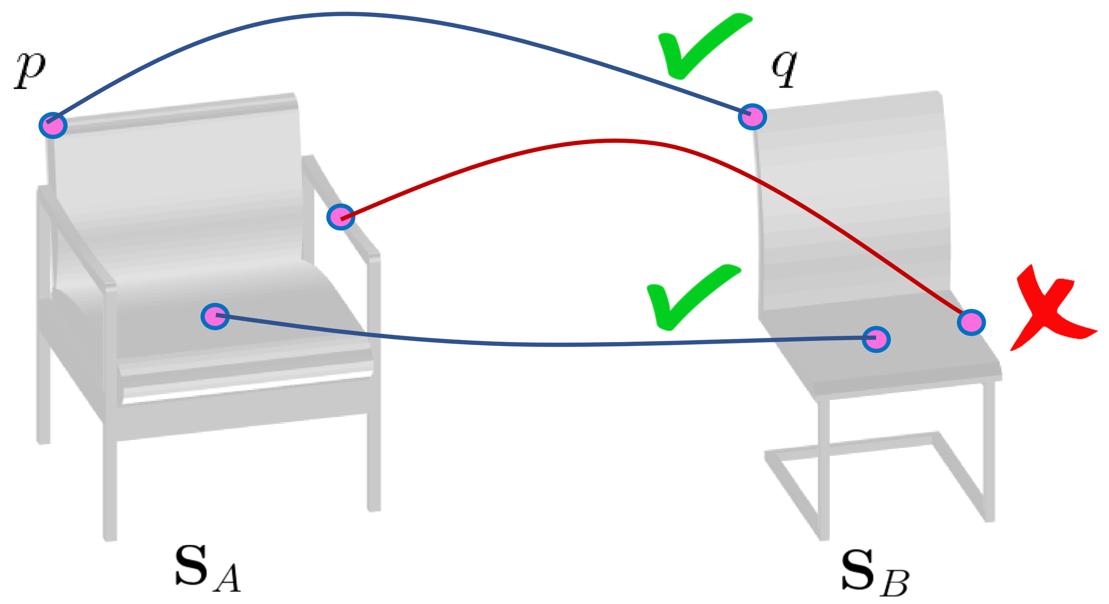
Learning Implicit Functions for Topology-Varying Dense 3D Shape Correspondence

Feng Liu Xiaoming Liu
Department of Computer Science and Engineering, Michigan State University

Dense 3D Correspondence for Man-made Objects

Definition:

Given two shapes $\mathbf{S}_A$ and $\mathbf{S}_B$ belonging in the same category, for an ***arbitrary*** point $p \in \mathbf{S}_A$, we are seeking its **semantically** equivalent point q on $\mathbf{S}_B$.



Challenge:

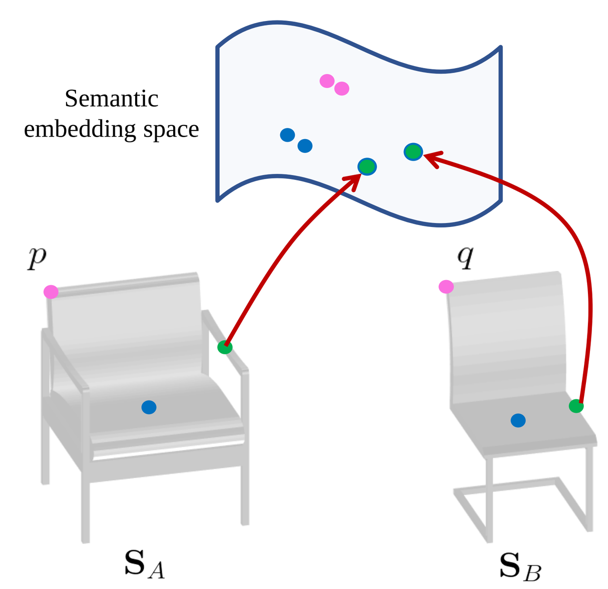
- Man-made objects often differ not only by geometric deformations, but also by ***part constitutions***;
- Prior methods have proven to be effective on organic shapes, e,g,. Human bodies and mammals, they become ***less suitable*** for generic object;
- Existing method for man-made object either performance fuzzy correspondence or predict a constant number of semantic points;
- The lack of annotations on dense correspondence often leaves ***unsupervised*** learning the only option.

Contributions:

- We propose a novel paradigm leveraging implicit functions for category-specific ***unsupervised*** dense 3D correspondence, which is suitable for ***topology-varying*** objects;
- We estimate a ***confidence score*** measuring if the predicted correspondence is valid or not;
- We demonstrate the superiority of our method in shape segmentation and 3D semantic correspondence.

Proposed Method

Formulation:



- We assume a semantic embedding function (SEF) $f : \mathbb{R}^3 \times \mathbb{R}^d \rightarrow \mathbb{R}^k$, the correspondence should satisfy:

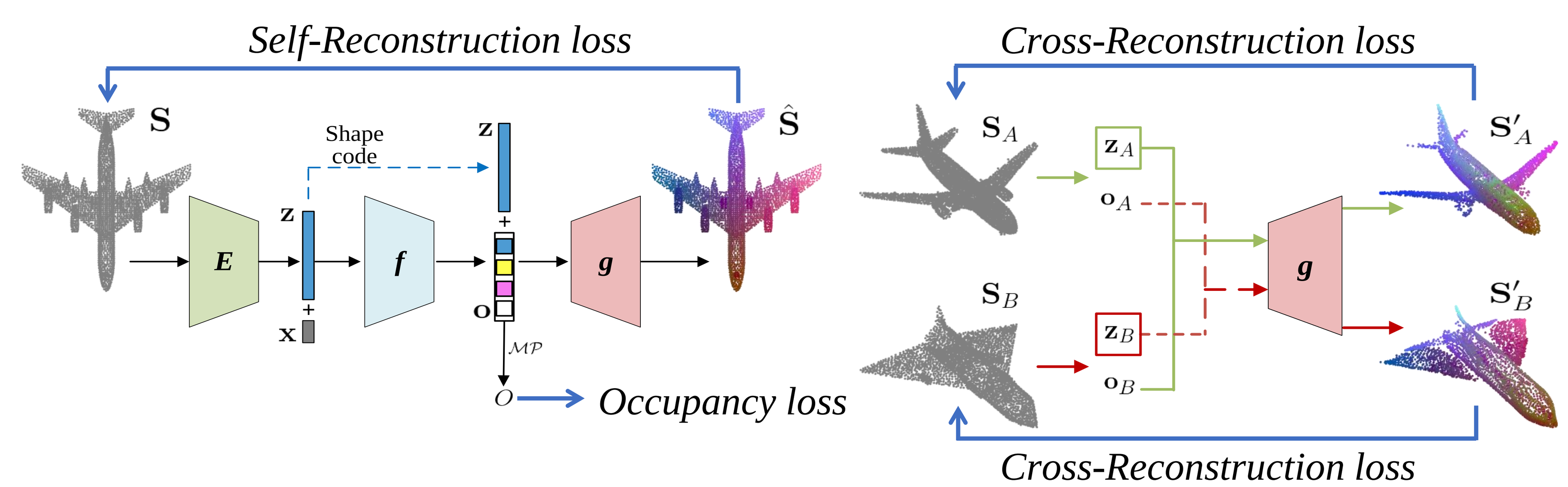
$$\left(\min_{q \in \mathbf{S}_B} ||f(p, \mathbf{z}_A) - f(q, \mathbf{z}_B)||_2\right) < \tau, \quad \forall p \in \mathbf{S}_A$$

shape latent code

- If the distance is too large ($\geq \tau$), there is no corresponding point in $\mathbf{S}_B$ for p ;
- If SEF could be learned, then $q = f^{-1}(f(p, \mathbf{z}_A), \mathbf{z}_B)$

inverse function

Solution:

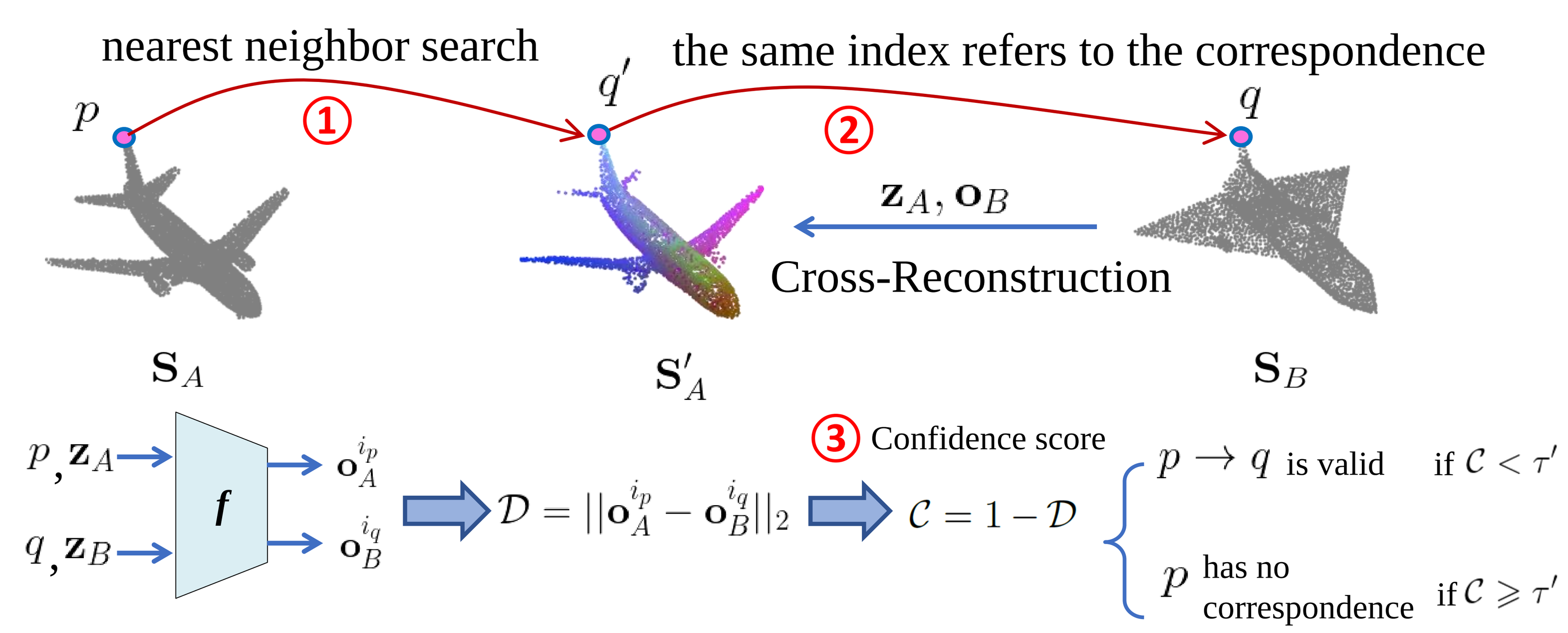


- A branched implicit function $\boxed{f}$ serves as the semantic embedding function (SEF).
- Design an inverse function $\boxed{g}$ mapping from the embedding space to 3D space: $g : \mathbb{R}^k \times \mathbb{R}^d \rightarrow \mathbb{R}^3$, so that the learning objectives can be defined in 3D space.
- Loss functions: $\mathcal{L}^{all} = \mathcal{L}^{occ} + \mathcal{L}^{SR} + \mathcal{L}^{CR}$

Occupancy loss *Self-Reconstruction loss* *Cross-Reconstruction loss*

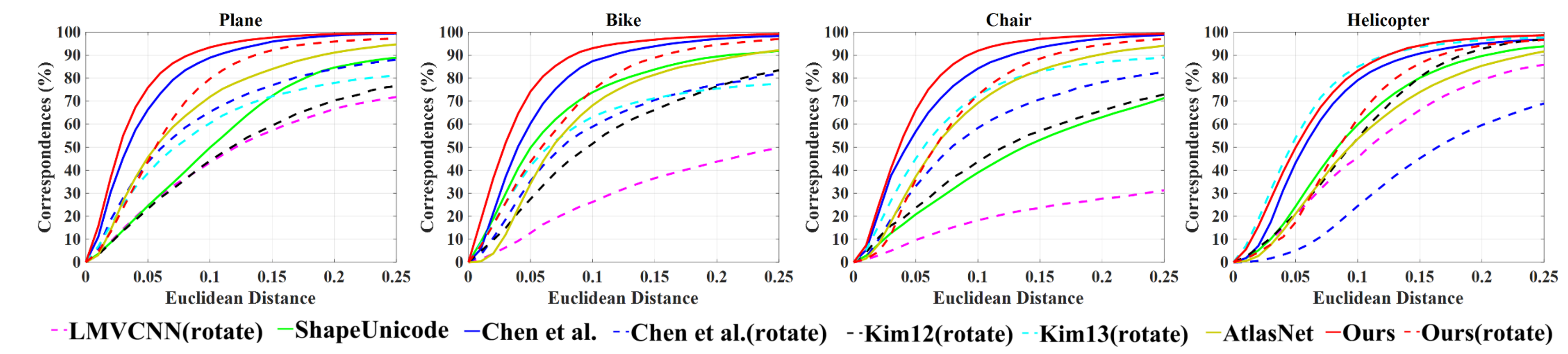
* Cross-Reconstruction loss enforces part embedding consistency across all shapes.

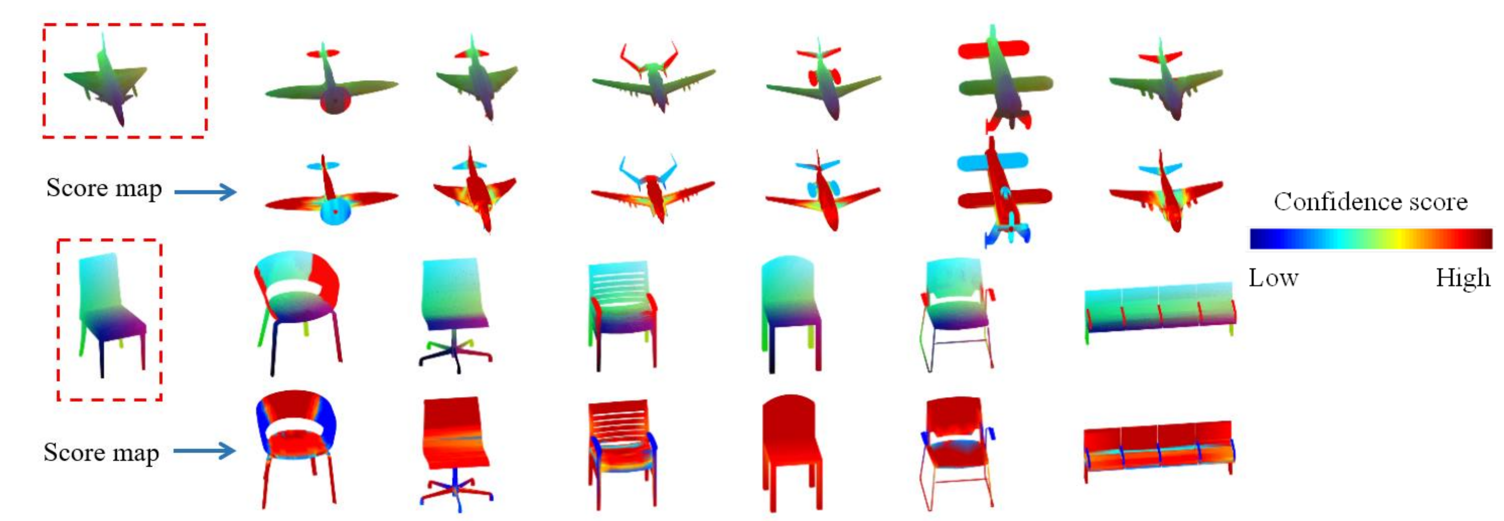
Inference:



Experimental Results:

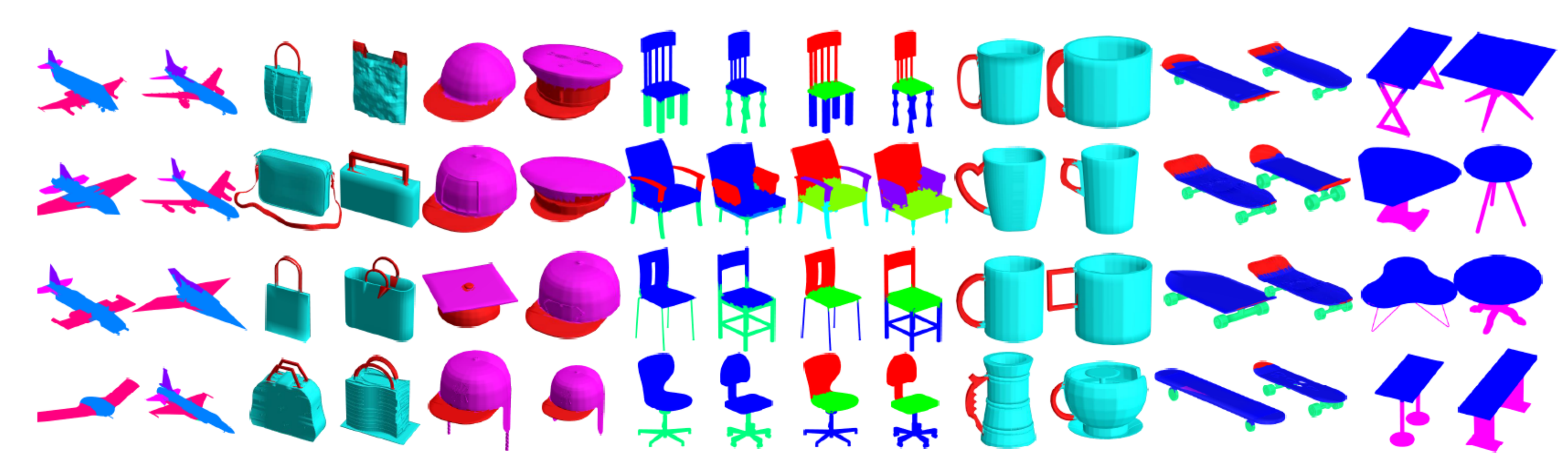
3D semantic Correspondence on BHCP





Unsupervised Shape Segmentation on ShapeNet

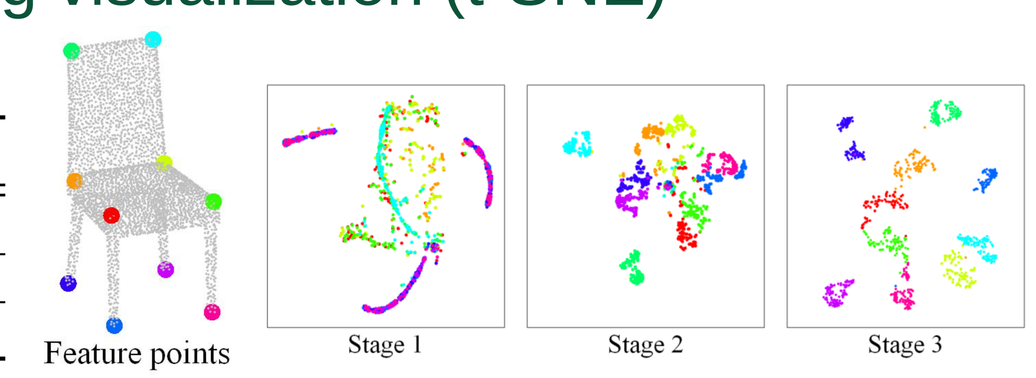
Shape (#parts)	plane (3)	bag (2)	cap (2)	chair (3)	chair* (4)	mug (2)	skateboard (2)	table (2)	Aver.
Segmented parts	body,tail, wing+engine	body, handle	panel, peak	back+seat, leg, arm	back, seat, leg, arm	body, handle	deck, wheel+bar	top, leg+support	
BAE-Net [9]	80.4	82.5	87.3	86.6	83.7	93.4	88.1	87.0	86.1
Proposed	81.0	85.4	87.9	88.2	86.2	94.7	91.6	88.3	**88.0**



Part embedding visualization (t-SNE)

Training stage

	Network	Loss
Stage 1	E, f	$\mathcal{L}^{occ}$
Stage 2	E, f, g	$\mathcal{L}^{occ}$ and $\mathcal{L}^{SR}$
Stage 3	E, f, g	$\mathcal{L}^{all}$



References

[1] Kaick et al. A Survey on Shape Correspondence. In Computer Graphics Forum, 2011.
[2] Huang et al. Functional Map Networks for Analyzing and Exploring Large Shape Collections. TOG, 2014.
[3] Chen et al. BAE-NET: Branched Autoencoder for Shape Co-Segmentation. In ICCV, 2019.
[4] Chen et al. Unsupervised Learning of Intrinsic Structural Representation Points. In CVPR, 2020.