An Application of Persistent Homology on Grassmann Manifolds to the Detection of Signals in Hyperspectral Imagery

Sofya Chepushtanova, Michael Kirby, and Chris Peterson*

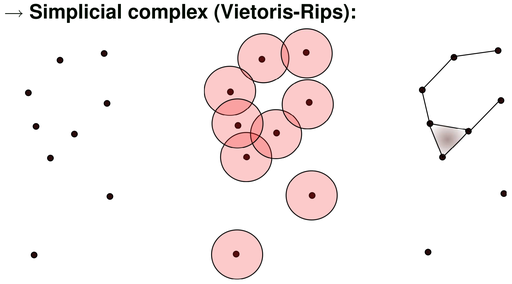
Department of Mathematics, Colorado State University, Fort Collins, CO

Introduction

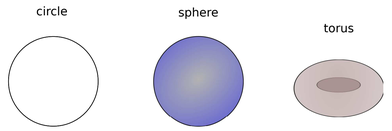
- **Topological data analysis (TDA)**
  Data points → *Geometric object* → *Topological summary*.
- **Persistent Homology (PH)** can be applied to a data set to capture the "persistence" of topological structure.
- We apply PH to **hyperspectral data** encoded as abstract points on a **Grassmann manifold**.
- $G(k, n)$ **framework** affords a form of compression while retaining topological structure → it becomes feasible to analyze **large volumes of hyperspectral data**.

Persistent Homology

- **Finite data set → Simplicial complex (Vietoris-Rips):**

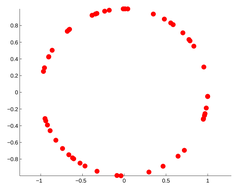


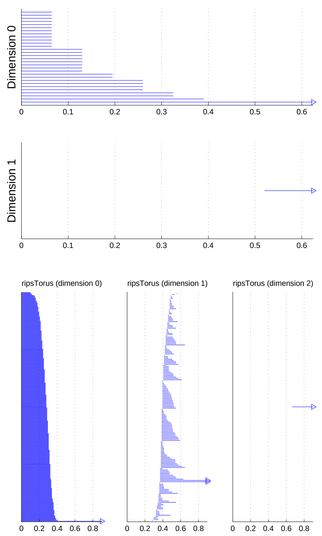
- **Betti numbers** describe the homology of a simplicial complex:
  $Betti_0$: 0th order holes, or clusters (connected components)
  $Betti_1$: 1st order holes, or holes (circles)
  $Betti_2$: 2nd order holes, or cavities (voids)

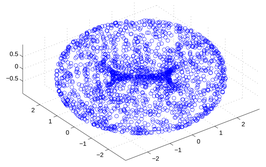


	$Betti_0$	$Betti_1$	$Betti_2$
circle	1	1	0
sphere	1	0	1
torus	1	2	1

- **Filtered simplicial complexes → persistent homology**

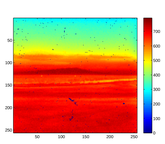


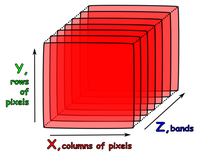




Data Set

- **Fabry-Pérot long-wavelength infrared data set:** Single scanning: the interferometer collects 20 images from different wavelengths ($8 - 11$ μm), each image is composed of 256×256 pixels.





- Triethyl Phosphate (TEP): 561 data cubes $256 \times 256 \times 20$.
- **Data collection:** a pre-determined quantity of TEP is released into the air to create an aerosol plume for **detection** against natural background. The 561 cubes are collected successively to record the entire event from 'pre-burst' to 'post-burst'.

Grassmannian Framework

The real **Grassmann manifold** (Grassmannian) $G(k, n)$ is the collection of all k-dimensional subspaces of $\mathbb{R}^n$, for fixed $k \leq n$.
A point on $G(k, n)$ can be represented by an $n \times k$ orthogonal matrix U ($U^T U = I_k$).

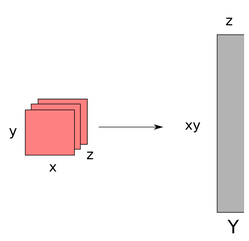
(Point representation on $G(k, n)$ is not unique: we say $U_1 = U_2$ if $span(U_1) = span(U_2)$.)

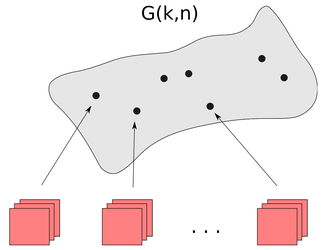
Data cube or region in data cube $x \times y \times z$, $z < xy$ → reshape → $xy \times z$ matrix Y → SVD: $Y = U\Sigma V^T$ → U - point on $G(k, n)$, with $k = z$ and $n = xy$

Note: If $z > xy$, then U - point on $G(k, n)$ with $k = xy$ and $n = z$.

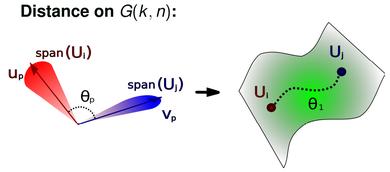
Examples:

- data cube $256 \times 256 \times 20$ will be mapped to a point on $G(20, 65536)$
- $4 \times 4 \times 20$ patch from a data cube will be mapped to a point on $G(16, 20)$.





Distance on $G(k, n)$:



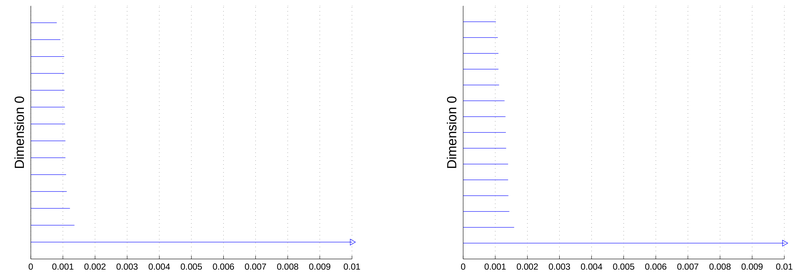
Pseudo metric: $D_{ii} = 0, D_{ij} = d_G(U_i, U_j) = \theta_1$ where θ_1 is the smallest principal angle between subspaces.

Experiment:

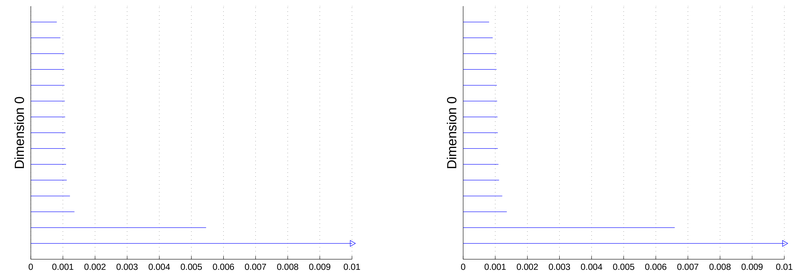
- We use 3 (out of 20) bands selected by sparse support vector machine algorithm.
- A patch $4 \times 8 \times 3$ from each cube is mapped to a point on $G(3, 32)$.
- The location of the patch (same for each cube) is determined by Adaptive Coherence/Cosine Estimator (ACE) such that it is a part of plume after TEP release.
- We generate $Betti_0$ barcodes to see clusters in data.

Barcodes for Fabry-Pérot Data Set

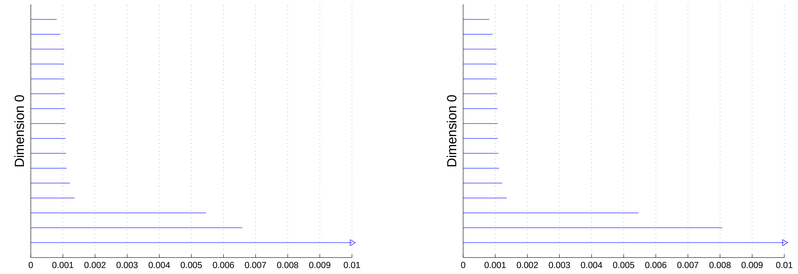
Points before (cubes 91-104, left) and after (545-549, right) TEP release



Points before release (91-104) and one TEP point (112 - left, and 114 - right)



Points before release (91-104) and two TEP points (112,114 - left and 112,116 - right)



Future Work: Strengthening the Topological Signal

- Employ different data settings.
- Use other distance (pseudo) metrics on the Grassmannian.
- Explore $Betti_1$ barcodes for analysis.

Bibliography

A. Edelman, T. A. Arias, and S. T. Smith, *The geometry of algorithms with orthogonality constraints*, SIAM Journal on Matrix Analysis and Applications, 20(2), pp. 303–353, 1998.

H. Edelsbrunner, D. Letscher, and A. Zomorodian, *Topological persistence and simplification*, Discrete and Computational Geometry, 28, pp. 511–533, 2002.

A. Zomorodian and G. Carlsson, *Computing persistent homology*, Discrete and Computational Geometry, 33, pp. 249–274, 2005.

B. R. Cosofret, et al., *Imaging Sensor Constellation for Tomographic Chemical Cloud Mapping*, Applied Optics, 48(10), pp. 1837–1852, 2009.

This material is based upon work supported by the National Science Foundation under Grants No. DMS-1228308 and DMS-1322508. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.