problem, we did linear regression previous to the spline.

7.5 Comparison of Spline Methods

Although the spline domain is compromised from the compact Riemannian manifold to convex bounding box, we still expect our method would improve results from TPS. We did several experiments with simulated shape on closed spherical domain and open plain domain. We then compared our results with spherical spline in [68] and TPS. Since the interpolation results depends on other factors such as the distribution of the landmarks, thus, we repeated 100 experiments for each example. For each experiment, we fixed the number of the random selected correspondent points and ranked each method for that experiment by the errors.

The result for spherical domain is shown in Table. 7.1. From this table, we see that among all the experiments, 62% time, the spherical spline achieved the best result; while 38% of the box spline achieved the best result; 97% of the TPS result ranked 3rd. One example of such comparison is shown in Fig. 7.2. From this example we can see, the result from Box spline (Fig. 7.2E) is similar to the result from spherical spline (Fig. 7.2D) and both methods show great improvement respect to the thin plate spline (Fig. 7.2C).

We also did experiments on domains with boundary. Because there is no explicit solution for the Laplace's equation on the open surface, we here only compared the TPS results with those from box spline. From Table. 7.2, we can see that, among 100 repeated experiments with randomized fixed number landmarks, 98% time box spline gave better result. One example of such comparison is shown in Fig. 7.3.

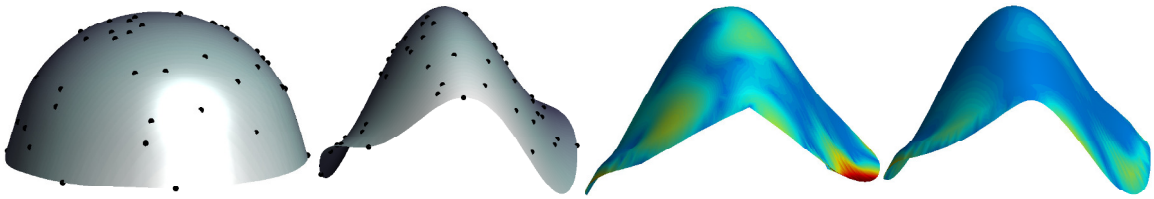


(A) Domain (B) Target Shape (C) TPS (D) Spherical Spline (E) Box Spline

Figure 7.2: One example of the spline results on the spherical domain. (A) The domain with the landmarks(black). (B) The target shape with the landmarks(black). The color of (A) and (B) shows the correspondence the points. (C) The result of the TPS. (D) The result of Spherical spline. (E) The result of Box spline. In (C, D, E), the color map shows the magnitude of error at each point, while red means large error, blue means small. The color map has been standardized to the same range.

Table 7.2: Rank of the TPS and box spline on open surface domain.

Method \ Rank	1st	2nd
TPS	2%	98%
Box	98%	2%



(A) Domain (B) Target Shape (C) TPS (D) Box Spline

Figure 7.3: One example of the spline results on the open surface domain. (A) The domain with the landmarks(black). (B) The target shape with the landmarks(black). The color of (A) and (B) shows the correspondence of the points. (C) The result of the TPS. (D) The result of Box spline. In (C, D), the color map shows the magnitude of error at each point, while red means large error, blue means small. The color map has been standardized to the same range.

From Fig. 7.3, the error color map shown in Figs. 7.3C and 7.3D, we can clearly notice that the error from Box spline was dramatically reduced from TPS method. This reduced error is more obvious near the boundary.

7.6 Dense Surface Model

Box spline described in Section 7.4 is used to find dense correspondence across samples by interpolating the sparse landmarks (Fig: 7.1) on the avian embryos.

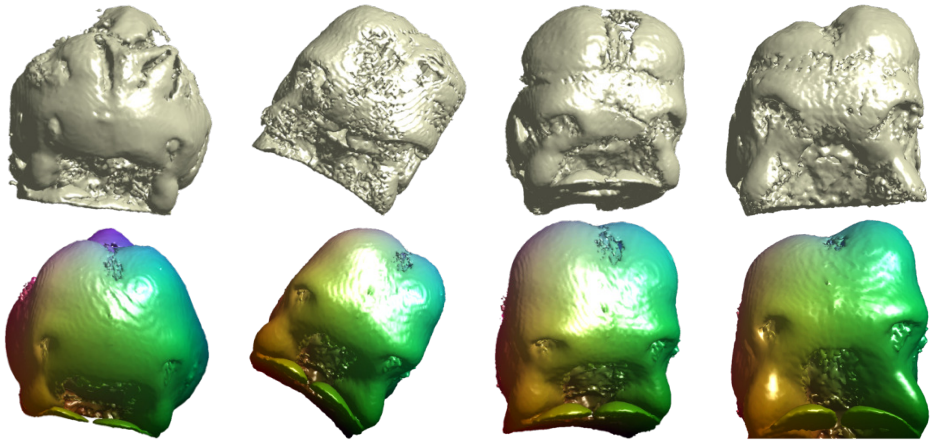


Figure 7.4: Examples of the mesh registration. First row shows the original mesh, second row shows the registered mesh. The colormap of the second row indicates the correspondence.

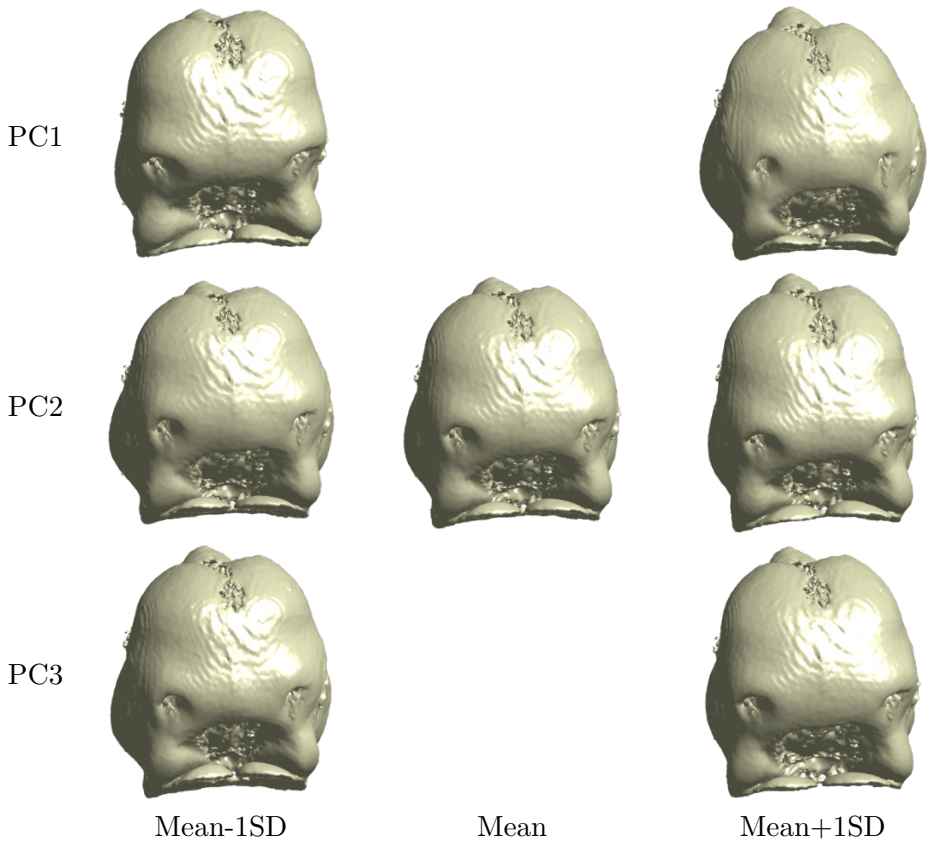
Since the embryos data suffer from a lot artifacts through the scanning process, our registration not only provide dense correspondence across sample but also complement the whole shape. Based on the sparse landmarks, if the surface data is in high quality, we can also use iterative methods to improve the dense correspondence. The algorithm is the following:

1. Choose an arbitrary mesh as the reference mesh.
2. Take all the meshes in the dataset as target meshes.
   (a) For each target mesh, interpolate reference mesh to target mesh using sparse landmarks by box spline. The result mesh is called splined mesh.
   (b) For randomly selected points from the splined mesh, find the closest point in the target mesh.
   (c) Interpolate the splined mesh to the target mesh using the guidance of correspondent points from previous step. The result mesh is called interpolated mesh.
   (d) Use interpolated mesh as reference mesh.
3. After each iteration for every mesh, compute the mean shape and final set of aligned mesh sets by Procrustes alignment [29].

4. If the distance between two consecutive mean shapes and the average distance of the correspondent aligned meshes are below certain threshold, stop; otherwise, go back to step 2.

From the dense correspondence, we conduct Procrustes analysis (PA) on the dense correspondence of all the samples.

Table 7.3: Embryo shape variation along the first three principal directions of all samples in the dataset

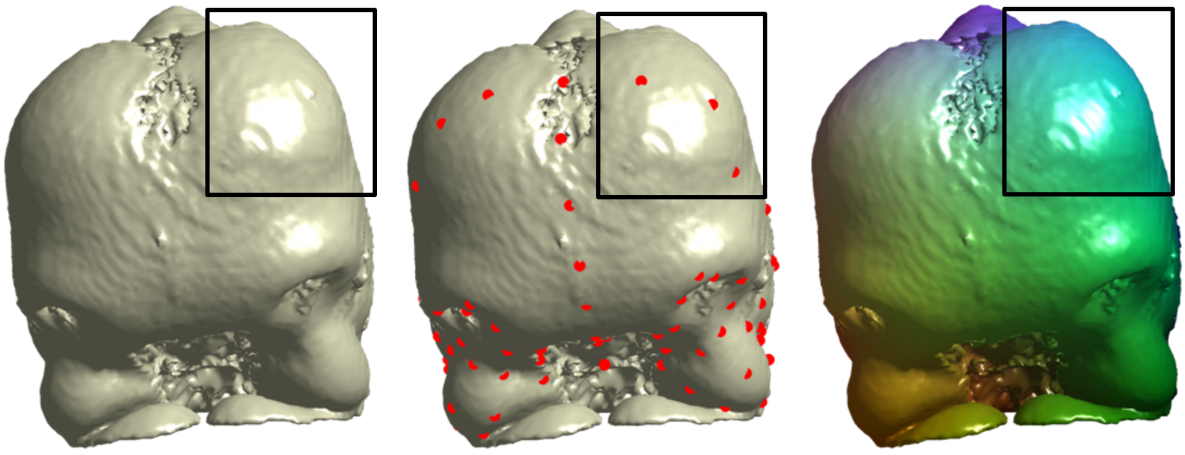


7.6.1 Patch Analysis

One aspect of necessity of patch analysis arises from the aspect of the shape analysis. Normally we tend to build relationship between shape space and function space or even gene space. This relationship was typically built using the leading several PC scores as the representation of the shapes. This approach will give us robust result. However, in certain application the localized variance holds the large interest, for instance, the analysis of cleft cleft and cleft palate. But, this localized variance maybe overlooked when the analysis was done on a larger scale.

Beside necessity of algorithm design, analyzing patches and studying the relationship between patches also has potential in biological merits. During growth, object changes dramatically in shape as well as in size. An organism comprises several interdependent bones/muscles that grow and develop under the influence of various local and systemic factors. Patch analysis enable us to further discover the pattern of individual region as well as the interaction between several regions.

Despite all the great applications, patch analysis is highly limited by the number of landmarks under the sparse landmark. In certain region, there are even no points that have an exact biological correspondence. The dense model provided by our spline method overcome this barrier. Since those dense surface models can established a correspondence of thousands of points across each 3D image.



(A) (B) (C)

Figure 7.5: Examples of the patch analysis. (A) One patch sample (in black box). (B) Sparse Model: patch Analysis is limited by the number of landmarks. Since there are only three landmarks (red) in this region, the PA is impossible to capture the real anatomical change within this region. (C) Dense Model: correspondence of thousands of points overcome the limitation.

CHAPTER 8

DISCUSSION

To accommodate newly emerging shape data, which differs from traditional data in size and type, the existing models have to be modified and new methods have to be developed. In this dissertation, we explored these ideas from two aspects.

- Develop Landmark-free Model

  In this dissertation, we also developed an ad hoc shape representation for quantifying amorphous shapes with no clearly defined form. Particularity, the analysis based on landmarks is not applicable to this problem. Motivated by the study of gene expression domains, we present a method that combines optical projection tomography scanning, a shape regularization technique and a landmark-free approach to quantify gene expression domains variation. A key strength of the method stems from the fact that it is difficult, or even impossible, to quantify such variation with the usual methods of geometric morphometrics because gene expression domains typically lack homologous landmarks. In addition, 3D image acquisition and processing introduce many artifacts that further exacerbate the problem. Our landmard-free approach quantifies variation in shape of seemingly amorphous gene expression domains, enhances their most salient morphological characteristics and is robust to uninformative local shape variation and irregularities associated with image acquisition.

  Our model is applied to quantify variation in the morphology of Sonic hedgehog expression domains in the frontonasal ectodermal zone (FEZ) of avians and investigate relationships with embryonic craniofacial shape. Combined with PCA and CCA methods, the model revealed axes in *Shh* expression and craniofacial morphospaces along which variation exhibits a strong linear relationship at high statistical significance. Although in this dissertation, we only applied this model to *Shh* expression domains, we believe that it has tremendous potential for advancing quantitative integration across the genotype-phenotype map for complex morphologies. The method should be particularly useful in quantitative analyses of 3D smooth, surface-like structures that have ill-defined shape.

- Construct Dense Correspondence from Sparse Landmarks

  To take full advantages of the high resolution geometrical information, we develop a spline method to construct dense correspondence across all shapes. Instead of using Euclidean domain, because of the nature of shape surface as a manifold embedded in Euclidean space, we

present a general theoretical framework of spline in which the Euclidean domain can be extended to manifold domain. In order to be practical applicable, we provide a computationally effective algorithm to compute such spline function based on bounded rectangular domain. The computation framework shows clear improvement respect to the thin plate spline method.

The proposed spline method can establish a correspondence of thousands of points across each shape. Compared with sparse landmarks, the dense correspondence not only provides dramatic visualization of the shape variation but also carries the potential for precisely identifying the shape features. Without the limitation of the number of landmarks, the dense model is also encouraging new studies, such as modular analysis on shape patches and correlation analysis between patches.

All the pictures included in section 7.6 demonstrate that dense surface models can generate information visualizations of embryonic morphology in 3D. Any benefit the dense model may provide in training of medical experts and disease diagnosis is yet to be evaluated. As the data gathering for these become disease orientated, it will also become possible to study the discriminate ability of dense models in disease diagnosis.

In conclusion, we believe, the integration of those newly developed shape analysis models with machine learning algorithms and statistical inference methods will allow biologists to explore how morphological variation correlates to biological variates and help advance various areas of research.

REFERENCES

[1] G. Abraham and M. Inouye. Fast principal component analysis of large-scale genome-wide data. *PLoS ONE*, 9(4), 2014.

[2] D. Bates, M. Lindstorm, G. Wahba, and B. Yandell. Gcvpack - routines for generalized cross validation. *Communication in Statistics- Simulation and Computation, 16, 263-297*, 1987.

[3] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 15:1373–1396, 2003.

[4] M. Belkin, J. Sun, and Y. Wang. Discrete laplace operator on meshed surfaces. *Proceedings of the twenty-fourth annual symposium on Computational geometry*, 2008.

[5] M. Belkin, J. Sun, and Y. Wang. Constructing laplace operator from point clouds in r^d. *Proceedings of the twenty-fourth annual symposium on Computational geometry*, 2009.

[6] F. L. Bookstein. Size and shape spaces for landmark data in two dimensions (with discussion). *Statistical Science*, 1:181–242, 1986.

[7] F. L. Bookstein. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge: Cambridge University Press, 1996.

[8] W. M. Boothby. *An Introduction to Differentiable Manifold and Riemannian Geometry*. Harcourt Brace Jovanovich, 1986.

[9] H.J. Chong, N.M. Young, D. Hu, J. Jeong, A.P. McMahon, B. Hallgrimsson, and R.S. Marcucio. Signaling by shh rescues facial defects following blockade in the brain. *Dev Dyn*, 241:247–256, 2012.

[10] I.L. Dryden and K.V. Mardia. *Statistical shape analysis*. John Wiley and Son, Chichester, 1998.

[11] C. Duchon. Splines minimizing rotation-invariant semi-norms in sobolev spaces. *Constructive Theory of Functions of Several Variables*, 1977.

[12] Y. Fan, D. Houle, and W. Mio. Learning metrics for shape classification and discrimination. *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 2652–2655, 2010.

[13] M. E. Fisher, A. K. Clelland, A. Bain, R. Baldock, P. Murphy, H. Downie, C. Tickle, D.R. Davidson, and R. Buckland. Integrating Technologies for Comparing 3D Gene Expression Domains in the Developing Chick Limb. *Developmental biology*, 317(1):13–23, May 2008.

[14] S. Fortune. A sweepline algorithm for voronoi diagrams. *Algorithmica*, 2:153–174, 1987.

[15] C.C. Fowlkes, C. L. Hendriks, S.V. Keränen, G. H. Weber, O. Rübel, M. Huang, S. Chatoor, A.H. DePace, L. Simirenko, C. Henriquez, A. Beaton, R. Weiszmann, S. Celniker, B. Hamann, D.W. Knowles, M.D. Biggin, M.B. Eisen, and J. Malik. A Quantitative Spatiotemporal Atlas of Gene Expression in the Drosophila Blastoderm. *Cell*, 133(2):364–74, April 2008.

[16] B. Fuche, M. amd Juttler, O. Scherzer, and H. Yang. Shape metrics based on elastic deformations. *Journal of Mathematical Imaging*, 35:86–102, 2009.

[17] L. Guibas and J. Stolfi. Primitives for the manipulation of general subdivisions and the computation of voronoi diagrams. *ACT TOG*, 4, 1985.

[18] B. Hallgrimsson, H. Jamniczky, N.M. Young, C. Rolian, T.E. Parsons, J.C. Boughner, and R.S. Marcucio. Deciphering the palimpsest:studying the relationship between morphological integration and phenotypic covariation. *Evolutionary Biology*, 36:355–376, 2009.

[19] J. L. Hendrikse, T. E. Parsons, and B. Hallgrimsson. Evolvability as the proper focus of evolutionary developmental biology. *Evolution & Development*, 9:393–401, 2007.

[20] A. N. Hirani. *Discrete Exterior Calculus*. Thesis for Degree of Doctor of Philosophy, 2003.

[21] H. Hotelling. Relations between two sets of variants. *Biometrika*, 28:321–377, 1936.

[22] D. Houle, D.R. Govindaraju, and S. Omholt. Phenomics: The next challenge. *Nature Reviews Genetics*, 11:855–866, 2010.

[23] D. Hu and R. S. Marcucio. A shh-responsive signaling center in the forebrain regulates craniofacial morphogenesis via the facial ectoderm. *Development*, 136:107–116, 2009.

[24] D. Hu and R. S. Marcucio. Assessing signaling properties of ectodermal epithelia during craniofacial development. *Journal of Visualized Experiments: JoVE*, 49:2557, 2011.

[25] D. Hu, R. S. Marcucio, and J. A. Helms. A zone of frontonasal ectoderm regulates patterning and growth in the face. *Development*, 130:1749–1758, 2003.

[26] D. Hu, N.M. Young, Q. Xu, H. Jamniczky, R.M. Green, W. Mio, R.S. Marcucio, and B. Hallgrimsson. Signals from the brain induce variation in avian facial shape. *Submitted for Publication*, 2015.

[27] S. Huckemann and H. Ziezold. Principal component analysis for riemannian manifolds, with an application to triangular shape spaces. *Advances in Applied Probability*, 38:299–319, 2006.

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

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

[30] D. G. Kendall, D. Barden, T. K. Carne, and H. Le. *Shape and Shape Theory*. Wiley,Chichester, New York, 1999.

[31] P. Kim. Splines on riemannian manifolds and a proof of a conjecture by wahba. 1999.

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

[33] W. J. Krzanowski. *Principles of Multivariate Analysis: A User's Perspective*. New York: Oxford University Press, 1988.

[34] T. Kudoh, M. Tsang, N. A. Hukriede, X. Chen, M. Dedekian, C. J. Clarke, A. Kiang, S. Schultz, J. A. Epstein, R. Toyama, and I. Dawid. Gene expression screen in zebrafish embryogenesis. *Genome Res.*, 11:1979–1987, 2001.

[35] W. Kuehnel. *Differential Geometry:Curves-Surfaces-Manifolds*. Originally published in the German language by Friedr, 2003.

[36] H. Le and D. L. Kendall. The riemannian structure of euclidean shape spaces: a novel environment for statistics. *Annals of Statistics*, 21(3):1225–1271, 1993.

[37] E. Lein, M. J. Hawrylycz, N. Ao, M. Ayres, A. Bensinger, A. Bernard, A.F. Boe, M. S. Boguski, K.S. Brockway, E.J. Byrnes, L. Chen, L. Chen, T. Chen, M. Chin, J. Chong, B. E. Crook, A. Czaplinska, C.N. Dang, S. Datta, N.R. Dee, and et al. Genome-wide Atlas of Gene Expression in the Adult Mouse Brain. *Nature*, 445(7124):168–76, January 2007.

[38] X. Liu, W. Mio, Y. Shi, I. Dinov, X. Liu, N. Lepore, F. Lepore, M. Fortin, P. Voss, and P. M. Thompson. Models of normal variation and local contrasts in hippocampal anatomy. In *MICCAI '08 Proceedings of the 11th International Conference on Medical Image Computing and Computer-Assisted Intervention, Part II*, 2008.

[39] R. S. Marcucio, D. Cordero, and J. A. Helms. Molecular interactions coordinating development of the forebrain and face. *Dev. Biol.*, 284:48–61, 2005.

[40] K. V. Mardia, J. T. Kent, and J. M. Bibby. Multivariate analysis. *Academic Press*, 1979.

[41] J. Meinguet. Multivariatien interpolationa at arbitrary points made simple. *Journal of Applied Mathematics and Physics,volume 30, 292-304*, 1979.

[42] J. Mercer. Functions of positive and negative type and their connection with the theory of integral equations. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. A 209 415-446*, 1909.

[43] P. Michor and D. Mumford. Riemannian geometries on spaces of plane curves. *Journal of the European Mathematical Society*, 8:1–48, 2006.

[44] P. Michor and D. Mumford. An overview of the riemannian metrics on spaces of curves using the hamiltonian approach. *Applied and Computational Harmonic Analysis*, 23:74–113, 2007.

[45] P. Michor, D. Mumford, J. Shah, and L. Younes. A metric on shape space with explicit geodesics. *Rend. Lincei Mat. Appl.*, 9:25–57, 2008.

[46] M. I. Miller and L. Younes. Group actions, homeomorphisms, and matching: A general framework. *International Journal of Computer Vision*, 41:61–84, 2001.

[47] W. Mio, J. C. Bowers, M. K. Hurdal, and X. Liu. Modeling brain anatomy with 3d arrangements of curves. *In Proc. Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA)*, 2007.

[48] W. Mio, J. C. Bowers, and X. Liu. Shape of elastic strings in euclidean space. *International Journal of Computer Vision(IJCV)*, 82:96–112, 2009.

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

[50] E. Myasnikova, A. Samsonova, K. Kozlov, M. Samsonova, and J. Reinitz. Registraion of the Expression Patterns of Drosophila Segmentation Genes by Two Independent Methods. *Bioinformatics*, 17(1):3–12, 2001.

[51] T. F. Oostendorp, A. Oosterom, and G. Huiskamp. Interpolation on a triangulated 3d surface. *Journal of Computational Physics*, 80, 1989.

[52] E. Parzen. On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33:1065–1076, 1962.

[53] K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2:559–572, 1901.

[54] L. Quintana and J. Sharpe. Optical projection tomography of vertebrate embryo development. *Cold Spring Harb Protoc*, pages 586–594, 2011.

[55] M. Reuter, S. Biasotti, D. Giorgi, G. Patane, and M. Spagnuolo. Discrete laplace-beltrami operators for shape analysis and segmentation. *Computer& Graphics*, 33(3):381–390, June 2009.

[56] S. Rosenberg. *The Laplacian on a Riemannian Manifold*. Cambridge University Press, 1997.

[57] G. A. Seber. *Multivariate Observations*. Hoboken, NJ: John Wiley and Sons, Inc., 1984.

[58] J. Sharpe. Optical projection tomography as a tool for 3d microscopy and gene expression studies. *Science*, 296:541–545, 2002.

[59] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. London: Chapman Hall/CRC, 1998.

[60] J. Talairach and P. Tornoux. *Co-Planar Stereotactic Atlas of the Human Brain*. Stuttgart, Germany: Georg Thieme Verlag, 1988.

[61] O. Tassy, D. Dauga, F. Daian, D. Sobral, F. Robin, P. Khoueiry, D. Salgado, V. Fox, D. Caillol, R. Schiappa, B. Laporte, A. Rios, G. Luxardi, T. Kusakabe, J. Joly, S. Darras, L. Christiaen, M. Contensin, H. Auger, C. Lamy, C. Hudson, U. Rothbächer, M.J. Gilchrist, K.W. Makabe, K. Hotta, S. Fujiwara, N. Satoh, Y. Satou, and P. Lemaire. The ANISEED database: Digital Representation, Formalization, and Elucidation of a Chordate Developmental Program. *Genome research*, 20(10):1459–68, October 2010.

[62] D. W. Thompson. *On Growth and Form*. John Wiley and Son, Chichester, 1998.

[63] K. Uhlenbeck. Eigenfunction of laplace operators. *Bulletin of The American Mathematical Society*, 78(6), 1972.

[64] K. Uhlenbeck. Generic properties of eigenfunctions. *American Journal of Mathematics*, pages 1060–1078, 1976.

[65] A. Visel, C. Thaller, and G. Eichele. Genepaint.org: An atlas of gene expression patterns in the mouse embryo. *Nucleic Acids Res*, 32:552–556, 2004.

[66] P. Wagner, M. Ruta, and M. Coates. Evolutionary patterns in early tetrapods. ii. differing constraints on available character space among clades. *Proc. R. Soc. London Ser. B*, 273:2113–2118, 2006.

[67] G. Wahba. Spline interpolation and smoothing on the sphere. *SIAM Journal of Scientific and Statistical Computing,2,5-16*, 1981.

[68] G. Wahba. *Spline Models for Observational Data*. Philadelphia: Society for Industrial and Applied Mathematics, 1990.

[69] G. Wahba. *Spline Models for Observational Data*. Philadelphia: Society for Industrial and Applied Mathematics, 1990.

[70] F. Wong, M. Welten, C. Anderson, A. Bain, J. Liu, M. N. Wicks, G. Pavlovska, M. G. Davey, P. Murphy, D. Davidson, C. Tickle, C. D. Stern, R. Baldock, and D. W. Burt. eChickAtlas: an introduction to the database. *Genesis (New York, N.Y. : 2000)*, 51(5):365–71, May 2013.

[71] A. J. Yezzi and A. Mennucci. Conformal metrics and true 'gradient flows' for curves. *Proceedings of the 10th IEEE International Conference on Computer Vision*, pages 913–919, 2005.

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

[73] L. Younes. Optimal matching between shapes via elastic deformations. *Journal of Image and Vision Computing*, 17:381–389, 1999.

[74] L. Younes, P. Michor, J. Shah, and D. Mumford. A metric on shape space with explicit geodesics. *Rend. Lincei, Mat. Appl.*, 9, 2007.

[75] N. M. Young, H. J. Chong, D. Hu, B. Hallgrimsson, and R.S. Marcucio. Quantitative analyses link modulation of sonic hedgehog signaling to continuous variation in facial growth and shape. *Development*, 137(20):3405–3409, 2010.

BIOGRAPHICAL SKETCH

Qiuping Xu grew up in Tianjin, China. She graduated from Nankai Univerisity in 2009 with a B.S. in Applied Mathematics. She came to US to continue her education at the same year. She is pursuing a Ph.D. in Biomedical Mathematics at Florida State University.

Qiuping's passion is about mathematical/statistical modeling and data analysis. Her current research is concerned with problems in shape and image analysis and their applications in biological and medical studies. In particular, she is interested in quantifying biological shapes that are fundamentally linked to underlying mechanisms and functions. This includes modeling morphological variation and analyzing relationships between genotype and phenotype, as well as between different phenotypes. She endeavors to build interpretable computational models that can be integrated with machine learning and statistical inference methods that allow biologists to explore how morphology correlates to other biological variates to help advance research in developmental and evolutionary biology and medical diagnosis.

She submitted two journal papers

[1] D. Hu, N.M. Young, Q. Xu, H. Jamniczky, R.M. Green, W. Mio, R.S. Marcucio, and B. Hallgrimsson. *Signals from the brain induce variation in avian facial shape.*

[2] Q. Xu, H. Jamniczky, D. Hu, R.M. Green, R.S. Marcucio, B. Hallgrimsson, and W. Mio. *Correlations between the morphology of sonic hedgehog expression domains and embryonic craniofacial shape.*