Generalized Data Depth and Applications

Mukund Raj, Mahsa Mirzargar, Robert M. Kirby and Ross T. Whitaker

MOTIVATION

- **Ensembles of data objects are common in many areas.**
- **Parametric methods of analysis require prior knowledge of the underlying distribution.**
- **Data depth is a nonparametric approach for characterizing ensembles.**

DATA DEPTH BACKGROUND

- **Salient Features of Data Depth:**
  - **Nonparametric**
  - **Robust**
  - **Descriptive Statistical Method**
  - **Derived order statistics can be used for visualization based on the classical boxplot.**

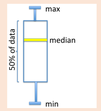

Fig. A classic boxplot for univariate data.

- **Data Depth Formulations for Multivariate Points:**
  - **Simplicial depth**
  - **Location depth**
  - **And many others..**
- **Data Depth Formulations for Complex Data:**
  - **Functions [1] and multivariate curves [3,4]**
  - **Sets and Isocontours [2]**

$$S \in sB(S_1, \dots, S_j) \iff \bigcap_{k=1}^{j} S_k \subset S \subset \bigcup_{k=1}^{j}$$

$$sBD_j = \sum_{j=2}^{J} \mathrm{P}[S \in sB(S_1, \dots, S_j)]$$

PATH BOXPLOT: CHARECTERIZING PATH ENSEMBLES ON GRAPHS

- **Definition of Band for Paths on a Graph**

Let graph $G = \{V, E, W\}$ be a set of vertices, edges and weights on edges. We denote a path p as $p : \mathcal{I} \mapsto V$ over an index set $\mathcal{I} = [1, 2, \dots, m]$.

The convex hull of a set of j vertices $\mathcal{V}_j$ is the smallest geodesic-convex set that contains $\mathcal{V}_j$ and is denoted as $\mathrm{H}[\mathcal{V}_j]$

Then, band formed by j paths is defined as follows:

$$p \in B[\mathcal{P}_j] \quad \text{iff} \quad \mathrm{p(l)} \in \mathrm{H}[\mathrm{p_1(l)}, \dots, \mathrm{p_j(l)}] \quad \forall \mathrm{l} \in \mathcal{I}.$$

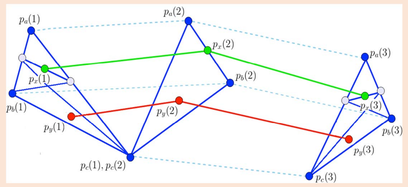

Fig. Illustration for band formed by 3 paths (shown using dotted edges).

- **Path Band Depth (pBD) for Paths on a Graph**

pBD for a path p is defined as follows:

$$\mathrm{pBD}(p) = E\left[\chi\left(p \in B(\mathcal{P}_j)\right)\right]$$

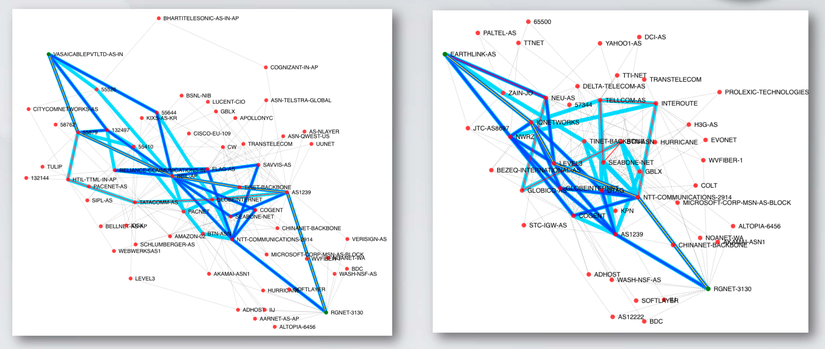

Fig. Path boxplot without without inherent embedding of vertices for ensembles of paths taken by packets over Autonomous System (AS) graph.

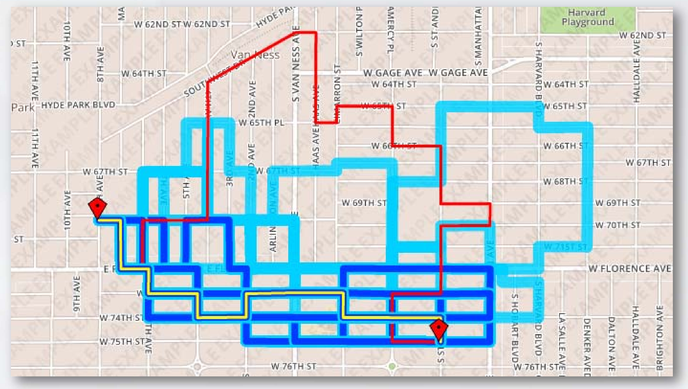

Fig. Path box with inherent embedding of vertices for paths on a road graph.

EVALUATING ALIGNMENT OF SHAPES

- **Evaluating alignment of shapes is important in many areas.**
- **Ensemble visualization through contour boxplot [2] can be an effective method to evaluate alignment of shapes.**

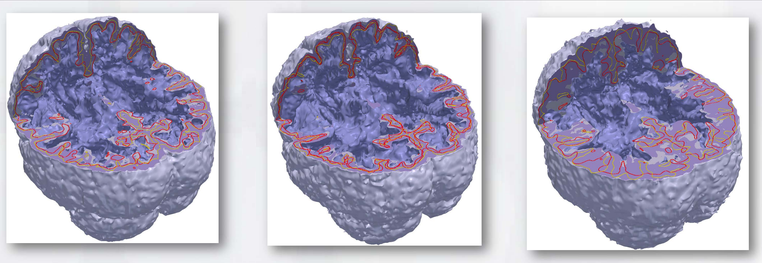

Fig. Comparison of atlases: Contour boxplot for ensembles of cortices in atlas space. Atlases constructed using different parameters / subject groups.

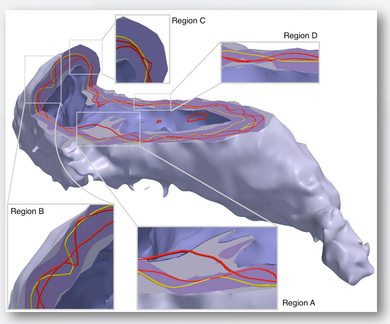

Fig. Evaluating local alignment: Contour boxplot for an ensemble of ventricles.

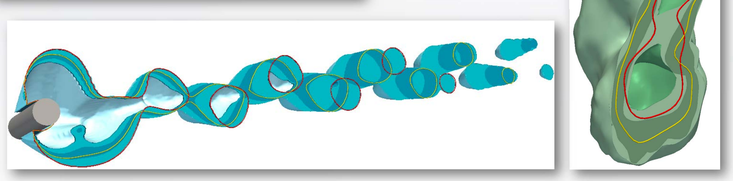

Fig. Contour boxplot for ensemble of isosurfaces of pressure field of fluid flow (left) and an ensemble of simulated HIV molecules (right).

CONTRIBUTIONS

- **Generalization of contour boxplot method for 3D shapes.**
- **Application of contour boxplot to evaluate alignment of shapes.**
- **Formulation of method to calculate data depth for paths on a graph and corresponding path boxplot visualization scheme.**

REFERENCES

- **[1] Pintado et al. JASA 2009**
- **[2] Whitaker et al. IEEE Visualization 2013**
- **[3] Pintado et al. ADAC 2014**
- **[4] Mirzargar et al. IEEE Visualization 2014**

ACKNOWLEDGEMENTS

This work was supported by NSF grant IIS-1212806.