Deep Polycuboid Fitting for Compact 3D Representation of Indoor Scenes

3DV 2025 International Conference on 3D Vision
Singapore
March 25-28, 2025
k-NN

Gahye Lee Hyejeong Yoon Jungeon Kim Seungyong Lee
{gahye0509, hjyoon02, jungeonkim, leesy}@postech.ac.kr

POSTECH COMPUTER GRAPHICS

Motivation and Contributions

Structured and interpretable 3D scene representation via **polycuboid** meshes, ensuring compactness and geometric coherence

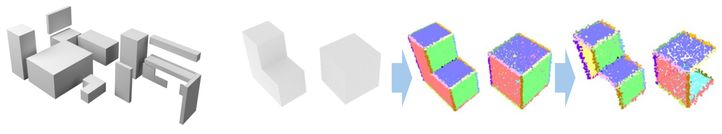
- Deep learning-based polycuboid fitting framework using **cuboid geometric priors** defined by six face types and fourteen spatial relationships

- Detecting **polycuboid faces** using a Transformer network and aggregating them into polycuboid instances using **spatial relationships** inferred via Graph Convolutional Network (GCN)

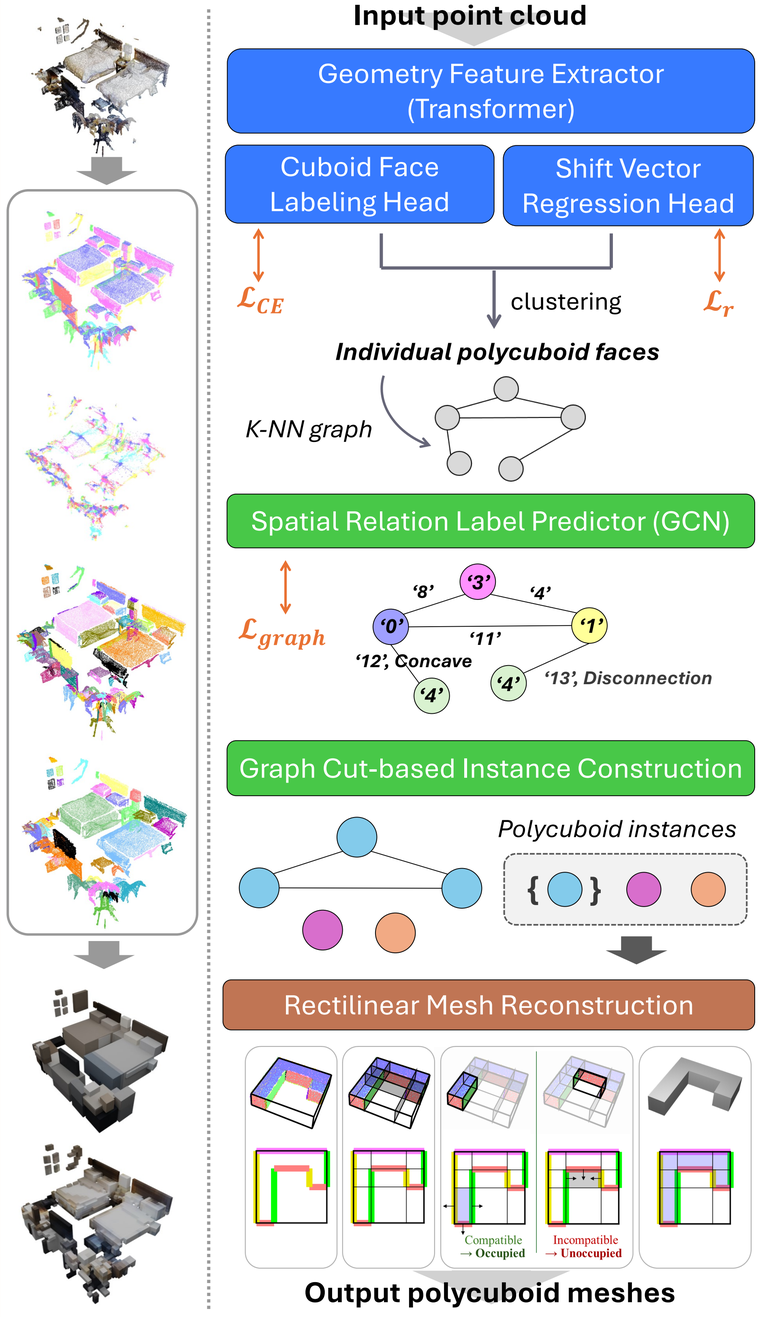
Dataset

Synthetic polycuboid dataset designed for robust generalization to real-world indoor scenes

- **Diverse spatial configurations** of cuboid and polycuboid meshes with 3D points sampled from surfaces
- **Noise and occlusion** added to mimic real-world scanning conditions
- Used for training **polycuboid face detection** and **spatial relation label prediction** networks



Deep Polycuboid Fitting



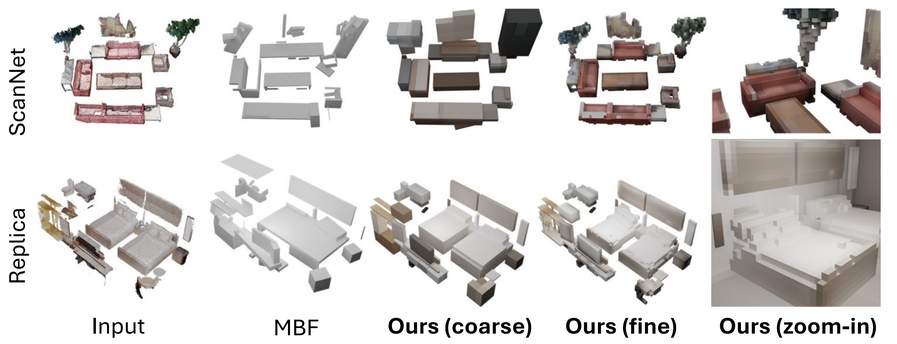
Results

▶ ***Quantitative and qualitative results on real-world dataset***

- Comparison with MBF [1], which fits cuboids to the point cloud, on real world datasets: ScanNet [2] and Replica [3]

1. Chamfer Distance, 2. Intersection over Union

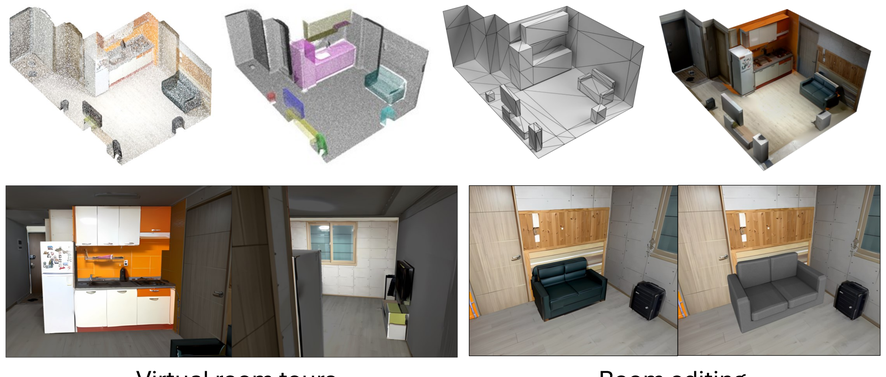
	Reconstruction fidelity (CD[1])		Object-level representation (IoU[2])			
	ScanNet	Replica	IoU25		Iou50	
			Recall	Precision	Recall	Precision
MBF	0.066	0.078	0.41	0.11	0.11	0.03
Ours	**0.040**	**0.044**	**0.64**	**0.15**	**0.43**	**0.10**



- Improved reconstruction quality with plausible object-level decomposition from noisy input data

▶ ***Potential applications: virtual room tours and room editing***

- Concise and editable living room reconstruction from iPhone-scanned data, combined with off-the-shelf texture mapping



[1] Ramamonjisoa et al., "MonteBoxFinder: Detecting and filtering primitives to fit a noisy point cloud", *in Proc. ECCV, Part XXVIII*, 161–177, 2022
[2] A. Dai et al., "ScanNet: Richly-annotated 3D reconstructions of indoor scenes", in *Proc. IEEE CVPR*, pp. 5828–5839, 2017
[3] Straub et al., "The Replica dataset: A digital replica of indoor spaces," *arXiv preprint arXiv:1906.05797*, 2019