# Geometric Analysis of Axonal Tree Structures

Adam Duncan[1]
a.duncan@stat.fsu.edu

Eric Klassen[2]
http://www.math.fsu.edu/~klassen/

Xavier Descombes[3]
http://stat.fsu.edu/~anuj/

Anuj Srivastava[1]
http://stat.fsu.edu/~anuj/

[1] Department of Statistics
Florida State University
Tallahassee, Florida, USA

[2] Department of Mathematics
Florida State University
Tallahassee, Florida, USA

[3] Morpheme, INRIA
Sophia Antipolis, France

**Abstract**

This paper develops a framework for shape analysis of tree-like structures with the following common features: (1) a main branch viewed as a parameterized curve in $\mathbb{R}^3$, and (2) a random number of secondary branches, each one of them a parameterized curve in $\mathbb{R}^3$, emanating from the main branch at arbitrary points. In this framework, comparisons of objects is based on shapes-scales-orientations of the curves involved, and locations and number of the side branches. The objects are represented as composite curves made up of: a main branch and a continuum of side branches along the main branch with each branch being a curve in $\mathbb{R}^3$ itself (including the null curve, or zero curve). Extending the previous work on elastic shape analysis of Euclidean curves, the space of these composite curves is endowed with a natural Riemannian metric, using the SRVF representation, and one computes geodesic paths in the quotient space of this representation modulo the re-parameterization function. As a result, appropriate geometric structures are optimally matched across trees, and geodesic paths show deformations of main branches into each other while either deforming/sliding/creating/destroying the side branches. We present some preliminary results using axonal trees taken from the Neuromorpho database.

## 1 Introduction

Shape analysis of objects is an important problem in many disciplines, including vision, biology, medicine, and manufacturing. There have been many efforts that study shapes of points sets, curves, surfaces, and other objects that have fixed topologies [5, 9, 17, 22]. These efforts often have a common theme – they start by choosing mathematical representations of objects and then quotient out shape-preserving transformations, such a rigid motions, global scaling, and parameterizations. In this paper we consider structures that can differ in topology also, in addition to their geometries. Specifically, we are interested in analyzing trees – configurations where a stem branches into side branches, side branches into tertiary structures, and so on. Such trees are important in a variety of scientific applications, including
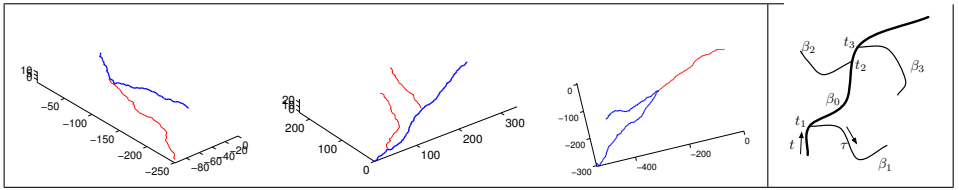
Figure 1: Left: Three examples of simple axon trees taken from *Neuromorpho* database. Right: A schematic of an axonal tree with the main branch and three side branches.

phylogeny [2], axon morphology [14], road networks, blood vessels, epidemiology models, etc. The properties of interest in a tree are: (1) geometries (size, shape, orientations) of different curves, including the main and the side branches, and (2) the locations and number of the side branches. There is a great need for tools that allow us to: (1) compare such trees in order to quantify their differences in terms of these features, (2) study statistical variability in a set of trees sampled from a population, e.g. by computing their means and variances, and (3) develop statistical models to capture the observed variability.

While a wide variety of topological structures can be studied in this context, we will focus on a specific family that is motivated by our interest in axonal trees [13, 14]. The topologies of these trees are characterized by a main branch – a stem or a backbone – and a number of side branches connected sequentially to a main branch. (There are tertiary branches too but they are ignored in this paper.) We will assume the availability of data in the following form. The main branch is represented as a parameterized curve $\beta_0 : [0,1] \to \mathbb{R}^3$; for the arc-length parameterization of $\beta_0$, we get the connection points – parameter values along $\beta_0$ – where the side branches are connected, as $\{t_1, t_2, \ldots t_n\} \subset [0,1]$, with $n$ being a variable itself. For each $i = 1, 2, \ldots, n$, $\beta_i$ represents a side branch and is a parameterized curve in $\mathbb{R}^3$ itself. Due to the connectivity of side branches with the main sbranch, we have $\beta_0(t_i) = \beta_i(0)$. Let $\mathcal{B}$ denote the set of absolutely continuous, parameterized curves in $\mathbb{R}^3$. A mathematical representation for an axonal tree is: $\{\beta_0, (t_1, \beta_1), (t_2, \beta_2), \ldots, (t_n, \beta_n)\}$, where $0 < t_1 < t_2 \ldots < t_n < 1$, $\beta_i \in \mathcal{B}$. Several examples of trees are shown in Fig. 1, along with a schematic in the rightmost panel.

## 1.1  Past Literature

Despite restricting to simpler tree-like structures, the analysis still remains difficult. The main difficulty comes from presence of both geometrical and topological variability in these structures. The past work on shape analysis, especially that relying on differential geometry, is usually restricted to objects with fixed topology – closed planar curves, spherical surfaces, etc. On the other hand, the past works on tree analysis often ignore the geometry (shapes) of the branches and focus only on the topology. Methods such as tree-edit distance [18], and its variations, simply focus on the numbers and locations of side branches, considering them as binary choices (branch or no branch). The shapes of of branches have important implications in scientific applications, and cannot be simply discarded in structural comparisons. To take into account both topological and geometrical variability requires advanced mathematical tools. Furthermore, the need for automatic matching of features, both geometrical and topological, makes this problem even harder.

There is substantial work in elastic shape analysis of curves in Euclidean spaces [17]. Its extension to *annotated* curves, where the analysis relies not only on the geometry of

curves but also on annotations defined as finite-dimensional functions defined along those curves [11], is closely related to the approach developed in the current paper. This technique has been used in the past to study *colored* curves where both shapes and colors are used in matching and comparisons. A similar idea was used to study protein backbones as with the location of side chains treated as annotations [12]. Mottini et al. [13, 14] was the first paper to apply this framework to axon morphology, viewing the side branches as discrete annotations along the main branch.

## 1.2 Our Approach: Composite Trajectories

Our approach here is to represent each tree by a trajectory in an infinite-dimensional space $\mathcal{P}$, denoting a certain set of parameterized curves in $\mathbb{R}^3$. More precisely, a trajectory consists of: (1) a continuous sequence of seed points in $\mathbb{R}^3$ forming the main branch $\beta_0$ and (2) attached to *each* seed point is a side branch (a curve) represented by an element of $\mathbb{L}^2([0,1], \mathbb{R}^3)$. Although, in practice, a tree has only a finite number of side-branches, we assume a *continuum* of side-branches in order to facilitate development of underlying theory. This is not absurd since a singleton point can also be viewed as a constant curve, or a *null* or a *trivial* curve, a valid element of the curve space, thus providing trivial side branches where the real one do not exist. These composite trajectories are then compared using a novel Riemannian structure that allows for elastic matching of features across trajectories. In particular, we use the elastic metric to register the main branches across trees, and to deform the shapes of these composite trajectories from one to another. The latter part implies that both the main and the side branches are deformed. Due to elastic metric, a combination of stretching and bending is used to match points across main branches. In cases where a side branch of one trajectory is matched to a trivial curve on the other trajectory – we will see a birth or an annihilation of a side branch along the geodesic. Otherwise, we will expect a deformation of a side branch of one trajectory to its matched counterpart on the other trajectory.

The rest of this paper is laid out as follows. Section 2 summarizes past work on elastic shapes analysis of Euclidean curves and its extension to trajectories in Hilbert spaces and their quotient spaces. Section 3 introduces a mathematical representation of trees as composite curves and lays out the techniques for computation of geodesics between arbitrary trees. Section 4 presents some preliminary experimental results and the paper ends with a short summary in Section 5.

# 2 Background: Elements of Shape Theory

We are interested in shapes of trees made up of branches of different kinds. We will treat all of these branches as *parameterized curves*, and will involve methods that allows us to compare their shapes and other features. Although there are many choices for this, we will use elastic shape analysis of Euclidean curves [10, 17]. While in traditional shape analysis one is often not interested in the location, scale, and orientation of the curve, treating them as nuisance variables, we are interested in shape, relative orientations and sizes of the branches, and will include them in comparisons. The locations of these branches will also be included in the analysis but separately.

**Elastic Comparison of Curves in $\mathbb{R}^3$:**
We briefly summarize ideas for elastic comparisons of curves in $\mathbb{R}^3$; for details please re-
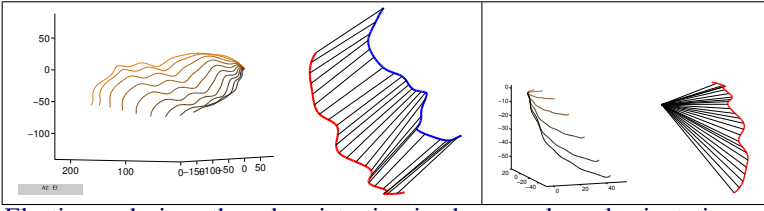
Figure 2: Elastic geodesic path and registration in shape-scale-and-orientation space. Left: between two non-trivial curves. Right: between a non-trivial curve and a trivial curve.

fer to [10, 17]. Let $\mathcal{B}$ be the space of absolutely-continuous curves in $\mathbb{R}^3$. A single point $p \in \mathbb{R}^3$, denoted as a constant curve $\beta(t) = p$, is also an element of $\mathcal{B}$ and is termed a *trivial* curve. For any $\beta_i \in \mathcal{B}$, we study its shape using its square-root-velocity function (SRVF) given by: $q_i(t) = \frac{\dot{\beta}_i(t)}{\sqrt{|\dot{\beta}_i(t)|}}$. In case $\dot{\beta}_i(t) = 0$, then $q_i(t) = 0$ also. Note that the $\mathbb{L}^2$ norm of $q_i$ equals the length of the curve: $\|q_i\| = \sqrt{\int_0^1 |q_i(t)|^2 dt} = \int_0^1 |\dot{\beta}_i(t)| dt = L[\beta_i]$. Let $\Gamma$ be the set of all re-parameterizations of curves (weakly increasing, absolutely continuous functions with appropriate boundary conditions). If a curve $\beta$ is re-parameterized to $\beta \circ \gamma$, then its SRVF changes to $(q \circ \gamma)\sqrt{\dot{\gamma}}$. The shape-scale-orientation of a curve is now represented by the equivalence class $[q_i] = \{(q_i \circ \gamma)\sqrt{\dot{\gamma}}|\gamma \in \Gamma\}$. The set of all such equivalence classes, $\mathbb{L}^2([0,1], \mathbb{R}^3)/\Gamma$, is called the *shape-scale-orientation space*. For any two curves $\beta_1, \beta_2 \in \mathcal{B}$, represented by their SRVFs $q_1, q_2$, respectively, the optimal registration is given by: $\gamma^* = \arginf_{\gamma \in \Gamma} \|q_1 - (q_2, \gamma)\|$. Consequently, the geodesic path between them in $\mathbb{L}^2([0,1], \mathbb{R}^3)/\Gamma$ is given by $\alpha(s) = (1-s)q_1 + s(q_2, \gamma^*)$, and the geodesic distance is $\|q_1 - q_2\|_{\mathbb{L}^2([0,1],\mathbb{R}^3)/\Gamma} = \|q_1 - (q_2, \gamma^*)\|$. Fig. 2 shows two examples of elastic geodesic between curves as elements of $\mathbb{L}^2([0,1], \mathbb{R}^3)/\Gamma$. In each case the left panel shows the geodesic and the right panel shows the registration. The left example uses two curves with arbitrary shapes and orientations. For the purpose of enhancing the display, these curves have been scaled to unit length in the right panel. In case one or both of the curves are trivial, these results still hold. If $\beta_i$ is a trivial curve, i.e. $\beta_i(t) = p \in \mathbb{R}^3$, for all $t \in [0,1]$, then its SRVF is a constant function with value $\mathbf{0} \in \mathbb{R}^3$. The geodesic between $\beta_i$ and a nontrivial curve $\beta_j$ is simply $\alpha(s) = sq_j$ and geodesic distance is $\|q_j\|$. The right side of Fig. 2 shows an example of a geodesic deformation going from a null curve to a nontrivial curve.

**Elastic Comparison of Trajectories in $\mathbb{L}^2([0,1], \mathbb{R}^n)$**
The framework for comparing Euclidean curves can naturally be extended to compare and analyze trajectories in function spaces, e.g. to $\mathbb{L}^2([0,1], \mathbb{R}^n)$. Even the computer implementations can be easily adapted to this case with minimal changes. This is because the $\mathbb{L}^2$ norm on the function space, $\|f\|^2 = \int_0^1 |f(t)|^2 dt \approx \delta \sum_{i=1}^k f(t_i)^2 = \delta \|v\|_{\mathbb{R}^k}^2$, where $\{t_i\}$ is a uniform partition of $[0,1]$ with bin size $\delta$ and $v \in \mathbb{R}^{n \times k}$ is a matrix of values $\{f(t_i)\}$. The $\mathbb{L}^2$ norm between two functions is approximated using the Frobenious norm of the corresponding difference matrix. Thus, in practice, one can treat the given functions as elements of $\mathbb{R}^{n \times k}$, for a large $k$, and then apply the setup for Euclidean curves.

Let $X : [0,1] \to \mathbb{L}^2([0,1], \mathbb{R}^n)$ be an absolutely-continuous trajectory in $\mathbb{L}^2([0,1], \mathbb{R}^n)$. We can also view $X$ as a mapping from $[0,1]^2$ to $\mathbb{R}^n$. For each $t \in [0,1]$, the mapping $X(t, \tau)$ denotes a curve in $\mathbb{R}^n$ with the parameter $\tau$; $t$ forms a continuous indexing of these curves.
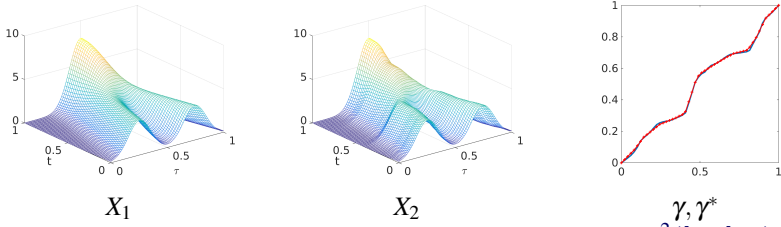
Figure 3: A simulated example of registration of two trajectories in $\mathbb{L}^2([0,1],\mathbb{R})$: here $X_2 = X_1 \circ \gamma$. The original $\gamma$ and the estimated $\gamma^*$ are shown in the right panel.

Define the SRVF of trajectory $X$ according to:

$$q_X(t,\tau) = \frac{\frac{dX(t,\tau)}{dt}}{\sqrt{\left\|\frac{dX(t,\cdot)}{dt}\right\|_{\mathbb{L}^2([0,1],\mathbb{R}^n)}}} \text{ , where } \left\|\frac{dX(t,\cdot)}{dt}\right\|_{\mathbb{L}^2([0,1],\mathbb{R}^n)} = \sqrt{\int_0^1 \left|\frac{dX(t,\tau)}{dt}\right|^2 d\tau} \text{ .}$$

We can reconstruct $X$ back from $q_X$ (up to a different translation for each $t$) using: $X(t,\tau) = \int_0^t q(s,\tau)\|q(s,\cdot)\|_{\mathbb{L}^2([0,1],\mathbb{R}^n)} ds$. We assume that $q_X$ is an element of $\mathbb{L}^2([0,1],\mathbb{L}^2([0,1],\mathbb{R}^n))$. Since they are square-integrable maps, we can extend the standard $\mathbb{L}^2$ norm on curve space to measure the differences between these maps:

$$\|q_{X_1} - q_{X_2}\|^2 = \int_0^1 \|q_{X_1}(t,\cdot) - q_{X_2}(t,\cdot)\|^2 dt = \int_0^1 \left(\int_0^1 (q_{X_1}(t,\tau) - q_{X_2}(t,\tau))^2 d\tau\right) dt \text{ .}$$

One can define a geodesic path between any two such SRVFs as a straight line: $\alpha : [0,1] \to \mathbb{L}^2([0,1],\mathbb{L}^2([0,1],\mathbb{R}))$, $\alpha(s) = (1-s)q_{X_1} + sq_{X_2}$. In order to perform elastic analysis of trajectories in $\mathbb{L}^2([0,1],\mathbb{R}^n)$, define the action of the re-parameterization group $\Gamma$ on a trajectory according to $(X,\gamma)(t) = X(\gamma(t))$ (that is, $(X(t,\tau),\gamma(t)) \equiv X(\gamma(t),\tau)$). The corresponding action on the SRVF of $X$ is given by: $(q_X,\gamma)(t) = q_X(\gamma(t))\sqrt{\dot{\gamma}(t)}$. To temporally register any two trajectories $X_1$ and $X_2$, we take the associated SRVFs $q_{X_1}$ and $q_{X_2}$, and solve the following optimization problem:

$$\inf_{\gamma \in \Gamma} \|q_{X_1} - (q_{X_2} \circ \gamma)\sqrt{\dot{\gamma}}\|^2 = \inf_{\gamma \in \Gamma}\left[\int_0^1\left(\int_0^1\left(q_{X_1}(t,\tau) - q_{X_2}(\gamma(t),\tau)\sqrt{\dot{\gamma}(t)}\right)^2 d\tau\right) dt\right] \text{ . } (1)$$

The infimum of this objective function provides a distance between the re-parameterization orbits of $q_{X_1}$ and $q_{X_2}$, and the geodesic equation mentioned above applies with $q_{X_2}$ replaced by $q_{X_2}(\gamma^*(t))\sqrt{\dot{\gamma}^*(t)}$. Fig. 3 shows an example of this idea for $n = 1$. The left panel shows a path $X_1$ in $\mathbb{L}^2$ – for small values of $t$, $X_1(t,\tau)$ is a bimodal function (as a function of $\tau$) that turns into a unimodal function as $t$ increases. We use a time-warping function $\gamma$ to create a new path $X_2(t,\tau) = X_1(\gamma(t),\tau)$ shown in the second panel. Then, we use the corresponding SRVFs to perform alignment via the dynamic programming algorithm. The estimated optimal $\gamma^*$ is shown in the top right panel, drawn over the original $\gamma$. The high degree of overlap implies a high accuracy in alignment of trajectories.

# 3    Shape Analysis of Axonal Trees

Now that we have tools for representing and comparing individual curves, both in $\mathbb{R}^3$ and $\mathbb{L}^2([0,1],\mathbb{R}^3)/\Gamma$, we return to the original problem of comparing shapes of trees. In this section, we specify a mathematical representation of axonal trees as *composite trajectories* and develop a framework for comparing the shapes of these composite trajectories. Our goal is to develop a metric that can be used to: (1) compare the geometries of any two axonal trees. i.e. quantify differences in their shapes, (2) find a geodesic path for morphing one tree into the other, and (3) generate statistical summaries for any finite collection of trees. In this paper we accomplish only the first two goals and leave the third for future.

## 3.1    Representing Trees as Trajectories

We start with representing a tree as a composite trajectory in a certain Hilbert space. We take the viewpoint that attached to the main branch is a continuum of branches – at each point $t \in [0,1]$, the corresponding branch is a curve with a certain shape-scale-orientation attributes.

**Definition 1** *(Composite Trajectory)*
*Define $\beta : [0,1]^2 \to \mathbb{R}^3$ to be a composite trajectory with the following components: (1) The base curve $\beta_0 : [0,1] \to \mathbb{R}^3$, defined by $\beta_0(t) \equiv \beta(t,0)$ is a parameterized curve in $\mathbb{R}^3$ and denotes the main branch. (2) For each t, $\beta^{(t)}(\tau) : [0,1] \to \mathbb{R}^3$ defined by $\beta^{(t)}(\tau) = \beta(t,\tau)$ is a parameterized curve in $\mathbb{R}^3$ representing a side branch attached to the main branch at the point $\beta_0(t)$. In other words, $\beta^{(t)}(0) = \beta_0(t) = \beta(t,0)$.*

For each $t$, we consider the side branch $\beta^{(t)}$ to be smooth and represent its shape-scale-orientation features by its SRVF: $q^{(t)}(\tau) = \frac{d\beta^{(t)}(\tau)}{d\tau} \Big/ \sqrt{\left| \frac{d\beta^{(t)}(\tau)}{d\tau} \right|}$; this $q^{(t)}$ lies in $\mathbb{L}^2([0,1],\mathbb{R}^3)$. Note that given the SRVF $q^{(t)}$ and the starting point $\beta_0(t)$, we can reconstruct the original side branch $\beta^{(t)}$ exactly. Therefore, there is no loss of information in representing the original tree by the set $\left\{ \left( \beta_0(t), q^{(t)} \right) \middle| t \in [0,1] \right\}$. Denoting $\mathcal{P} \equiv (\mathbb{R}^3 \times \mathbb{L}^2([0,1],\mathbb{R}^3))$, the product space of all base points in $\mathbb{R}^3$ and SRVFs of all curves as elements of $\mathbb{L}^2([0,1],\mathbb{R}^2)$. We can impose a metric on it using a weighted sum of metrics on its components. For any $P_1, P_2 \in \mathcal{P}$, where $P_i = (P_i^{(1)}, P_i^{(2)})$ and $P_i^{(1)} \in \mathbb{R}^3$, $P_i^{(2)} \in \mathbb{L}^2([0,1],\mathbb{R}^3)$, we define:
$\|P_1 - P_2\|_{\mathcal{P}} = \sqrt{(1-\lambda) \left\| P_1^{(1)} - P_2^{(1)} \right\|_{\mathbb{R}^3}^2 + \lambda \left\| P_1^{(2)} - P_2^{(2)} \right\|_{\mathbb{L}^2([0,1],\mathbb{R}^3)}^2}$. Here $\lambda \in [0,1]$ is a parameter which controls the relative importance of the base point versus side branches. With this notation, we can represent a tree as a smooth trajectory: $Y : [0,1] \to \mathcal{P}$, $Y(t) = (\beta_0(t), q^{(t)})$; $Y$ is a parameterized curve, or a trajectory, in $\mathcal{P}$ and we are interested comparing such trajectories. (We assume $Y$ to be smooth. If it is not, we can use a smoothing technique to make it smooth.) Towards this goal, we will use the same idea as in the previous section, except the trajectories have an additional information in form of the base curve $\beta_0$. We form SRVFs of these trajectories according to: $Q_Y(t) = (\left\| \frac{dY(t)}{dt} \right\|_{\mathcal{P}})^{-1/2}(\frac{dY(t)}{dt}) = (\|(\frac{d\beta_0(t)}{dt}, \frac{dq^{(t)}}{dt})\|_{\mathcal{P}})^{-1/2}(\frac{d\beta_0(t)}{dt}, \frac{dq^{(t)}}{dt})$.

We will denote the space of such trajectories by $\mathcal{Q} \equiv \mathbb{L}^2([0,1],\mathcal{P})$. The weighted metric on $\mathcal{P}$ can be extended to impose a metric on space $\mathcal{Q}$ as follows. For any $Q_1, Q_2 \in \mathcal{Q}$, where
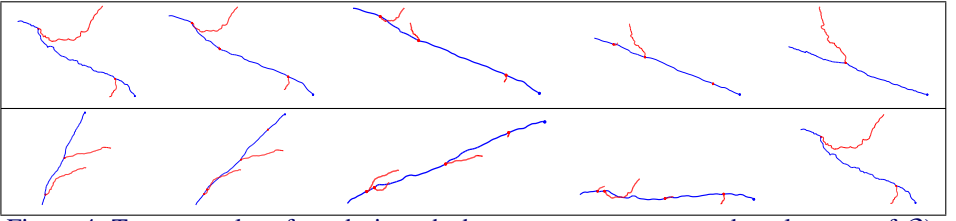
Figure 4: Two examples of geodesic paths between trees represented as elements of $\mathcal{Q}$).

$Q_i(t) = (Q_i^{(1)}(t), Q_i^{(2)}(t)) \in \mathcal{P}$ (i.e. $Q_i^{(1)}(t) \in \mathbb{R}^3$ and $Q_i^{(2)}(t) \in \mathbb{L}^2([0,1], \mathbb{R}^3))$, define the norm:

$$
\begin{aligned}
\|Q_1 - Q_2\|_{\mathcal{Q}}^2 &= \int_0^1 \|Q_1(t) - Q_2(t)\|_{\mathcal{P}}^2 dt = \int_0^1 \left( |Q_1^{(1)}(t) - Q_2^{(1)}(t)|_{\mathbb{R}^3}^2 + \|Q_1^{(2)}(t) - Q_2^{(2)}(t)\|_{\mathbb{L}^2}^2 \right) dt \\
&= \int_0^1 \left( |Q_1^{(1)}(t) - Q_2^{(1)}(t)|_{\mathbb{R}^3}^2 + \left( \int_0^1 |Q_1^{(2)}(t)(\tau) - Q_2^{(2)}(t)(\tau)|_{\mathbb{R}^3}^2 \, d\tau \right) \right) dt .
\end{aligned}
$$

The geodesic path between any two trees, represented by $Q_1, Q_2 \in \mathcal{Q}$ is given by a straight line: $\alpha : [0,1] \to \mathcal{Q}$, $\alpha(s) = (1-s)Q_1 + sQ_2$, For each time point $s$, $\alpha(s) \in \mathcal{Q}$ is a trajectory in $\mathcal{P}$; it can be written as $\alpha(s)(t)$ where $t$ is the parameter of the trajectory in $\mathcal{P}$. Furthermore, for each $s$ and $t$, $\alpha(s)(t)$ has two components $(\alpha^{(1)}(s)(t), \alpha^{(2)}(s)(t))$ where the first component is a point in $\mathbb{R}^3$ denoting the starting point of a side chain and the second component denotes the shape-scale-orientation of that side chain.

Figs. 4 shows two examples of the geodesic paths between two trees in $\mathcal{Q}$. In case when the corresponding points across the two trees – $Q_1(t)$ and $Q_2(t)$ – have non-trivial branches, the geodesic will show one branch deforming into the other. When one of these branches is trivial (length zero), and other is not, we see a growth (or attrition) from a point into a branch. Finally, when both the branches are zero, then there is no change in the structure along the geodesic.

## 3.2 Removing Reparameterization: Elastic Matching

An important strength of this approach is the ability to match curves using nonlinear registrations. We discuss this idea in the context of comparing side branches first and then extend the idea to the comparisons of full trees.

**Matching of Side Branches**: Each (indexed) side branch has been represented an an element of the set $\mathcal{P} = \mathbb{R}^3 \times \mathbb{L}^2([0,1], \mathbb{R}^3)$. In order to study the shape-scale-orientation of this curve, we need to remove the re-parameterization group $\Gamma$. Therefore, we will replace the second component of the set $\mathcal{P}$ by $\mathbb{L}^2([0,1], \mathbb{R}^3)/\Gamma$ to obtain: $\tilde{\mathcal{P}} = \mathbb{R}^3 \times \mathbb{L}^2([0,1], \mathbb{R}^3)/\Gamma$. The distance on this space is given by:

$$
\|P_1 - P_2\|_{\tilde{\mathcal{P}}} = \left( \|P_1^{(1)} - P_2^{(1)}\|_{\mathbb{R}^3}^2 + \|P_1^{(2)} - P_2^{(2)}\|_{\mathbb{L}^2([0,1], \mathbb{R}^3)/\Gamma}^2 \right)^{1/2} .
$$

The geodesic paths between elements of $\tilde{\mathcal{P}}$ are straightforward. The only change is in the second component, $\mathbb{L}^2([0,1], \mathbb{R}^3)/\Gamma$, and geodesics there are given previously.
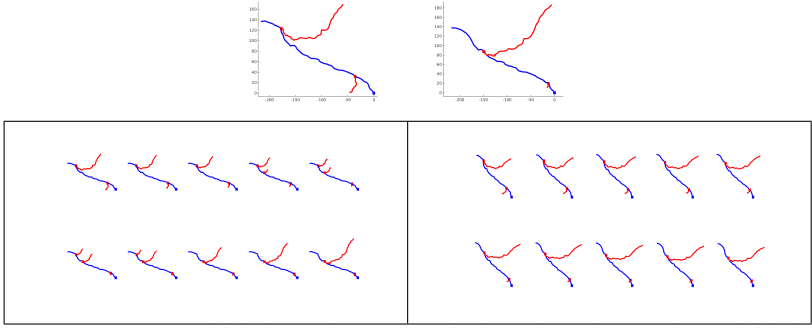
Figure 5: Top: Two trees which have the same main branch and same-shaped side branches, but at different locations Geodesic between them using $\lambda = 0.9$ (left) and 0.999 (right).

**Matching of Main Branch**: The representation of a tree is modified to: $Y : [0,1] \rightarrow \tilde{\mathcal{P}}$, $Y(t) = (\beta_0(t), [q^{(t)}])$, where $\beta_0(t) \in \mathbb{R}^3$, and $[q^{(t)}] \in \mathbb{L}^2([0,1], \mathbb{R}^3)/\Gamma$. Now we are interested in trajectories in the space $\tilde{\mathcal{P}}$ and therefore define a new space $\tilde{\mathcal{Q}} = \mathbb{L}^2([0,1], \tilde{\mathcal{P}})$. We reiterate that for any element $Q \in \tilde{\mathcal{Q}}$ and a time $t \in [0,1]$, $Q(t)$ has two components – one for $\mathbb{R}^3$ denoting the base point and one for $\mathbb{L}^2([0,1], \mathbb{R}^3)/\Gamma$ denoting the elastic shape-scale-orientation space of a side chain. In order to impose a metric on $\tilde{\mathcal{Q}}$ and to compare elements of this space, we will use the same idea as in the previous section, with the obvious modification that $\mathcal{P}$ element is replaced by $\tilde{\mathcal{P}}$. Everything else, including the expressions for the norm, geodesic paths, and geodesic distances, remain same. For each time point $s$, $\alpha(s) \in \tilde{\mathcal{Q}}$ is a trajectory in $\tilde{\mathcal{P}}$; it can be written as $\alpha(s)(t)$ where $t$ is the parameter of the trajectory in $\tilde{\mathcal{P}}$. For each $s$ and $t$, $\alpha(s)(t)$ has two components $(\alpha^{(1)}(s)(t), \alpha^{(2)}(s)(t))$, where the first component is a point in $\mathbb{R}^3$ denoting the starting point of a side chain and the second component denotes the elastic shape-scale-orientation of that side chain.

The re-parameterization group $\Gamma$ acts on the set $\mathcal{Q}$ according to the mapping $(Y, \gamma)(t) = Y(\gamma(t))$. The SRVF of the re-parameterized $Y$ is given by $(Q \circ \gamma)\sqrt{\dot{\gamma}}$. This suggests the matching of full trees according to: $\inf_{\gamma \in \Gamma} \|Q_1 - (Q_2 \circ \gamma)\sqrt{\dot{\gamma}}\|_{\tilde{\mathcal{Q}}}$. One again, the geodesic path is a straight line between the matched trajectories $Q_1$ and $(Q_2, \gamma^*)$, and the geodesic distance is the minimum value obtained in the optimization above.

# 4    Experimental Results

Now we present some preliminary results that demonstrate the elastic tree alignment/comparison outlined above.

**Example 1**: This case, shown in the top of Figure 5, the two trees are identical except that the side branches of the second tree have been artificially moved to a new location, so that they each start at a different location on the main branch. We examine the geodesic between the two trees with two different weights $\lambda = 0.9$ and $\lambda = 0.999$. The two geodesics are shown in the bottom row of Figure 5. Although $\lambda = 0.9$ gives more weight to the side branches, it's low enough in this case that the optimal alignment does not match the larger side branches to each other, so the geodesic simultaneously shrinks one side branch to a null curve while growing a new one. At $\lambda = 0.999$, the side branches are matched to each other and they smoothly deform into one another.
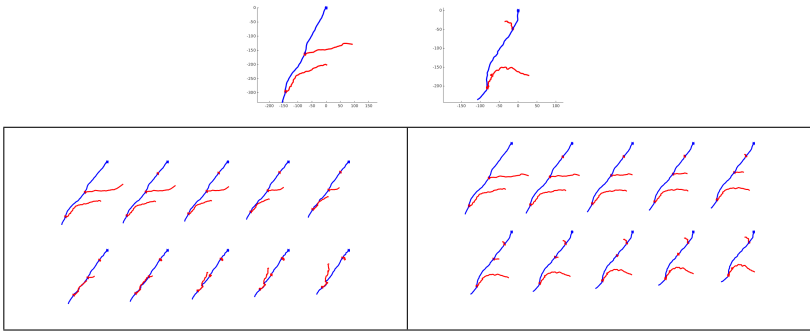
Figure 6: Top: two trees from [15]. Bottom row: Geodesic between them using $\lambda = 0.5$ (left) and 0.999 (right).
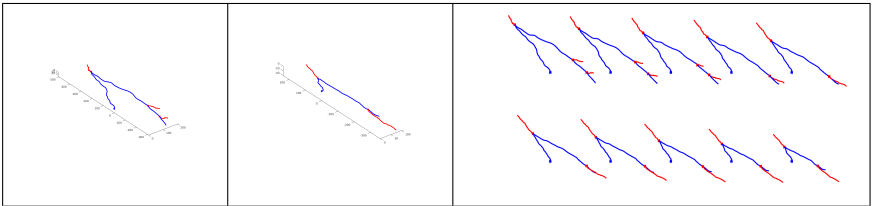


Figure 7: The first two plots are trees from [15] – one with three side branches and one with two. On the right, the array of ten plots is a geodesic between the two trees using $\lambda = 0.999$.

**Example 2**: Here we align two different axonal trees from real data taken from neuromorpho database. The top row of Figure 6 shows the two initial trees and the bottom row shows geodesics using $\lambda = 0.5$ and $\lambda = 0.999$.

**Example 3**: Next, we show an optimal alignment and geodesic path between two trees with different numbers of branches – one with three branches and the other with two. The two trees and their optimal geodesic are depicted in Figure 7. In the optimal alignment, two pairs of branches are matched to each other, and in the geodesic, the remaining unmatched branch is smoothly transformed into the null branch.

# 5 Summary

We have introduced an extension of the previous work, on elastic shape analysis of Euclidean curves, to include curves with infinite-dimensional annotations. We represent an axonal tree by a composite trajectory in the joint position and shape-scale-orientation space of curves, and develop a framework for elastic comparisons of such trajectories. This results in geodesic paths across trees with arbitrary shapes, numbers, and locations of side branches. Some preliminary results are presented.

# References

[1] Giorgio A. Ascoli, Duncan E. Donohue, and Maryam Halavi. Neuromorpho.org: A central resource for neuronal morphologies. *The Journal of Neuroscience*, 27(35): 9247–9251, 2007. URL http://www.jneurosci.org/content/27/35/9247.short.

[2] B. Aydin, G. Pataki, H. Wang, E. Bullitt, and J. S. Marron. A principal component analysis for trees. *Ann. Appl. Stat.*, 3(4):1597–1615, 12 2009. URL http://dx.doi.org/10.1214/09-AOAS263.

[3] B. Aydin, G. Pataki, H. Wang, A. Ladha, E. Bullitt, and J.S. Marron. Visualizing the structure of large trees. *Electron. J. Statist.*, 5:405–420, 2011.

[4] L. J. Billera, S. P. Holmes, and K. Vogtmann. Geometry of the space of phylogenetic trees. *Advances in Applied Mathematics*, 27(4):733–767, 2001.

[5] I. L. Dryden and K. V. Mardia. *Statistical Shape Analysis*. Wiley series in probability and statistics. Wiley, 1998.

[6] A. Feragen, F. Lauze, and M. Nielsen. Fundamental geodesic deformations in spaces of treelike shapes. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 2089–2093, Aug 2010.

[7] A. Feragen, S. Hauberg, M. Nielsen, and F. Lauze. Means in spaces of tree-like shapes. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 736–746, Nov 2011.

[8] A. Feragen, P. Lo, M. de Bruijne, M. Nielsen, and F. Lauze. Toward a theory of statistical tree-shape analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(8):2008–2021, Aug 2013.

[9] S. Kurtek, E. Klassen, Zhaohua Ding, Sandra W. Jacobson, Joseph B. Jacobson, M.J. Avison, and A Srivastava. Parameterization-invariant shape comparisons of anatomical surfaces. *IEEE Transactions on Medical Imaging*, 30(3):849–858, March 2011.

[10] S. Kurtek, A. Srivastava, E. Klassen, and Z. Ding. Statistical modeling of curves using shapes and related features. *Journal of American Statistical Association*, 107(499): 1152–1165, 2012.

[11] W. Liu, A. Srivastava, and E. Klassen. Joint shape and texture analysis of objects boundaries in images using a riemannian approach. In *Asilomar Conference on Signals, Systems, and Computers*, 2008.

[12] W. Liu, A. Srivastava, and J. Zhang. A mathematical framework for protein structure comparison. *PLoS Computational Biology*, 7(2):e1001075, 2011.

[13] A. Mottini. *Axon Morphology Analysis: From Image Processing to Modelling*. PhD thesis, University of Nice Sophia Antipolis, 2014.

[14] A. Mottini, X. Descombes, and F. Besse. From curves to trees: A tree-like shapes distance using the elastic shape analysis framework. *Neuroinformatics*, 13(2):175–191, 2015.

[15] L. F. Santiago, E. G. Rocha, C. L. Santos, and C. W. Picanço-Diniz A. Pereira Jr, J. G. Franca. {S1} to {S2} hind- and forelimb projections in the agouti somatosensory cortex: Axon fragments morphological analysis. *Journal of Chemical Neuroanatomy*, 40(4):339 – 345, 2010.

[16] T.B. Sebastian, P.N. Klein, and B.B. Kimia. Recognition of shapes by editing their shock graphs. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26 (5):550–571, May 2004.

[17] A. Srivastava, E. Klassen, S. H. Joshi, and I. H. Jermyn. Shape analysis of elastic curves in euclidean spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(7):1415–1428, July 2011.

[18] K.-C. Tai. The tree-to-tree correction problem. *Journal of the Association for Computing Machinery*, 26(3):422–433, 1979.

[19] N.H. Trinh and B.B. Kimia. Learning prototypical shapes for object categories. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pages 1–8, June 2010.

[20] H. Wang and J. S. Marron. Object oriented data analysis: Sets of trees. *Ann. Statist.*, 35(5):1849–1873, 10 2007.

[21] Y. Wang, J. S. Marron, B. Aydin, A. Ladha, E. Bullitt, and H. Wang. A nonparametric regression model with tree-structured response. *Journal of the American Statistical Association*, 107(500):1272–1285, 2012. URL http://dx.doi.org/10.1080/01621459.2012.699348.

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