Helicality: An Isomap-based Measure of Octave Equivalence in Audio Data

Sripathi Sridhar[1] **and Vincent Lostanlen**[1,2] **sripathi.sridhar@nyu.edu**

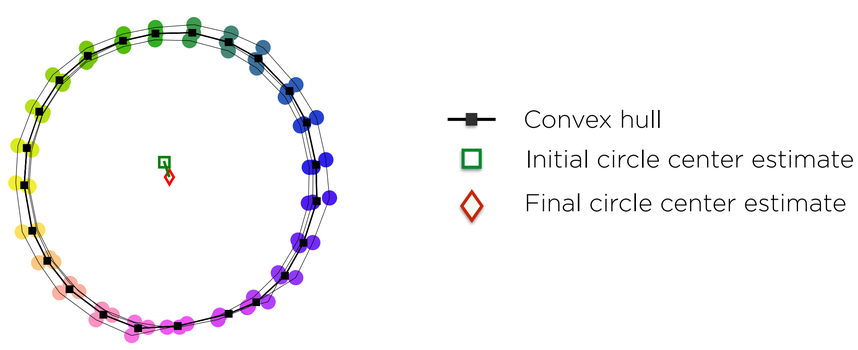
[1]Music and Audio Research Laboratory, New York University, [2]Cornell Lab of Ornithology

- Motivation: **Octave equivalence** in audio data isomap embedding
- Context: Frequency sub-bands in audio data represented as points in 3-D space; distance corresponds inversely to strength of correlations (**Isomap**)
- Visual inspection needed to assess octave equivalence, not scalable
  - Propose an algorithm to quantify "**helicality**"
- **Fit a helix** to the embedding point cloud
  - Parametrically- using circle and line estimates

Isomap embedding protocol [Lostanlen]

- CQT: octave equivariant **time–frequency representation**
- **Pearson correlations** between frequency sub-bands
- *k*-**nearest neighbor graph** → shortest path distances
- Isomap eigenbasis gives low-dimensional embedding

Helix fit- circle estimate, line estimate



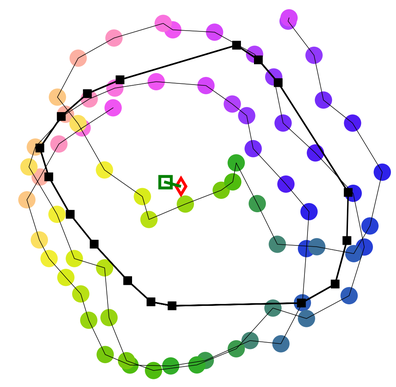
- Octave-average point cloud in two dimensions
- Estimate **convex hull** using **Quickhull** algorithm [Barber]
- **Circle fit** using custom **Frank-Wolfe** conditional gradient [Jaggi]
- **Linear regression** on third dimension
- Fit helix using circle and line estimates
- Helicality- **Inverse of square Euclidean distance** between embedding point cloud $\psi[p]$ and helix estimate in 3-D $\psi'[p]$

$$H = \frac{1}{\frac{1}{P}\sum_{p=1}^{P} \left\| \psi[p] - \psi'[p] \right\|_2^2}$$

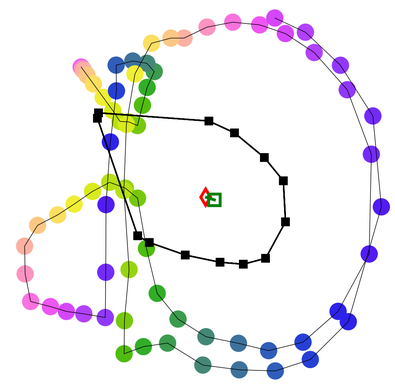
Datasets- Music, Speech and Drums

- TinySOL: 2913 recordings. *Instruments* Acc, ASax, Bn, Fl, ClBb, Ob, TpC, Tbn, Hn, BTb, Vn, Va, Vc, Cb
- North Texas Vowel Dataset: 3190 recordings, 50 speakers
- ENST-drums: 107 isolated drum hits, 3 drummers

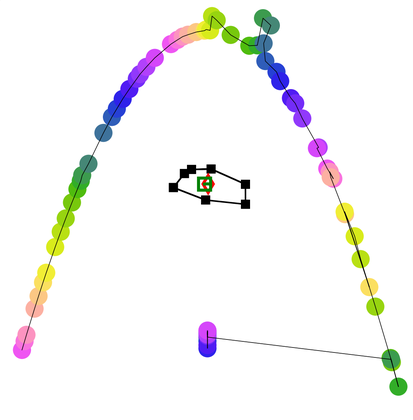
Cross-dataset: TinySOL, NTVow, ENST-drums



Music. Helicality:0.54



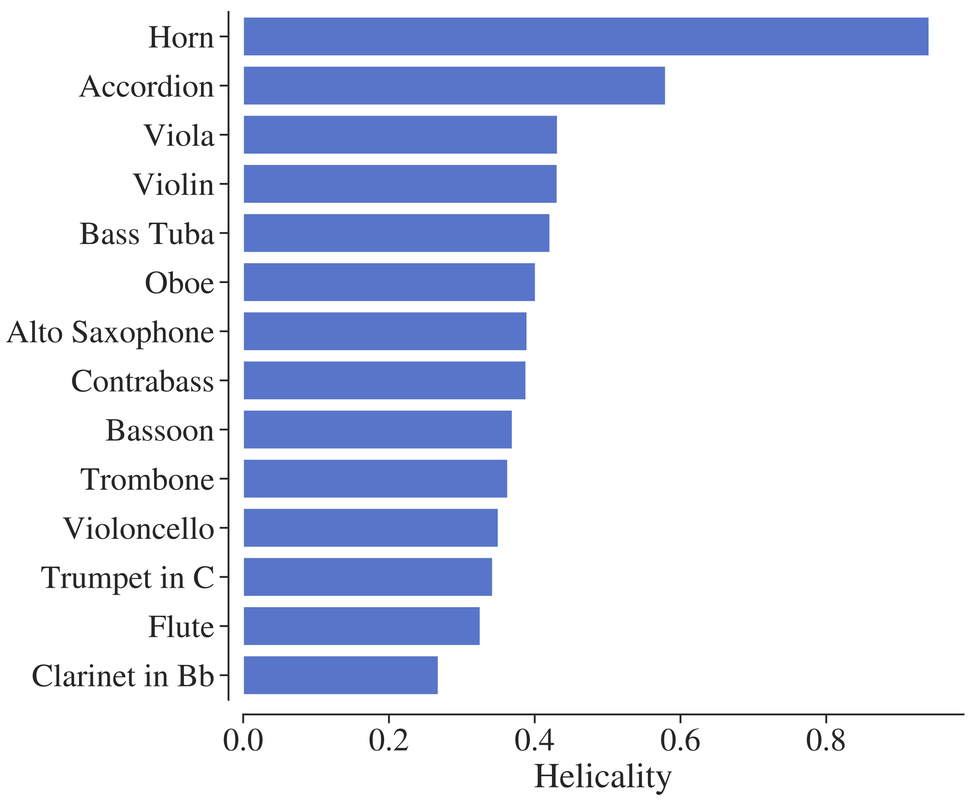
Speech. Helicality:0.30



Drums. Helicality:0.28

- Music, speech, drums embedding topologies
- $\text{Helicality}_{\text{music}} > \text{Helicality}_{\text{speech}} > \text{Helicality}_{\text{drums}}$
- In line with domain knowledge

TinySOL: Helicality of instrument classes



- **Hn** has highest helicality
- Low **TpC** helicality despite its harmonic nature

Future work

- Does it match perception?

NEW YORK UNIVERSITY

TheCornellLab of Ornithology