Computational Geometry Learning

Jean-Daniel Boissonnat
Frédéric Chazal

Geometrica, INRIA
http://www-sop.inria.fr/geometrica

Lectures at MPRI

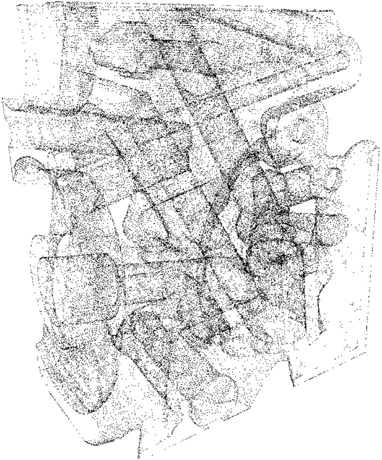
Computational Geometry Learning

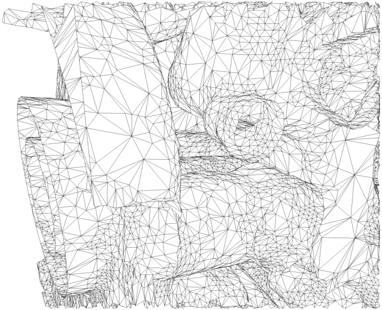
Jean-Daniel Boissonnat
Frédéric Chazal

Geometrica, INRIA
http://www-sop.inria.fr/geometrica

Lectures at MPRI

Reconstructing surfaces from point clouds

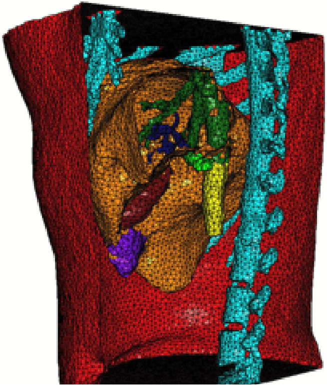


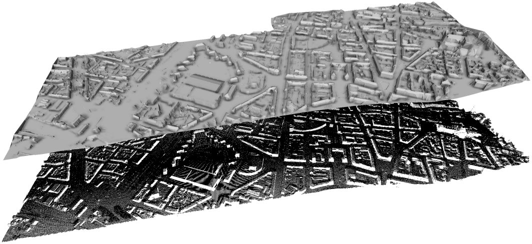


One can reconstruct a surface from 10^6 points within 1mn [CGAL]

CGAL-mesh

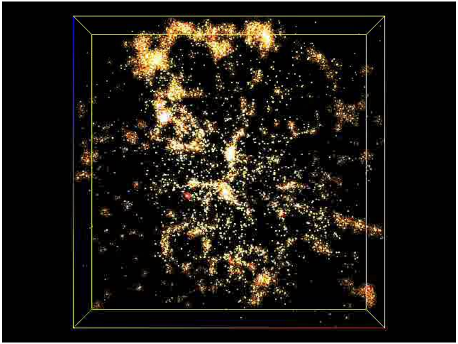
GeometryFactory, Acute3D





Geometric data analysis

Images, text, speech, neural signals, GPS traces,...



Geometrisation : Data = points + distances between points

Hypothesis : Data lie close to a structure of "small" intrinsic dimension

Problem : Infer the structure from the data

Image manifolds

An image with 10 million pixels
$\rightarrow$ a point in a space of 10 million dimensions!

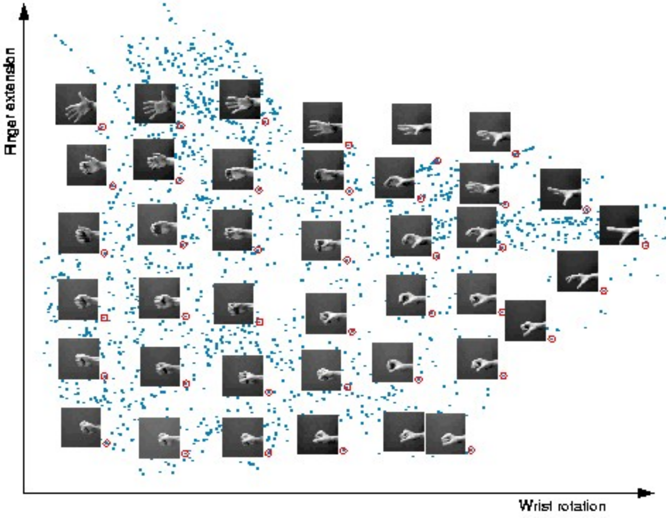
camera : 3 dof
light : 2 dof

The image-points lie close to a structure of intrinsic dimension 5 embedded in this huge ambient space

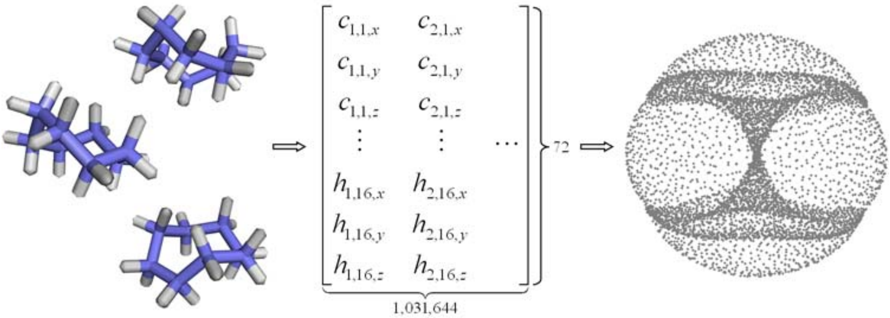
Motion capture

Typically $N = 100$, $D = 100^3$, $d \leq 15$

Dimensionality reduction



Conformation spaces of molecules e.g. C_8H_{16}



- Each conformation is represented as a point in $\mathbb{R}^{72}$ ($\mathbb{R}^{24}$ when neglecting the H atoms)
- The intrinsic dimension of the conformation space is 2
- The geometry of C_8H_{16} is highly nonlinear

Issues in high-dimensional geometry

- Dimensionality severely restricts our intuition and ability to visualize data
  $\Rightarrow$ need for automated and provably correct methods methods

- Complexity of data structures and algorithms rapidly grow as the dimensionality increases

  $\Rightarrow$ no subdivision of the ambient space is affordable

  $\Rightarrow$ data structures and algorithms should be sensitive to the intrinsic dimension (usually unknown) of the data

- Inherent defects : sparsity, noise, outliers

Course overview : some keywords

- Computational geometry and topology
- Triangulations, simplicial complexes
- Algorithms in high dimensions

- Shape reconstruction
- Geometric inference
- Topological data analysis

- Algorithmic geometry of triangulations [JDB]

  1. Simplicial complexes in metric spaces (26/09)
  2. Delaunay-type complexes (03/10)
  3. Weighted Delaunay and alpha complexes (10/10)
  4. Thickness and relaxation (17/10)
  5. Reconstruction of submanifolds (24/10)

- Geometric inference [FC]

  6. Distance functions, sampling, stability of critical points (31/10)
  7. Noise and outliers, distance to a measure (07/11)
  8. Computational homology (14/11)
  9. Topological persistence (21/11)
  10. Multi-scale inference and applications (28/11)

Further reading

Theses at Geometrica

- Persistent Homology : Steve Oudot (HDR, 26/11)
- Distance to a measure : Q. Mérigot (HDR, 17/11)
- Triangulation of manifolds : A. Ghosh (2012)
- Data structures for computational topology : C. Maria (2014)

Course Notes

`www-sop.inria.fr/geometrica/courses/supports/CGL-poly.pdf`

Colloquium J. Morgenstern

`www-sop.inria.fr/colloquium`

Vin de Silva : Point-clouds, sensor networks, and persistence: algebraic topology in the 21st century 26/3/2009

Projects

- European project Computational Geometric Learning (CGL)
  `cgl.uni-jena.de/Home/WebHome`
- ANR TopData
  Geometry meets statistics
- ERC Sdvanced Grant GUDHI
  Geometry Understanding in Higher Dimensions
- On the industrial side
  Californian Startup : www.ayasdi.com