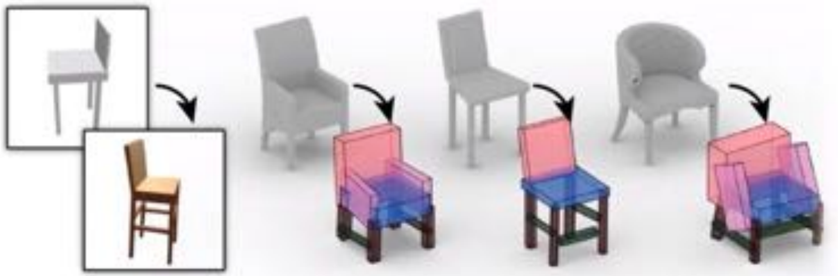
Part-level and Structural Understanding for 3D Shape Perception, Synthesis and Editing

Kaichun Mo
Stanford University



Leonidas Guibas Laboratory

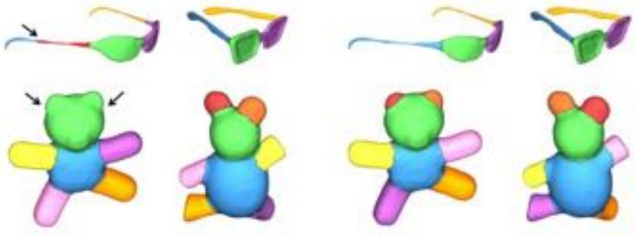
Geometric Computing

Short Bio

- Kaichun Mo 莫凯淳
- 4th year CS Ph.D. at Stanford
- Advisor: Prof. Leonidas Guibas
- https://cs.stanford.edu/~kaichun/
- kaichun@cs.stanford.edu

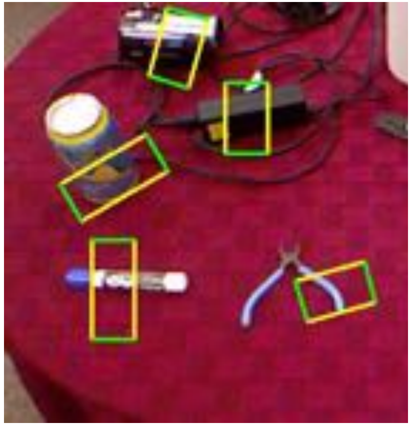
- Research Area
  - 3D computer vision, graphics, robotics

Part-level and Structural Shape Understanding

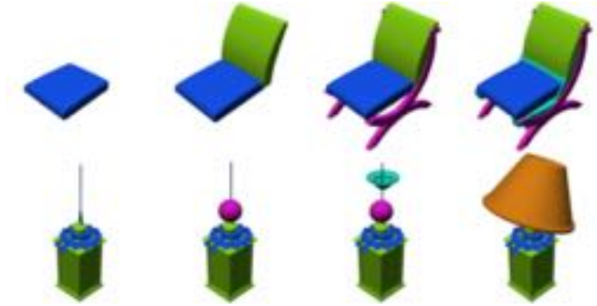


Part Segmentation

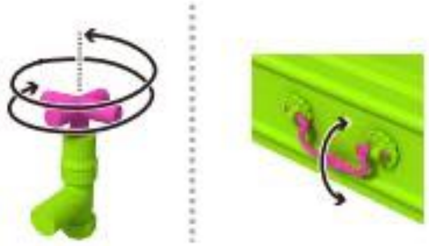
Structure Abstraction



Graspable Part



Shape Synthesis



Part Mobility

Part / Object State

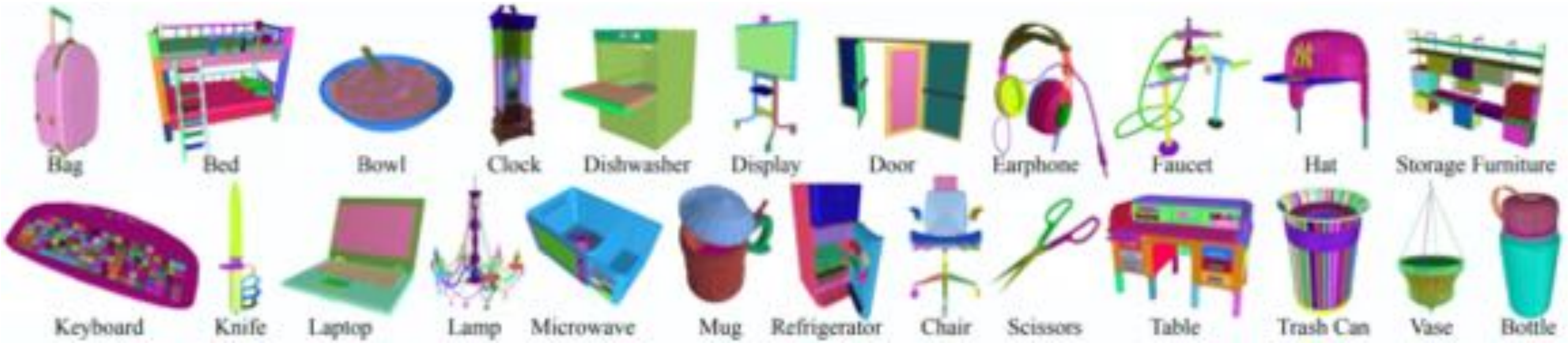
Content for Today

- PartNet (*CVPR 2019*)
  - A Large-scale Benchmark for Fine-grained and Hierarchical Part-level 3D Object Understanding

- StructureNet (*Siggraph Asia 2019*)
  - Hierarchical Graph Networks for 3D Shape Generation

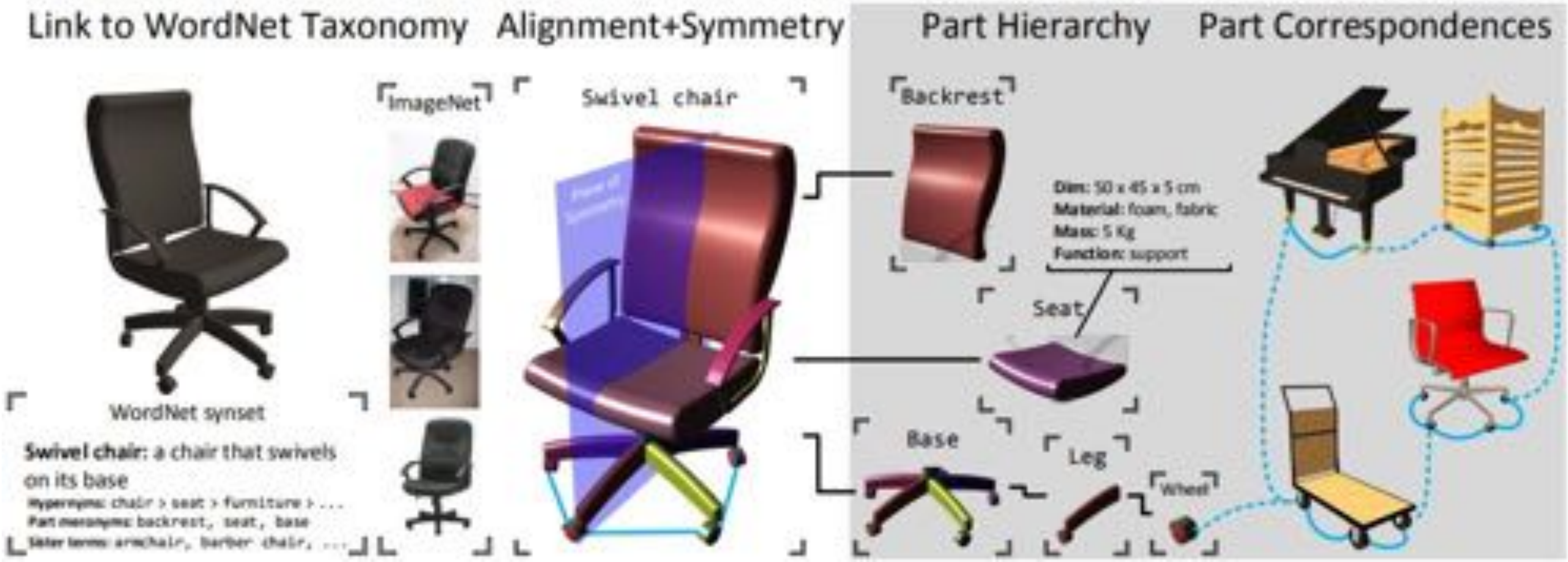
- StructEdit (*CVPR 2020*)
  - Learning Structural Shape Variations

PartNet: A Large-scale Benchmark for Fine-grained and Hierarchical Part-level 3D Object Understanding



Kaichun Mo, Shilin Zhu, Angel X.Chang, Li Yi, Subarna Tripathi,
Leonidas J. Guibas and Hao Su
CVPR 2019

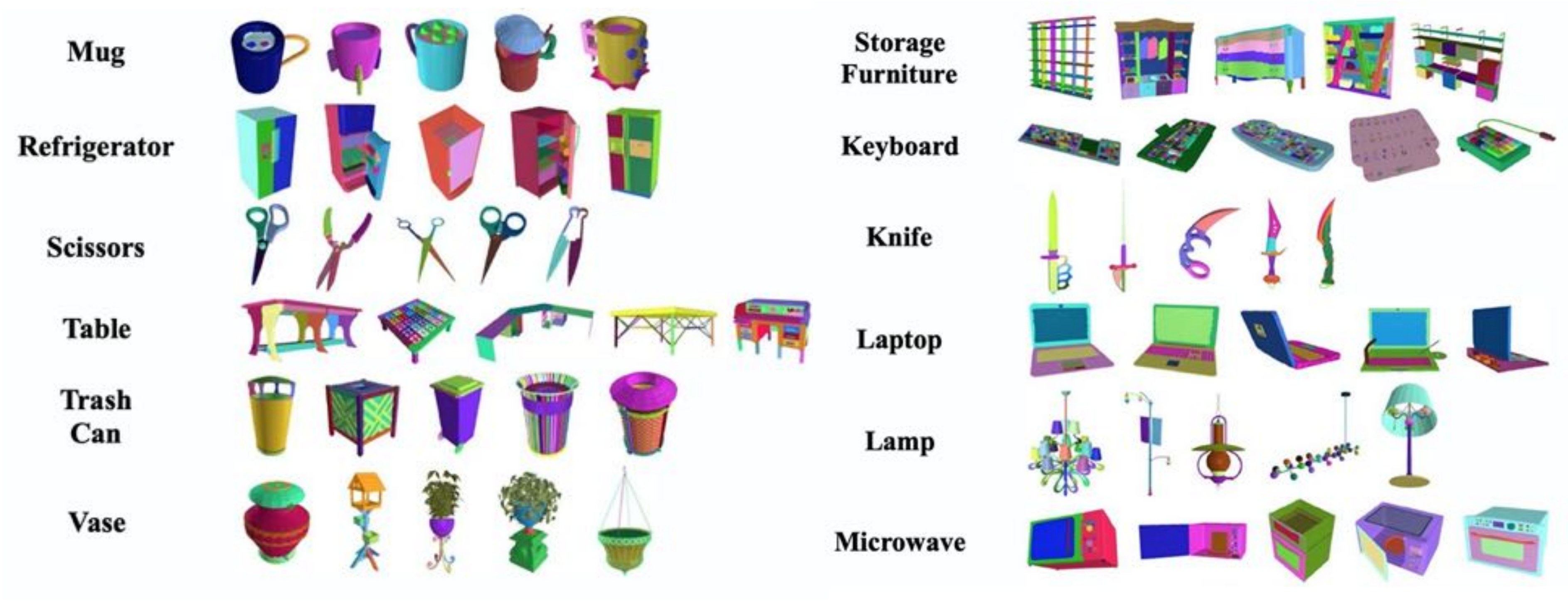
ShapeNet



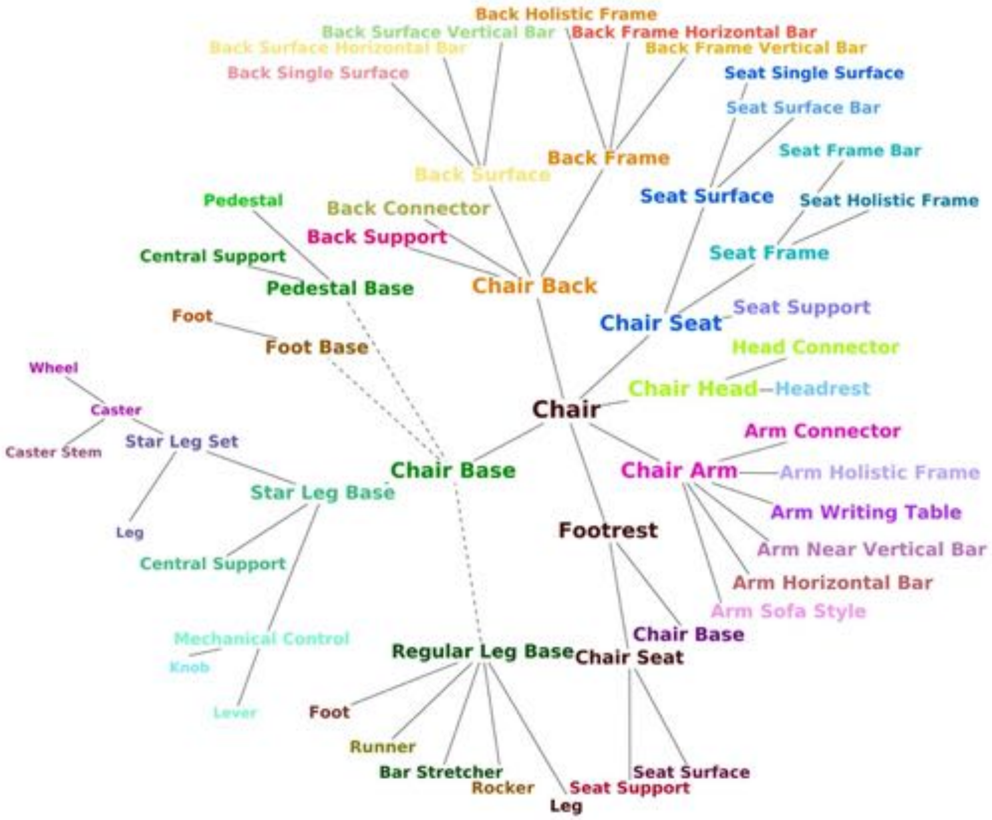
The contribution of PartNet

ShapeNet: An Information-Rich 3D Model Repository

Fine-grained and Instance-level Part Annotation



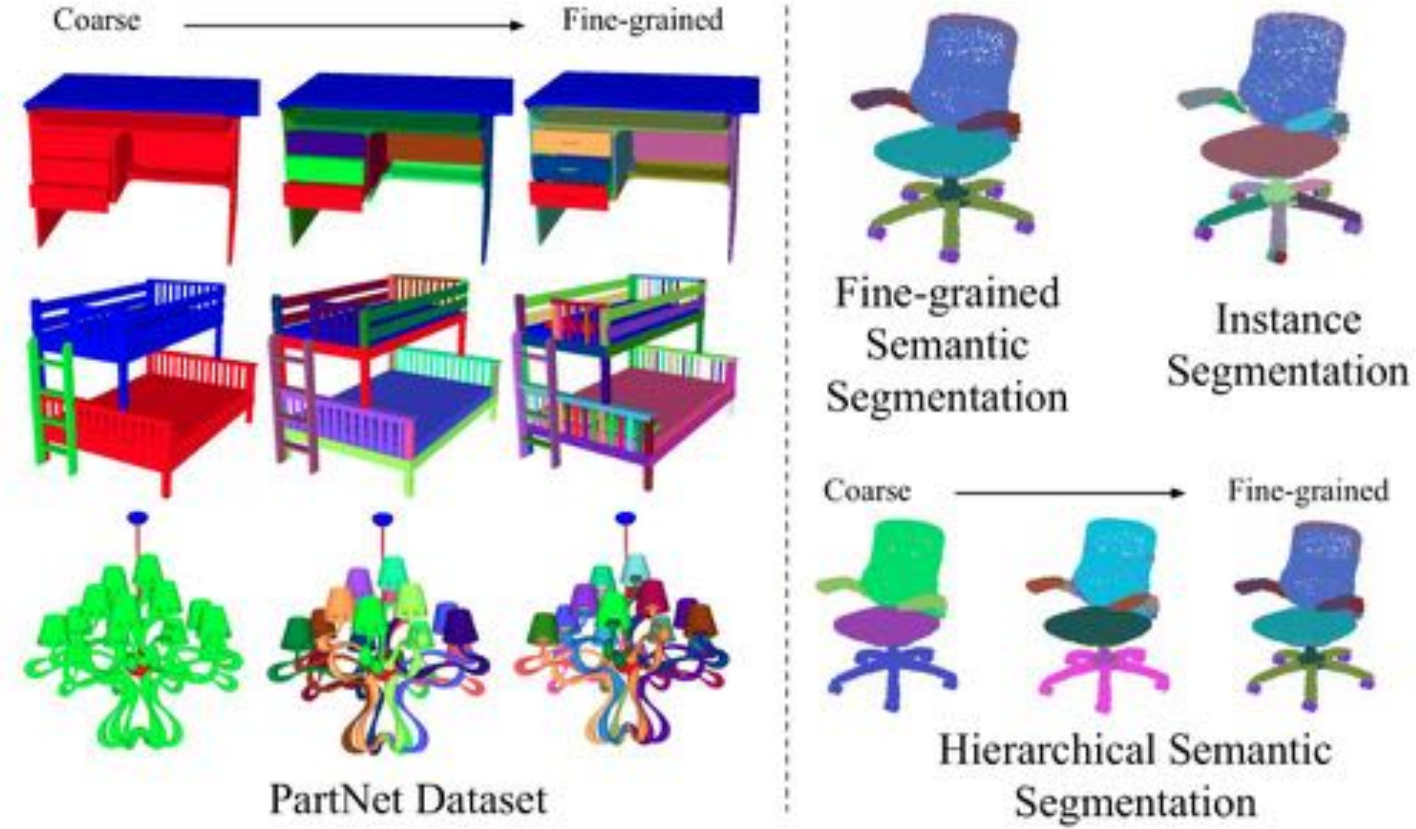
Hierarchical Part Annotation



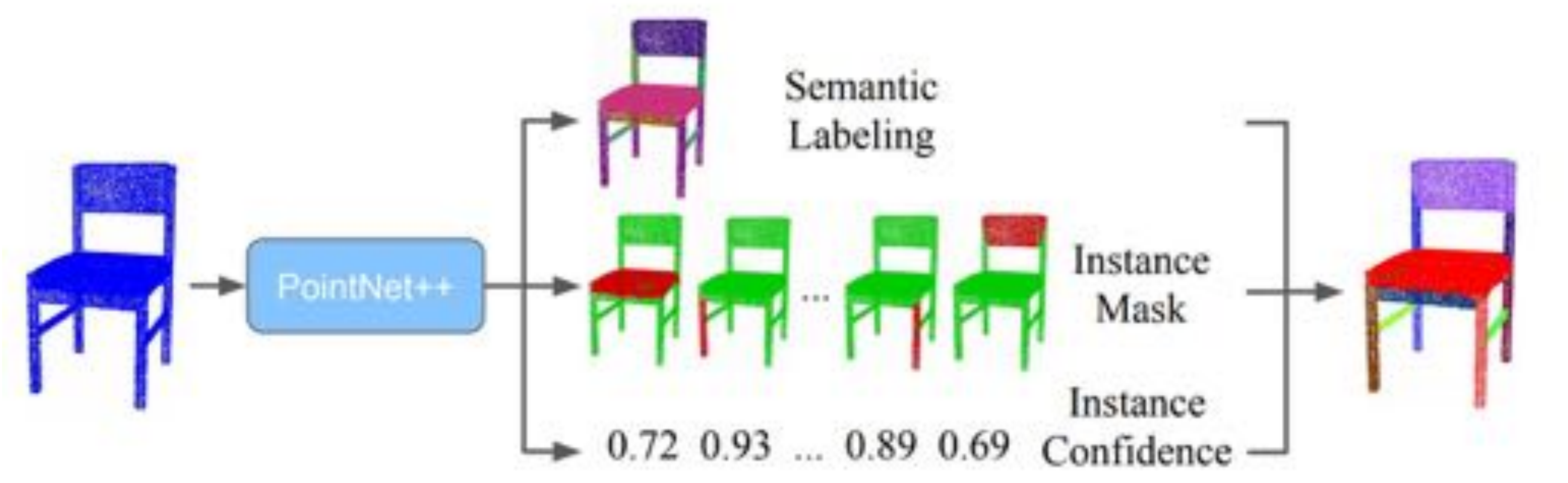
Dataset Statistics

- 24 Object Categories
- 26,671 Different Shapes
- 573,585 Different Parts
- Avg 18 Part/shape, Max 230

Benchmarking Part Segmentation Tasks



Part Instance Segmentation

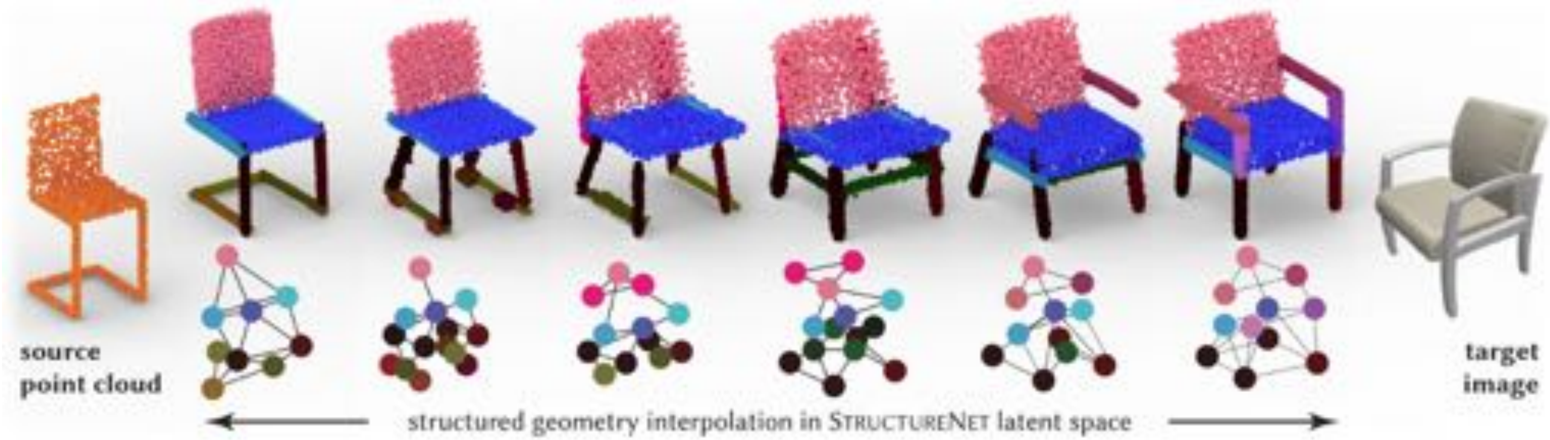


PartNet InsSeg Baseline

Part Instance Segmentation

	Avg
S1	55.7
S2	29.7
S3	29.5
Avg	46.8
O1	62.6
O2	37.4
O3	36.6
Avg	54.4

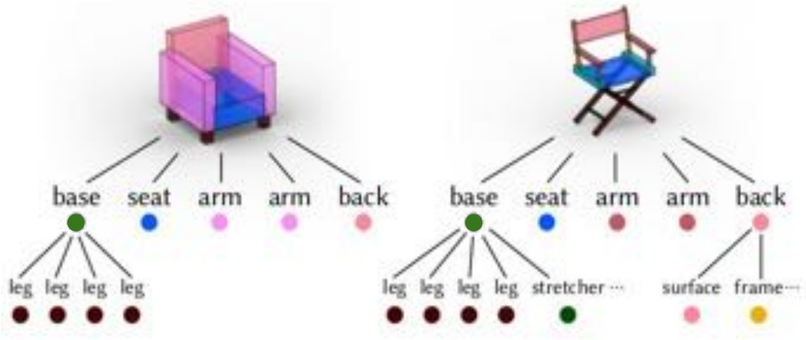
StructureNet: Hierarchical Graph Networks for 3D Shape Generation

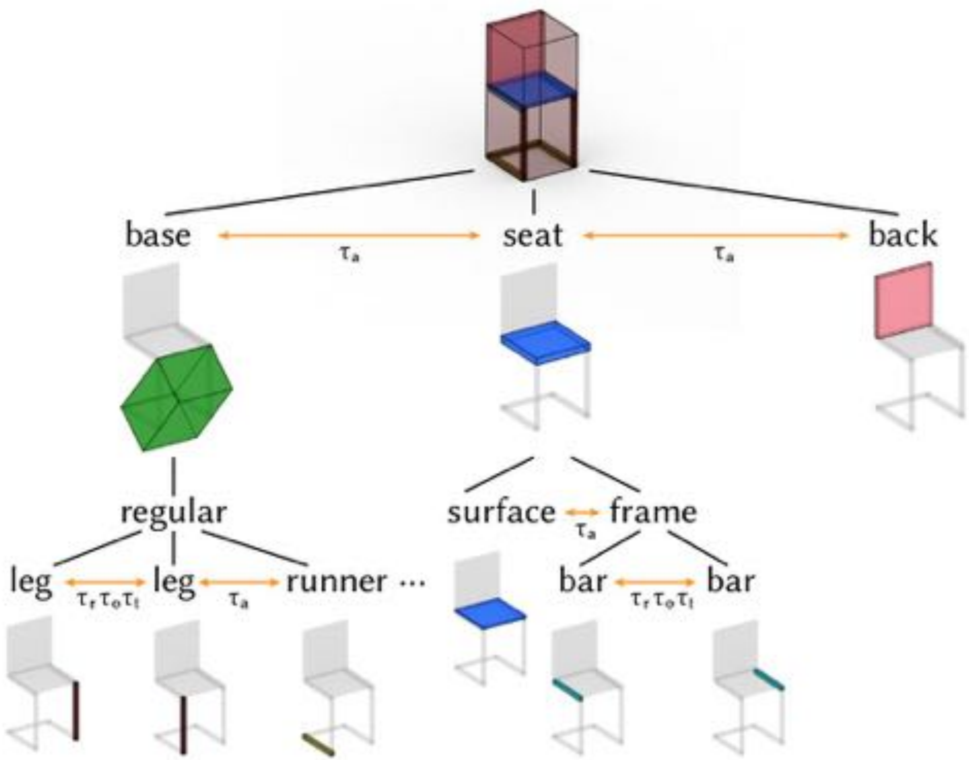


Kaichun Mo[*], Paul Guerrero[*], Li Yi, Hao Su,
Peter Wonka, Niloy Mitra, Leonidas Guibas

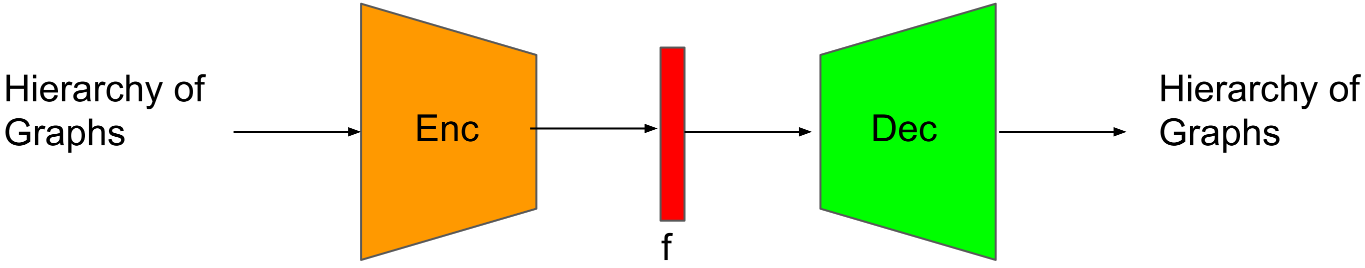
Siggraph Asia 2019

Consistent Hierarchy of Graphs



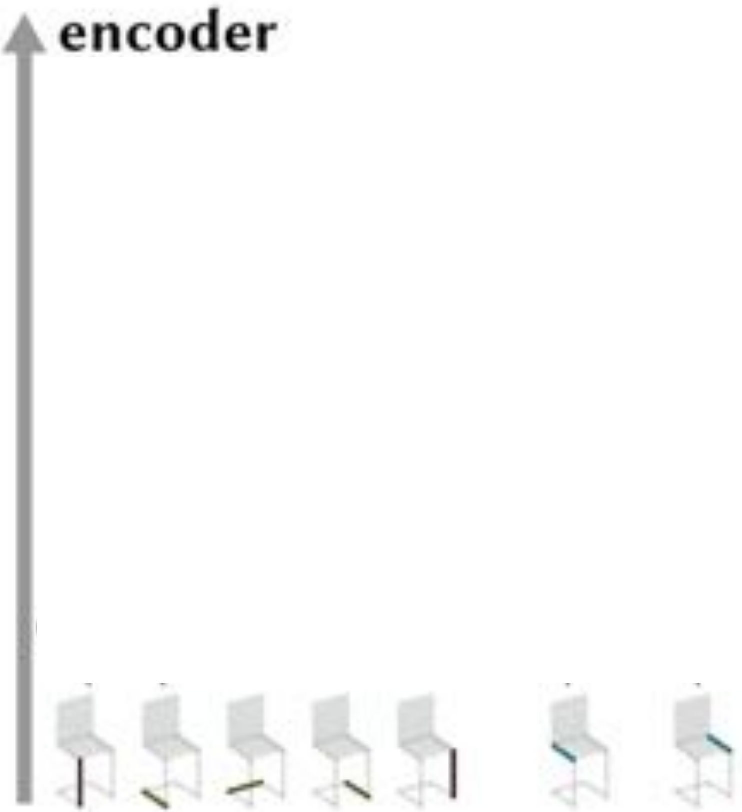


Overall Architecture

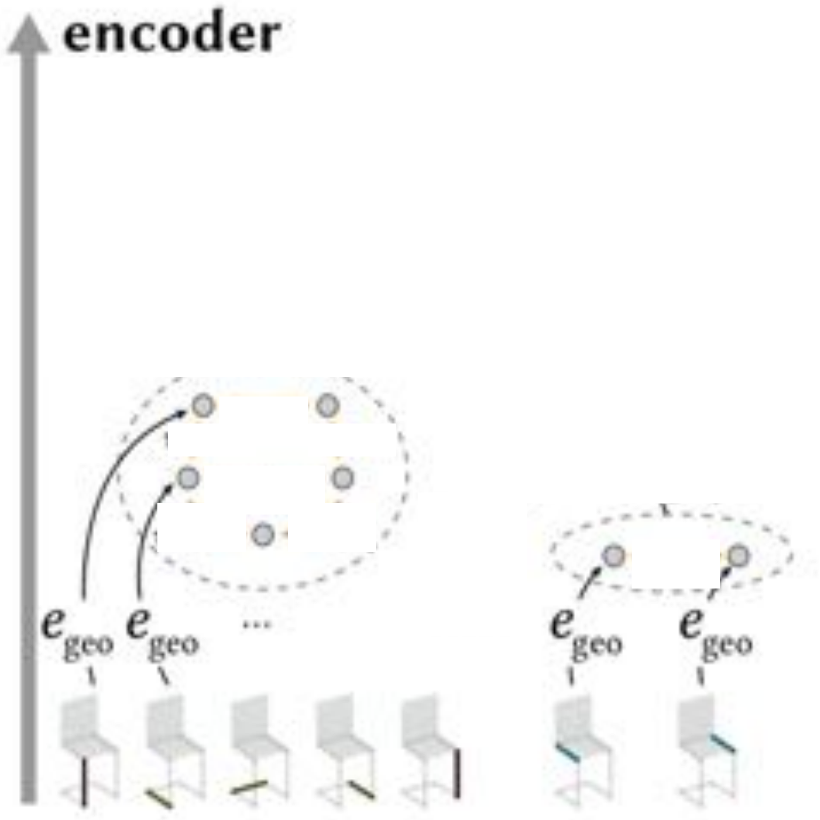


Variational Auto-Encoder (VAE)

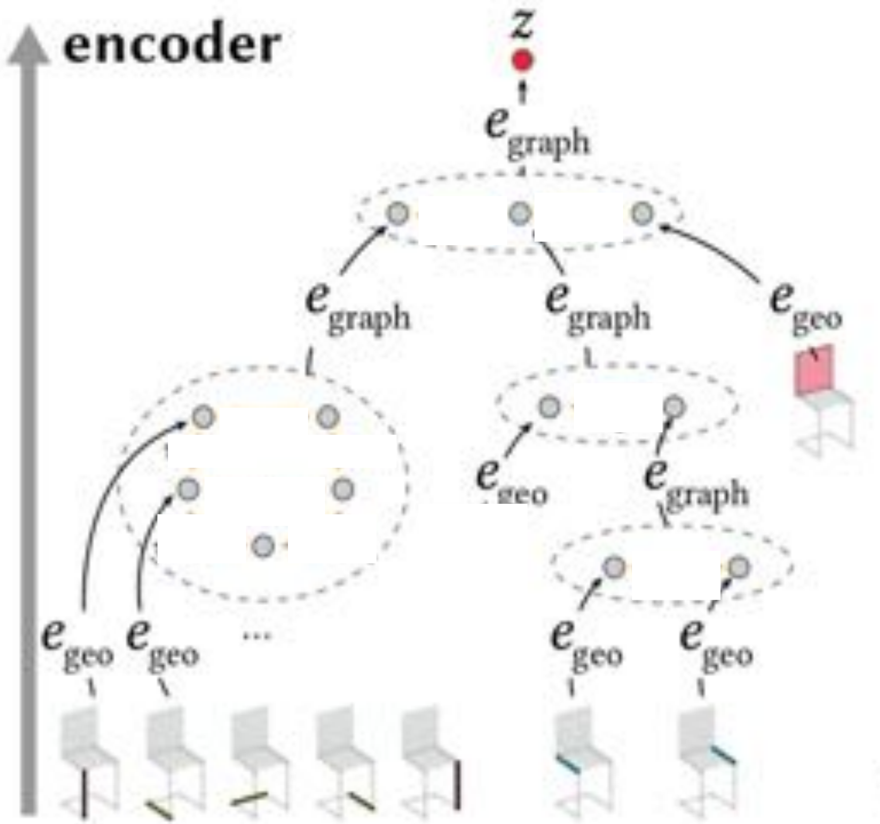
Hierarchical Graph Network as a VAE



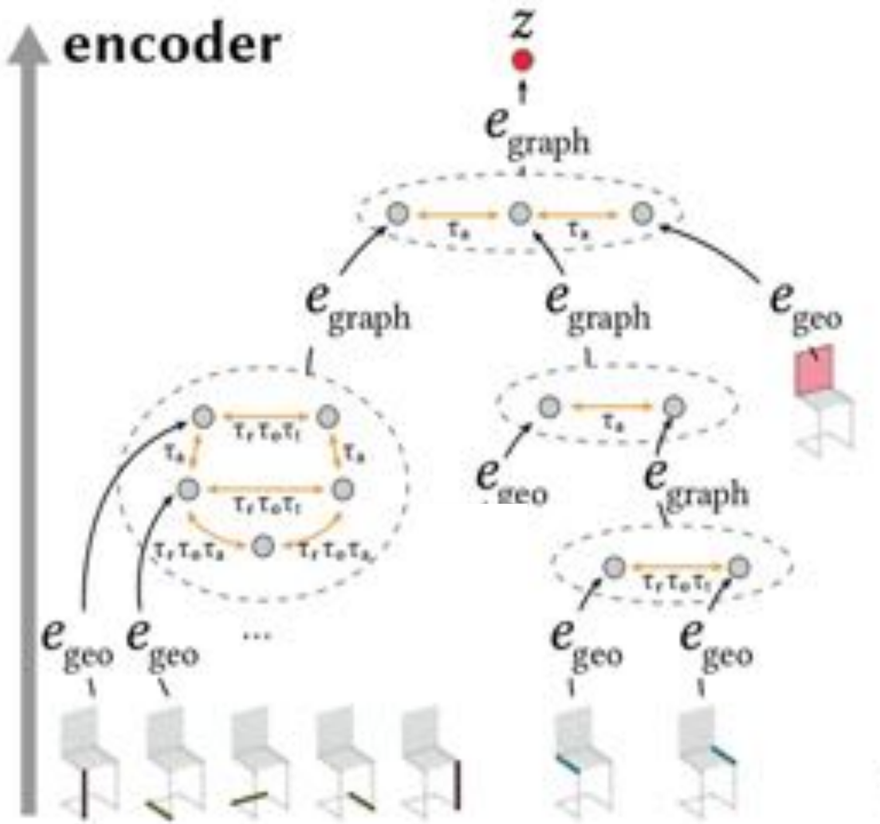
Hierarchical Graph Network as a VAE



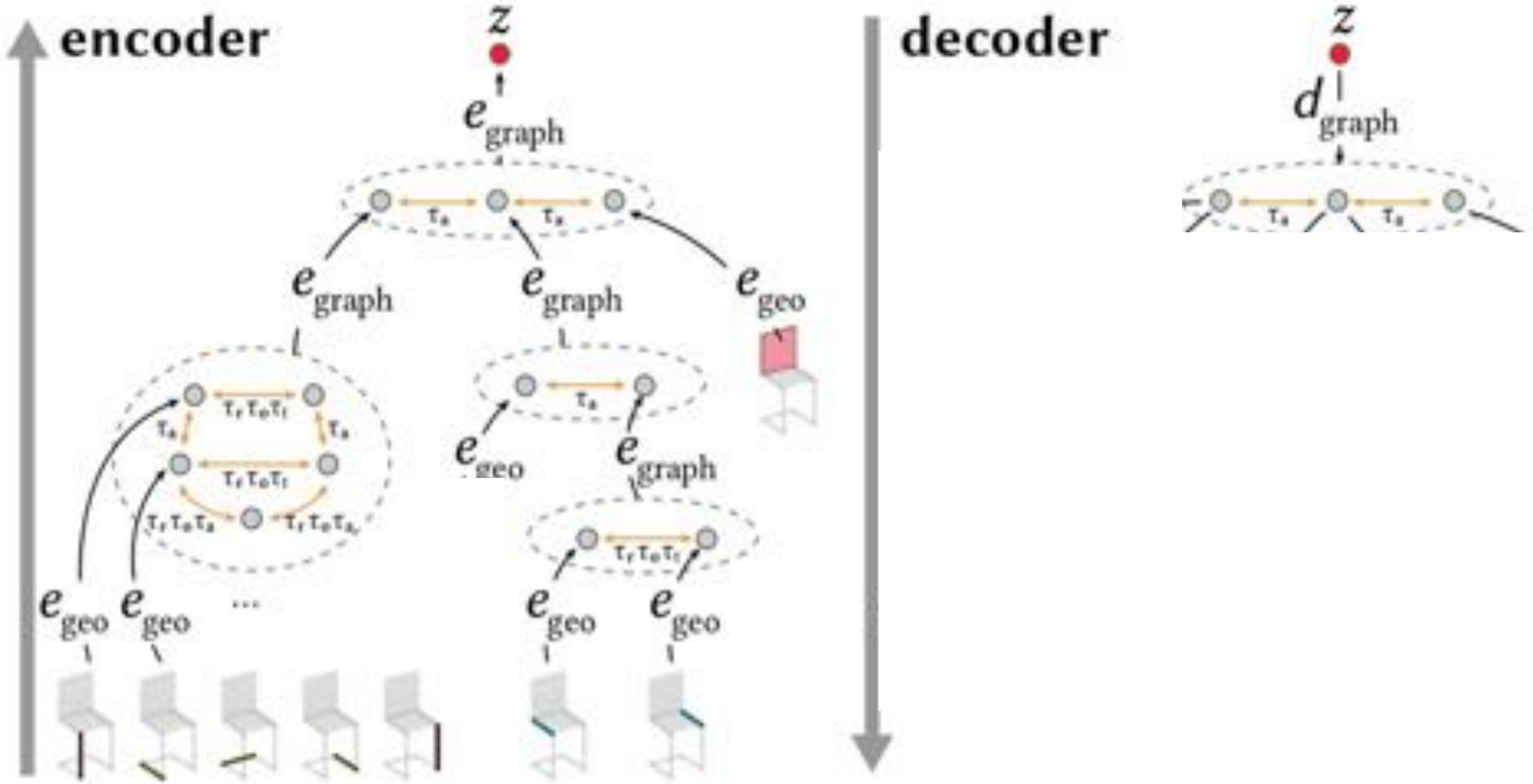
Hierarchical Graph Network as a VAE



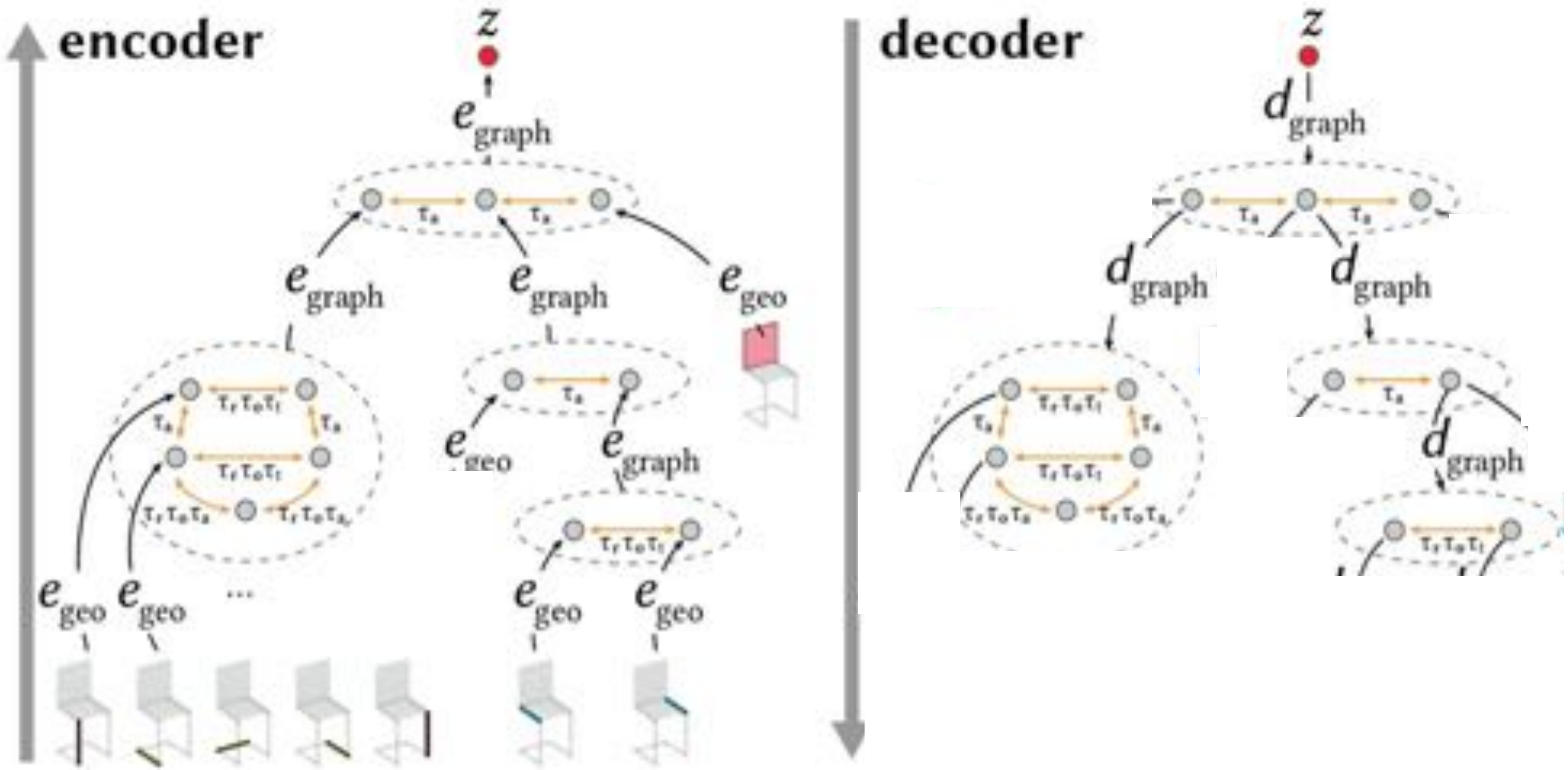
Hierarchical Graph Network as a VAE



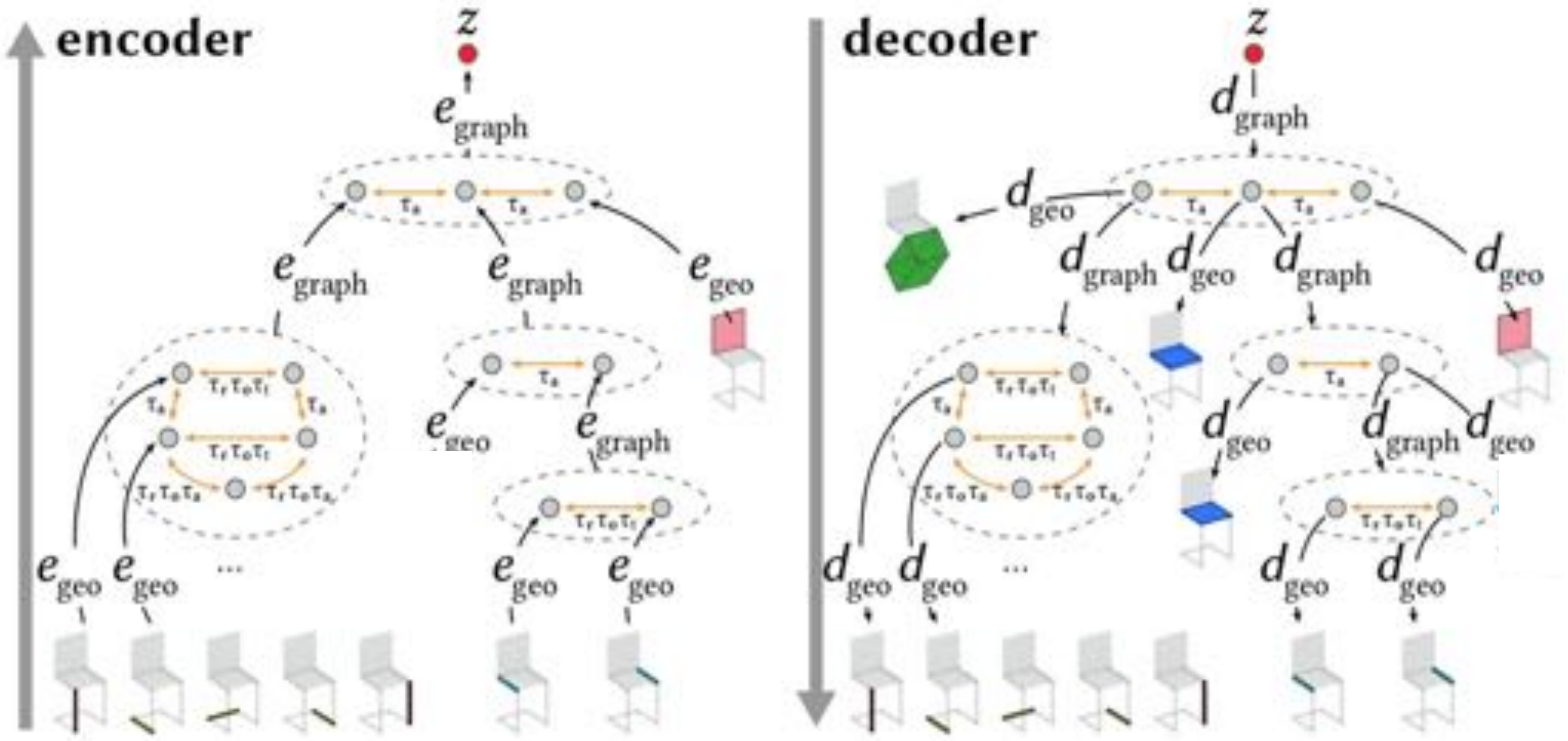
Hierarchical Graph Network as a VAE



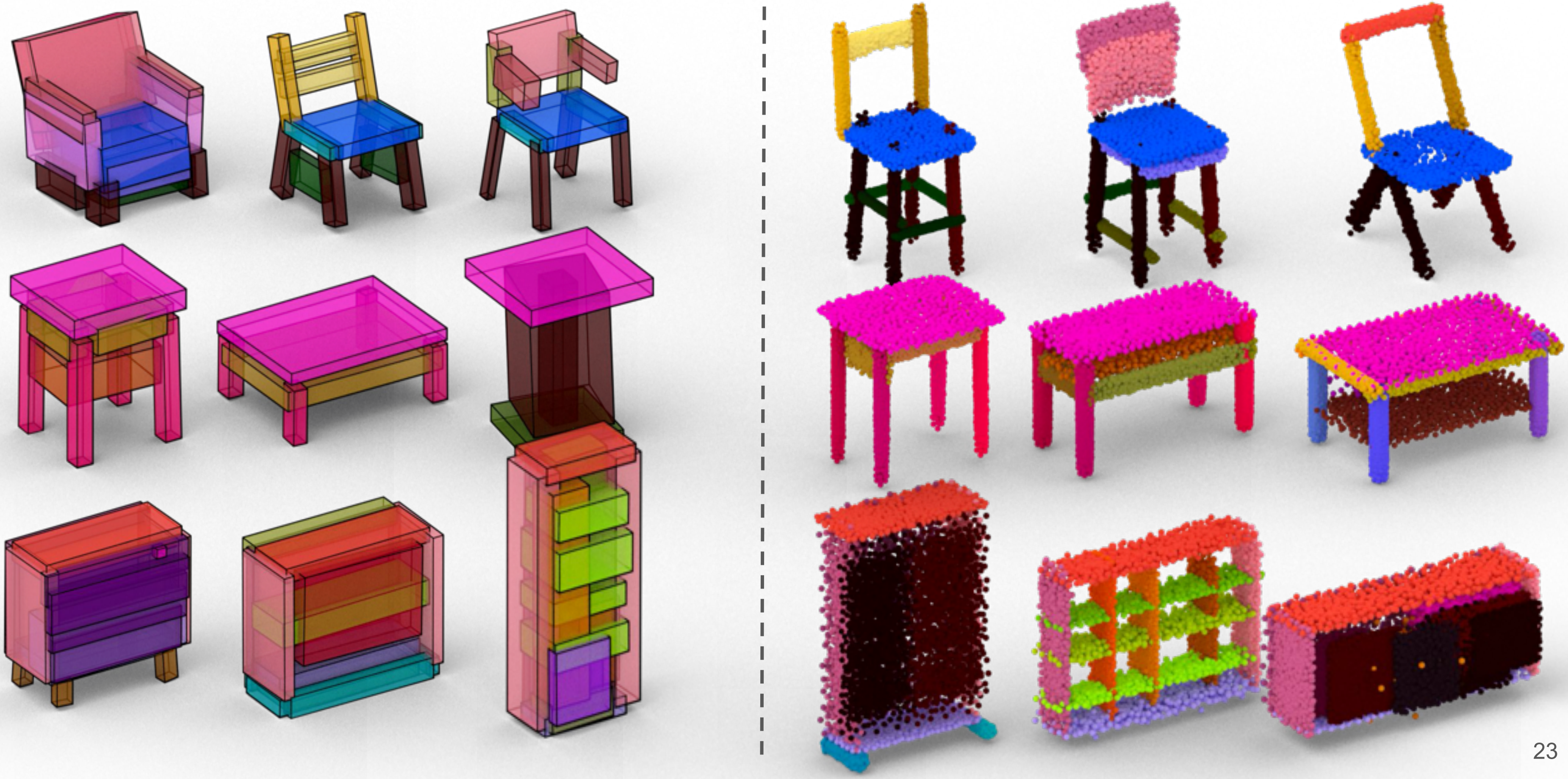
Hierarchical Graph Network as a VAE



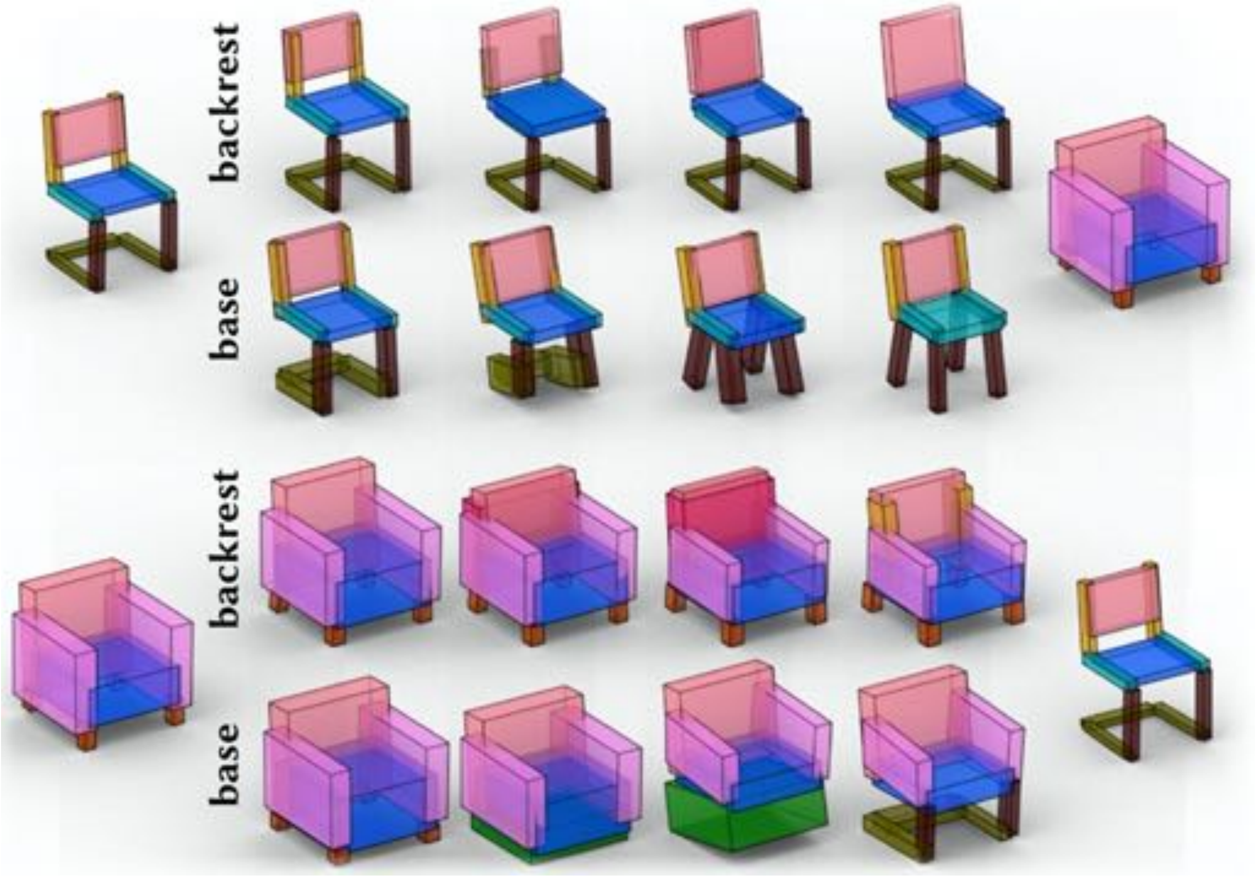
Hierarchical Graph Network as a VAE



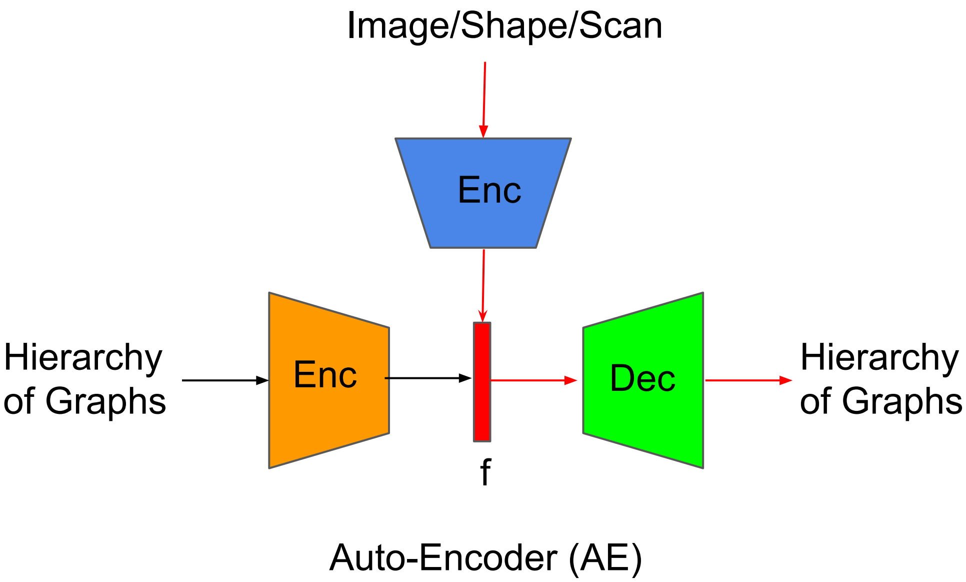
Results: Free Shape Generation

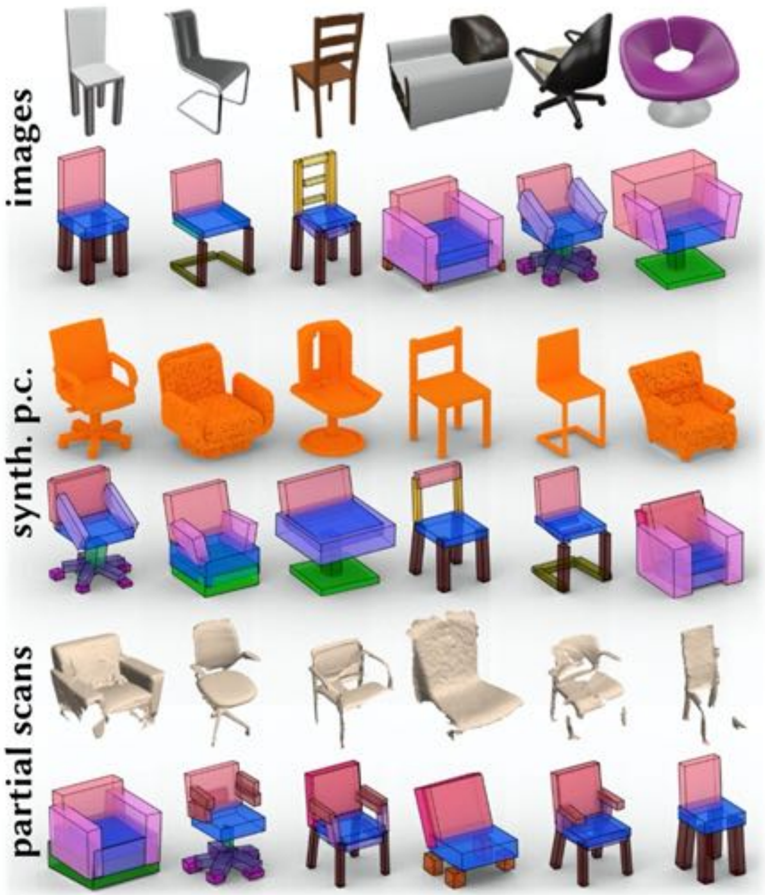


Results: Part Interpolation



Results: Shape Abstraction

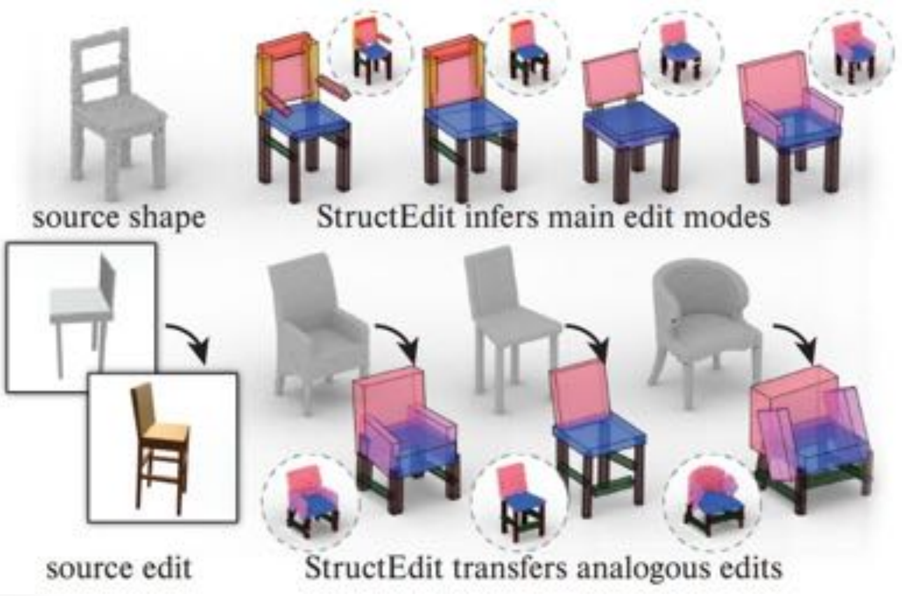




StructEdit: Learning Structural Shape Variations

Kaichun Mo[*], Paul Guerrero[*],
Li Yi, Hao Su,
Peter Wonka, Niloy Mitra,
Leonidas Guibas

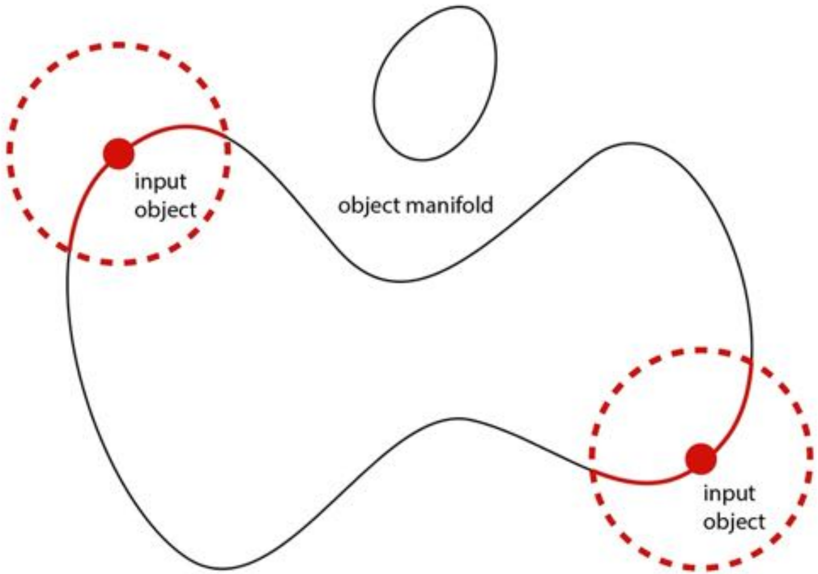
CVPR 2020



Learn to Embed Local Shape Neighborhoods

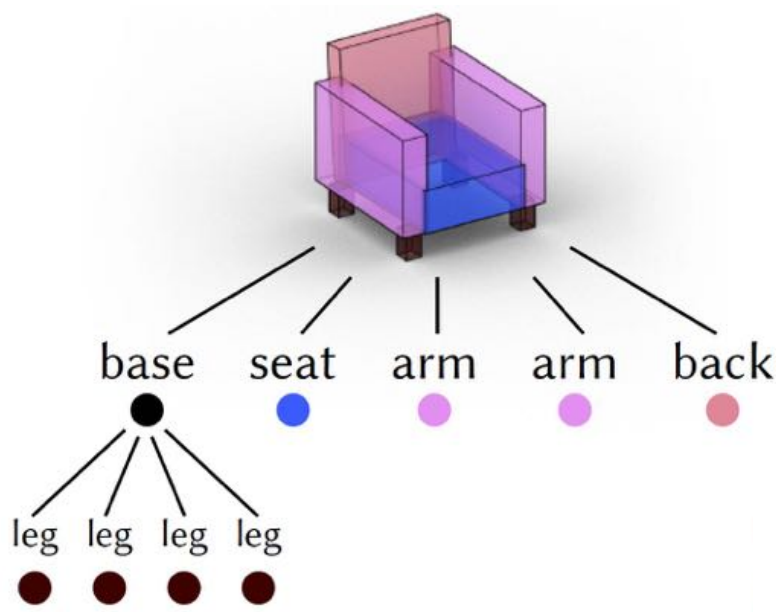
Learn a Structural Shape-diff Space, which is

- **Unified**: learn one space that can be applied for any input shape
- **Specialized**: suggest different plausible shape-diff's for different input shapes
- **Coherent**: suggest similar plausible shape-diff's for similar input shapes

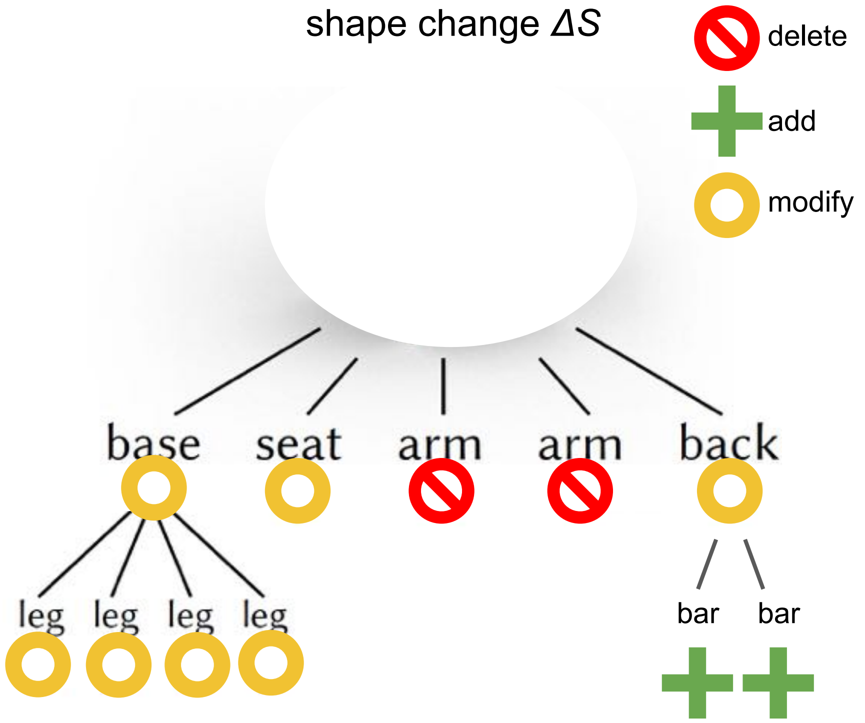


Shape Difference (Structure + Deformation)

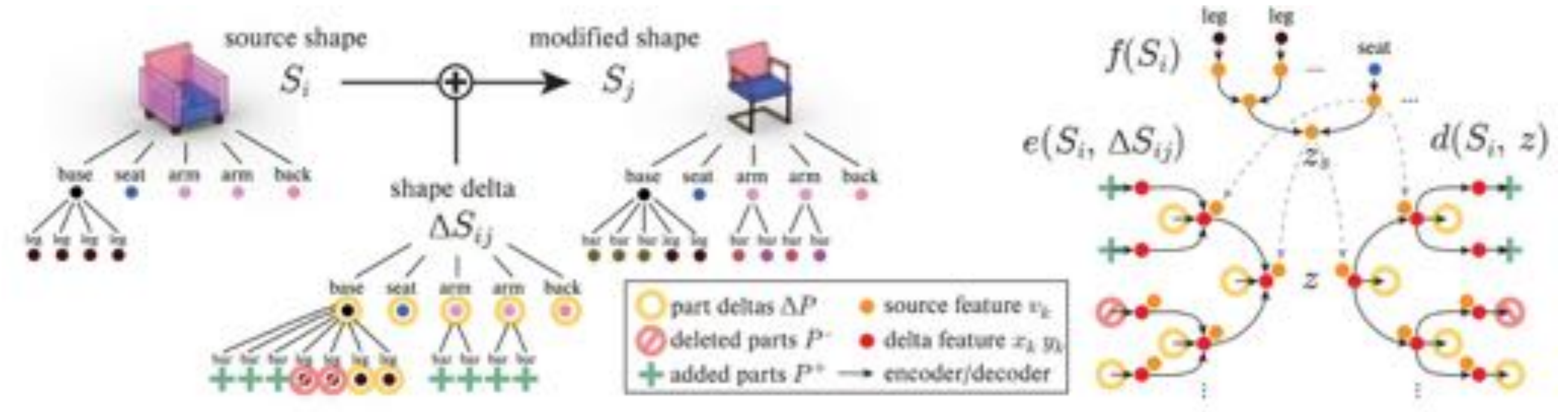
input shape S



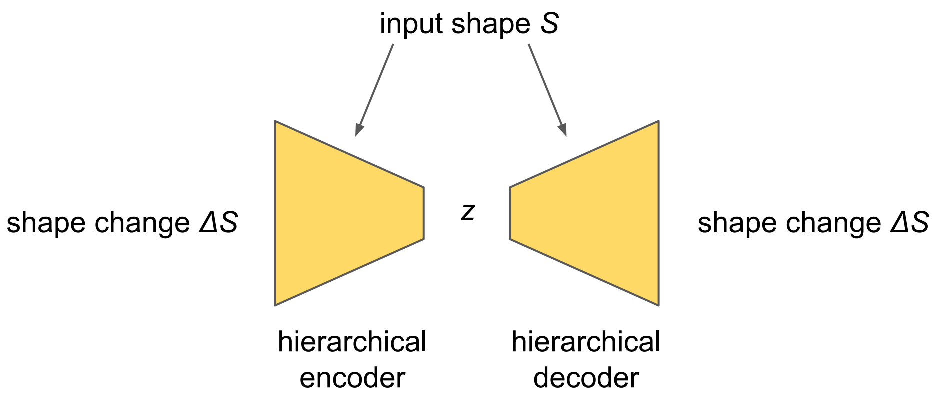
shape change ΔS



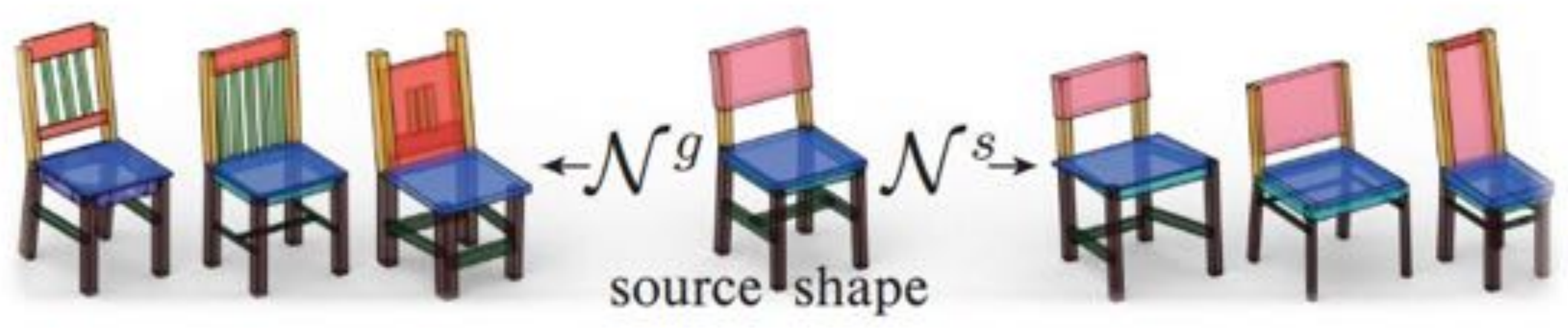
Network Architecture



StructEdit (cVAE)



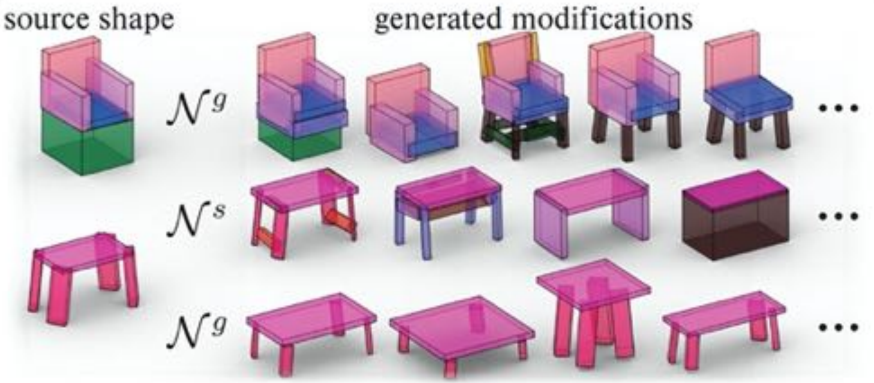
Support Different Types of Shape Neighborhood

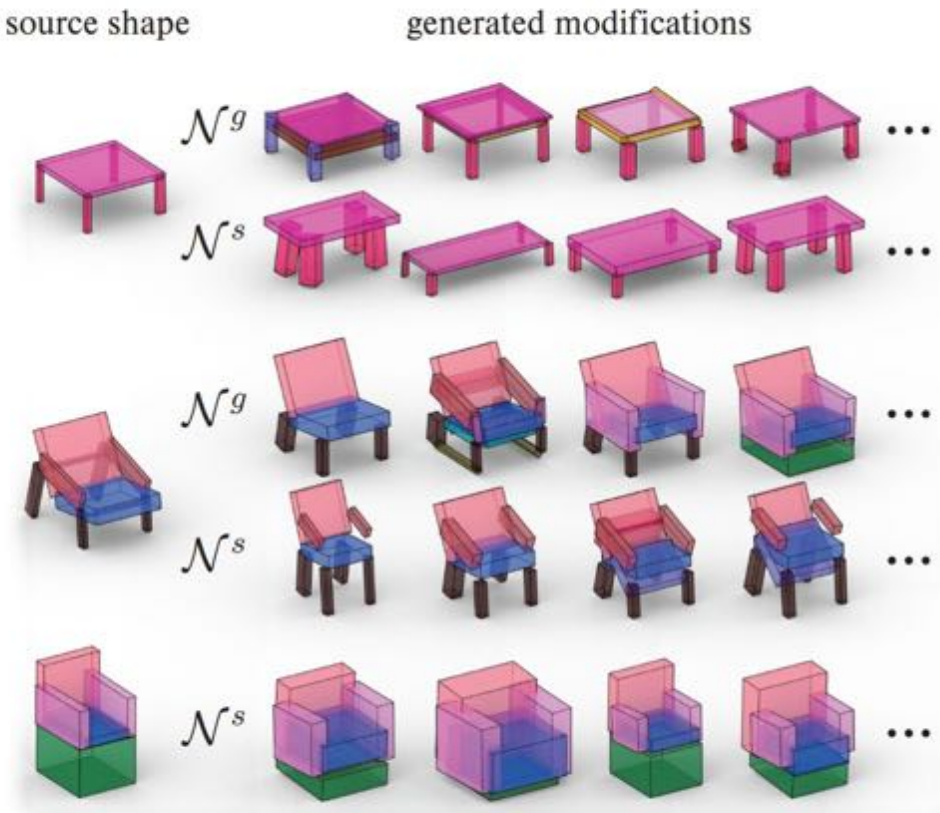


Or, any other neighborhood types defined in the data

Results: Edit Generation

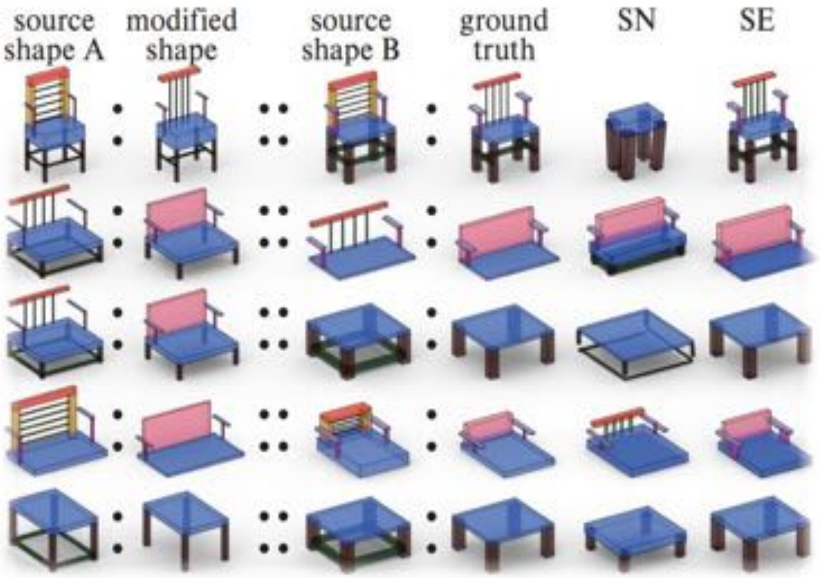
		$\mathcal{N}^g$				$\mathcal{N}^s$			
		chair	table	furn.	avg.	chair	table	furn.	avg.
E_{qc}^{geo}	ID	1.822	1.763	1.684	1.756	1.629	1.479	1.446	1.518
	$SN_{0.2}$	1.760	2.076	1.626	1.821	1.308	1.208	1.243	1.253
	$SN_{0.5}$	1.722	2.068	1.558	1.783	1.241	1.103	1.135	1.160
	$SN_{1.0}$	1.768	2.189	**1.554**	1.837	1.232	1.057	1.017	1.102
	SE	**1.593**	**1.655**	1.561	**1.603**	**1.218**	**1.000**	**1.015**	**1.078**
E_{qc}^{st}	ID	1.281	1.215	1.288	1.261	1.437	1.303	1.442	1.394
	$SN_{0.2}$	1.081	0.878	1.015	0.991	1.466	3.484	1.414	2.121
	$SN_{0.5}$	0.871	0.729	0.873	0.824	1.373	3.300	1.204	1.959
	$SN_{1.0}$	0.751	0.667	**0.726**	0.715	1.763	3.622	1.167	2.184
	SE	**0.559**	**0.524**	0.741	**0.608**	**0.609**	**0.451**	**0.676**	**0.579**

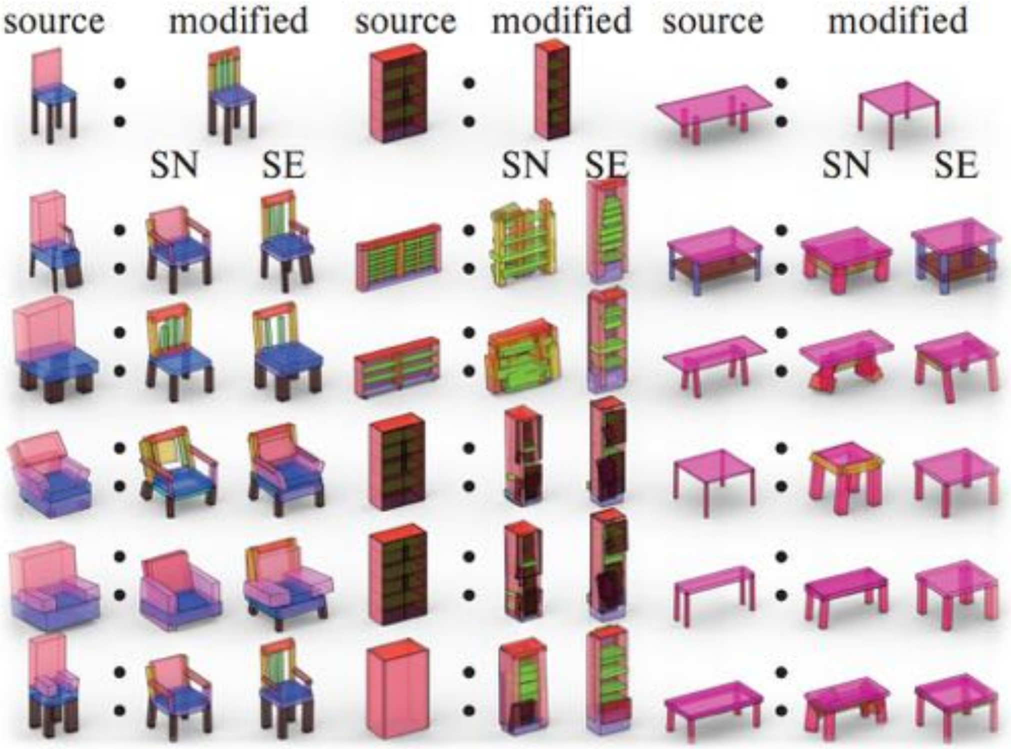




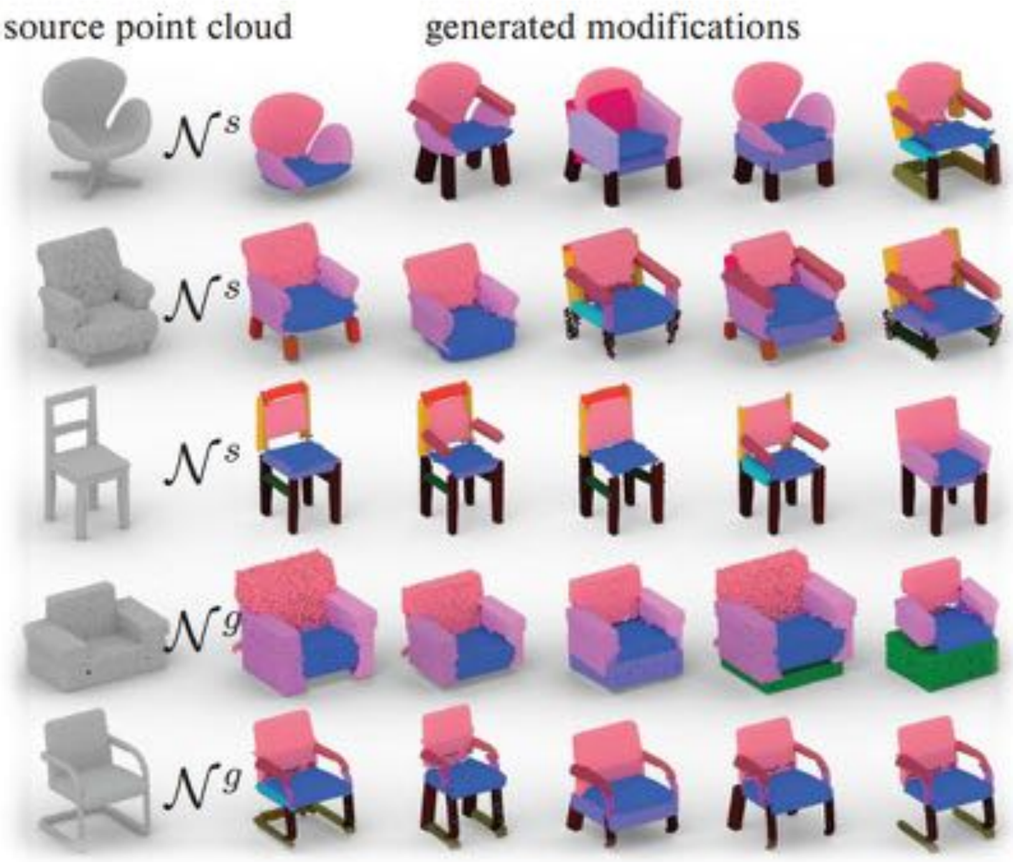
Results: Edit Transfer

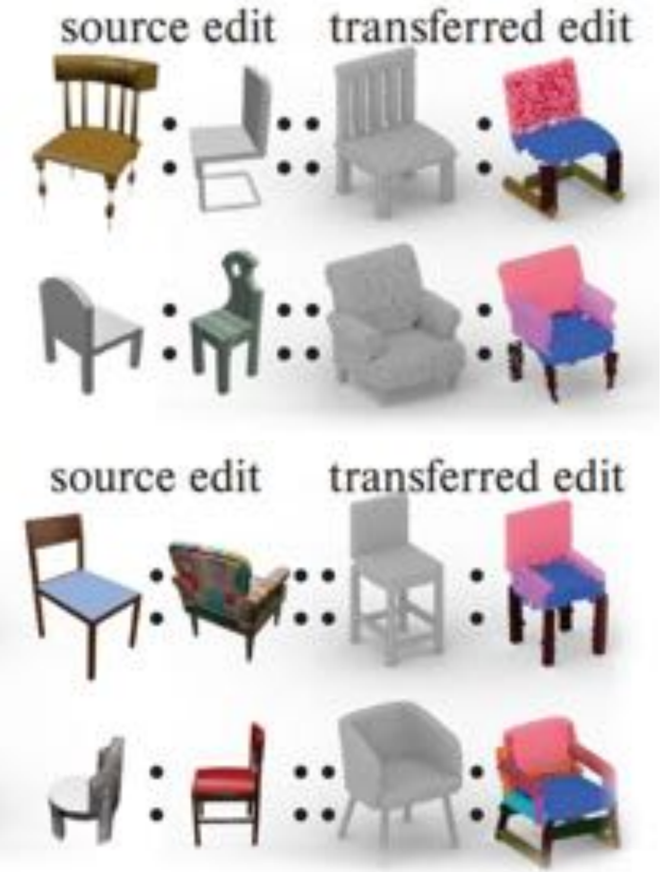
$\mathcal{N}$		chair	sofa	stool	c. → s.	c. → st.	avg.
E_t^{geo}	Identity	1.002	0.938	0.892	0.892	0.938	0.932
	StructureNet	0.868	0.764	0.721	0.888	1.307	0.910
	StructEdit	**0.586**	**0.566**	**0.599**	**0.572**	**0.698**	**0.604**
E_t^{st}	Identity	0.941	1.328	0.333	0.333	1.328	0.853
	StructureNet	0.208	0.161	0.025	0.671	0.871	0.387
	StructEdit	**0.005**	**0.001**	**0.003**	**0.002**	**0.123**	**0.027**



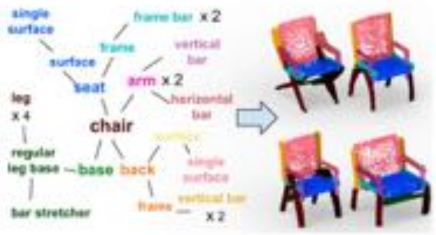


Applications





More Recent Works

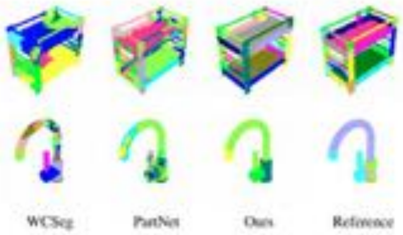


PT2PC: Learning to Generate 3D Point Cloud Shapes from Part Tree Conditions

Kaichun Mo, He Wang, Xinchen Yan and Leonidas J. Guibas
arXiv:2003.08624 [cs.CV]

This paper investigates the novel problem of generating 3D shape point cloud geometry from a symbolic part tree representation. In order to learn such a conditional shape generation procedure in an end-to-end fashion, we propose a conditional GAN "part tree"-to-"point cloud" model (PT2PC) that disentangles the structural and geometric factors.

[Paper] [Project] [Bibtex]

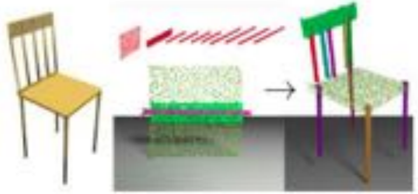


Learning to Group: A Bottom-Up Framework for 3D Part Discovery in Unseen Categories

Tiange Luo, **Kaichun Mo**, Zhiao Huang, Jiarui Xu, Siyu Hu, Liwei Wang and Hao Su
ICLR 2020

We address the problem of learning to discover 3D parts for objects in unseen categories under the zero-shot learning setting. We propose a learning-based iterative grouping framework which learns a grouping policy to progressively merge small part proposals into bigger ones in a bottom-up fashion. On PartNet, we demonstrate that our method can transfer knowledge of parts learned from 3 training categories to 21 unseen categories.

[Paper] [Project] [Video] [Bibtex]



Learning 3D Part Assembly from a Single Image

Yichen Li*, **Kaichun Mo***, Lin Shao, Minhyuk Sung and Leonidas J. Guibas
arXiv:2003.09754 [cs.CV]

We introduce a novel problem, single-image-guided 3D part assembly, that assembles 3D shapes from parts given a complete set of part point cloud scans and a single 2D image depicting the object. The task is motivated by the robotic assembly setting and the estimated per-part poses serve as a vision-based initialization before robotic planning and control components.

[Paper] [Bibtex]

Part-level and Structural Understanding for 3D Shape Perception, Synthesis and Editing

Kaichun Mo
Stanford University

Thank you

Questions?

Leonidas Guibas Laboratory
Geometric Computing