Dept. of Computer Science, University of Copenhagen

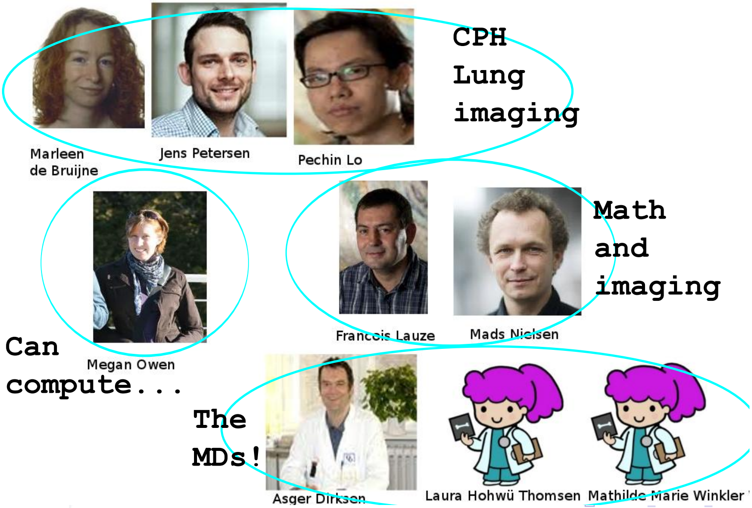
Anatomical tree-shape analysis

Martinos Center for Biomedical Imaging, Boston
May 29 2012

Aasa Feragen

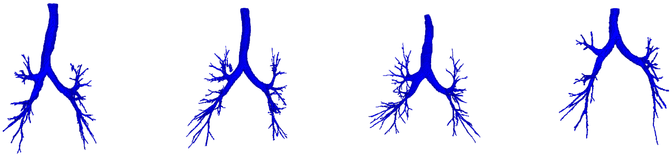
aasa@diku.dk

In collaboration with!



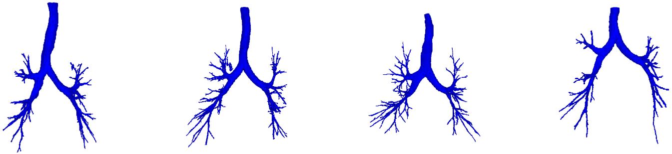
Airway shape modeling

Starting point: What does the average human airway tree look like?



Wanted: Parametric statistical model for trees, allowing variations in branch count, tree-topological structure and branch geometry

Airway shape modeling



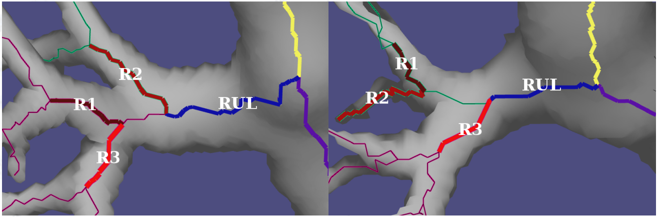
- Smoker's lung (COPD) is caused by inhaling damaging particles.
- Likely that damage made depends on airway geometry
- Reversely: COPD changes the airway geometry, e.g. airway wall thickness.
- $\rightsquigarrow$ Geometry can help diagnosis/prediction.

aasa@diku.dk,

Airway shape modeling

Properties of airway trees:

- Topology, branch shape, branch radius
- Somewhat variable topology (combinatorics) in *anatomical* tree
- Substantial amount of noise in *segmented* trees (missing or spurious branches), especially in COPD patients, *i.e. inherently incomplete data*



Airway shape modeling

Wanted properties:

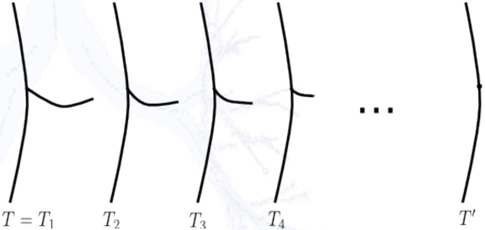


Figure: Tolerance of structural noise.

Airway shape modeling

Wanted properties:

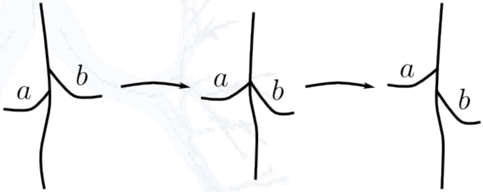
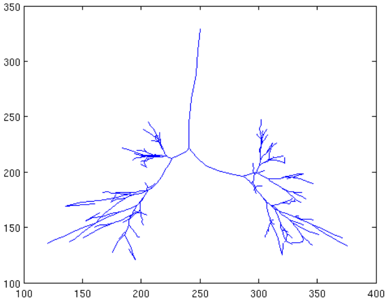


Figure: Handling of internal structural differences.

Airway shape modeling

We shall consider airway centerline trees embedded in $\mathbb{R}^3$.



aasa@diku.dk,

Not just airways....

Data with an underlying tree- or graph-structure appear in all kinds of applications:

- Anatomical structures such as airways, blood vessels and other vascularization systems;
- Skeletal structures such as medial axes;
- Descriptors of hierarchical structures (genetics, scale space)

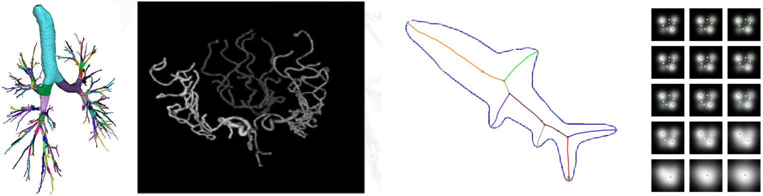
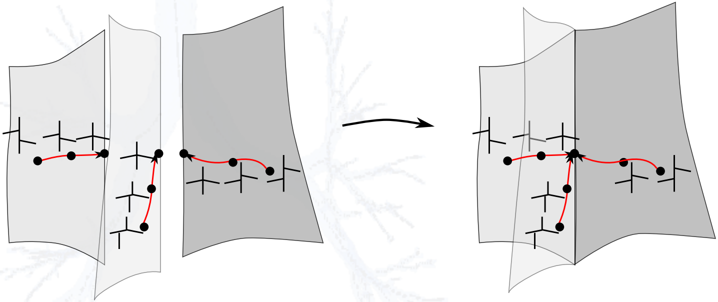


Figure: Figures from Lo; Wang et al.; Sebastian et al.; Kuijper

A space of tree-like shapes: Intuition

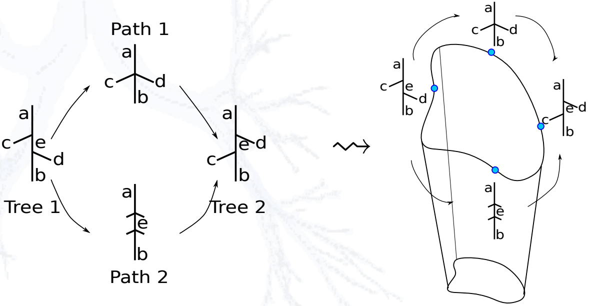
What would a path-connected space of deformable trees look like?



- Easy: Trees with same topology in their own "component"
- Harder: How are the components connected?
- Solution: glue collapsed trees, deforming one topology to another
- $\rightsquigarrow$ Stratified space, self intersections

A space of tree-like shapes: Intuition

The tree-space has conical "bubbles"

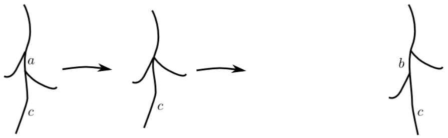


Classical example: Tree edit distance (TED)

- TED is a classical, algorithmic distance
- tree-space with TED is a nonlinear metric space
- dist(T_1, T_2) is the minimal total cost of changing T_1 into T_2 through three basic operations:
- Remove edge, add edge, deform edge.

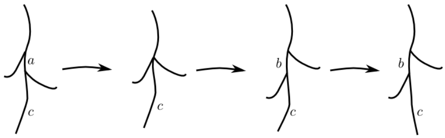
Classical example: Tree edit distance (TED)

- TED is a classical, algorithmic distance
- tree-space with TED is a nonlinear metric space
- dist(T_1, T_2) is the minimal total cost of changing T_1 into T_2 through three basic operations:
- Remove edge, add edge, deform edge.



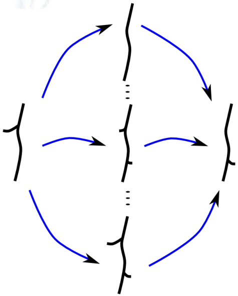
Classical example: Tree edit distance (TED)

- TED is a classical, algorithmic distance
- tree-space with TED is a nonlinear metric space
- dist(T_1, T_2) is the minimal total cost of changing T_1 into T_2 through three basic operations:
- Remove edge, add edge, deform edge.



Classical example: Tree edit distance (TED)

- Tree-space with TED is a geodesic space, but almost all geodesics between pairs of trees are non-unique (infinitely many).



- Then what is the average of two trees? Many!
- Tree-space with TED has everywhere unbounded curvature.
- TED is *not* suitable for statistics.

Classical example: Tree edit distance (TED)

Many state-of-the-art approaches to distance measures and statistics on tree- and graph-structured data *are* based on TED!

- Ferrer, Valveny, Serratosa, Riesen, Bunke: Generalized median graph computation by means of graph embedding in vector spaces. Pattern Recognition 43 (4), 2010.
- Riesen and Bunke: Approximate Graph Edit Distance by means of Bipartite Graph Matching. Image and Vision Computing 27 (7), 2009.
- Trinh and Kimia, Learning Prototypical Shapes for Object Categories. CVPR workshops 2010.

Classical example: Tree edit distance (TED)

The problems can be "solved" by choosing specific geodesics.
OBS! Geometric methods can no longer be used for proofs, and one risks choosing problematic paths.

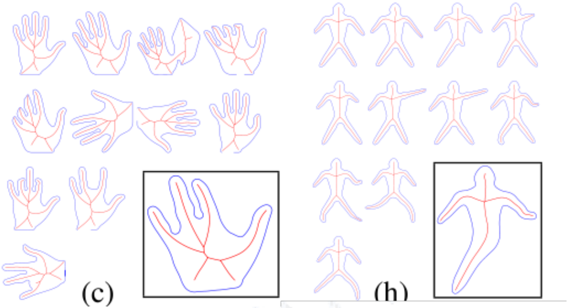


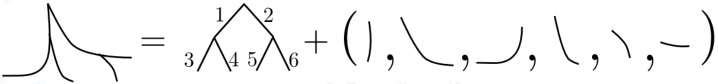
Figure: Trinh and Kimia (CVPR workshops 2010) compute average shock graphs using TED with the simplest possible choice of geodesics.

Build a tree-space: Tree representation

How to represent geometric trees mathematically?

Tree-like (pre-)shape = pair $(\mathscr{T}, x)$

- $\mathscr{T} = (V, E, r, <)$ rooted, ordered/planar binary tree, describing the tree topology (combinatorics)
- $x \in \bigoplus_{e \in E} A$, each coordinate in an attribute space A describing edge shape



Build a tree-space: Tree representation

We are allowing collapsed edges, which means that

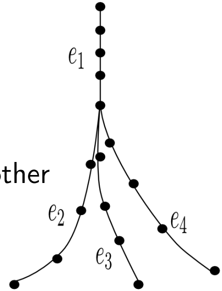
- we can represent higher order vertices
- we can represent trees of different sizes using the same combinatorial tree $\mathscr{T}$

(dotted line = collapsed edge = zero/constant attribute)

aasa@diku.dk,

Build a tree-space: Tree representation

- Edge representation through landmark points:
- Edge shape space is $(\mathbb{R}^d)^n$, $d = 2, 3$.
- (For most results, this can be generalized to other vector spaces)



aasa@diku.dk,

The space of tree-like preshapes

First: $\mathcal{T}$ an infinite, ordered (planar), rooted binary tree

Definition

Define the space of tree-like *pre*-shapes as the direct sum

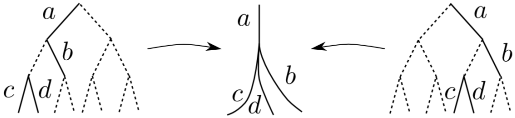
$$\bigoplus_{e \in E} (\mathbb{R}^d)^n$$

where $(\mathbb{R}^d)^n$ is the edge shape space.

This is just a space of *pre-shapes*.

From pre-shapes to shapes

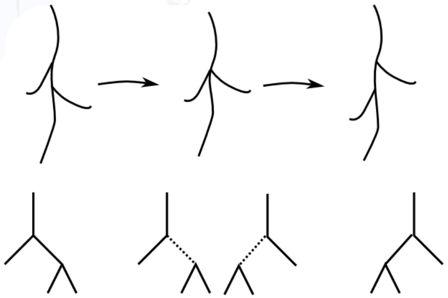
Many shapes have more than one representation



aasa@diku.dk,

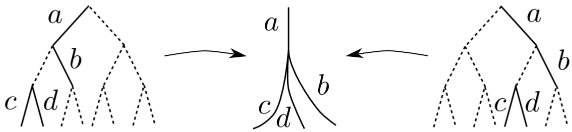
From pre-shapes to shapes

Not all shape deformations can be recovered as natural paths in the pre-shape space:



Shape space definition

- Start with the pre-shape space $X = \bigoplus_{e \in E} (\mathbb{R}^d)^n$.
- Define an equivalence $\sim$ by identifying points in X that represent the same tree-shape.



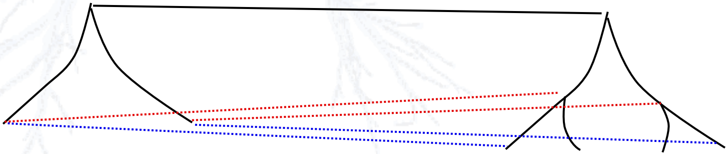
- This corresponds to identifying, or gluing together, subspaces $\{x \in X | x_e = 0 \text{ if } e \notin E_1\}$ and $\{x \in X | x_e = 0 \text{ if } e \notin E_2\}$ in X.
- The space of ordered (planar) tree-like shapes $\bar{X} = X/\sim$ is a folded vector space.

aasa@diku.dk,

Shape space definition

Remark

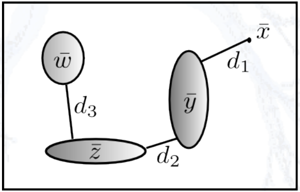
- Tree-shape definition a little unorthodox: we do not factor out scale and rotation of the tree.
- Our data (segmented airway trees) are incomplete; the number of segmented branches is unstable and depends on the health of the patient.



Definition of metric on tree-space

Given a metric d on the vector space $X = \bigoplus_{e \in E} (\mathbb{R}^d)^n$ we define the quotient pseudometric $\bar{d}$ on the quotient space $\bar{X} = X/\sim$ by setting

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



$$\bar{d}(\bar{x}, \bar{w}) = d_1 + d_2 + d_3$$

Theorem

The quotient pseudometric $\bar{d}$ is a metric on $\bar{X}$. □

Definition of metric on tree-space

- Two metrics on $\bar{X}$ from two product norms on $X = \bigoplus_{e \in E} (\mathbb{R}^d)^n$:

$$\begin{aligned} l1 \text{ norm:} \quad & d_1(x, y) = \sum_{e \in E} \|x_e - y_e\| \\ l2 \text{ norm:} \quad & d_2(x, y) = \sqrt{\sum_{e \in E} \|x_e - y_e\|^2} \end{aligned}$$

- $\bar{d}_1$ = Tree Edit Distance
- Terminology: $\bar{d}_2$ = QED (Quotient Euclidean Distance) metric.

Theorem

Let $\bar{d} = \bar{d}_1$ or $\bar{d}_2$. Then $(\bar{X}, \bar{d})$ is a geodesic space. □

Unordered trees

- Give each tree a random order
- Denote by G the group of reorderings of the edges (in $\mathscr{T}$) that do not alter the connectivity of the tree.
- The space of spatial/unordered trees is the space $\bar{\bar{X}} = \bar{X}/G$
- Give $\bar{\bar{X}}$ the quotient pseudometric $\bar{\bar{d}}$.
- $\bar{\bar{d}}(\bar{\bar{x}}, \bar{\bar{y}})$ chooses the order that minimizes $\bar{d}(\bar{\bar{x}}, \bar{\bar{y}})$.

Theorem

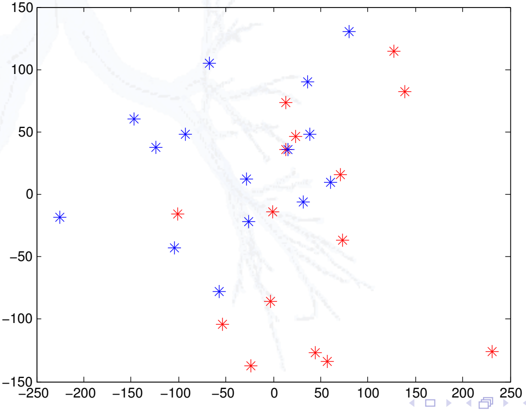
For the quotient pseudometric $\bar{\bar{d}}$ induced by either $\bar{d}_1$ or $\bar{d}_2$, the function $\bar{\bar{d}}$ is a metric and $(\bar{\bar{X}}, \bar{\bar{d}})$ is a geodesic space.

aasa@diku.dk,

Distances between airways?

Evaluation of metric:

MDS based on approximate geodesic distances between 30 airways of healthy individuals and individuals with moderate COPD.

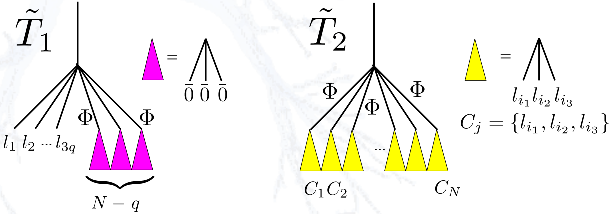


Complexity of computing tree-space geodesics?

Assume edge attributes have dimension > 1
(for dim $= 1$, Scott Provan).

Theorem

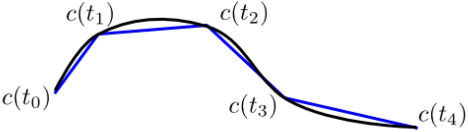
Computing QED geodesics is NP complete.



A little metric geometry – geodesics

- Let (X, d) be a metric space. The length of a curve $c\colon [a,b] \to X$ is

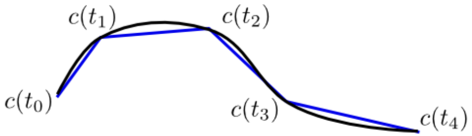
$$l(c) = sup_{a=t_0 \leq t_1 \leq \ldots \leq t_n = b} \sum_{i=0}^{n-1} d(c(t_i, t_{i+1})).$$



A little metric geometry – geodesics

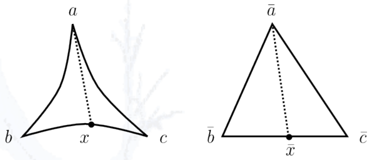
- Let (X, d) be a metric space. The length of a curve $c\colon [a, b] \to X$ is

$$l(c) = sup_{a=t_0 \leq t_1 \leq \ldots \leq t_n = b} \sum_{i=0}^{n-1} d(c(t_i, t_{i+1})).$$



- A *geodesic* from x to y in X is a path $c\colon [a, b] \to X$ such that $c(a) = x, c(b) = y$ and $l(c) = d(x, y)$.
- (X, d) is a *geodesic space* if all pairs x, y can be joined by a geodesic.

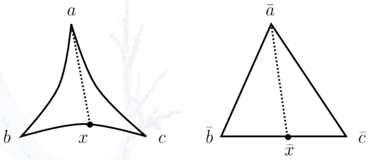
Curvature in metric spaces



- A *CAT*(0) space is a metric space in which geodesic triangles are "thinner" than for their comparison triangles in the plane; that is, $d(x, a) \leq d(\bar{x}, \bar{a})$.

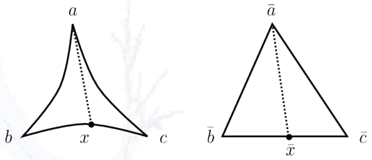
aasa@diku.dk,

Curvature in metric spaces



- A $CAT(0)$ space is a metric space in which geodesic triangles are "thinner" than for their comparison triangles in the plane; that is, $d(x,a) \leq d(\bar{x},\bar{a})$.
- A space has non-positive curvature if it is locally $CAT(0)$.

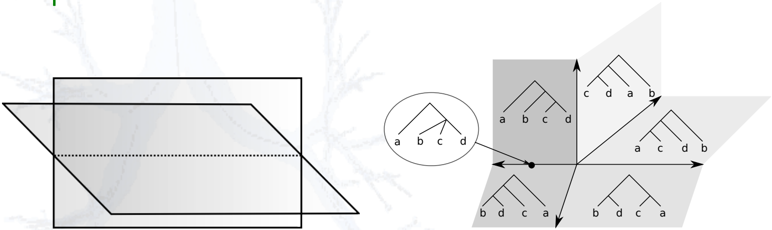
Curvature in metric spaces



- A $CAT(0)$ space is a metric space in which geodesic triangles are "thinner" than for their comparison triangles in the plane; that is, $d(x, a) \leq d(\bar{x}, \bar{a})$.
- A space has non-positive curvature if it is locally $CAT(0)$.
- (Similarly define curvature bounded by κ by using comparison triangles in hyperbolic space or spheres of curvature κ.)

Curvature in metric spaces

Example



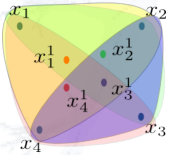
Theorem (see e.g. Bridson-Haefliger)

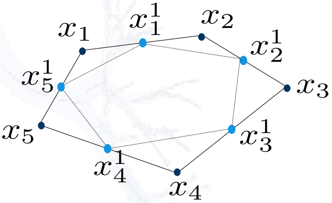
Let (X, d) be a $CAT(0)$ space; then all pairs of points have a unique geodesic joining them. The same holds locally in $CAT(\kappa)$ spaces, $\kappa \neq 0$. □

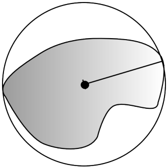
Statistics in metric spaces?

Theorem (Sturm)

Frechet means are unique in $CAT(0)$ spaces.





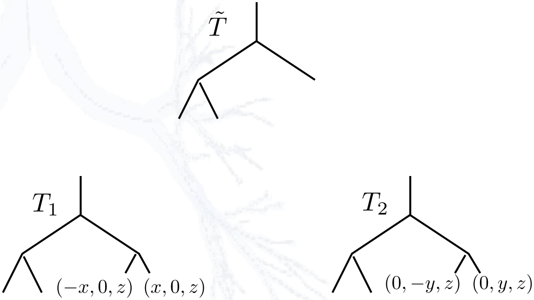


Other midpoint-finding algorithms also converge in CAT(0) spaces:

- centroid
- Birkhoff shortening
- circumcenters

Curvature of shape space?

- This space has everywhere unbounded curvature!

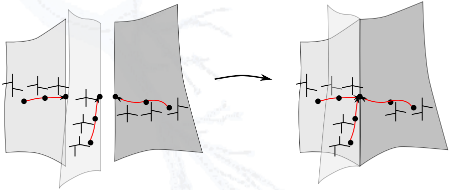


- Bad news for statistics?

aasa@diku.dk,

Regularize via extra assumptions on the trees:

- **Restrict to:** all representations of certain restricted tree topologies.
- **Example 1:** Restrict to the set $\bar{\bar{X}}_N$ of trees with N leaves.



- **Example 2:** Restrict to all topologies occuring in airway trees.

Curvature of shape space

Theorem

- Consider $(\bar{X}, \bar{d}_2)$ and $(\bar{\bar{X}}, \bar{\bar{d}}_2)$, ordered/unordered tree-shape space.
- At generic points, the space is locally $CAT(0)$.
- Its geodesics are locally unique at generic points.
- At non-generic points, the curvature is unbounded. □

Curvature of shape space

In fact, curvature is one of:

- $+\infty$
- 0
- $-\infty$

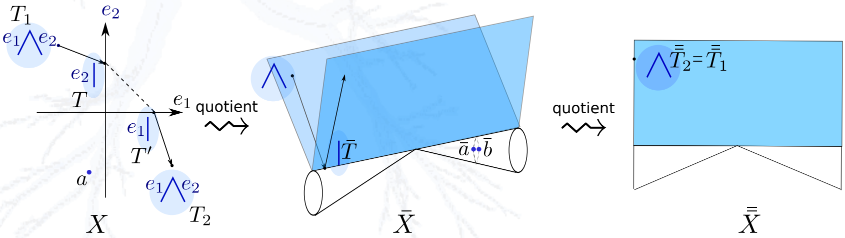


Figure: Space of ordered/unordered trees with at most 2 edges

aasa@diku.dk,

How much better did it get?

- Reasonably large $CAT(0)$ neighborhoods are now larger containing several top-dimensional tree topologies.
- Computational complexity? Still NP complete.

We can compute means!

Leaf vasculature data:

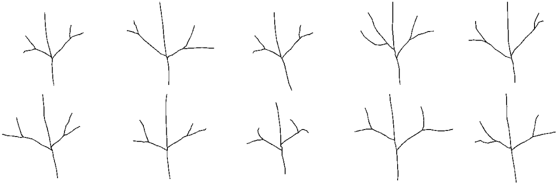


Figure: A set of vascular trees from ivy leaves form a set of planar tree-shapes.

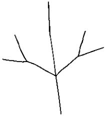


Figure: a): The vascular trees are extracted from photos of ivy leaves. b) The mean vascular tree.

aasa@diku.dk,

We can compute means!

The mean upper airway tree[1]

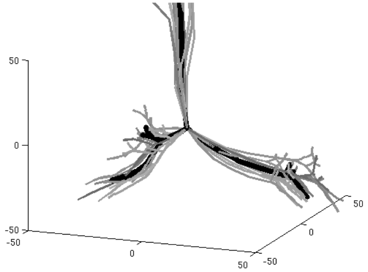


Figure: A set of upper airway tree-shapes along with their mean tree-shape.

[1] Feragen et al, *Means in spaces of treelike shapes*, ICCV2011

aasa@diku.dk,

We can compute means!

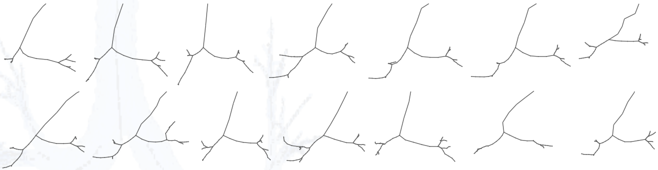


Figure: A set of upper airway tree-shapes (projected).[1]

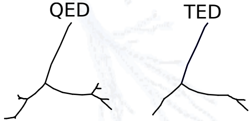
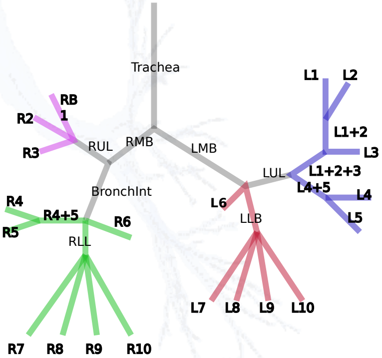


Figure: The QED and TED (algorithm by Trinh and Kimia) means.

[1] Feragen et al, *Towards a theory of statistical tree-shape analysis*, submitted

Dealing with NP - Useful property of airways

The first 6-8 generations of the airway tree are "similar" in different people.

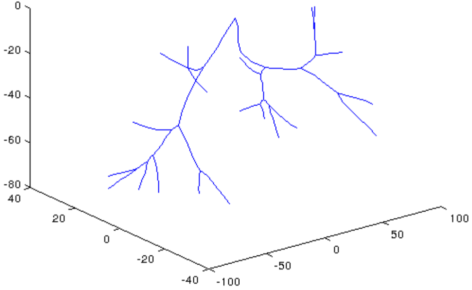


NB!: Not all present in all people; **not all present in all segmentations.**

Regularize via second set of assumptions

- **Label** the "leaves" of your trees and insist that all trees have the same leaf label set.
- Polynomial time distance algorithms (Owen, Provan)
- **Also:** Factor out leaf labels via leaf permutation group $\rightsquigarrow$ NP complete.

Statistics on larger trees: Mean airway

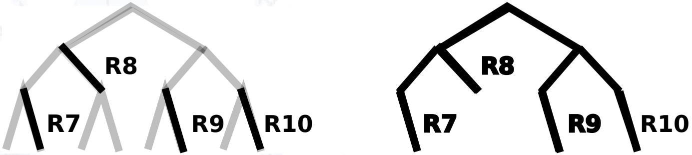


Joint with Megan Owen.

Application: Geodesic airway branch labeling[2]

Idea:

- Generate leaf label configurations and the corresponding tree spanning the labels



- Evaluate configuration in comparison with training data using geodesic deformations between leaf-labeled airway trees (Owen, Provan)

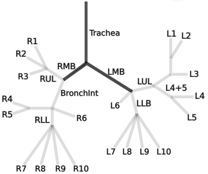
[2]Feragen, Petersen, Owen, Lo, Thomsen, Dirksen, Wille, de Bruijne, *A hierarchical scheme for geodesic anatomical labeling of airway trees*, MICCAI 2012.

aasa@diku.dk,

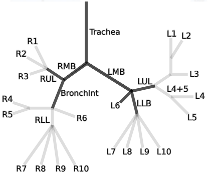
Application: Geodesic airway branch labeling[2]

Idea:

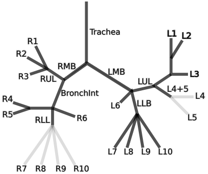
- Make tractable using a hierarchical labeling scheme



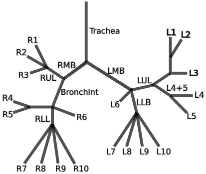
search 3 generations



search 2 and 2 generations



search 2, 2, 2 and 3 generations



search 3 and 2 generations

[2] Feragen, Petersen, Owen, Lo, Thomsen, Dirksen, Wille, de Bruijne, *A hierarchical scheme for geodesic anatomical labeling of airway trees*, MICCAI 2012.

Application: Geodesic airway branch labeling[2]

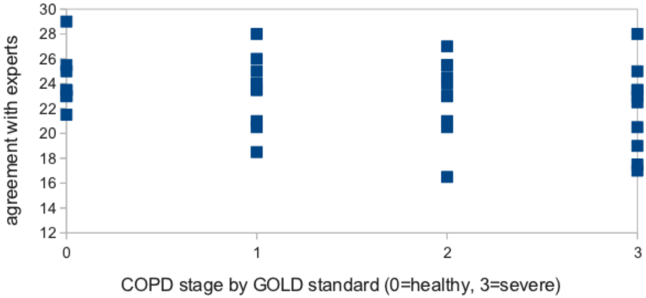
Performance:

- 40 airway trees from 20 subjects with different stages of COPD, hand labeled by 3 experts in pulmonary medicine.
- All 20 segmental labels were assigned (segmental = most distal branches) at an average success rate of 72.8%.
- **Performance:** as good as the performance of an expert in pulmonary medicine.
  (measured in terms of ability to agree with the two other experts)

[2]Feragen, Petersen, Owen, Lo, Thomsen, Dirksen, Wille, de Bruijne, *A hierarchical scheme for geodesic anatomical labeling of airway trees*, MICCAI 2012.

Application: Geodesic airway branch labeling[2]

Performance:



Spearman: ($\rho = 0.22$, $p = 0.18$)

[2] Feragen, Petersen, Owen, Lo, Thomsen, Dirksen, Wille, de Bruijne, *A hierarchical scheme for geodesic anatomical labeling of airway trees*, MICCAI 2012.

Application: Geodesic airway branch labeling[2]

Performance:
2 scans per subject, registered for label transfer. Reproducible segmental labels on average:

- 14.0 (expert 1)
- 15.1 (expert 2)
- 15.2 (algorithm)

[2] Feragen, Petersen, Owen, Lo, Thomsen, Dirksen, Wille, de Bruijne, *A hierarchical scheme for geodesic anatomical labeling of airway trees*, MICCAI 2012.

Conclusion and discussion

- We have introduced a geometric framework for analysis of tree-shapes such as airways
- We have made proof-of-concept statistical experiments
- Distance computations are generally NP hard; we use heuristics
- We have utilized the tree-shape framework to automatically assign labels to anatomical airway trees
- This gives a fast and robust procedure with very few tuning parameters, which performs well in presence of COPD.

Future work

- Development of heuristics for tree geodesic computation
- More extensive statistical analysis of airway trees
- Statistics on individual branches based on branch labeling
- Kernels on anatomical trees (for speed)
- Kernels on anatomical graphs

aasa@diku.dk,