# Introduction to Persistent Homology

Matthew L. Wright

**Department of Mathematics, Statistics, and Computer Science, St. Olaf College, Northfield, USA**
wright5@stolaf.edu

## Abstract

This video presents an introduction to persistent homology, aimed at a viewer who has mathematical aptitude but not necessarily knowledge of algebraic topology. Persistent homology is an algebraic method of discerning the topological features of complex data, which in recent years has found applications in fields such as image processing and biological systems. Using smooth animations, the video conveys the intuition behind persistent homology, while giving a taste of its key properties, applications, and mathematical underpinnings.

**1998 ACM Subject Classification** F.2.2 Geometrical problems and computations

**Keywords and phrases** Persistent Homology, Topological Data Analysis

**Digital Object Identifier** 10.4230/LIPIcs.SoCG.2016.72

**Category** Multimedia Contribution

## 1 Persistent Homology

In recent years, persistent homology has become a popular mathematical tool for discerning the geometric structure of complex, often high-dimensional data. Not only has persistent homology been a subject of intense theoretical interest [4, 7] and computational work [6, 11], but it has also been applied to applications in digital imaging [2, 8], biological systems [9], neuroscience [1], and more.

Persistent homology detects topological features—such as clusters, holes, and voids—of data. Generally speaking, the data is often presented as a set of points (perhaps in a high-dimensional space), values of a function at a set of sample points, or a discrete metric space. The basic methodology involves associating to the data a sequence of topological objects, usually simplicial complexes, constructed from the data for an increasing sequence of scale parameters (real numbers). Homology is computed for each object in this sequence, and topological features that persist across a range of scale parameters are identified. Thus, persistent homology is useful when one is unable to predict the scale at which features of interest will appear, or when features exist at various scales.

Persistent homology provides invariants, often called *barcodes*, that represent the persistent topological features of data. A barcode is a collection of intervals of real numbers, often drawn as parallel line segments. Yet a barcode is really a visualization of an algebraic structure called a *persistence module*, which is a graded module over the polynomial ring $k[x]$ for some field $k$. The structure theorem for finitely-generated modules over principal ideal domains says that such a module decomposes into a direct sum of interval modules. These intervals give the bars in the barcode.

Persistence barcodes are computable via standard algorithms involving linear algebra. The worst-case runtime of these algorithms is $O(n^3)$, where $n$ is the number of simplices in the simplicial complex built from the data, although the computation is rarely worst-case. In


32nd International Symposium on Computational Geometry (SoCG 2016).
Editors: Sándor Fekete and Anna Lubiw; Article No. 72; pp. 72:1–72:3
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

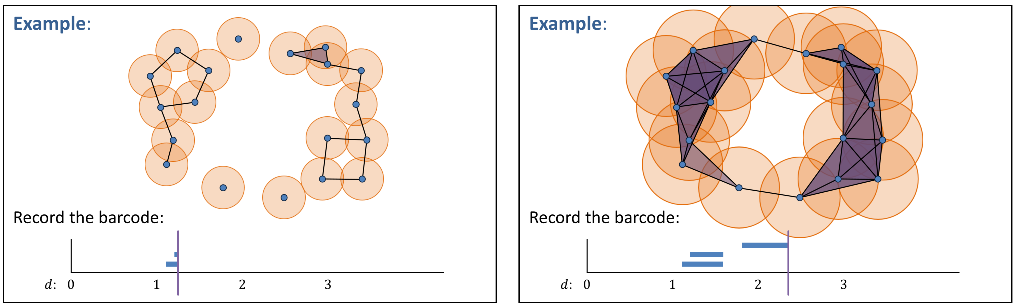


**Figure 1** Two frames from the video, showing simplicial complexes constructed from point-cloud data for different scale parameters. The barcode is drawn as the scale parameter increases.

addition, clever data structures and topological simplification can speed up the computation significantly.

Persistent homology has been successfully applied to various problems. For example, Carlsson and others studied the collection of high-contrast patches, 3-by-3 pixels in size, from a large set of grayscale photographs. Treating the patches as 9-dimensional vectors and applying persistent homology, they found that the resulting point cloud was close to the surface of a Klein bottle embedded in the higher-dimensional space [2]. This observation helped create a dictionary useful for image-recognition applications [8].

For another example, Perea and others used persistent homology to detect periodicity in time-series data. Using sliding-window embeddings, the researchers converted measurements of a periodic signal into a point cloud in a high-dimensional space. Persistent homology, applied to the point cloud, detects cycles, which correspond to periodicity in the original signal. Applying this methodology to gene expression time-series data, the researchers detected periodicity that was missed by other methods—especially periodicity in the presence of damping [9].

Persistent homology has also been applied to neuroscience [1], bone morphology [10], virus evolution [3], and sensor networks [5], and more.

## 2 The Video

Given the increasing interest in persistent homology among mathematicians, scientists, and others, it is desirable to have resources to help teach this subject to those who are new to topological data analysis. The present video is designed to do just that: to make persistent homology accessible to a wide audience. In its first eight months on YouTube, the video has been viewed more than 2500 times by viewers around the globe.

While the mathematics of algebraic topology are formidable, the intuition behind persistent homology is easily conveyed in video format. The author has found that animations depicting the evolution of a simplicial complex and the corresponding barcode as the scale parameter increases, are an effective way introducing persistent homology to non-specialists. Such an animation is found in the present video, beginning at about 4 minutes and 40 seconds after the start of the video; two frames of the animation appear in Figure 1. The author has used this animation, as well as other clips from the video, to teach persistent homology to students, colleagues, and friends.

The video is designed to be accessible and intuitive. It is not the author's intent to minimize the rich mathematical theory of homology, nor to hide the computational challenges involved in using persistent homology in real applications, as the video points to the theory and challenges that one can appreciate with further study. Though the video does not dwell on technical details, but merely skims the surface of persistent homology, the author has attempted to convey that great and fascinating depths exist under that surface. It is the author's hope that this video will motivate viewers to embark on further study of mathematical theory, computation, and applications—or at least will lead them to a greater appreciation of the elegance and utility of mathematics.

The animations seen in the video were prepared using the animation tools in Microsoft PowerPoint 2013. The audio was captured with a Blue Snowball microphone, and the in-person video was recorded with a Canon Digital Rebel T3i camera. Screen capture and editing were performed using ActivePresenter by Atomi Systems.

**Acknowledgements.** This video was prepared while the author was a Postdoctoral Fellow at the Institute for Mathematics and its Applications (IMA) at the University of Minnesota, supported by funding from the National Science Foundation. This submission was prepared while the author was a Visiting Assistant Professor at St. Olaf College. The author is partially supported by NSF DMS-1606967. The author thanks the IMA, St. Olaf College, and the NSF for their support of mathematics research and education.

## References

1 P. Bendich, J.S. Marron, E. Miller, A. Pieloch, and S. Skwerer. Persistent homology analysis of brain artery trees. *Annals of Applied Statistics*, to appear.

2 G. Carlsson, T. Ishkhanov, V. de Silva, and A. Zomorodian. On the local behavior of spaces of natural images. *International Journal of Computer Vision*, 76(1):1–12, 2008.

3 J.M. Chan, G. Carlsson, and R. Rabadan. Topology of viral evolution. *Proceedings of the National Academy of Sciences*, 110(46), November 2013.

4 D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. *Discrete and Computational Geometry*, 37(1):103–120, 2007.

5 V. de Silva and R. Ghrist. Coverage in sensor networks via persistent homology. *Algebraic and Geometric Topology*, 7:339–358, 2007.

6 H. Edelsbrunner and J. Harer. Persistent homology: a survey. In *Surveys on discrete and computational geometry: twenty years later: AMS-IMS-SIAM Joint Summer Research Conference, June 18-22, 2006, Snowbird, Utah*, volume 453, page 257. American Mathematical Society, 2008.

7 H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete and Computational Geometry*, 28(4):511–533, 2002.

8 J. Perea and G. Carlsson. A klein-bottle-based dictionary for texture representation. *International Journal of Computer Vision*, 107(1):75–97, 2014.

9 J. Perea, A. Deckard, S.B. Haase, and J. Harer. Sw1pers: Sliding windows and 1-persistence scoring; discovering periodicity in gene expression time series data. *BMC Bioinformatics*, 16(1), 2015.

10 K. Turner, S. Mukherjee, and Doug M Boyer. Sufficient statistics for shapes and surfaces. *preprint*, 2013.

11 A. Zomorodian and G. Carlsson. Computing persistent homology. *Discrete and Computational Geometry*, 33(2):249–274, 2005.