

Towards a theory of statistical tree-shape analysis

Aasa Feragen, Pechin Lo, Marleen de Bruijne, Mads Nielsen, François Lauze
E-mail: {aasa, pechin, marleen, madsn, francois}@diku.dk

Abstract—In order to develop statistical methods for shapes with a tree-structure, we construct a shape space framework for tree-shapes and study metrics on the shape space. This shape space has singularities, which correspond to topological transitions in the represented trees. We study two closely related metrics on the shape space, TED and QED. QED is a quotient Euclidean distance arising naturally from the shape space formulation, while TED is the classical tree edit distance. Using Gromov’s metric geometry we gain new insight into the geometries defined by TED and QED. We show that the new metric QED has nice geometric properties which are needed for statistical analysis: geodesics always exist, and are generically locally unique. Following this we can also show existence and generic local uniqueness of average trees for QED. TED, while having some algorithmic advantages, does not share these advantages. Along with the theoretical framework we provide experimental proof-of-concept results on synthetic data trees as well as small airway trees from pulmonary CT scans. This way, we illustrate that our framework has promising theoretical and qualitative properties necessary to build a theory of statistical tree-shape analysis.

Index Terms—Trees, Tree metric, Graph models, Shape, Anatomical structure, Pattern matching, Pattern recognition, Geometry

1 INTRODUCTION

Tree-shaped objects are fundamental in nature, where they appear, e.g., as delivery systems for gases and fluids [24], as skeletal structures, or hierarchies. Examples encountered in image analysis and computational biology are airway trees [14], [25], [37], vascular systems [7], shock graphs [3], [31], [36], scale space hierarchies [8] and phylogenetic trees [5], [17], [28].

Statistical methods for tree-structured data would have a wide range of applications. For instance, more consistent studies of changes in airway geometry and structure related to airway disease [33], [40] could improve tools for computer aided diagnosis and prognosis.

Due to the wide range of applications, extensive work has been done in the past 20 years on comparison of trees and graphs in terms of matching [14], [25], object recognition [8] and machine learning [15], [18], [29] based on inter-tree distances. However, most existing tree-distance frameworks are algorithmic rather than geometric. Very few attempts [2], [27], [39] have been made to build analogues of the theory for landmark point shape spaces using manifold statistics and Riemannian submersions [20], [21], [32]. There exists no principled general approach to studying the space of tree-structured data, and as a consequence, the standard statistical properties are not well-defined. As we shall see, difficulties appear even in the basic problem of finding the average of two tree-shapes. This paper starts to fill the gap by

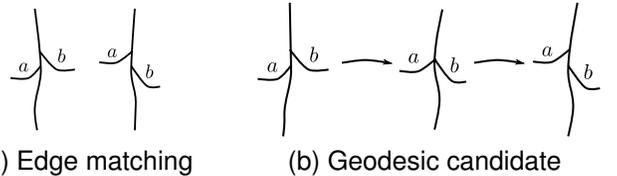


Fig. 1. A good metric must handle edge matchings which are inconsistent with tree topology.

introducing a shape-theoretical framework for geometric trees, which is suitable for statistical analysis.

Most statistical measurements are based on a concept of distance. The most fundamental statistic is the *mean* (or prototype) m for a dataset $\{x_1, \dots, x_n\}$, which can be defined as the minimizer of the sum of squared distances to the data points:

$$m = \operatorname{argmin}_x \sum_{i=1}^n d(x, x_i)^2. \quad (1)$$

This definition of the mean, called the *Fréchet mean*, assumes a space of tree-shapes endowed with a distance d , and is closely connected to geodesics, or shortest paths, between tree-shapes. For example, the midpoint of a geodesic from x_1 to x_2 is a mean for the two-point dataset $\{x_1, x_2\}$. Hence, if there are multiple geodesics connecting x_1 to x_2 , with different midpoints, then there will also be multiple means. As a consequence, without (local, generic) uniqueness of geodesics, most statistical properties are fundamentally ill-posed!

Thus, geometry enters the picture, and the idea of a geodesic tree-space gives a constraint on the possible geometric structure of tree-space. In a shape space

A. Feragen, P. Lo, M. de Bruijne, M. Nielsen and F. Lauze are with the eScience Center, Dept. of Computer Science, University of Copenhagen, Denmark; and M. de Bruijne is with the Biomedical Imaging Group Rotterdam, Depts of Radiology & Medical Informatics, Erasmus MC – University Medical Center Rotterdam, The Netherlands

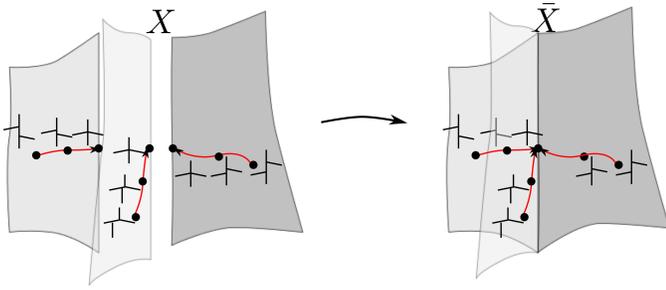


Fig. 2. Continuous transitions in tree topology: Tree-shapes with different tree topology live in different components of tree-space. Paths in different components can terminate in identical collapsed tree-shapes. To identify these, the tree-space components are glued together along subsets with the collapsed tree-shapes.

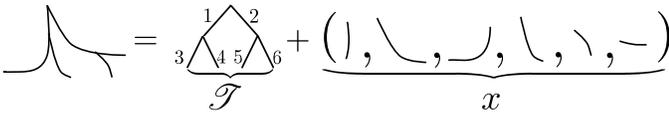


Fig. 3. Tree-shapes are encoded by an ordered binary tree and a set of attributes describing edge shape.

where distances are given by shortest path length, we must be able to continuously deform any given tree-shape into any other by traveling along the shortest path that connects the two shapes. The deformation-paths are easy to describe when only branch shape is changed, while tree topology (branch connectivity) is fixed. Such deformations take place in portions of tree-space where all trees have the same topological structure. It is more challenging to describe deformation-paths in which the tree-topological structure is changed, for instance through a collapsed internal branch as in fig. 1b. We model topologically intermediate trees as collapsed versions of trees with differing tree topology. Different portions of tree-space are glued together along subspaces that correspond to collapsed trees, as in fig. 2. As a consequence, tree-space has self-intersections and is not smooth, but has singularities.

The main theoretical contributions of this paper are the construction of a mathematical tree-shape framework along with a geometric analysis of two natural metrics on the shape space. One of these metrics is the classical Tree Edit Distance (TED), into which we gain new insight. The second metric, called Quotient Euclidean Distance (QED), is induced from a Euclidean metric. Using Gromov’s approach to metric geometry [16], we show that QED generically gives locally unique geodesics and means, whereas finding geodesics and means for TED is ill-posed even locally. We explain why using TED for computing average trees must always be accompanied by a carefully engineered choice of edit paths in order to give unique results; choices which can yield average trees which are substantially different from the trees in the dataset [36]. The QED approach, on

the contrary, allows us to investigate statistical methods for tree-like structures which have previously not been possible, like different well-defined concepts of average tree. This is our motivation for studying the QED metric.

The paper is organized as follows: In Section 1.1 we discuss related work. The tree-space is defined in Sections 2 and 3, and the geometric and statistical properties of tree-space are analyzed in Section 4. In Section 5, we discuss practical computation of geodesics in both metrics, and present a simple QED approximation. In Section 6, we illustrate the properties of QED by computing geodesics and different types of average tree for synthetic planar data trees as well as by computing QED means for sets of small 3D airway trees from human lungs. We finish with a discussion and conclusions in Section 7.

1.1 Related work

Metrics on sets of tree-structured data have been studied by different research communities for the past 20 years. The best-known approach is perhaps Tree Edit Distance (TED), which has been used extensively for shape matching and recognition based on medial axes and shock graphs [31], [36]. TED and, more generally, graph edit distance (GED), are also popular in the pattern recognition community, and are still used for distance-based pattern recognition approaches to trees and graphs [15], [29]. The TED and GED metrics will nearly always have infinitely many shortest edit paths, or geodesics, between two given trees, since edit operations can be performed in different orders and increments. As a result, even the problem of finding the average of two trees is not well posed. With no kind of uniqueness of geodesics, it becomes hard to meaningfully define and compute average shapes or modes of variation. This problem can be solved to some extent by choosing a preferred edit path [15], [29], [36], but there will always be a risk that the choice has negative consequences in a given setting. Trinh and Kimia [36] face this problem when they use TED for computing average medial axes using the simplest possible edit paths, leading to average shapes which can be substantially different from most of the dataset shapes.

Statistics on tree-shaped objects receive growing interest in the statistical community. Wang and Marron [39] study metric spaces of trees and define a notion of average tree called the median-mean as well as a version of PCA, which finds modes of variation in terms of *tree-lines*, encoding the maximum amount of structural and attributal variation. Aydin et al. [2] extend this work by finding efficient algorithms for PCA. This is applied to analysis of brain blood vessels. The metric defined by Wang and Marron does not give a natural geodesic structure on the space of trees, as it places a large emphasis on the tree-topological structure of the trees. The metric has discontinuities in the sense that a sequence of trees with a shrinking branch will not converge to a tree that

does not have that branch. Such a metric is not suitable for studying trees with continuous topological variations and noise, such as anatomical tree-structures extracted from medical images, since the emphasis on topology makes the metric sensitive to structural noise.

A more general problem is analysis of graphs, where a substantial amount of work has also been done. Baloch and Krim [4] analyze the geometry and topology of shapes from their skeletal graphs using Morse theory. Jain and Obermayer [18] study metrics on attributed graphs, represented as incidence matrices. The space of graphs is defined as a quotient of the Euclidean space of incidence matrices by the group of vertex permutations. The graph-space inherits the Euclidean metric, giving it the structure of an orbifold. This construction is similar to the tree-space presented in this paper in the sense that both spaces are constructed as quotients of a Euclidean space. The graph-space framework does not, however, give continuous transitions in internal graph-topological structure, which leads to large differences between the geometries of the tree- and graph-spaces.

Trees also appear in genetics. Hillis et al. [17] visualize sets of phylogenetic trees using multidimensional scaling. Billera et al. [5] invented a geodesic phylogenetic tree-space for analysis of phylogenetic trees, and Owen and Provan [28] have developed fast algorithms for computing geodesics in phylogenetic tree-space. Nye [27] developed the first notion of PCA in phylogenetic tree-space, but makes strict assumptions on principle components being “simple lines” for the sake of computability. Phylogenetic trees are not geometric, and have fixed, labeled leaf sets, making the space of phylogenetic trees much simpler than the space of tree-shapes.

We have previously [12], [13] studied geodesics between small tree-shapes in the same type of singular shape space as studied here, but most proofs have been left out. In [11], we study different algorithms for computing average trees based on the QED metric. This paper extends and continues [12], giving proofs, in-depth explanations and more extensive examples illustrating the potential of the QED metric.

2 THE SPACE OF ORDERED TREE-SHAPES

Let us discuss which properties are desirable for a tree-shape model. We require some form of existence and uniqueness of geodesics in order to compute average trees and analyze variation in datasets. When geodesics exist, we want the topological structure of the intermediate trees along the geodesic to reflect the resemblance in structure of the trees being compared. In particular, *a geodesic passing through the trivial one-vertex tree should indicate that the trees being compared are maximally different*. We want to compare trees where the desired edge matching is inconsistent with tree topology, as in fig. 1a, which requires geodesic deformations in which the tree topology changes, e.g., as in fig. 1b.

2.1 Representation of trees

In this paper, a “tree-shape” is an embedded tree in \mathbb{R}^2 or \mathbb{R}^3 , and consists of a series of edge embeddings, connected as determined by a rooted, ordered combinatorial tree. Tree-shapes are invariant to translation, but our definition of tree-shape does not remove scale and rotation. However, tree-shapes can always be aligned with respect to scale and rotation prior to comparison.

Any tree-shape is represented as a pair (\mathcal{T}, x) consisting of a rooted, planar tree \mathcal{T} with edge attributes x . In $\mathcal{T} = (V, E, r)$, V is the vertex set, $E \subset V \times V$ is the edge set, and r is the root vertex. The tree \mathcal{T} describes the tree-shape topology, and the attributes describe edge shape, as illustrated in fig. 3. The shape attributes, represented by a point $x = (x_e)$ in the product space $\prod_{e \in E} A$, is a concatenation of edgewise attributes from an attribute space A . The attributes could, e.g., be edge length, landmark points or edge parametrizations. In this work, we use open curves translated to start at the origin, described by a fixed number of landmark points along the edge. Thus, throughout the paper, the attribute space A is $(\mathbb{R}^d)^n$ where $d = 2$ or 3 and n is the number of landmark points per edge. Collapsed edges are represented by a sequence of origin points. In some illustrations, we shall use scalar attributes for the sake of visualization, in which case $A = \mathbb{R}$.

In order to compare trees of different sizes and structures, we need to represent them in a unified way. We describe all shapes using the *same* tree \mathcal{T} to encode tree topology. By choosing a sufficiently large \mathcal{T} , we can represent all the trees in our dataset by filling out with empty (collapsed) edges. We call \mathcal{T} *the maximal tree*.

We model tree-shapes using *binary* maximal trees \mathcal{T} . Tree-shapes which are not binary are represented by the binary tree \mathcal{T} in a natural way by allowing constant, or collapsed, edges, represented by the zero scalar or vector attribute. In this way *an arbitrary attributed tree can be represented as an attributed binary tree*, see fig. 4a. This is geometrically very natural. Binary trees are geometrically *stable* in the sense that small perturbations of a binary tree-shape do not change the tree-topological structure of the shape. Conversely, a trifurcation or higher-order vertex can always be turned into a series of bifurcations sitting close together by an arbitrarily small perturbation. In our representation, thus, trifurcations are represented as two bifurcations sitting infinitely close together, etc.

Trees embedded in the plane have a natural left-right order on the children of any edge. We say that a tree is *ordered* whenever each set of sibling edges in the tree is endowed with such a total order. Conversely, an ordered combinatorial tree always has a unique, implicit embedding in the plane where siblings are ordered from left to right. For this reason we use the terms “planar tree” and “ordered tree” interchangeably. We first study metrics on the set of ordered binary trees, which are heavily studied in pattern recognition and vision [3], [31], [36]. Next, we induce distances between unordered

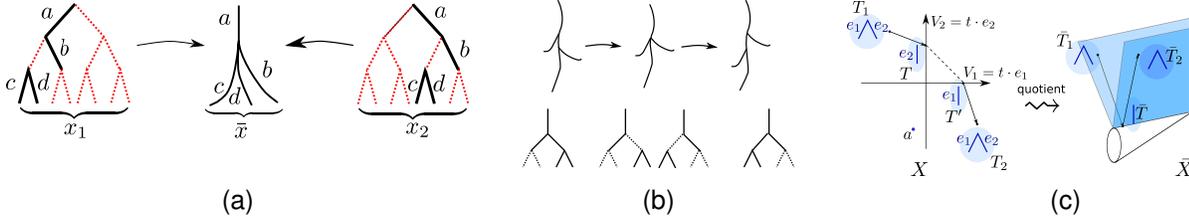


Fig. 4. (a) Higher-degree vertices are represented by collapsing internal edges (dotted line = zero attribute = collapsed edge). We identify those tree pre-shapes whose collapsed structures are identical: The pre-shapes x_1 and x_2 represent the same tree-shape \bar{x} . (b) The tree deformation shown in the top row does not correspond to a path in X , as the two representations of the intermediate tree are found at distinct points in X . (c) The simplest non-trivial tree-shape space, consisting of ordered trees with two edges with scalar attributes. Along the x - and y -axes are trees with a single branch. For each real number a , the tree-shape found at $T' = (a, 0)$ is also represented at $T = (0, a)$. We build the tree-shape space by gluing the different representations of the same tree-shapes (e.g., T and $T' \in X$) together, obtaining the shape space shown on the right. Note the path from \bar{T}_1 to \bar{T}_2 through \bar{T} in \bar{X} on the right; the corresponding path in the pre-shape space X involves a "teleportation" between the representations T and T' of \bar{T} .

trees by considering all possible orders, allowing us to model trees in \mathbb{R}^3 . Considering all orders leads to computational challenges, which are discussed in Section 5.

In order to build a space of tree-shapes, fix an ordered maximal binary tree \mathcal{T} with edges E , which encodes the connectivity of all our trees. Any attributed tree T is now represented by a point $x = (x_e)_{e \in E}$ in $X = \prod_{e \in E} (\mathbb{R}^d)^n$, where the coordinate x_e describes the shape of the edge e . Since we allow zero-attributed edges, some tree-shapes will be represented by several points in X (fig. 4a). As a result, some natural tree-deformations are not found as continuous paths in X . In figs. 4b and 4c, the paths in X corresponding to the indicated deformations require a "teleportation" between two representations of the intermediate tree-shape. We tackle this by using a refined tree-shape space \bar{X} , where different representations of the same tree-shape are identified as being the same point. The original space X is called the tree pre-shape space, analogous to Kendall's terminology [20].

2.2 The singular space of ordered tree-shapes

We go from pre-shapes to shapes by identifying those pre-shapes which define the same shape.

Consider two ordered tree-shapes with collapsed edges. Replace their binary representations by collapsed representations, where the zero attributed edges have been removed. The orders of the original trees induce well-defined orders on the collapsed trees. We say that two ordered tree-shapes are *the same* when their collapsed ordered, attributed versions are identical, as in fig. 4a. More precisely, given a pre-shape tree $x \in X = \prod_E (\mathbb{R}^d)^n$, denote by

$$E_x = \{e \in E | x_e \neq 0\} \quad (2)$$

the set of non-collapsed edges with non-zero attributes of x . Given two pre-shape trees $x, y \in X$ we say that x is equivalent to y , and write $x \sim y$, if there exists an

order-preserving bijection φ preserving edge attributes:

$$\varphi : E_x \rightarrow E_y, \quad y_{\varphi(e)} = x_e. \quad (3)$$

Showing that \sim is actually an equivalence relation is trivial. The quotient space

$$\bar{X} = X / \sim \quad (4)$$

of equivalence classes \bar{x} of ordered-pre-shape trees is the space of ordered trees.

Quotient spaces are standard constructions from topology and geometry [6, Chapter 1.5]. The geometric interpretation of the tree-space quotient is that we "glue" the pre-shape space along the identified subspaces; i.e., when $x_1 \sim x_2$, we "glue" the two points x_1 and x_2 together. See Fig. 4c for an illustration.

Identification of pre-shape trees extends to entire subspaces of X . Indeed, associated to E_x is the subspace V_x of X , spanned by axes of X and defined by

$$V_x = \{z \in X | z_e = 0 \text{ if } e \notin E_x\}. \quad (5)$$

Now, given an ordered bijection $\varphi : E_x \rightarrow E_y$, it induces a linear isomorphism

$$\Phi : V_x \rightarrow V_y, \quad \Phi((z_e)) = z_{\varphi(e)}. \quad (6)$$

which connects each representation z in V_x to a unique representation in V_y .

2.3 Metrics on the space of ordered trees

Given a metric d on the pre-shape-space $X = \prod_{e \in E} (\mathbb{R}^d)^n$, there is a quotient pseudometric [6] \bar{d} induced on the quotient space $\bar{X} = X / \sim$, defined by

$$\bar{d}(\bar{x}, \bar{y}) = \inf \left\{ \sum_{i=1}^k d(x_i, y_i) | x_1 \in \bar{x}, y_i \sim x_{i+1}, y_k \in \bar{y} \right\}. \quad (7)$$

This corresponds to finding the optimal path from \bar{x} to \bar{y} , consisting of any number k of concatenated Euclidean lines, passing through $k - 1$ identified subspaces, as

shown in fig. 4c. It is clear from the definition that the distance function \bar{d} is symmetric and satisfies the triangle inequality. Since (7) is an infimum, though, the distance between two distinct tree-shapes may be zero, as occurs with some intuitive shape distances [26]. This is why \bar{d} is called a *pseudometric*, and it remains to prove that it actually is a metric; i.e., that $\bar{d}(\bar{x}, \bar{y}) = 0$ implies $\bar{x} = \bar{y}$.

We prove this for two specific metrics on X , which come from two different ways of combining individual edge distances: The metrics d_1 and d_2 on $X = \prod_{e \in E} (\mathbb{R}^d)^n$ are induced by the norms

$$\|x - y\|_1 = \sum_{e \in E} \|x_e - y_e\|, \quad (8)$$

$$\|x - y\|_2 = \sqrt{\sum_{e \in E} \|x_e - y_e\|^2}. \quad (9)$$

From now on, d and \bar{d} will denote either the distance functions d_1 and \bar{d}_1 , or d_2 and \bar{d}_2 . In Section 2.4 we prove:

Theorem 1 *The distance function \bar{d} is a metric on \bar{X} , which is a contractible, complete, proper geodesic space.*

Thus, given any two trees, we can always find a geodesic between them in both metrics \bar{d}_1 and \bar{d}_2 .

2.3.1 Geometric interpretation of the metrics

It follows from the definition that the metric \bar{d}_1 coincides with the classical tree edit distance (TED) metric for ordered trees. In this way the abstract, geometric construction of tree-space gives a new way of viewing the intuitive TED metric.

The metric \bar{d}_2 is a descent of the Euclidean metric on \bar{X} , and geodesics in this metric are concatenations of straight lines in flat regions. A qualitative comparison of the TED and QED metrics is made in Section 5.

2.4 Proof of Theorem 1

We now construct a proof of Theorem 1. The rest of the article is independent of this section, and during the first read, the impatient reader may skip to Section 3. However, while the proof is technical, Theorem 1 is a fundamental building block for the shape space framework. We shall assume that the reader has a good knowledge of metric geometry or general topology [6], [9]. It is crucial for the proof that we are only identifying subspaces of the Euclidean space X which are spanned by Euclidean axes, and these are finite in number. This induces a well-behaved projection $p: X \rightarrow \bar{X}$, which carries many properties from X to \bar{X} .

2.4.1 Details and terminology

We say that ordered tree-shapes whose (collapsed) ordered structure is the same, belong to the same *combinatorial tree-shape type*. For each combinatorial type of ordered tree-shape C_j ($j = 1, \dots, K$) which can be represented by collapsing edges in the maximal tree $\mathcal{T} = (V, E, r, <)$, there is a family (E_j^i) of subsets of

E , the E_j^i corresponding to the i^{th} representation of any tree-shape of type C_j . That is, any E_j^i will induce the particular type C_j when the edges in E_j^i ($i = 1, \dots, n_j$) are endowed with nonzero attributes, leaving all other edges with zero attributes. These subsets are characterized by the properties

- the cardinality $|E_j^i| = |E_j^l|$ for all $i = 1, \dots, n_j, j = 1, \dots, K$.
- there is a depth-first order on each E_j^i induced by the depth-first order on E , such that the combinatorial tree-shape type defined by any E_j^i coincides with that defined by E_j^l .

That is, the subset E_j^i for any i lists the set of edges in \mathcal{T} which have nonzero attributes for the i^{th} representation of any shape of type C_j . Corresponding to each E_j^i is the linear subspace V_j^i of X given by

$$V_j^i = \left\{ (x_e) \in \prod_{e \in E} (\mathbb{R}^d)^n \mid x_e = 0 \text{ if } e \notin E_j^i \right\} \quad (10)$$

and by condition b) we can define isometries $\phi_j^i: V_j^i \rightarrow \prod_{e \in E_j^i} \mathbb{R}^d$ by ignoring the zero entries in V_j^i and keeping the depth-first coordinate order. We generate the equivalence \sim on X by asking that $z \sim w$ whenever $\phi_j^i(z) = \phi_j^l(w)$ for some i, j, l . We now define the space of ordered tree-shapes as the quotient $\bar{X} = X / \sim$, and define the quotient map $p: X \rightarrow \bar{X}$.

2.4.2 The pseudometric is a metric

Proposition 1 *Let d denote d_1 or d_2 . The pseudometric \bar{d} is a metric on \bar{X} .*

Proof: It is easy to show that $d_1(\bar{x}, \bar{y}) \geq d_2(\bar{x}, \bar{y})$ for any $\bar{x}, \bar{y} \in \bar{X}$, so it suffices to show that $\bar{d}_2(\bar{x}, \bar{y}) = 0$ implies $\bar{x} = \bar{y}$. Hence, from now on, write \bar{d} for \bar{d}_2 , and assume that $\bar{d}(\bar{x}, \bar{y}) = 0$ for two tree-shapes \bar{x} and \bar{y} .

Choose $\epsilon > 0$ such that

$$\epsilon \ll \min \left\{ \begin{array}{l} \|x_e\| > 0, \\ \|y_e\| > 0, \\ \|x_e - x_{\bar{e}}\| > 0, \\ \|y_e - y_{\bar{e}}\| > 0, \end{array} \middle| \begin{array}{l} x = (x_e) \in \bar{x}, \\ y = (y_e) \in \bar{y} \end{array} \right\}, \quad (11)$$

that is, ϵ is smaller than the size of any of the non-zero edges in \bar{x} and \bar{y} .

We may assume that $x, y \in \bigcup_{i,j} V_j^i$ since otherwise we may assume by symmetry that $\bar{x} = \{x\}$ and $\bar{d}(\bar{x}, \bar{y}) \geq \min\{d(x, \bar{y}), d(x, \bigcup_{i,j} V_j^i)\} > 0$.

Denote by \bar{X}_j the image of V_j^i under the quotient projection $p: X \rightarrow \bar{X}$ for any i .

We may assume that \bar{x} and \bar{y} belong to the same identified subspace; that is, there exist i, j such that

$$\bar{x} \cap V_j^i \neq \emptyset, \quad \bar{y} \cap V_j^i \neq \emptyset \quad (12)$$

since otherwise,

$$\bar{y} \cap \left(\bigcup \{V_j^i \mid \bar{x} \cap V_j^i \neq \emptyset\} \right) = \emptyset \text{ for all } i, j. \quad (13)$$

Because \bar{y} is a finite set, and $\bigcup \{V_j^i \mid \bar{x} \cap V_j^i \neq \emptyset\}$ is a closed set, (13) implies $d(\bar{y}, \bigcup \{V_j^i \mid \bar{x} \cap V_j^i \neq \emptyset\}) > 0$.

In this case, the path will have to go through some V_j^i which does not contain points equivalent to y , and $d(\bar{y}, \bigcup\{V_j^i | \bar{y} \cap V_j^i = \emptyset\}) > \epsilon$, since in order to reach $\bigcup\{V_j^i | \bar{y} \cap V_j^i = \emptyset\}$, we need to remove edge attributes which are nonzero in \bar{y} , and $\epsilon \ll \|y_e\|$ for all $y_e \neq 0$. Thus, eq. 12 holds, and in fact, it holds for all the intermediate path points \bar{x}_i from eq. 7.

But if the path points stay in \bar{X}_j , then the path consists of changing the nonzero edge attributes of the trees in question. This will only give a sum $< \epsilon$ if the trees are identical and the path is constant. \square

2.4.3 Topology of the space of tree-shapes

Here, we prove the rest of Theorem 1, namely that the tree-shape space (\bar{X}, \bar{d}) is a complete, proper geodesic space, and \bar{X} is contractible. First, we note that although \bar{X} is not a vector space, there is a well-defined notion of size for elements of \bar{X} , induced by the norm on X :

Lemma 1 *Note that if $x \sim y$, we must have $\|x\| = \|y\|$; hence we can define $\|\bar{x}\| := \|x\|$.*

Proof: The equivalence is generated recursively from the conditions $x \sim y$ whenever either $x = y$, indicating $\|x\| = \|y\|$; or $\phi_j^i(x) = \phi_j^i(y)$, indicating $\|x\| = \|\phi_j^i(x)\| = \|\phi_j^i(y)\| = \|y\|$ since the ϕ are isometries. Hence, the lemma holds by recursion. \square

We will prove that (\bar{X}, \bar{d}) is a proper geodesic space using the Hopf-Rinow theorem for metric spaces [6], which states that every complete locally compact length space is a proper geodesic space. A length space is a metric space in which the distance between two points can always be realized as the infimum of lengths of paths joining the two points. Note that this is a weaker property than being a geodesic space, as the geodesic joining two points does not have to exist; it is enough to have paths that are arbitrarily close to being a geodesic. It follows from [6, Ch. I, L. 5.20] that (\bar{X}, \bar{d}) is a length space for *any* metric d on X where \bar{d} is a metric.

The tree-shape space is locally compact: Since the projection $p: X \rightarrow \bar{X}$ is finite-to-one, so any open subset U of \bar{X} has as pre-image a finite union $\bigcup_{i=1}^N U_i$ of open subsets of X . Thus, $\text{diam}(U_i) = \text{diam}(U)$ and $p(\bigcup_i \bar{U}_i) = \bar{U}$ is compact whenever U is bounded.

We also need to prove that (\bar{X}, \bar{d}) is complete:

Proposition 2 *Let \bar{d} denote either of the metrics \bar{d}_1 and \bar{d}_2 . The shape space (\bar{X}, \bar{d}) is complete.*

The proof needs a lemma from general topology:

Lemma 2 [9, Chapter XIV, Theorem 2.3] *Let (Y, d) be a metric space and assume that the metric d has the following property: There exists $\epsilon > 0$ such that for all $y \in Y$ the closed ball $\bar{B}(y, \epsilon)$ is compact. Then \bar{d} is complete.* \square

Using the projection $p: X \rightarrow \bar{X}$, we can prove:

Lemma 3 *Bounded closed subsets of \bar{X} are compact.*

Proof: Since Lemma 1 defines a notion of size in \bar{X} , any closed, bounded subspace C in \bar{X} is contained

in a closed ball $\bar{B}_{\bar{d}}(\bar{0}, R)$ in \bar{X} for some $R > 0$, where $\bar{0}$ is the image $p(0) \in \bar{X}$. Since $\|x\| = \|\bar{x}\|$, it follows that $p^{-1}(\bar{B}_{\bar{d}}(\bar{0}, R)) = \bar{B}_d(0, R)$, which is a closed and bounded ball in X . Since closed, bounded subsets of X are compact $\bar{B}_d(0, R)$ is compact. By continuity of p , $\bar{B}_{\bar{d}}(\bar{0}, R)$ is compact. Then C is compact too. \square

It is now very easy to prove Proposition 2:

Proof of Proposition 2: By Lemma 3, all closed and bounded subsets of \bar{X} are compact, but then by Lemma 2 the metric \bar{d} must be complete. \square

Using the Hopf-Rinow theorem [6, Chapter I, Proposition 3.7] we thus prove that (\bar{X}, \bar{d}) is a complete, proper geodesic space. We still miss contractibility:

Lemma 4 *Let B be a normed vector space and let \sim be an equivalence on B such that $a \sim b$ implies $t \cdot a \sim t \cdot b$ for all $t \in \mathbb{R}$. Then $\bar{B} = B / \sim$ is contractible.*

Proof: Define a map $H: \bar{B} \times [0, 1] \rightarrow \bar{B}$ by setting $H(\bar{x}, t) = t \cdot \bar{x}$. Now H is well defined because of the condition on \sim , and $H(\bar{x}, 0) = 0 \forall \bar{x} \in \bar{B}$ so H is a homotopy from $\text{id}_{\bar{B}}$ to the constant zero map. \square

Combining the results of Section 2.4, we see that the proof of Theorem 1 is complete.

3 THE SPACE OF UNORDERED TREE-SHAPES

The world is not two-dimensional, and in most applications trees are embedded in \mathbb{R}^3 . As opposed to planar trees, these trees have no canonical edge order. The order on a planar tree restricts the set of possible edge matchings, whereas for unordered trees, we need to consider all orderings of the same tree and choose orders which minimize the distance. This gives a straightforward mathematical derivation of the space of unordered trees from the space of ordered trees.

The space of unordered 3D tree-shapes is the quotient $\bar{\bar{X}} = \bar{X}/G$, where G is the group of reorderings of the maximal binary tree \mathcal{T} . The metric \bar{d} on \bar{X} induces a quotient pseudometric $\bar{\bar{d}}$ on $\bar{\bar{X}}$. We can prove:

Theorem 2 *For $\bar{\bar{d}}$ induced by either \bar{d}_1 or \bar{d}_2 , the function $\bar{\bar{d}}$ is a metric and the space $(\bar{\bar{X}}, \bar{\bar{d}})$ is a contractible, complete, proper geodesic space.* \square

Theorem 2 can be proven using standard results on compact transformation groups along with similar techniques as for Theorem 1. While considering all different possible orderings of the tree is easy from the point of view of geometric analysis, this becomes a computationally impossible task when the size of the trees grow beyond a few generations. In Section 5 we discuss computational heuristics and approximations in real applications.

4 CURVATURE OF TREE-SPACE AND REDUCED TREE-SPACES

Having proven Theorems 1 and 2, we may now pass to studying the geometry of the tree-shape space through

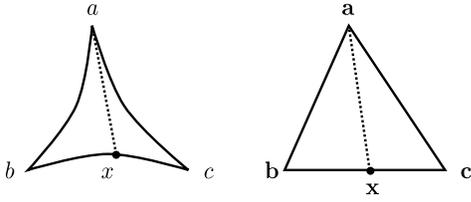


Fig. 5. A metric space is $CAT(0)$ if, for a geodesic triangle abc and for any point x on the triangle, the distance from x to the opposite vertex is not longer than the corresponding distance in the planar comparison triangle abc .

its geodesics. Uniqueness of geodesics and means is closely connected to the geometric notion of *curvature*, a concept which fundamentally depends on the underlying metric. We shall start out by reviewing the concept of curvature in metric geometry [6], [16] before we investigate the curvature of the tree-shape space with the QED and TED metrics. Using curvature, we obtain well-posed local statistical methods for QED. We discuss the implications of locality and explain how locality issues can be dealt with by restricting to smaller tree-spaces that are particularly relevant for a given problem. As part of the exposition we discuss two conjectures on statistics for global generic datasets.

4.1 Curvature in metric spaces

Since general metric spaces can have all kinds of anomalies, the concept of curvature in such spaces is defined through a comparison with spaces that are well understood. Metric spaces are studied using *geodesic triangles*, which are compared with *comparison triangles* in model spaces with a fixed curvature κ . The model spaces are spheres ($\kappa > 0$), the plane \mathbb{R}^2 ($\kappa = 0$) and hyperbolic spaces ($\kappa < 0$), where the metric space curvature is bounded by the curvature κ of the model space. In this paper we shall use comparison with planar triangles, which gives us curvature bounded from above by 0. For an extensive review of metric geometry we refer to [6].

Given a geodesic metric space Y , a *geodesic triangle* abc in Y consists of three points a, b, c and geodesic segments joining them. A planar *comparison triangle* abc for the triangle abc consists of a triangle with vertices a, b, c in the plane, whose side lengths are identical to those of abc , see fig. 5.

A $CAT(0)$ space is a metric space in which geodesic triangles are "thinner" than their planar comparison triangles. That is, $d(x, a) \leq \|x - a\|$ for any x on the edge bc , where x is the unique point on the edge bc such that $d(b, x) = \|b - x\|$ and $d(x, c) = \|x - c\|$. If the planar comparison triangle is replaced by a spherical or hyperbolic comparison triangle of curvature κ , we get a $CAT(\kappa)$ space.

A space Y has non-positive curvature if it is locally $CAT(0)$, i.e., if any point $x \in Y$ has a radius r such that the ball $B(x, r)$ is $CAT(0)$. Similarly, define curvature bounded by κ as being locally $CAT(\kappa)$.

- Example 1**
- The so-called *open book* U obtained by gluing a family of Euclidean spaces U_i together along isomorphic affine subspaces $V_i \subset U_i$ is a $CAT(0)$ space. At any smooth point in U , the local curvature is 0, since the space is locally isomorphic to the corresponding U_i . At any singular point, the local curvature is $-\infty$ ($CAT(\kappa)$ for every κ).
 - The Generalized PCA construction by Vidal et al. [38] defines a $CAT(0)$ space, giving a potential use of $CAT(0)$ spaces and metric geometry in machine learning.
 - As we are about to see, the space of tree-shapes is locally a $CAT(0)$ space almost everywhere.

One of the main reasons why $CAT(\kappa)$ spaces (and $CAT(0)$ in particular) are attractive, is because for $\kappa > 0$, geodesics exist and are unique locally. For $\kappa \leq 0$, geodesics exist and are unique globally.

There are many competing ways of defining central elements given a dataset in a metric space. We consider the *circumcenter* considered in [6], the *centroid* considered, among other places, in [5], and the *mean* [19].

- Definition 1**
- Circumcenters.** Consider a metric space (Y, d) and a bounded subset $S \subset Y$. There exists a unique smallest closed ball $\bar{B}(c_S, r_S)$ in Y which contains S ; the center c_S of this ball is the *circumcenter* of S .
 - The centroid of a finite set.** Let X be a uniquely geodesic metric space (any two points are joined by a unique geodesic). The centroid of a set $S \subset X$ of n elements is defined recursively as a function of the centroids of subsets with $n - 1$ elements as follows: Denote the elements of S by s_1, \dots, s_n . If $|S| = 2$, the centroid $c(S)$ of S is the midpoint of the geodesic joining s_1 and s_2 . If $|S| = n$, define $c^1(S) = \{c(S') : |S'| = n - 1\}$ and $c^k(S) = c^1(c^{k-1}(S))$ for $k > 1$; these are sets with n elements each. If the elements of $c^k(S)$ converge to a point $c \in X$ as $k \rightarrow \infty$, then $c = c(S)$ is the *centroid* of S in X .

It is well-known that centroids [5] and circumcenters [6] are unique in $CAT(0)$ -spaces. Uniqueness of means follows from a more general theorem by Sturm [34, Proposition 4.3]; we include an elementary proof here for completeness.

The problem of existence and uniqueness of means can be attacked using convex functions. Recall that a function $f: [a, b] \rightarrow \mathbb{R}$ is convex if $f((1-s)t + st') \leq (1-s)f(t) + sf(t')$ for all $s \in [0, 1]$ and $t, t' \in [a, b]$. If we can replace \leq with $<$ whenever $s \in]0, 1[$, then f is *strictly convex*. Convex coercive functions have minimizers, which are unique for strictly convex functions. Hence, existence and uniqueness of means can be proven by expressing them as minimizers of strictly convex functions.

We say that a function $f: X \rightarrow \mathbb{R}$ on a geodesic metric space X is (strictly) convex if for any two points $x, y \in X$ and any geodesic $\gamma: [0, l] \rightarrow X$ from x to y , the function $f \circ \gamma$ is (strictly) convex. We shall make use of

the following standard properties of convex functions:

- Lemma 5** i) If $f: \mathbb{R} \rightarrow \mathbb{R}$ and $g: \mathbb{R} \rightarrow \mathbb{R}$ are both convex, g is monotonous and increasing, and g is strictly convex, then $g \circ f$ is strictly convex.
- ii) If $f: \mathbb{R} \rightarrow \mathbb{R}$ and $g: \mathbb{R} \rightarrow \mathbb{R}$ are both convex, then $g + f: \mathbb{R} \rightarrow \mathbb{R}$ is also convex. If either f or g is strictly convex, then $g + f$ is strictly convex as well. \square

The mean of a finite subset $\{x_1, \dots, x_s\}$ in a metric space (X, d) is defined as in eq. 1, and is called the *Fréchet mean*. Local minimizers of eq. 1 are called *Karcher means*.

Theorem 3 Means exist and are unique in $CAT(0)$ -spaces.

Proof: The function $d_y: Y \rightarrow \mathbb{R}$ given by $d_y(x) = d(x, y)$ is convex for any fixed $y \in Y$ by [6, Proposition II.2.2], so the function d_y^2 is strictly convex by Lemma 5 i). But then $D = \sum_{i=1}^s d_{x_i}^2$ is strictly convex by Lemma 5 ii), and a mean is just a minimizer of the function D . The function D is coercive, so the minimizer exists. Since D is strictly convex, the minimizer is unique. \square

4.2 Genericity

Definition 2 (Generic property) A *generic property* in a metric space (X, d) is a property which holds on an open, dense subset of X .

Some of our results, e.g., uniqueness of means in tree-space, will not hold in general, but we shall see that local uniqueness of means is a generic property. This is a topological way of saying that our results hold for a randomly selected dataset with limited variation. In tree-space, generic properties hold almost surely, or with probability one, with respect to natural probability measures. Thus, for a random tree-shape, we can safely assume that it satisfies generic properties, e.g., that the tree-shape is a binary tree. A *non-generic property* is a property whose “not happening” is generic. This is similarly interpreted as a property that may *not* hold for randomly selected tree-shapes. A detailed discussion of the relation between genericity and probability is found in Appendix A in the supplemental material.

Example 2 Being a truly binary tree-shape (i.e. the corresponding collapsed tree-shape is a binary tree) is a generic property. To see this, let \tilde{T} be a tree-shape in \bar{X} or \bar{X} which is not truly binary, represented by a maximal binary tree \mathcal{T} . By adding small noise to the zero attributes on edges of \mathcal{T} , we can obtain truly binary tree-shapes \tilde{T}' which are arbitrarily close to \tilde{T} . Thus, the set of full truly binary tree-shapes in \bar{X} or \bar{X} is open and dense and contained in the set of binary tree-shapes.

One common misunderstanding is that “generic tree” refers to a particular class of trees with a particular generic property. It is important to note that many different properties, which do not necessarily happen on the same subsets of tree-space, may all be generic at the same time. For instance, there are generic properties which do not hold for all binary trees. However, any

finite set of generic properties *will* combine to a new, more restrictive, generic property.

4.3 Curvature in the space of tree-shapes

We now reach our first conclusions on the curvature of the space of tree-shapes.

Theorem 4 (Curvature for ordered tree-shapes)

- i) Endow \bar{X} with the TED metric \bar{d}_1 . The metric space (\bar{X}, \bar{d}_1) does not have locally unique geodesics anywhere, and the curvature of (\bar{X}, \bar{d}_1) is $+\infty$ everywhere.
- ii) Endow \bar{X} with the QED metric \bar{d}_2 . Having local non-positive curvature is a generic property in \bar{X} . That is, the set of points $\bar{x} \in \bar{X}$ which have a neighborhood $U \subset \bar{X}$ in which the curvature is non-positive contain an open, dense set. At the remaining points, the curvature of (\bar{X}, \bar{d}_2) is $+\infty$, i.e. unbounded from above.

Local properties are defined as properties which hold within a sufficiently small neighborhood.

Proof: Case i): Consider any tree-shape $x \in \bar{X}$, where trees in \bar{X} are spanned by $\mathcal{T} = (V, E, r)$ with at least two edges $e_1, e_2 \in E$, and induce a second tree-shape $x' \in \bar{X}$, such that $x'_e = x_e$ for all $e \in E \setminus \{e_1, e_2\}$ and $x'_{e_1} = x_{e_1} + \epsilon$ and $x'_{e_2} = x_{e_2} + \epsilon$. Now there are at least two geodesics from x to x' : One which first deforms x_{e_1} to x'_{e_1} before deforming x_{e_2} to x'_{e_2} , and one which first deforms x_{e_2} to x'_{e_2} before deforming x_{e_1} to x'_{e_1} . Moreover, the TED distance between (\mathcal{T}, x') and (\mathcal{T}, x) can be arbitrarily small. Since this is possible anywhere in tree-space, geodesics are nowhere locally unique and the curvature of tree-space with TED is unbounded everywhere [6, Prop. II 1.4].

Case ii): Recall from Section 2.2 how \bar{X} was formed by identifying subspaces $V_i \subset X$, defined in eq. 6. These identified subspaces corresponded to different representations in X of the same shapes $\bar{x} \in \bar{X}$. The points in \bar{X} can now be divided into three categories:

- Points $\bar{x}_1 \in \bar{X}$ which do not belong to the image of an identified subspace because they only have one representative in X .
- Points $\bar{x}_2 \in \bar{X}$ at which the space \bar{X} is locally homeomorphic to an open book singularity, as in example 1 a).
- Points $\bar{x}_3 \in \bar{X}$ which are preimages of points $x_3 \in X$ at which identified subspaces V_i, V_j intersect in X . An example of such points is the image of the origin in fig. 4c. Infinitely close to these trees, we will find pairs of trees in \bar{X} between which geodesics are not unique, as in fig. 4c.

These three classes of points correspond to local curvature 0, $-\infty$ and $+\infty$, respectively. That is, the space is locally $CAT(0)$ at points in category a); at points from b) it is $CAT(\kappa)$ for every $\kappa \in \mathbb{R}$, so has curvature $-\infty$; and at points from c) it is not $CAT(\kappa)$ for any $\kappa \in \mathbb{R}$; hence the curvature is $+\infty$. It thus suffices to show that the set of points in categories i)-ii) contains an open, dense subset. This follows from the fact that the points

in category *iii*) sit in a lower-dimensional subspace of \bar{X} . \square

It is easy to prove that the same results also hold for unordered tree-shapes. Recall that the space of unordered tree-shapes $\bar{X} = \bar{X}/G$ is the quotient of \bar{X} with respect to the action of a finite group G , which means that \bar{X} is locally well-behaved almost everywhere. In particular, off fixed-points for the action of reorderings $g \in G$ on \bar{X} , the projection $\bar{p}: \bar{X} \rightarrow \bar{X}$ is a local isometry, i.e., it is distance preserving within a neighborhood. Hence, the geometry from \bar{X} is preserved off the fixed points. Geometrically, a fixed point in \bar{X} is an ordered tree-shape where a reordering g of certain branches does not change the ordered tree-shape; that is, some pair of sibling edges must have the same shape attributes. In particular, being a fixed point is a non-generic property in \bar{X} , because the fixed points belong to the lower-dimensional subset of \bar{X} where two sibling edges have identical shape. Thus, we have:

Theorem 5 (Curvature for unordered tree-shapes)

Non-positive curvature is a generic property in the space (\bar{X}, d_2) of unordered trees with the QED metric. With the TED metric \bar{d}_1 , however, \bar{X} has everywhere unbounded curvature, and geodesics are nowhere locally unique. \square

As we shall see in Section 4.5, the practical meaning of Theorems 4 and 5 is that for datasets contained in a sufficiently small neighborhood, there exist unique means, centroids and circumcenters for the QED metric. Moreover, the same techniques cannot be used to prove uniqueness of prototypes for the TED metric. In fact, any geometric method which requires bounded curvature [5], [6], [19] will fail for the TED metric. This motivates our study of the QED metric.

4.4 The locality assumption

We have shown that the local curvature is nonpositive almost everywhere in \bar{X} , which at a first glance makes \bar{X} well suited for geometric definitions of statistical properties. However, our notion of “almost everywhere” is tied to (maximal) dimensionality, which again is tied to the topological structure of the maximal tree \mathcal{T} . As a consequence, the $CAT(0)$ neighborhoods are sometimes small. Below, we shall discuss one example relevant to our applications, where this is not a problem, and we shall discuss two examples where this is a problem. After investigating the locality constraint, we shall discuss natural restrictions on tree-space that allow us to increase the size of the $CAT(0)$ neighborhoods.

Consider the following three examples; we thank one of the anonymous reviewers for the second one.

Example 3 1) Consider the space of tree-shapes spanned by the rooted binary tree T with 3 leaves. This represents, for instance, the subtree of the central airways which spans the segments of the left upper lobe. In this space, the trees at which the tree-space curvature is $+\infty$ are missing two

consecutive edges, which rarely happens in airway trees. Moreover, the spatial orientations of the airway segmental branches in different individuals are fairly similar, although the branching order of given branches varies. This means that a typical dataset is quite local. Thus, in this tree-space, the $CAT(0)$ neighborhoods are not locally Euclidean, and our airway subtrees will usually be contained in a $CAT(0)$ neighborhood.

- 2) Consider the tree-shapes T_1 and T_2 shown in fig. 6a. These two tree-shapes are joined by two geodesics, and thus T_1 and T_2 are not contained in the same $CAT(0)$ neighborhood in \bar{X} . However, in a suitably chosen smaller tree-space \bar{Z} , as defined in Def. 3 below, they can be.
- 3) Consider the tree-shape \tilde{T} in the space of unordered tree-shapes with attributes in \mathbb{R}^2 , spanned by the maximal tree \mathcal{T} as in fig. 6b. Arbitrarily close to \tilde{T} we will find two trees T_1 and T_2 which are joined by two geodesics, as shown in the figure.

As illustrated by examples 2-3, the $CAT(0)$ neighborhoods in the shape-space \bar{X} can be quite small, mainly containing trees whose topology is the same as or similar to \mathcal{T} . On the other hand, both these examples rely on orthogonality or symmetry, as they use trees whose attributes are orthogonal to each other. If one of the trees is slightly perturbed, the geodesics are again unique.

The tree-spaces \bar{X} and \bar{X} are very natural spaces for geometric trees, and based on the symmetry properties in all known examples of non-unique geodesics, we conjecture that having a unique connecting geodesic is a generic property for pairs of trees in \bar{X} or \bar{X} without any locality assumption. For now, however, we propose making application-specific assumptions that restrict to natural reduced tree-spaces $\bar{Z} \subset \bar{X}$ and $\bar{Z} \subset \bar{X}$, whose exact definition gives modeling flexibility, and where $CAT(0)$ neighborhoods can be much larger. This allows us to do well-defined statistics on trees appearing in real applications.

Definition 3 (Reduced tree-shape space) Consider a subset $Z \subset X$, which only contains all representations of trees of certain particular topologies. The i^{th} topology is characterized by a subset $E_i \subset E$ consisting of the edges in the maximal tree \mathcal{T} which have nonzero attributes. Associated to E_i is a linear subspace $Z_i = \prod_{e \in E_i \subset E} (\mathbb{R}^n)^d \subset \prod_{e \in E} (\mathbb{R}^n)^d = X$ containing representations of all the trees of this particular topology. We include all representations of each tree topology (otherwise some shape space paths may disappear), and obtain a reduced preshape space

$$Z = \bigcup_i Z_i \subset \prod_{e \in E} (\mathbb{R}^n)^d = X,$$

containing all the trees that have of one of the considered topologies. The equivalence relation \sim on X restricts to an equivalence relation \sim_Z on Z , from which we obtain a reduced tree-shape space $\bar{Z} = Z / \sim_Z = p(Z) \subset \bar{X}$. The

QED metric d on X induces a metric d_Z on Z which induces a quotient pseudometric \bar{d}_Z on \bar{Z} .

Example 4 Denote by Z the space of all trees in X with n leaves, now \bar{Z} is the space of ordered tree-shapes with n leaves.

All the results that we have proven for \bar{X} and \bar{X} now carry over directly to \bar{Z} and \bar{Z} :

Theorem 6 • *The pseudometric \bar{d}_Z on the reduced space of ordered tree-shapes \bar{Z} is a metric, and (\bar{Z}, \bar{d}_Z) a contractible, complete, proper metric space.*

- *The same holds for the reduced space of unordered tree-shapes: Assume that Z is saturated with respect to the reordering group¹ G . For the corresponding reduced tree-space $\bar{Z} = \bar{Z}/G \subset \bar{X}$, the quotient pseudometric \bar{d}_Z is a metric and the space (\bar{Z}, \bar{d}_Z) is a contractible, complete, proper geodesic space.*
- *Local non-positive curvature is a generic property in \bar{Z} and \bar{Z} . At points in \bar{Z} and \bar{Z} where the curvature is not ≤ 0 , it is $+\infty$.*

Proof: First, we show that the pseudometric is a metric. The pseudometric in equation (7) defines the distance $\bar{d}(\bar{x}, \bar{y})$ as the infimum of lengths of paths in \bar{X} connecting \bar{x} and \bar{y} . Any path in \bar{Z} is also a path in \bar{X} , so if $\bar{d}_{\bar{Z}}(\bar{x}, \bar{y}) = 0$, then $\bar{d}(\bar{x}, \bar{y}) = 0$ as well, so $\bar{x} = \bar{y}$.

The proofs of the other claims follow the proof of Theorems 1, 2, 4 and 5. \square

Example 5 Denote by Z the space of all trees in X with a feasible airway tree topology, as defined in Appendix C. Now \bar{Z} is the reduced space of airway trees.

4.5 Means and related statistics for tree-shapes

In this section we use what we learned in the previous section to show that, given the QED metric on a space of tree-shapes, we can find various forms of average shape in the space of ordered tree-shapes, assuming that the data lie within a $CAT(0)$ neighborhood. We have:

Theorem 7 *Consider the tree-spaces (\bar{X}, \bar{d}_2) and (\bar{X}, \bar{d}_2) . Local uniqueness of means, centroids and circumcenters is generic in both spaces. That is, the set of points with a neighborhood U such that sets contained in U have unique means, centroids and circumcenters, contains an open, dense subset. For the TED metric, these statistics are not locally unique anywhere. The same results hold in the reduced tree-spaces \bar{Z} and \bar{Z} , defined in Def. 3.*

Proof: First consider \bar{X} and \bar{X} with the QED metric. By Theorems 4 and 5, \bar{X} and \bar{X} are both locally $CAT(0)$ spaces, and by Theorems 1 and 2 they are both complete metric spaces. We have seen in Theorem 3 that means exist and are unique in $CAT(0)$ spaces, so the statement holds for means. By [6, Proposition 2.7], any subset Y

1. Z is saturated if, for each tree topology appearing in Z , all reorderings of the same tree topology also appear in Z . Equivalently, $Z = p^{-1}(p(Z))$, where $p: \bar{X} \rightarrow \bar{X}/G$

of a complete $CAT(0)$ space has a unique circumcenter. Hence, the statement holds for circumcenters. Similarly, by [5, Theorem 4.1], finite subsets of $CAT(0)$ spaces X have centroids (unique by definition), so the statement holds for centroids.

For the TED metric, note that for any 2-point dataset, all these notions of mean coincide with the midpoint of a geodesic connecting the points. We know that geodesics and midpoints are not unique in the TED metric.

The proofs transfer directly to the spaces \bar{Z} and \bar{Z} . \square

All the results from Section 4.5 hold generically, within a neighborhood where the $CAT(0)$ property holds. We have seen examples of points where local $CAT(0)$ -property must fail, illustrating the limitations of $CAT(0)$ techniques for analyzing general trees. However, as discussed the examples where the $CAT(0)$ property fails rely on non-generic symmetries, and we conjecture that having a unique mean is a generic property for finite datasets in \bar{X} and \bar{X} . Analogous results for manifolds have appeared very recently [1], and we expect this extension from existing models to be nontrivial. In the meantime, as discussed in the previous section, application-specific reduced tree-spaces can be designed where $CAT(0)$ -neighborhoods are larger.

In manifold statistics, means are unique [19] in any convex neighborhood in which geodesics are unique. In tree-space, any convex neighborhood which does not contain points of curvature $+\infty$ will be a $CAT(0)$ neighborhood. In tree-space, radius is not a good measure for the size of a $CAT(0)$ neighborhood, as a tree may contain both small branches, leading to a small radius, as well as large branches, giving much larger $CAT(0)$ neighborhoods than indicated by the radius.

5 GEODESICS AND THEIR COMPUTATION

Computation of tree-space geodesics is not trivial, and it is an open problem to find a QED heuristic, exact or not, which is realistically computable for trees of moderate size. In this section we first compare the qualitative properties of QED and TED geodesics, and then discuss strategies for approximation which we use for computing geodesics between small trees. We stress that the main focus of this article is to investigate the potential of QED for tree-statistics. In this setting, proof-of-concept experiments are important for judging whether finding computable algorithms for QED between trees of arbitrary size is worthwhile, as we consider this a nontrivial research problem. The main problem is that the QED distance is optimal among an exponential number of cases, each of which involves a nontrivial optimization term. In order to make proof-of-concept experiments, we use approximation and semi-labeling to restrict the search space, and stick to small trees where a complete enumeration of the remaining options is feasible.

For a qualitative comparison of TED and QED geodesics, we compare edge matchings implied by the TED and QED geodesics between small, simple trees.

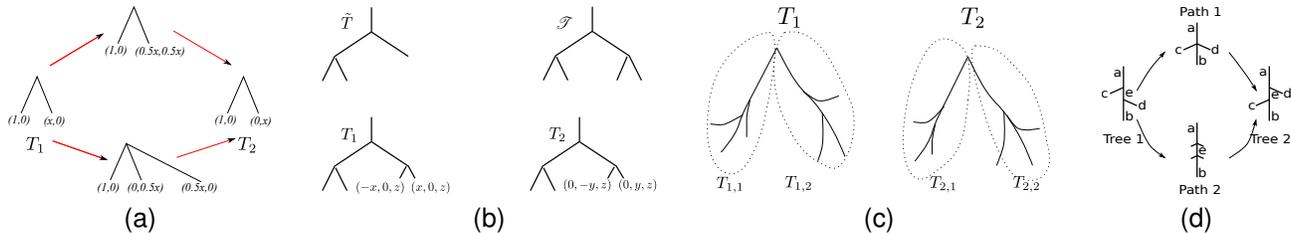


Fig. 6. (a) Two trees joined by two different geodesics in \bar{X} . (b) Arbitrarily close to \tilde{T} are two other trees T_1 and T_2 who cannot be joined by a unique geodesic. (c) Local-to-global properties of TED: If the trees T_1 and T_2 are decomposed into subtrees $T_{1,1}, T_{1,2}$ and $T_{2,1}, T_{2,2}$ as in fig. 6c, such that the geodesic from T_1 to T_2 restricts to geodesics between $T_{1,1}$ and $T_{2,1}$ as well as $T_{1,2}$ and $T_{2,2}$, then $d_1(T_1, T_2) = d_1(T_{1,1}, T_{2,1}) + d_1(T_{1,2}, T_{2,2})$. (d) Two options for structural transition in a path from Tree 1 to Tree 2.

Consider the two tree-paths in fig. 6d, where the edges are endowed with scalar attributes $a, b, c, d, e \in \mathbb{R}_+$ describing edge length. Computing the TED and QED costs of the two different paths, we find the shortest path.

Path 1 indicates a matching left and right side edges c and d , while Path 2 does not make the match. The cost of Path 1 is $2e$ in both metrics, while the cost of Path 2 is $2\sqrt{c^2 + d^2}$ in the QED metric and $2(c + d)$ in the TED metric. TED will identify the c and d edges whenever $e^2 \leq c^2 + 2cd + d^2$, while QED makes the identification whenever $e^2 \leq \frac{1}{2}(c^2 + d^2)$. Thus, in these cases, TED is more prone to internal structural change than QED.

The same occurs in the comparison of TED and QED matching in figs. 7a and 7b. Although the TED is more prone to matching trees with different tree-topological structures, the edge matching results are similar, as is expected, since the metrics are closely related.

Computing TED or QED distances between unordered tree-shapes is NP-complete [10], but we can often use geometry and prior knowledge (e.g., anatomy) as a heuristic to reduce the search space. For instance, trees appearing in applications are often not completely unordered, but are semi-labeled.

Example 6 (Semi-labeling of the central airway tree)

The airway tree has about 30 branches which have anatomical names and can be identified by experts. For instance, the root edge is the trachea; the second generation edges are the left and right main bronchi. As these top-generation branches are easily identified, we use their identification to simplify computations of inter-airway geodesics in Section 6.2.

The TED metric has nice decomposition properties which are used in dynamic programming algorithms for TED, see fig. 6c. QED does not share these decomposition properties, due to the square root in the Euclidean distance, and thus we have chosen to approximate the QED in our experiments. In order to approximate the QED distance, we take advantage of the fact that many types of biological trees have relatively few tree-topological variations. The airways is an example of such a type of tree, as explained in Example 6. For

such tree-types, it is safe to assume that the number of internal structural transitions found in a geodesic deformation is low, and that the geodesics pass through identified subspaces of low codimension. For instance, for the airway trees studied in Section 6.2 below, we find empirically that it is enough to allow for one structural change in each lung subtree, which has at most a trifurcation. Recall the definition of the metric from eq. 7; the approximation consists of imposing upper bounds K on the number k and D on the degrees of internal vertices appearing in eq. 7, respectively, giving:

$$\inf_{k \leq K} \left\{ \sum_{i=1}^k d(x_i, y_i) \mid x_1 \in \bar{x}, y_i \sim x_{i+1}, y_k \in \bar{y} \right\}, \quad (14)$$

where all x_i and y_i have vertex degrees at most D . Geometrically, k is the number of Euclidean segments concatenated to form the geodesic. Bounding k is equivalent to bounding the number of topological transitions throughout the geodesic. In fact, $k = 1 + \text{the number of topological transitions in the geodesic}$. Note that this approximation is similar to that employed for TED in [35].

We approximate the QED metric using Algorithm 1; see Appendix B for details. For the unordered airway trees in Section 6.2, Algorithm 1 is combined with a search through all tree orders. The algorithm was implemented for trees with depth at most 3 by listing all options for intermediate tree topologies within the approximation, optimizing for the branch attributes found in the intermediate topologies, choosing the optimal path among the considered ones. This is possible since the trees are small and the enumeration of options can be done manually. Finding more exact and efficient heuristics is an important but nontrivial research problem.

6 EXPERIMENTAL RESULTS

The QED metric is new, whereas the properties of the TED metric are well known [31]. Our experimental results on real and synthetic data illustrate the geometric properties of the QED metric. The experiments on airway trees in Section 6.2 show, in particular, that it is feasible to compute geodesics between real, 3D data trees. In all our

origin tree	endpoint tree/ distance from origin			
	 273	 282	 405	 407
	 273	 276	 313	 448
	 180	 282	 313	 395
	 276	 348	 395	 407
	 180	 348	 405	 448

(a) Matching in the QED metric.

origin tree	endpoint tree/ distance from origin			
	 455	 606	 625	 832
	 455	 549	 656	 783
	 472	 606	 656	 678
	 549	 625	 678	 746
	 472	 746	 783	 832

(b) Matching in the TED metric.

Fig. 7. Given a set of five data trees, we match each to the four others in both metrics.

Algorithm 1 Computing approximate QED distances between ordered, rooted trees with up to $k = K - 1$ structural transitions through trees of order at most D

- 1: x, y planar rooted binary trees in \bar{X} .
- 2: $\mathbf{V} = \{V\}$ set of ordered isometric pairs $V = \{V_1, V_2\}$ of subspaces of X in which the same ordered tree-shapes are representing, as in eq. (6).
- 3: **for** $\tilde{\mathbf{V}} = \{V^1, \dots, V^s\} \subset \mathbf{V}$ with $s \leq k$ **do**
- 4: $p_{\tilde{\mathbf{V}}} = \operatorname{argmin}\{f(p) | p_i^1 \in V_1^i, p_i^2 \in V_2^i, p_0^2 \in x, p_s^1 \in y\}$, where $p = (p_0^2, p_1^1, p_1^2, \dots, p_s^1)$ and $f(p) = \sum_{i=0}^s d(p_i^2, p_{i+1}^1)$.
- 5: $d_{\tilde{\mathbf{V}}} = f(p_{\tilde{\mathbf{V}}})$.
- 6: **end for**
- 7: $\tilde{\mathbf{V}}_{\min} = \operatorname{argmin}\{d_{\tilde{\mathbf{V}}} | \tilde{\mathbf{V}} \subset \mathbf{V}, |\tilde{\mathbf{V}}| \leq k\}$
- 8: $d = d_{\tilde{\mathbf{V}}_{\min}}$
- 9: $p = p_{\tilde{\mathbf{V}}_{\min}}$
- 10: geodesic $= g = \{x \ni p_0^2 \rightarrow p_1^1 \sim p_1^2 \rightarrow p_2^1 \sim p_2^2 \rightarrow \dots \rightarrow p_s^1 \sim p_s^2 \in y\}$
- 11: **return** d, g

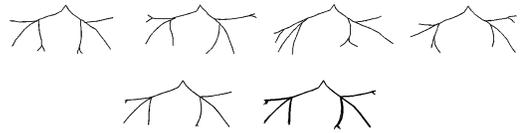


Fig. 8. Top: A small set of synthetic planar trees. Bottom left: The centroid tree. Bottom right: The mean tree.

experiments, edges are translated to start at $0 \in \mathbb{R}^d$, and are represented by 6 landmark points evenly distributed along the edge, the first at the origin.

6.1 Synthetic planar trees

Movies illustrating geodesics between planar depth 3 trees, as well as a table illustrating a tree-shape matching experiment for a set of synthetic planar depth 3 trees, are found on the web page http://image.diku.dk/aasa/tree_shape/planar.html. The movies and matchings show the geometrically intuitive behavior of the geodesic deformations. We see that the intermediate structures resemble the geodesic endpoint trees in a reasonable manner, and show the ability of the geodesics to handle internal topological differences. These are among the wanted properties from the model listed in the beginning of Section 2.

QED mean (computed with the weighted midpoints algorithm of [11], [34]) and centroid [5] trees are shown in fig. 8. The mean and centroid trees clearly represent the main common properties of the dataset trees.

6.2 Results in 3D: Pulmonary airway trees

As proof-of-concept experiments on real data we compute means for sets of pulmonary airway trees from the EXACT'09 airway segmentation competition [23]. The airway trees were first segmented from low dose screening computed tomography (CT) scans of the chest using a voxel classification based airway tree segmentation algorithm [22]. The centerlines were extracted from the segmented airway trees using a modified fast marching algorithm based on [30]. The method gives a tree structure directly through connectivity of parent and children branches. For simplicity, we only consider the central airways down to the lobar branches.

We compute TED means [36] and QED means (using the weighted midpoints algorithm [11], [34]). The dataset is shown in fig. 9a; the results in fig. 9b. We clearly see how the choice of TED geodesic makes the TED mean vulnerable to missing branches in the dataset trees, giving a mean shape whose structure is too simple. The QED mean, on the other hand, represents the data well.

7 DISCUSSION AND CONCLUSION

We define a tree-shape space, starting from a purely geometric point of view. This geometric framework allows us to simultaneously take both global tree-topological

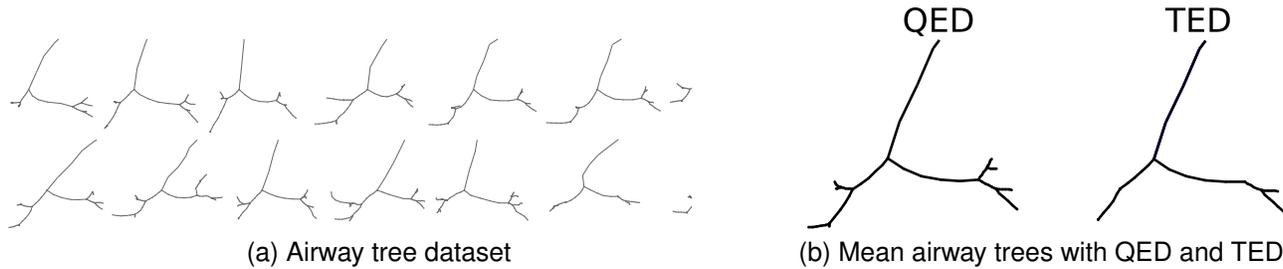


Fig. 9. (a) Subtrees of 14 airway trees from the EXACT'09 test set [23]. (b) The QED and TED mean airway trees. The TED mean misses the upper lobar branches on both sides, caused by the fact that the 13th data tree misses these.

structure and local edgewise geometry into account. We study two metrics on this shape space, TED and QED, which both give the shape space a geodesic structure (Theorems 1 and 2). The framework is developed for tree-shapes, but carries over to other geometric trees with vector attributes.

QED is the geometrically most natural metric, and turns out to have essential properties which allows statistical shape analysis for local data. Through a geometric analysis of the tree-shape framework, we show that the QED metric gives local uniqueness of geodesics and local existence and uniqueness for three versions of average shape, namely the mean, the circumcentre and the centroid (Theorem 7). Our analysis gives new insight into the metric space defined by TED, and explains why the problem of finding TED-averages is ill-posed.

In order to obtain unique averages with the help of $CAT(0)$ spaces, we assume that our data are contained in a neighborhood which is "sufficiently small" to be $CAT(0)$. We have seen that in some cases, these neighborhoods are small. While we conjecture that having a unique mean is a generic property for finite datasets without any locality assumption, we also show how to obtain larger $CAT(0)$ neighborhoods by restricting to reduced tree-spaces, where only trees of certain topologies, chosen in an application-specific manner, are found. Even with such an assumption, our results significantly extend the state-of-the-art, which either uses the TED metric [31], [36], or is unable to simultaneously handle topology and geometry [39].

Both metrics are generally NP complete to compute for 3D trees. We explain how approximation and semi-labeling can be used to compute geodesics, and illustrate this by computing QED means for sets of trees extracted from pulmonary airway trees as well as synthetic planar data trees. The QED lacks an obvious decomposition strategy for dynamic programming, and involves an optimization term which makes it more time consuming to compute than the TED metric. The two metrics have similar matching properties, and thus, for applications which do not require a unique geodesic (e.g., clustering) the TED metric is just as suitable as the QED metric. However, for statistical computations, the QED metric is the only suitable metric. This is emphasized by our experiments, where computed QED and TED means illustrate that the TED means miss important structure.

Future research will center around two points: Development of nonlinear statistical methods for the singular tree-shape spaces, and finding more efficient approximations and heuristics for the QED metric using both the tree geometry and the tree-space geometry. The latter will pave the road for computing averages and modes of variation for large, real 3D data trees, as well as considering more sophisticated edge shape spaces from modern shape analysis [21], [26]. Construction of efficient algorithms is by no means trivial, due to the non-smooth structure of the tree-shape space and the complexity of computing exact geodesics.

ACKNOWLEDGEMENTS

We thank the anonymous reviewers for their helpful comments, which improved the quality of the article. This work is partly funded by the Lundbeck foundation, the Danish Council for Strategic Research (NABIIT projects 09 – 065145 and 09 – 061346) and the Netherlands Organisation for Scientific Research (NWO). We thank Søren Hauberg for helpful discussions during the preparation of the article.

REFERENCES

- [1] M. Arnaudon and L. Miclo. Means in complete manifolds: Uniqueness and approximation. *arxiv.org*, 2012.
- [2] B. Aydin, G. Pataki, H. Wang, E. Bullitt, and J. S. Marron. A principal component analysis for trees. *Ann. Appl. Statist.*, 3(4):1597–1615, 2009.
- [3] X. Bai, X. Wang, L. J. Latecki, W. Liu, and Z. Tu. Active skeleton for non-rigid object detection. In *ICCV*, pages 575–582, 2009.
- [4] S.H. Baloch and H. Krim. 2d shape modeling using skeletal graphs in a morse theoretic framework. In *Statistics and Analysis of Shapes*, pages 61–80. 2006.
- [5] L. J. Billera, S. P. Holmes, and K. Vogtmann. Geometry of the space of phylogenetic trees. *Adv. in Appl. Math.*, 27(4):733–767, 2001.
- [6] M. R. Bridson and A. Haefliger. *Metric spaces of non-positive curvature*. Springer, 1999.
- [7] C. Chalopin, G. Finet, and I. E. Magnin. Modeling the 3D coronary tree for labeling purposes. *Med. Im. Anal.*, 5:301–315, 2001.
- [8] M. Demirci, B. Platel, A. Shokoufandeh, L. Florack, and S. Dickinson. The representation and matching of images using top points. *J. Math. Im. Vis.*, 35:103–116, 2009.
- [9] J. Dugundji. *Topology*. Allyn and Bacon Inc., 1966.
- [10] A. Feragen. Complexity of computing distances between geometric trees. In *SSPR/SPR*, volume 7626 of *Lecture Notes in Computer Science*, pages 89–97, 2012.
- [11] A. Feragen, S. Hauberg, M. Nielsen, and F. Lauze. Means in spaces of tree-like shapes. In *ICCV*, 2011.
- [12] A. Feragen, F. Lauze, P. Lo, M. de Bruijne, and M. Nielsen. Geometries on spaces of treelike shapes. In *ACCV 2010*, volume 6493 of *LNCS*, pages 160–173. 2011.

- [13] A. Feragen, F. Lauze, and M. Nielsen. Fundamental geodesic deformations in spaces of treelike shapes. In *Proc ICPR*, pages 2089–2093, 2010.
- [14] A. Feragen, J. Petersen, M. Owen, P. Lo, L.H. Thomsen, M.M.W. Wille, A. Dirksen, and M. de Bruijne. A hierarchical scheme for geodesic anatomical labeling of airway trees. In *MICCAI, in press*, 2012.
- [15] M. Ferrer, E. Valveny, F. Serratos, K. Riesen, and H. Bunke. Generalized median graph computation by means of graph embedding in vector spaces. *Pattern Rec.*, 43(4):1642–1655, 2010.
- [16] M. Gromov. Hyperbolic groups. In *Essays in group theory*, volume 8 of *Math. Sci. Res. Inst. Publ.*, pages 75–263. Springer, 1987.
- [17] D. M. Hillis, T. A. Heath, and K. St. John. Analysis and visualization of tree space. *System. Biol.*, 54(3):471–482, 2005.
- [18] B. J. Jain and K. Obermayer. Structure spaces. *J. Mach. Learn. Res.*, 10:2667–2714, 2009.
- [19] H. Karcher. Riemannian center of mass and mollifier smoothing. *Comm. Pure Appl. Math.*, 30(5):509–541, 1977.
- [20] D. G. Kendall. Shape manifolds, Procrustean metrics, and complex projective spaces. *Bull. London Math. Soc.*, 16(2):81–121, 1984.
- [21] S. Kurtek, E. Klassen, J.C. Gore, Z. Ding, and A. Srivastava. Elastic geodesic paths in shape space of parameterized surfaces. *TPAMI*, 34(9):1717–1730, 2012.
- [22] P. Lo, J. Sporning, H. Ashraf, J.J.H. Pedersen, and M. de Bruijne. Vessel-guided airway tree segmentation: A voxel classification approach. *Med. Im. Anal.*, 14(4):527–538, 2010.
- [23] P. Lo, B. van Ginneken, J.M. Reinhardt, and M. de Bruijne. Extraction of Airways from CT (EXACT09). *IEEE TMI*, 2012.
- [24] B. B. Mandelbrot. *The fractal geometry of nature*. W. H. Freedman and Co., 1982.
- [25] J. H. Metzen, T. Kröger, A. Schenk, S. Zidowitz, H-O. Peitgen, and X. Jiang. Matching of anatomical tree structures for registration of medical images. *Im. Vis. Comp.*, 27:923–933, 2009.
- [26] P. W. Michor and D. Mumford. Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc. (JEMS)*, 8:1–48, 2004.
- [27] T. M. W. Nye. Principal components analysis in the space of phylogenetic trees. *Ann. Statist.*, 39(5):2716–2739, 2011.
- [28] M. Owen and J. S. Provan. A fast algorithm for computing geodesic distances in tree space. *IEEE/ACM Trans. Comp. Biol. Bioinf.*, 8:2–13, 2011.
- [29] K. Riesen and H. Bunke. Graph classification based on vector space embedding. *Int. J. Pattern Rec. & Artif. Int.*, 23(6):1053–1081, 2009.
- [30] T. Schlathölter, C. Lorenz, I. C. Carlsen, S. Renisch, and T. Deschamps. Simultaneous segmentation and tree reconstruction of the airways for virtual bronchoscopy. In *SPIE Med. Im.*, volume 4684, pages 103–113, 2002.
- [31] T. B. Sebastian, P.N. Klein, and B.B. Kimia. Recognition of shapes by editing their shock graphs. *TPAMI*, 26(5):550–571, 2004.
- [32] C. G. Small. *The statistical theory of shape*. Springer, 1996.
- [33] L. Sørensen, P. Lo, A. Dirksen, J. Petersen, and M. de Bruijne. Dissimilarity-based classification of anatomical tree structures. In *Inf. Proc. Med. Im.*, LNCS, pages 475–485, 2011.
- [34] K.-T. Sturm. Probability measures on metric spaces of nonpositive curvature. volume 338 of *Contemp. Math.*, pages 357–390. 2003.
- [35] H. Touzet. A linear tree edit distance algorithm for similar ordered trees. In *Proc. 16th Annual Symposium on Combinatorial Pattern Matching (CPM)*, LNCS, pages 334–345, 2005.
- [36] N.H. Trinh and B.B. Kimia. Learning prototypical shapes for object categories. In *CVPR Workshops*, pages 1–8, 2010.
- [37] B. van Ginneken, W. Baggerman, and E. van Rikxoort. Robust segmentation and anatomical labeling of the airway tree from thoracic CT scans. In *MICCAI 2008*, pages 219–226, 2008.
- [38] R. Vidal, Y. Ma, and S. Sastry. Generalized Principal Component Analysis (GPCA). *TPAMI*, 27:1945–1959, 2005.
- [39] H. Wang and J. S. Marron. Object oriented data analysis: sets of trees. *Ann. Statist.*, 35(5):1849–1873, 2007.
- [40] G. R. Washko, T. Dransfield, R. S. J. Estepar, A. Diaz, S. Matsuoka, T. Yamashiro, H. Hatabu, E. K. Silverman, W. C. Bailey, and J. J. Reilly. Airway wall attenuation: a biomarker of airway disease in subjects with COPD. *J. Appl. Physiol.*, 107(1):185–191, 2009.



Aasa Feragen received her PhD in mathematics from the University of Helsinki, Finland, in 2010. She is currently a Freja Fellow and Assistant Professor at the Department of Computer Science, University of Copenhagen, Denmark. Her research interests include mathematical modeling and computational anatomy for biomedical image analysis, as well as statistics and machine learning for structured data.



Pechin Lo received his PhD degree in computer science from the University of Copenhagen in 2010. Currently, he is a postdoctoral researcher in University of California, Los Angeles. His research interests include image segmentation, pattern recognition and machine learning applied to the field of medical image processing.

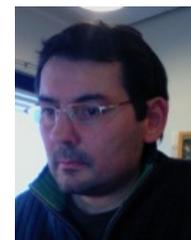


Marleen de Bruijne is associate professor of medical image analysis at Erasmus MC Rotterdam and at University of Copenhagen. She received the MSc degree in physics (1997) and PhD degree in medical imaging (2003) from Utrecht University. Dr. de Bruijne has (co-)authored over 100 peer-reviewed papers in international conferences and journals. She was on the program committee of many international conferences in medical imaging and computer vision and is a member of the editorial board of *Medical Image Analysis*. Her research interests are model based and quantitative analysis of medical images in various applications.



of the editorial board of *IJCV* and *JMIV*.

Mads Nielsen received his PhD in computer science in 1995 from the University of Copenhagen. In 1995-96 was post doc at Utrecht University and The University of Copenhagen, where he was assistant professor 1997-99. He became associate professor in 1999 and professor in 2002 at the IT University of Copenhagen, where he headed the Image Analysis Group. In 2007 he became professor and head of the Image Group, University of Copenhagen. He was general chair of MICCAI 2006 and is a member



and geometric methods for Image Analysis.

François Lauze studied Mathematics in France, University of Nice-Sophia-Antipolis where he was awarded a PhD in Algebraic Geometry in 1994. He spent some years in Burkina Faso, West Africa, where he taught Mathematics at the University of Ouagadougou. He then moved to Denmark and engaged in yet another PhD, awarded in 2004 at the IT University of Copenhagen, including work on Variational Methods for Motion Compensated Inpainting and motion recovery. He has since worked with variational