A Comparison of Relative Position Descriptors for 3D Objects

Andrew R. Buck[a], Derek T. Anderson[a], James M. Keller[a],
Robert H. Luke III[b], and Grant Scott[a]

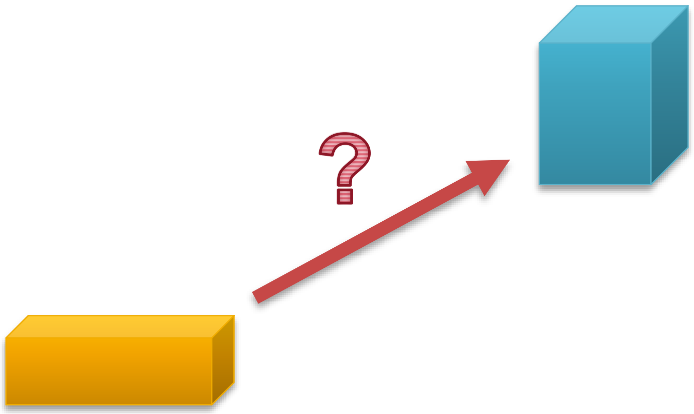
[a]Department of Electrical Engineering and Computer Science, University of Missouri, Columbia MO, USA
[b]U.S. Army DEVCOM C5ISR Center, Fort Belvoir, VA, USA

WCCI FUZZ-IEEE 2022
July 22, 2022

Presented by Andrew Buck

Outline

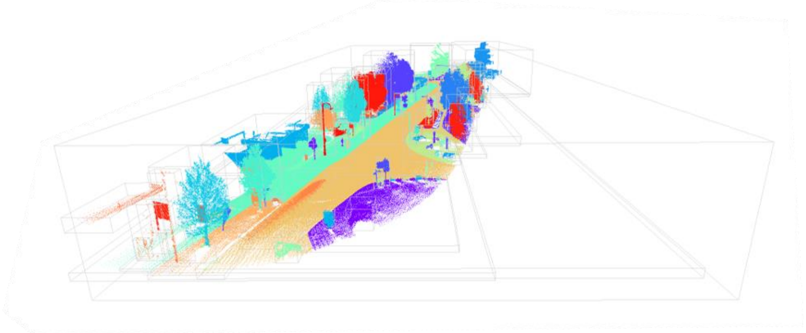
- Motivation
  - 3D relative position descriptor
  - Example NPM3D dataset
- Histogram of forces descriptors
  - 2D projections
- Triangular fuzzy number descriptors
  - Single-axis methods
  - Axis-aligned bounding boxes
- Comparison and analysis



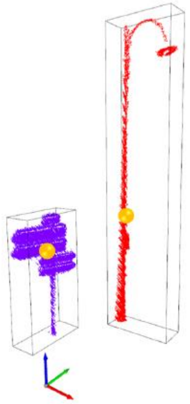
Motivation

- How to describe the relative position of two objects?

- Specifically interested in large outdoor scenes
  - Generated as 3D point clouds
  - Individual object segmentations

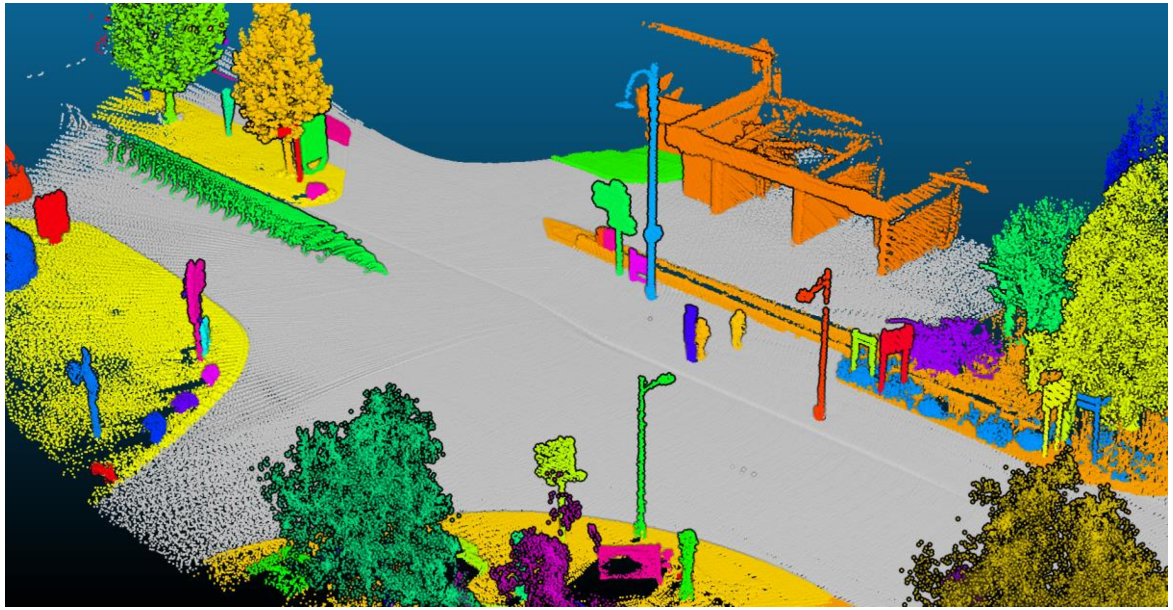


- What types of descriptors can be used?
  - Axis-aligned bounding boxes (TFN)
  - Histogram of forces



NPM3D Dataset

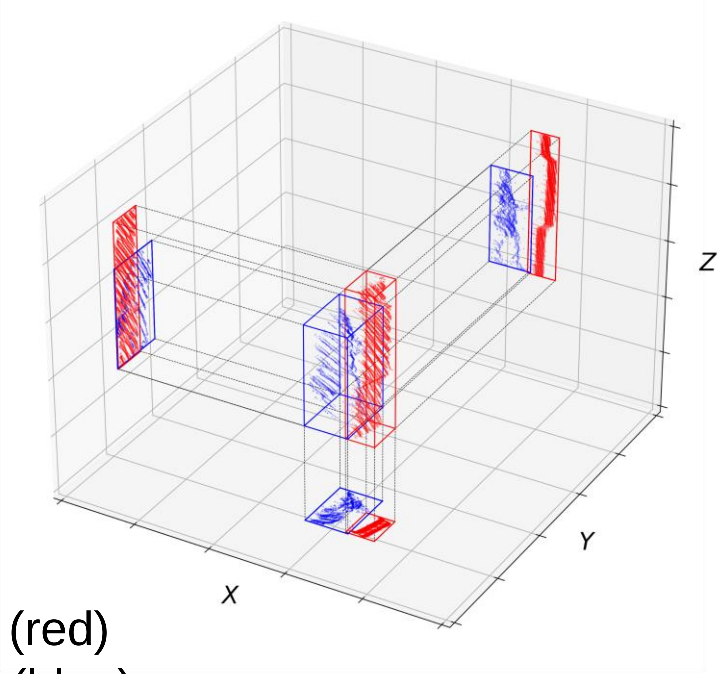
- **We use the Paris-Lille-3D dataset as an example**
  - **Outdoor urban scene captured with LiDAR**
  - **Human labeled object instances**



Example Object Pairs

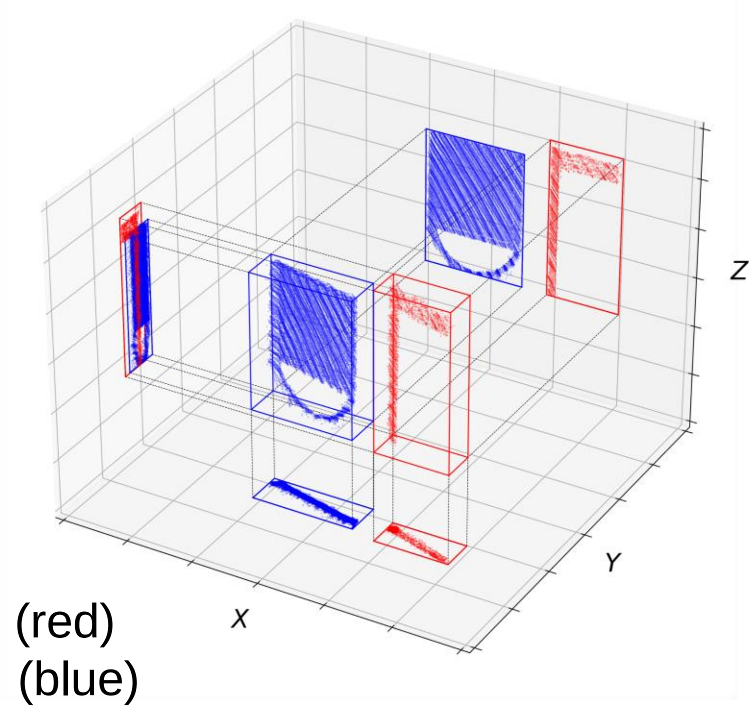
- **How to describe the relative position of $A \rightarrow B$?**
- **Is it similar to $A' \rightarrow B'$?**

Person standing near an information sign



Object A (red)
Object B (blue)

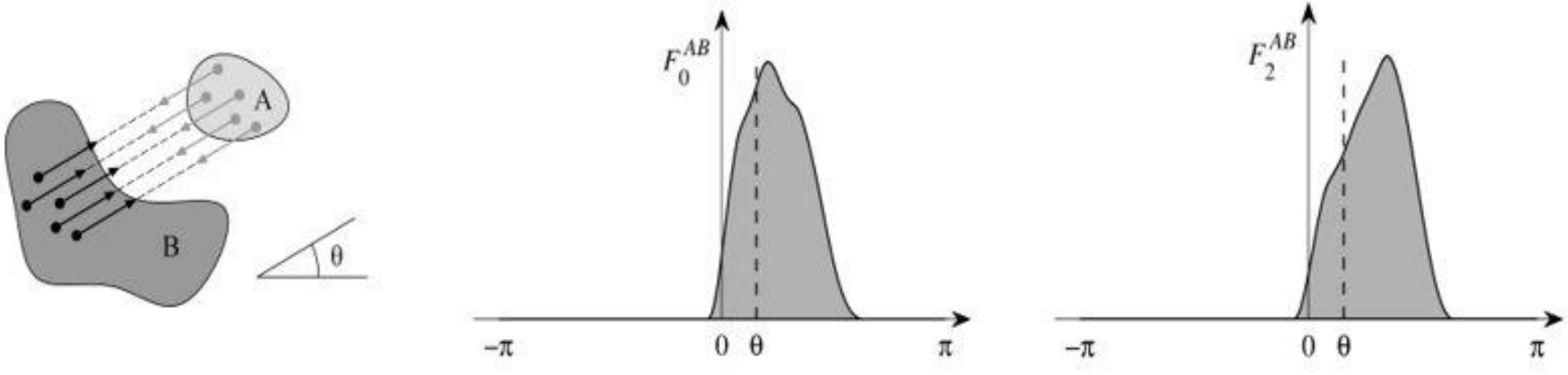
Two other signs in a similar spatial configuration



Object A' (red)
Object B' (blue)

Histograms of Forces (2D)

- **A force histogram $F_r^{AB}(\theta)$ represents the degree of truth of the statement "A is in direction θ from B."**
- **Variants:**
  - Constant forces (F_0) is independent of distance
  - Gravitational forces (F_2) is independent of scale
  - Hybrid forces (F_{02}) blends F_0 and F_2 and handles overlapping objects

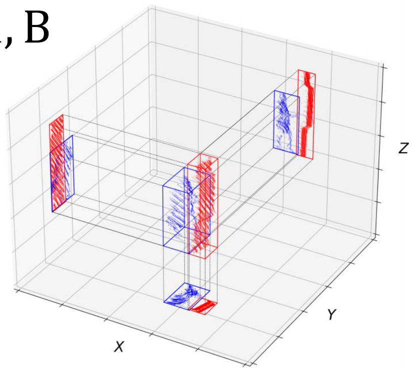


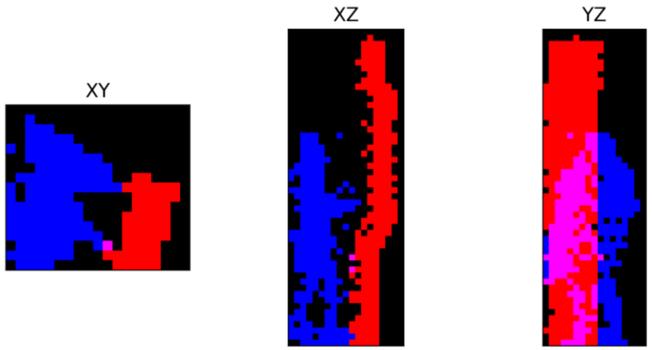
HoF Relative Position Descriptor

The HoF-RPD is comprised of F_0 and F_{02} histograms for each principle-axis 2D projection

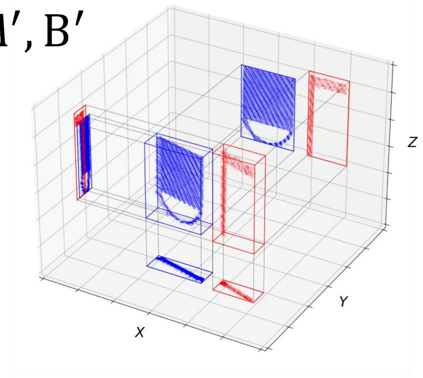
Axis projections

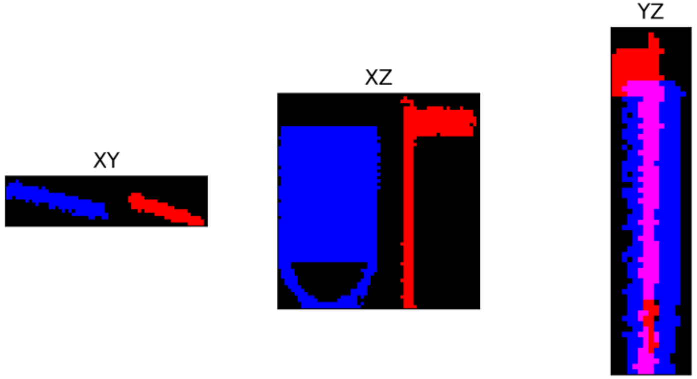
A, B



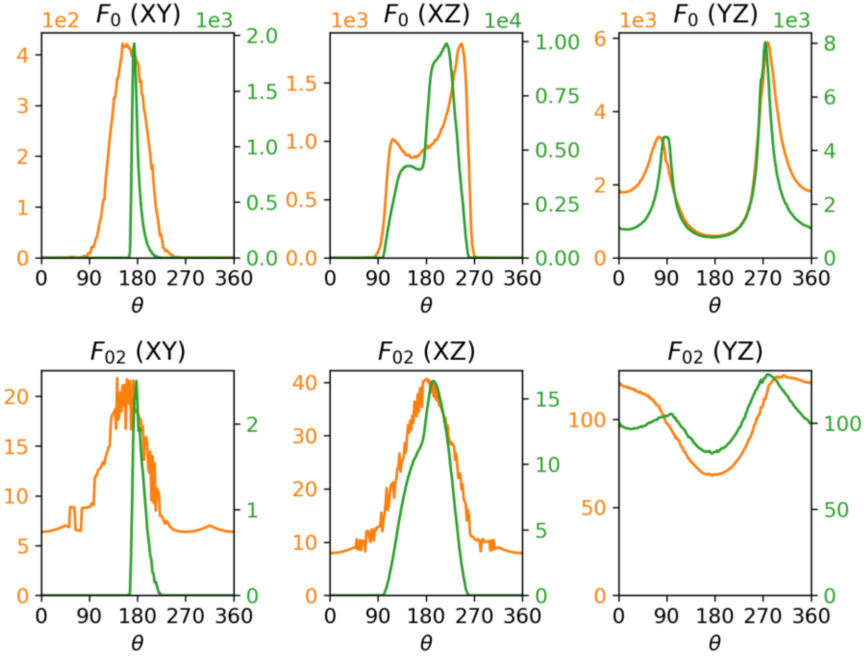


A', B'





$$\text{HoF-RPD}(A,B) = \begin{pmatrix} F_0^{AB_{XY}} & F_0^{AB_{XZ}} & F_0^{AB_{YZ}} \\ F_{02}^{AB_{XY}} & F_{02}^{AB_{XZ}} & F_{02}^{AB_{YZ}} \end{pmatrix}$$



Constant Forces

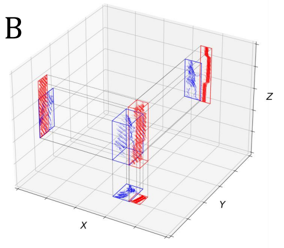
$A \to B$: orange
$A' \to B'$: green

Hybrid Forces

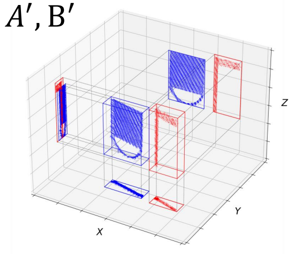
HoF-RPD Similarity

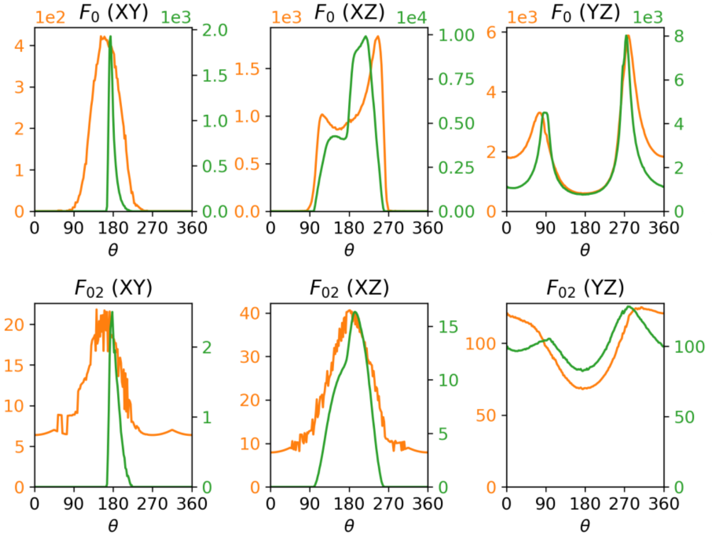
A measure $S_{HoF}(A, B, A', B')$ evaluates the similarity between the relationship $A \rightarrow B$ and $A' \rightarrow B'$

A, B



A', B'





Histogram similarity

$$\mu_T(h_1, h_2) = \frac{\sum_\theta \min\left(h_1(\theta), h_2(\theta)\right)}{\sum_\theta \max\left(h_1(\theta), h_2(\theta)\right)}$$

$$\mu_P(h_1, h_2) = 1 - \frac{\sum_\theta \left|h_1(\theta) - h_2(\theta)\right|}{\sum_\theta \left|h_1(\theta) + h_2(\theta)\right|}$$

$$\mu_C(h_1, h_2) = \frac{\sum_\theta h_1(\theta) h_2(\theta)}{\sqrt{\sum_\theta h_1^2(\theta)}\sqrt{\sum_\theta h_2^2(\theta)}}$$

Object pair similarity

$$s_\mu^k(A, B, A', B') = \frac{\mu(F_0^{AB_k}, F_0^{A'B'_k})}{2} + \frac{\mu(F_{02}^{AB_k}, F_{02}^{A'B'_k})}{2}$$

Projection aggregation

$$S_{\text{HoF,min},\mu}(A, B, A', B') = \min\left(s_\mu^{XY}, s_\mu^{XZ}, s_\mu^{YZ}\right)$$

$$S_{\text{HoF,mean},\mu}(A, B, A', B') = \frac{1}{3}\left(s_\mu^{XY} + s_\mu^{XZ} + s_\mu^{YZ}\right)$$

TABLE I
FORCE HISTOGRAM SIMILARITIES

	F_0			F_{02}		
	XY	XZ	YZ	XY	XZ	YZ
μ_T	0.236	0.201	0.722	0.016	0.207	0.868
μ_P	0.381	0.334	0.839	0.031	0.343	0.929
μ_C	0.548	0.859	0.948	0.482	0.899	0.989

TABLE II
OBJECT PAIR SIMILARITIES USING THE HOF-RPD

	s^{XY}	s^{XZ}	s^{YZ}	S_{min}	S_{mean}
μ_T	0.126	0.204	0.795	0.126	0.375
μ_P	0.206	0.338	0.884	0.206	0.476
μ_C	0.515	0.879	0.968	0.515	0.787

6 different HoF similarity methods
(defined here)

The TFN-RPD is based on the triangular fuzzy number difference vector D along each axis

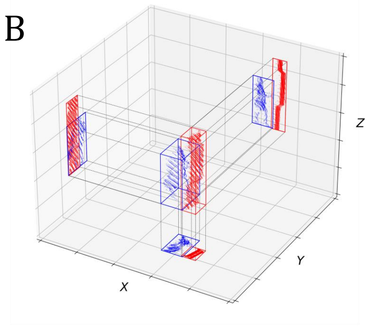
$$D_x = B_x - A_x$$
$$D_y = B_y - A_y$$
$$D_z = B_z - A_z$$

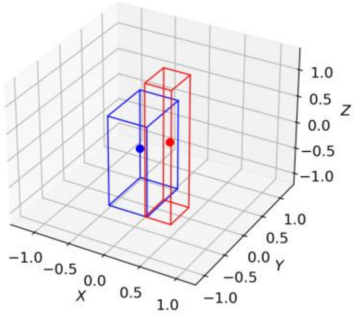
$$\text{TFN-RPD}(A, B) = (D_x, D_y, D_z)$$

Axis-aligned bounding boxes with object centroids

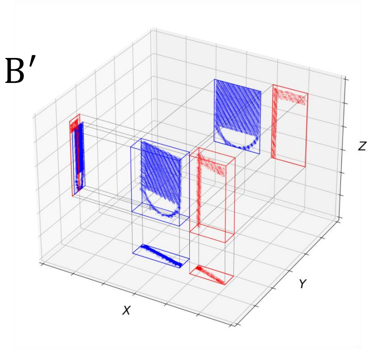
TFN and differences for each axis

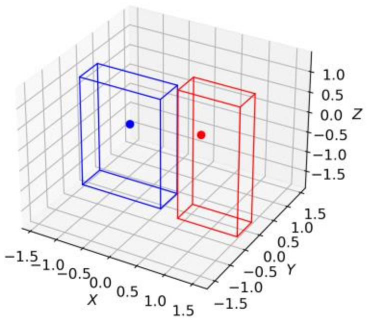
A, B

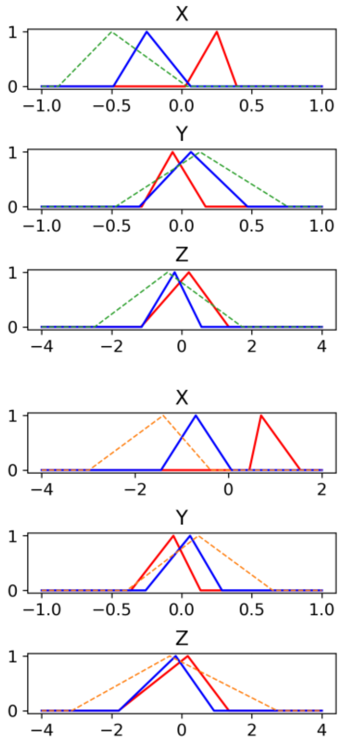




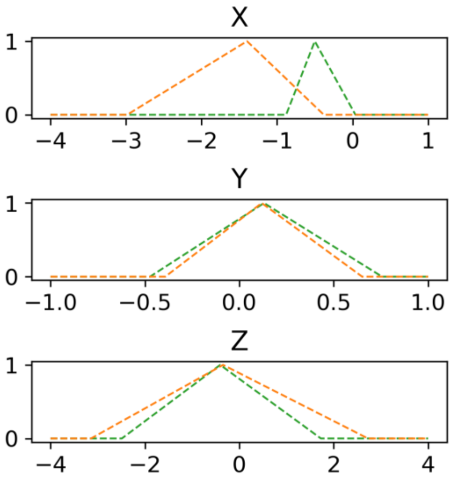
A', B'







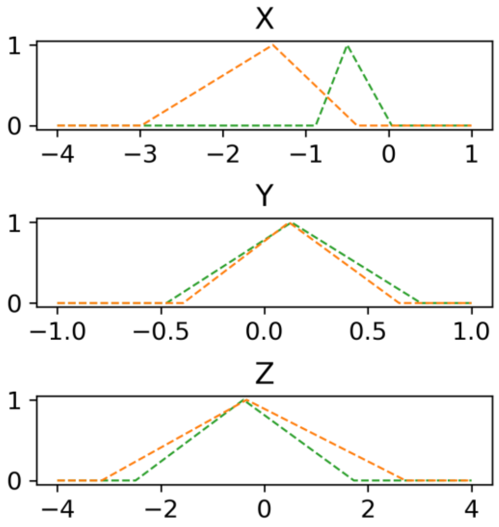
Difference TFNs



TFN-RPD Single Axis Similarity

The measure S_{TFN-SA} evaluates object pair similarity using single axis triangular fuzzy numbers

Difference TFNs



$$D = (d_1, d_2, d_3)$$
$$D' = (d'_1, d'_2, d'_3)$$

Difference TFN similarity

$$\mu_{\max}(D, D') = \max_{x \in \mathbb{R}} \left\{ \min \left(m_D(x), m_{D'}(x) \right) \right\}$$

$$\mu_{\text{IoU}}(D, D') = \frac{|D \cap D'|}{|D \cup D'|}$$

$$\mu_{\text{PD}}(D, D') = \frac{1}{1 + \text{PD}(D, D')}$$

$$\text{PD}(D, D') = \sum_{i=1}^{3} \frac{|d_i - d'_i|}{d_3 - d_1 + d'_3 - d'_1}$$

Aggregation along each axis

$$S_{\text{TFN}-\text{SA},\min,\mu}(D, D') = \min \left(s^X_\mu, s^Y_\mu, s^Z_\mu \right)$$

$$S_{\text{TFN}-\text{SA},\text{mean},\mu}(D, D') = \frac{1}{3} \left(s^X_\mu + s^Y_\mu + s^Z_\mu \right)$$

TABLE III
OBJECT PAIR SIMILARITIES USING THE TFN-RPD SINGLE-AXIS METHODS

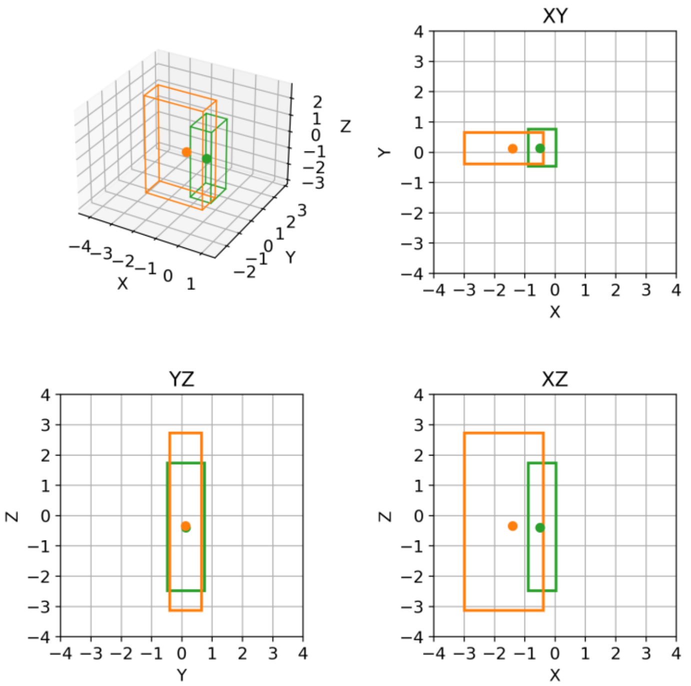
	s^X	s^Y	s^Z	$S_{\min}$	S_{mean}
$\mu_{\max}$	0.358	0.990	0.988	0.358	0.778
μ_{IoU}	0.053	0.846	0.717	0.053	0.539
μ_{PD}	0.507	0.920	0.855	0.507	0.761

6 different single-axis TFN similarity methods (defined here)

TFN-RPD Bounding Box Similarity

The measure S_{TFN-BB} evaluates object pair similarity using axis-aligned bounding boxes and TFNs

Difference TFNs as bounding boxes and centroids



Difference bounding box similarity

$$\mu_{\mathrm{IoU}}(D, D') = \frac{|D \cap D'|}{|D \cup D'|}$$

$$\mu_{\mathrm{GIoU}}(D, D') = \frac{1}{2}\left(\frac{|D \cap D'|}{|D \cup D'|} - \frac{|C \setminus (D \cup D')|}{|C|} + 1\right)$$

$$\mu_{\mathrm{PD}}(D, D') = \frac{1}{1 + \sum_{j \in \{x,y,z\}} \mathrm{PD}(D_j, D'_j)}$$

Projection aggregation

$$S_{\mathrm{TFN-BB,min},\mu}(D, D') = \min\left(s_\mu^{XY}, s_\mu^{XZ}, s_\mu^{YZ}\right)$$

$$S_{\mathrm{TFN-BB,mean},\mu}(D, D') = \frac{1}{3}\left(s_\mu^{XY} + s_\mu^{XZ} + s_\mu^{YZ}\right)$$

Full 3D similarity

$$S_{\mathrm{TFN-BB,3D},\mu}(D, D') = s_\mu^{XYZ}$$

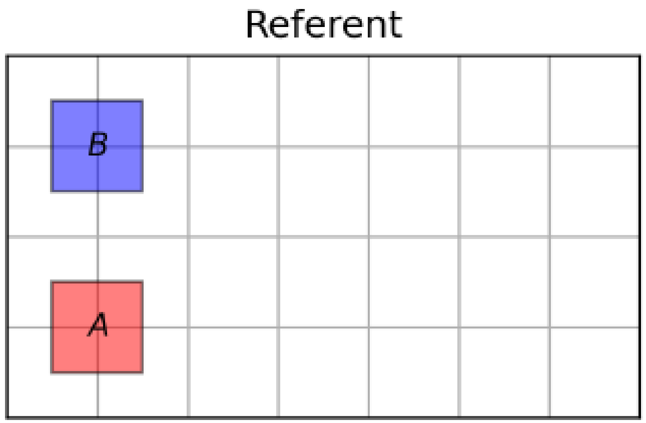
TABLE IV
OBJECT PAIR SIMILARITIES USING THE TFN-RPD BOUNDING BOX METHODS

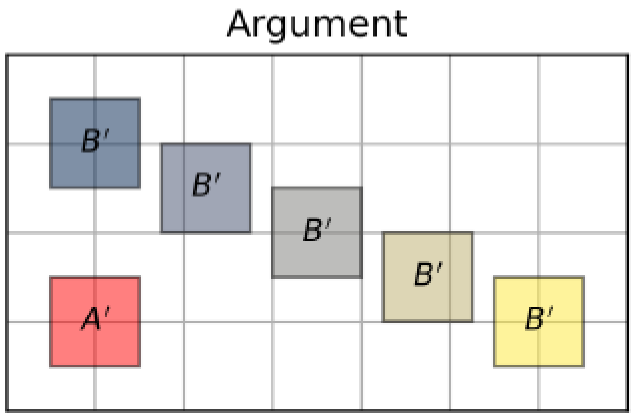
	s^{XY}	s^{XZ}	s^{YZ}	S_{min}	S_{mean}	S_{3D}
μ_{IoU}	0.157	0.123	0.637	0.123	0.306	0.119
μ_{GIoU}	0.526	0.542	0.797	0.526	0.622	0.484
μ_{PD}	0.486	0.467	0.796	0.467	0.583	0.449

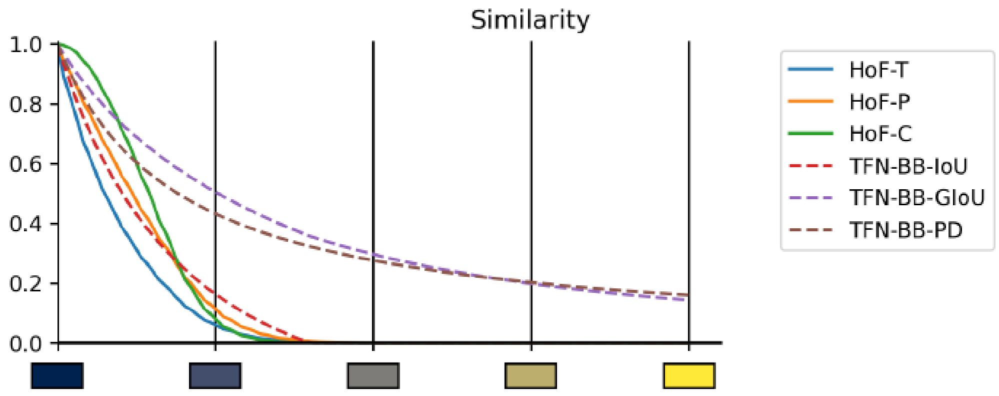
9 different bounding box TFN similarity methods (defined here)

2D Comparison Example 1

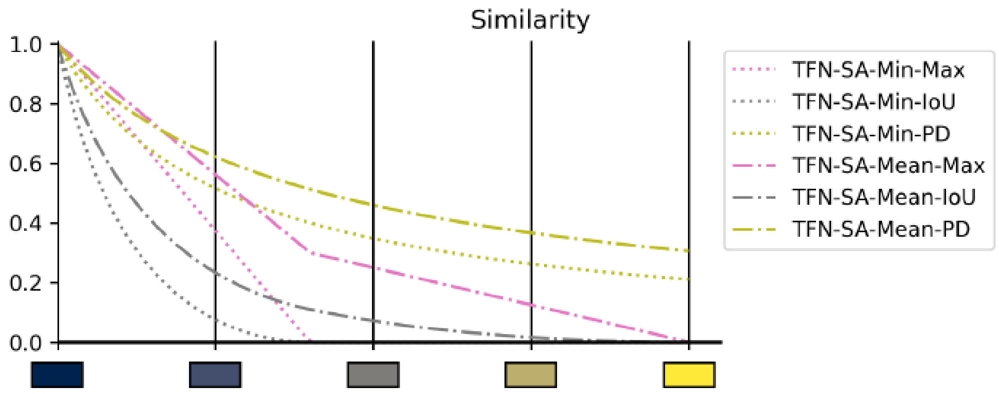
Similarity of $A' \rightarrow B'$ to $A \rightarrow B$ decreases as B' moves down and to the right.







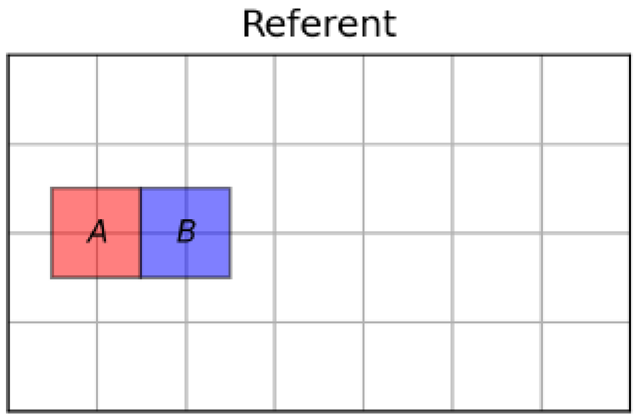
(b)

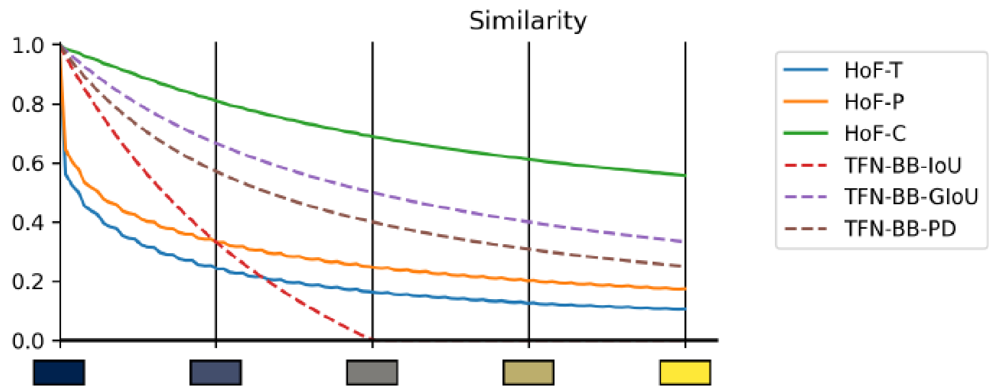


(c)

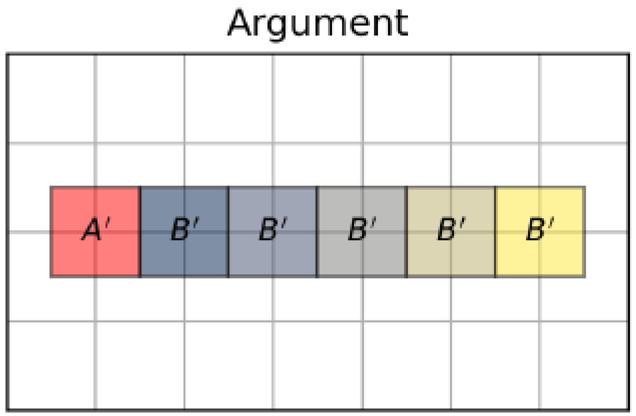
2D Comparison Example 2

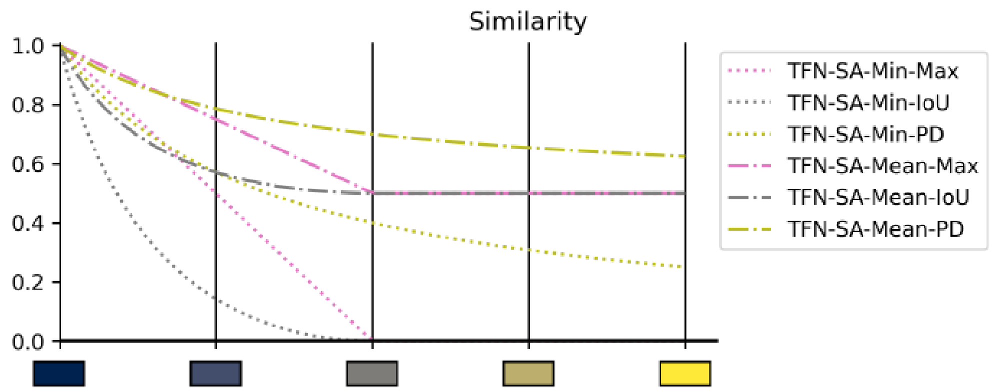
Similarity of $A' \rightarrow B'$ to $A \rightarrow B$ decreases as B' moves farther away.





(b)

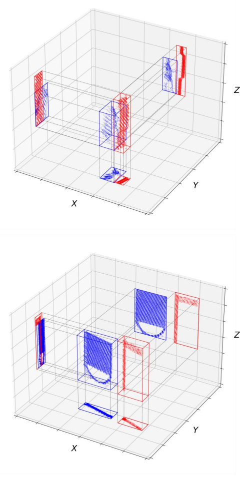




(c)

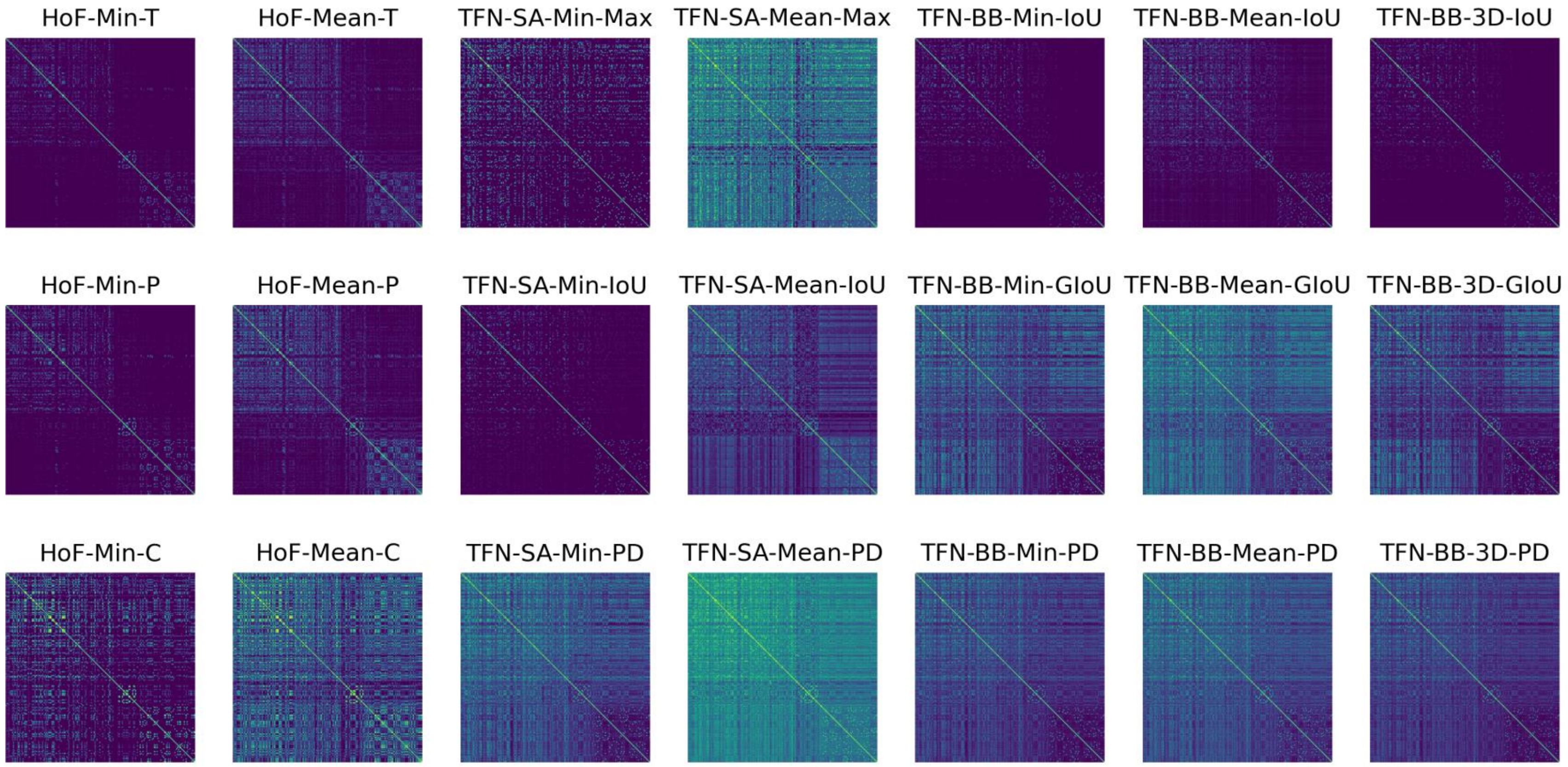
3D Comparison Experiments

We extract 202 object pairs from the NPM3D dataset (no more than 5 meters apart) and evaluate all 21 similarity measures against all combinations.



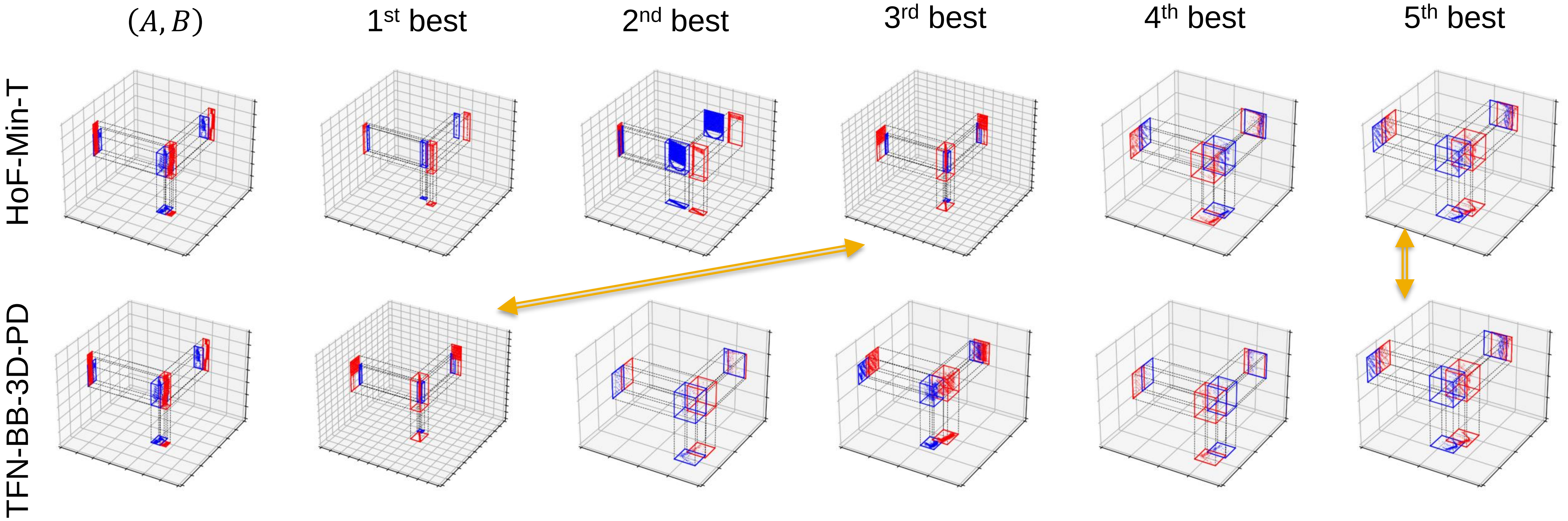
⋮

Dataset of 202 object pairs



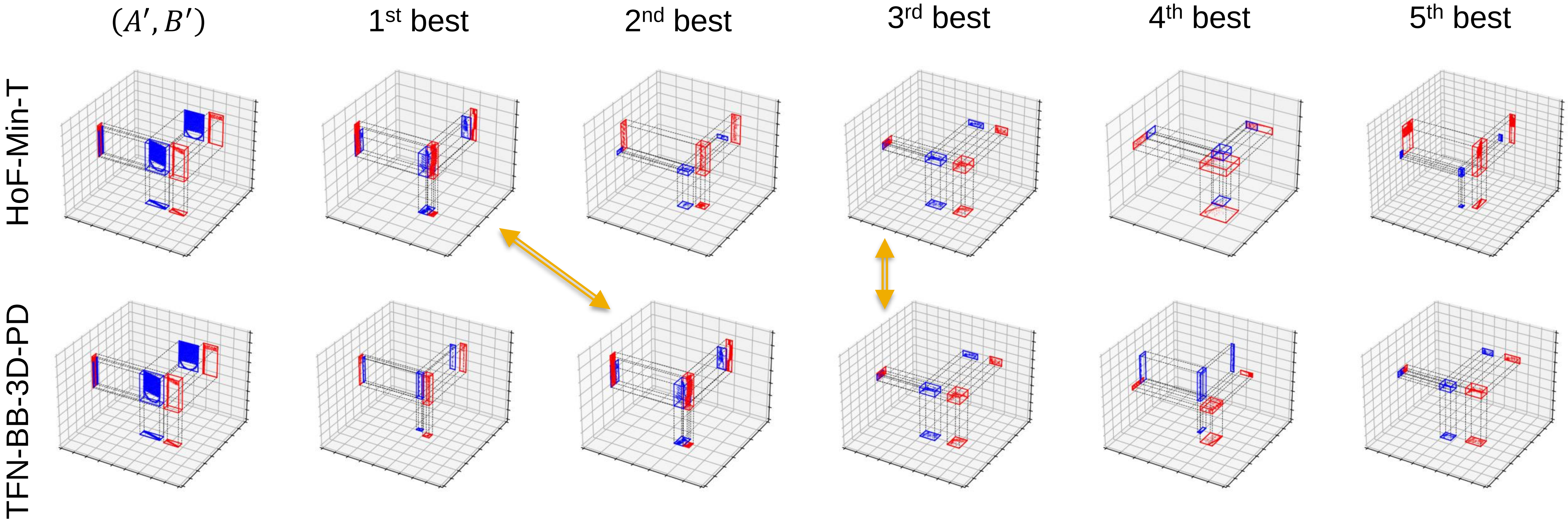
Top matches for pair (A, B)

Comparing the HoF-Min-T and TFN-BB-3D-PD methods on the example object pair (A, B), 2 of the top 5 (non-self) matches were the same.



Top matches for pair (A’, B’)

Comparing the HoF-Min-T and TFN-BB-3D-PD methods on the example object pair (A', B'), 2 of the top 5 (non-self) matches were the same.



Comparing Similarity Measures

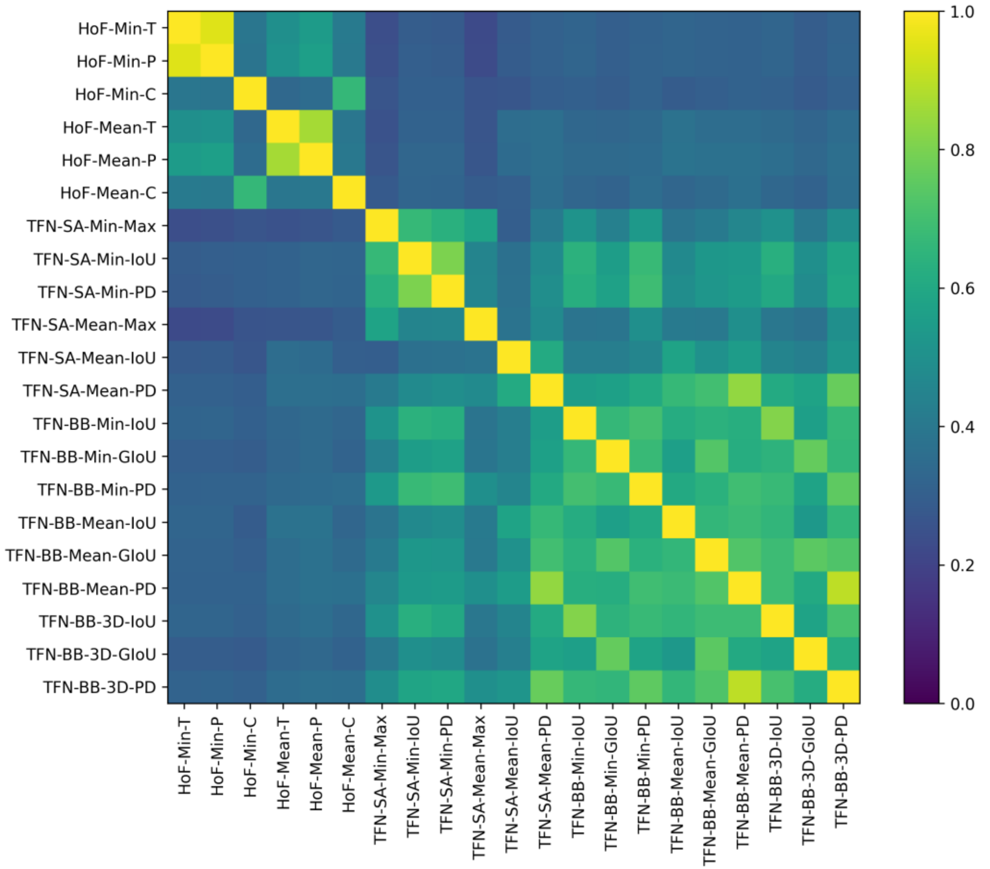
We compare all 21 measures by counting how often they identify the same set of similar object pairs.

For a pair of methods (i, j) and object pair index k,

Let M_{ik} be the set of 5 object pairs that most closely match object pair k according to method i,

Let M_{jk} be the set of 5 object pairs that most closely match object pair k according to method j.

$$s_m(i,j) = \frac{1}{|K|} \sum_{k \in K} \frac{|M_{ik} \cap M_{jk}|}{|M_{ik} \cup M_{jk}|}$$



Conclusions

- **There are many ways to represent the spatial relationship between two 3D objects**
  - No single best method; Each has strengths and weaknesses
- **HoF methods are more descriptive and good for recognizing shape differences between close objects**
  - Currently requires rasterization in 2D; Could explore 3D HoF
- **TFN methods apply approximations to be fast and are better suited for far apart objects**
  - Good when description is mostly direction and distance
- **Not currently accounting for rotation and scale invariance**
  - We assume a common frame of reference for both objects