Topological Data Analysis

and its application to image data

We make use of 3 key concepts in topology.

- Coordinate Invariance:
  The coordinate system in which we study objects of interest is not important; only their shape is.
- Deformation Invariance:
  The properties studied by topology are not affected by minor changes to the shape of an object; we can stretch or squash an object as we like, and so long as it is not 'torn' or 'reglued' in some place, its topological properties are invariant.
- Compressed Representations:
  Instead of looking at an object or surface directly, we study an approximation; losing data such as curvature is not an issue for this sort of analysis, as long as we maintain the fundamental topological properties e.g. a loop.

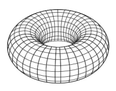


Figure 1:
A torus; the lack of axes and mathematical structure surrounding the torus demonstrate the lack of need for a coordinate system

Figure 2:
An image demonstrating the famous equivalency between the surfaces of a coffee cup and a torus; this arises due to deformation invariance

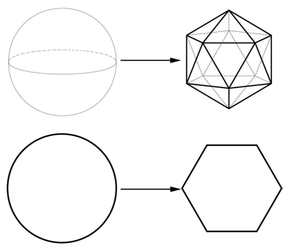
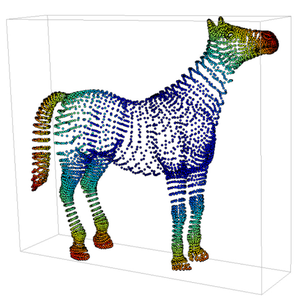


Figure 3:
This figure demonstrates that we may approximate a sphere with an icosahedron, or a circle with a hexagon. In either case, we reduce the data from an infinite surface to a finite number of vertices, edges and faces whilst maintaining the fundamental features of the surfaces



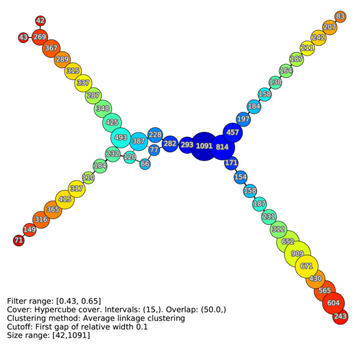
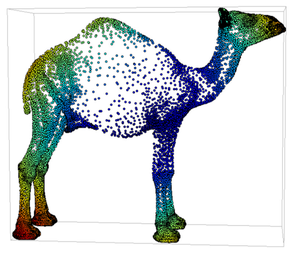
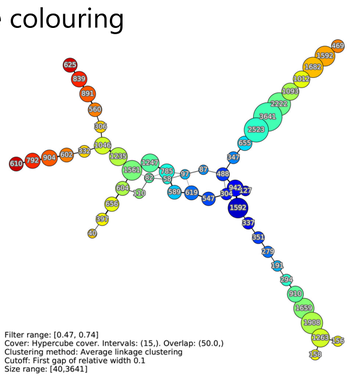


Figure 4a (above) and 4b (below):
Sample input and output data from Python Mapper, using the same filter (eccentricity, measuring distance from the 'centre' of the data). The input data is point-cloud shapes of a horse (a) and a camel (b) and the node colouring corresponds to that seen in the output.





What do we want to do?

- Use a tool such as Python Mapper:
  Developed by Daniel Müllner, this applies the Mapper algorithm, itself developed for use in TDA by Gurjeet Singh, Facundo Mémoli and Gunnar Carlsson in 2007. The algorithm can be applied to data partially clustered by any clustering algorithm, and Python Mapper implements the chain of clustering or filtering the data, applying the Mapper algorithm and visualising the results.
- Investigate image data:
  We want to find out if and how such tools can provide useful information about images, in terms of their structure, similarity, or lack thereof, extending or supplanting what traditional techniques can tell us.

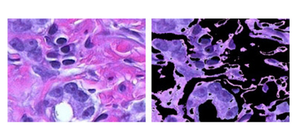


Figure 5:
An image and the result after using K-means clustering to perform colour-based segmentation. We hope that TDA will be able to extend methods like this, or maybe even reveal deeper information about an image.

Credits and References:

Figure 1: https://commons.wikimedia.org/wiki/File:Simple_Torus.svg
Figure 2: https://i.stack.imgur.com/qa0Ja.jpg
Figure 3: https://i.stack.imgur.com/D2pH8.jpg
https://i.stack.imgur.com/GS9AD.jpg
Figure 4: *From Python Mapper,* Daniel Müllner
http://danifold.net/mapper/index.html
Figure 5: https://uk.mathworks.com/discovery/image-segmentation.html