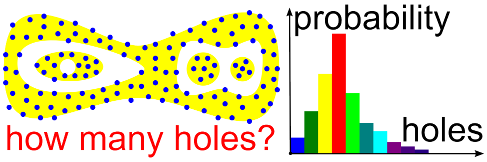
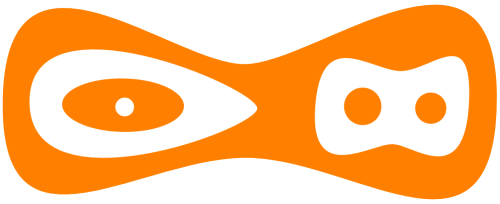
Topological Data Analysis: Applications to Computer Vision

Vitaliy Kurlin, kurlin.org, Microsoft Research Cambridge and Durham University, UK



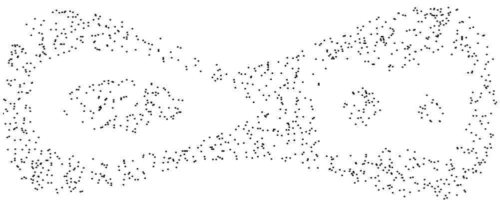
2D Hole Counting Problem

Count holes in a shape given by a noisy sample without parameters.



Shape X with 3 bounded holes: small disk, annulus, figure-eight.

Input: noisy point cloud C without user-defined parameters.



Noisy sample C of the shape X.

An *ε-sample* C of a shape X is a point cloud with $C \subset X^\varepsilon$, $X \subset C^\varepsilon$.

Auto-completion Problem

Given only a noisy sparse sample C of an unknown graph $G \subset \mathbb{R}^2$, complete closed contours that are boundaries of holes in $\mathbb{R}^2 - G$.

Motivation: real holes can be

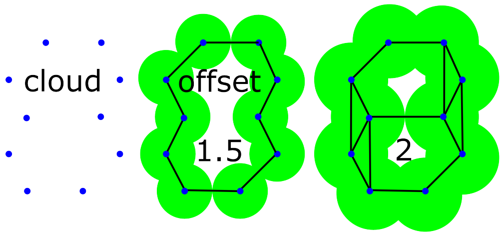
- hidden parts of *occluded images* used for reconstruction or search;
- empty regions in sparse *laser scans* e.g. windows in buildings;

kos.informatik.uni-osnabrueck.de/3Dscans

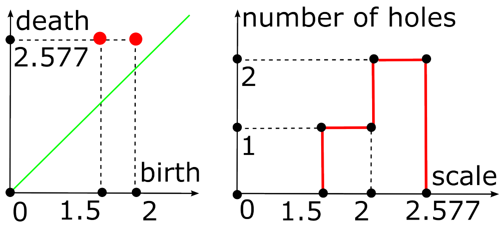
1D persistence of α-offsets

The *α-offset* X^α is the union of balls $\cup B(x;\alpha)$ of a radius α over $p \in X$.

When α is increasing, the offset C^α is growing: holes of C^α are born, die at critical values of the scale α.

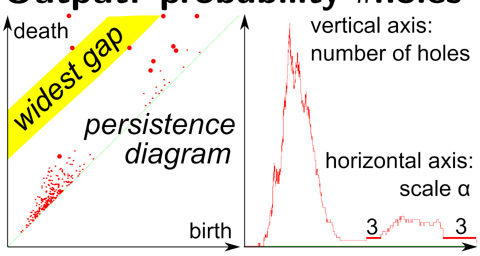


Holes are born at $\alpha = 1.5$, $\alpha = 2$.



The *persistence diagram* PD$\{C^\alpha\}$ is all pairs (birth, death) found from a Delaunay triangulation of $C \subset \mathbb{R}^2$.

Output: probability #holes



Left: a widest diagonal gap in PD$\{C^\alpha\}$ for the sample C of X. *Right:* holes of C^α depending on α.

The cloud C has $P(3 \text{ holes}) \approx 24\%$, $P(2 \text{ holes}) \approx 13\%$, $P(8) \approx 11\%, \ldots$

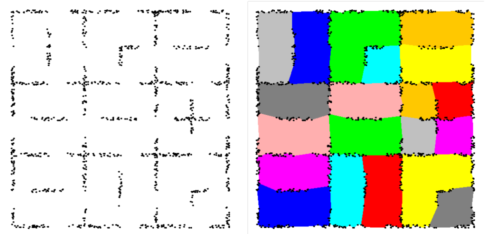
Guarantees for true #holes

$\alpha_{\min}(X)$:= min scale α when a hole is born or dies in an offset X^α.

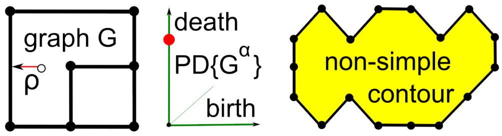
$\alpha_{\max}(X)$:= max scale α after which no holes are born or die in X^α.

Th (VK, CVPR'14). Let X satisfy $\alpha_{\min}(X) > \frac{1}{2}\alpha_{max}(X) + 4\varepsilon$ and no new holes appear in X^α when α is increasing. For any ε-sample C of X, the most likely #holes in $\{C^\alpha\}$ is the true number of holes of X.

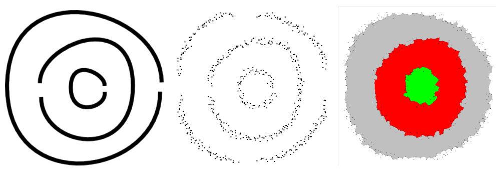
Output: all coloured holes



A graph G is *simple* if any hole in $\mathbb{R}^2 - G$ has the boundary L with $\rho(L)$ such that L^α is circular for $\alpha < \rho(L)$ and $L^\alpha \sim \cdot$ for $\alpha \geq \rho(L)$, so the hole in L^α dies at $\alpha = \rho(L)$.



Th (VK, CTIC'14). Let $G \subset \mathbb{R}^2$ be a simple graph with $\rho_1 \leq \ldots \leq \rho_m$ and $\rho_1 > 7\varepsilon + \max\{\rho_{i+1} - \rho_i\}$. Then for any ε-sample C of G, the algorithm finds m expected contours, they are *in the* 2ε*-offset* $G^{2\varepsilon}$.



Summary and References

input: only a noisy point cloud;

output: probability distribution for #holes, auto-completed contours;

running time: $O(n \log n)$ for any n points with real coordinates in 2D;

more details: a blog with examples and video at http://kurlin.org.

[1] Attali *et al* Persistence-sensitive simplification of functions on surfaces in linear time. *TopoInVis'09*.

[2] H. Edelsbrunner and J. Harer. *Computational topology*. AMS 2010.

[3] V. Kurlin. A fast and robust algorithm to count topologically persistent holes in clouds. *CVPR'14*.

[4] V. Kurlin. Auto-completion of contours in sketches, maps and sparse 2D images. *CTIC 2014*.