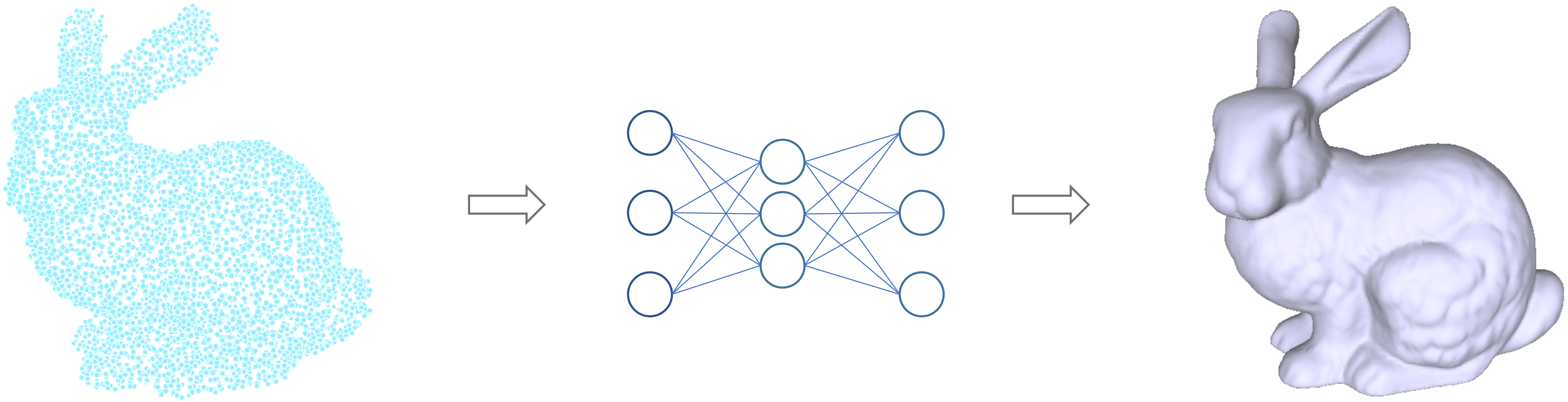
Dual Octree Graph Networks for Learning Adaptive Volumetric Shape Representations

Peng-Shuai Wang Yang Liu Xin Tong

Microsoft Research Asia

ACM Transactions on Graphics (SIGGRAPH), 2022

Learning to Reconstruct Continuous Surfaces



Point cloud

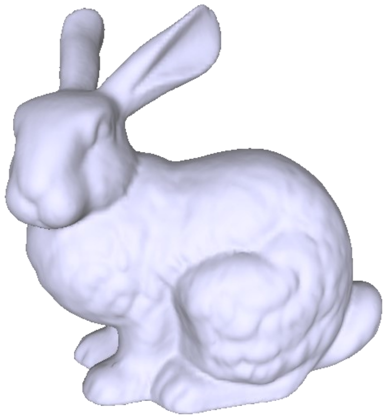
Mesh

Key Challenges

- Shape representation

  Discrete ↔ Continuous

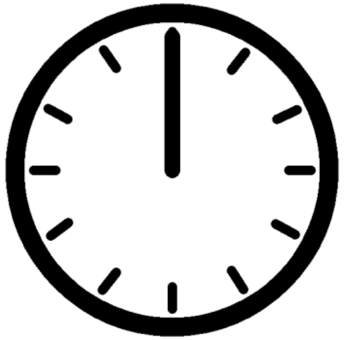
Point cloud



Mesh

- Network

  Efficient & Effective

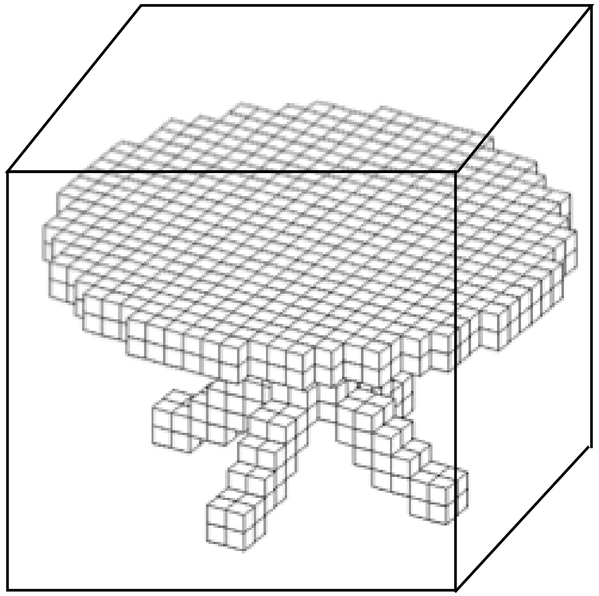


Time cost

Performance

Full-Voxel-based CNNs

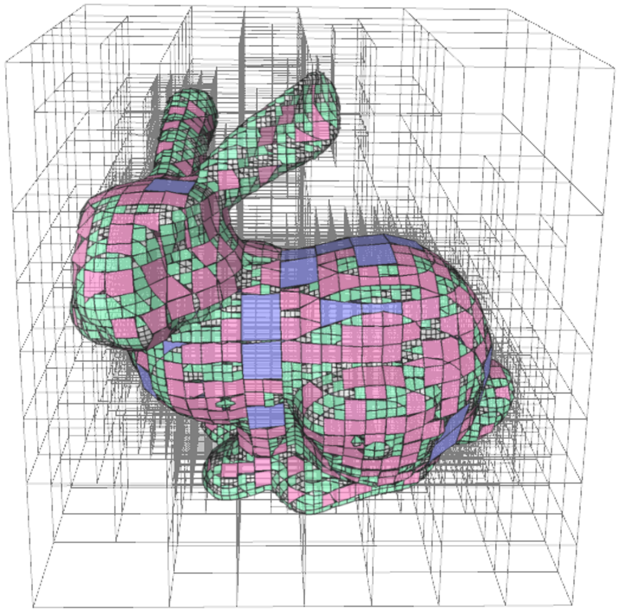
- Related work: [Wu et al. 2016; Choy et al. 2016; Dai et al. 2017]
- ✔ Natural extension of 2D CNNs
- ✘ Low efficiency



3D ShapeNet
[Wu et al. 2015]

Sparse-Voxel-based CNNs

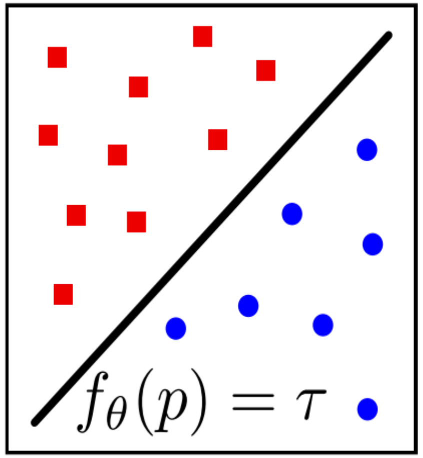
- Related work: [Wang et al. 2017; Graham et al. 2018; Shao et al. 2018;]
- ✔ High efficiency
- ✘ Hard to produce continuous surfaces



Adaptive O-CNN
[Wang et al. 2018]

Coordinate-based MLPs

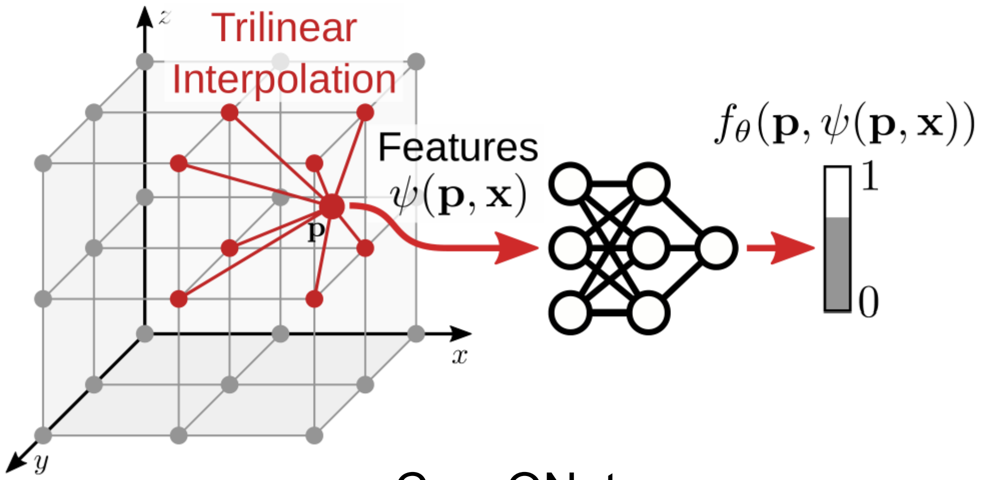
- Related work: [Chen et al. 2019; Mescheder et al. 2019; Park et al. 2019]
- ✓ A compact and continuous shape representation
- ✗ Hard to design encoding networks with MLPs



Occupancy Networks
[Mescheder et al. 2019]

Hybrid Methods

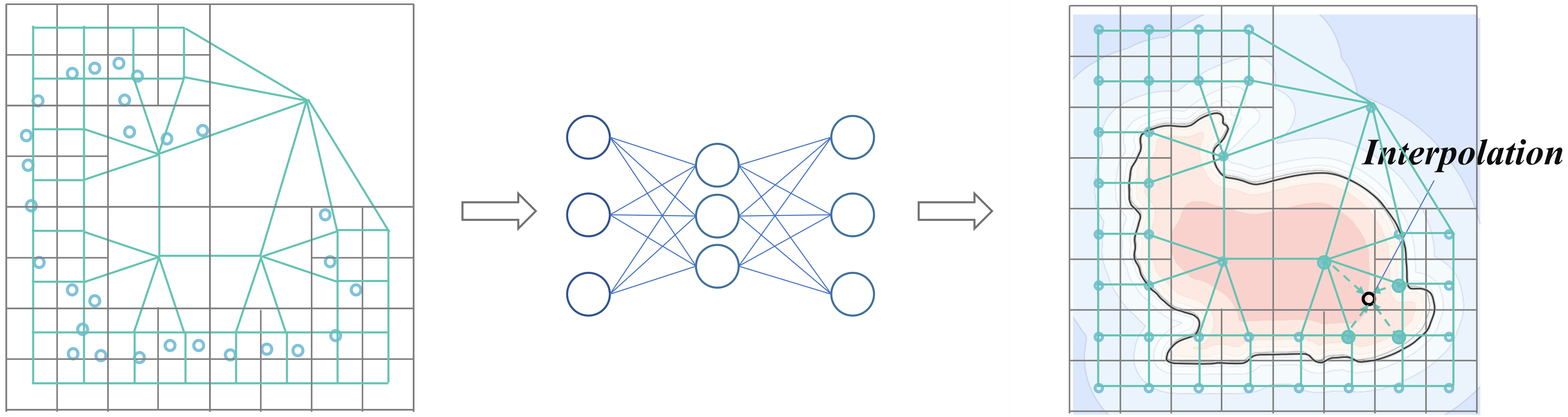
- Related works: [Peng et al. 2020; Martel et al. 2021; Takikawa et al. 2021]
- ✔ Partially overcome limitations of MLP-based methods
- ✘ Low resolution feature volume (ConvONet [Peng et al. 2020])
- ✘ No encoder networks (ACORN & NGLOD [Martel et al. 2021; Takikawa et al. 2021])



ConvONet
[Peng et al. 2020]

Key Idea

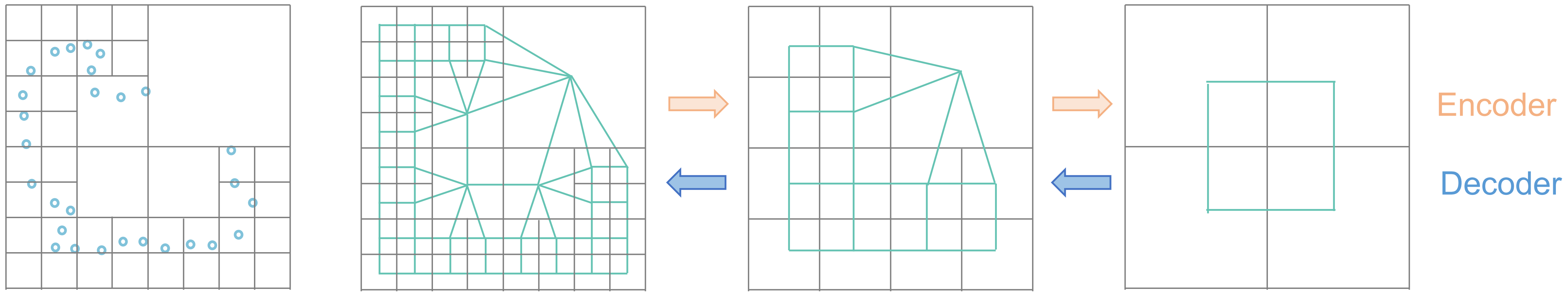
- Represent both volumetric fields and point clouds with Octrees
- Learn features with graph networks on dual Octrees



Dual Octree Graph

Adaptive Features

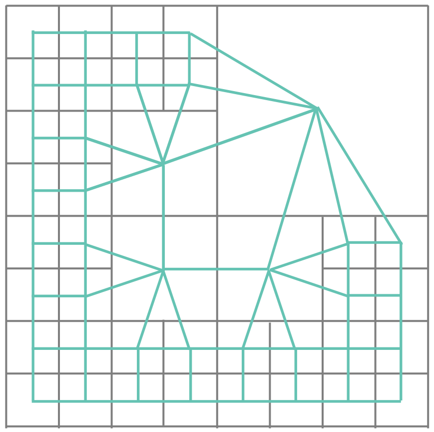
Dual Octree Graphs



Octree

Multi-Resolution Dual Octree Graphs

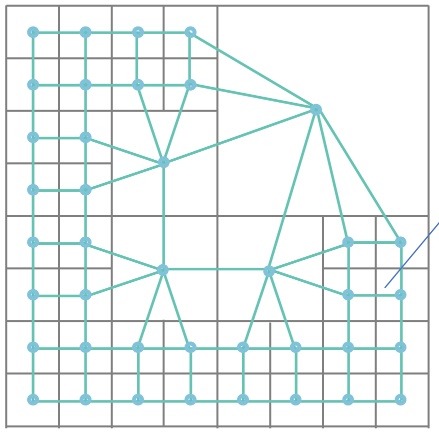
Graph Convolution on Dual Octrees



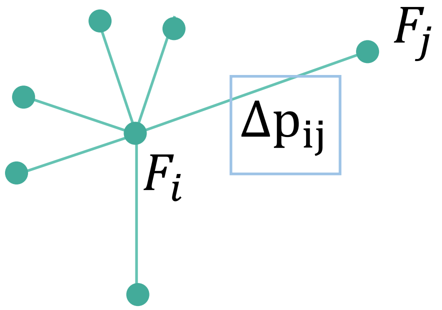
Message passing:

$$F_i = \sum_{j \in N_i} W(\Delta p_{ij}) \times F_j$$

Graph Convolution on Dual Octrees



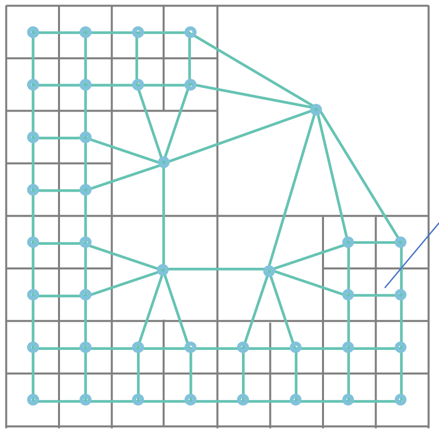
Δp_{ij} has finite number of values:



Message passing:

$$F_i = \sum_{j \in N_i} \boxed{W\left(\Delta p_{ij}\right)} \times F_j$$

Graph Convolution on Dual Octrees



Δp_{ij} has finite number of values:

$$F_i = \sum_{j \in N_i} \boxed{W_{I(\Delta p_{ij})}} \times F_j \quad \Leftarrow$$

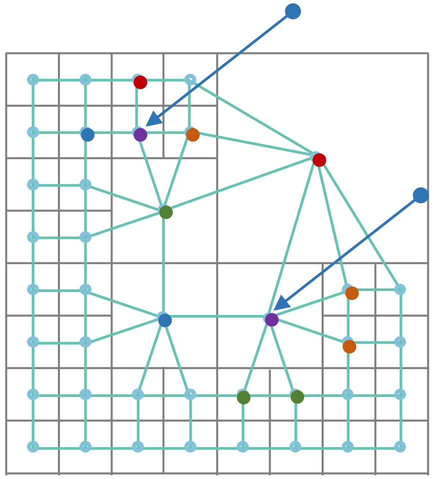
where $I(p_{ij}) \in \{0, 1, 2, \ldots, 7\}$

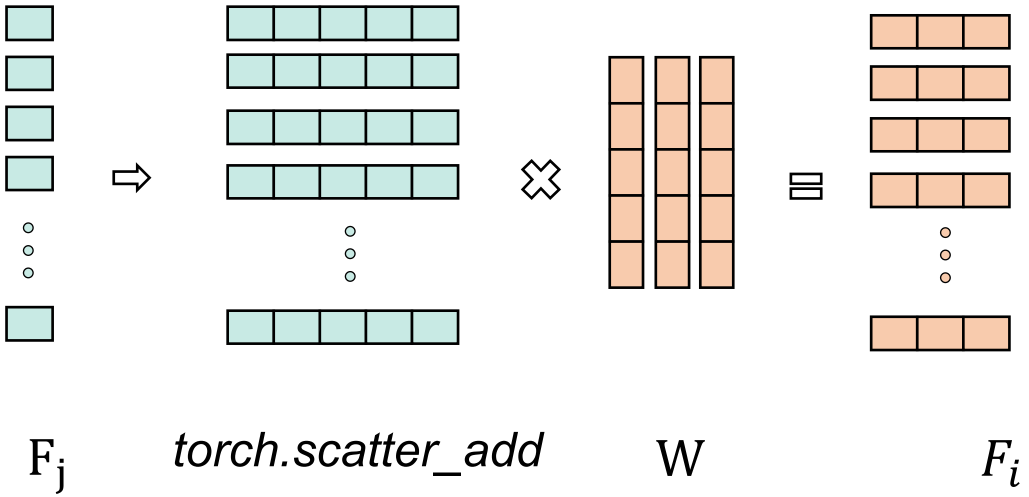
Message passing:

$$F_i = \sum_{j \in N_i} \boxed{W(\Delta p_{ij})} \times F_j$$

Graph Convolution on Dual Octrees

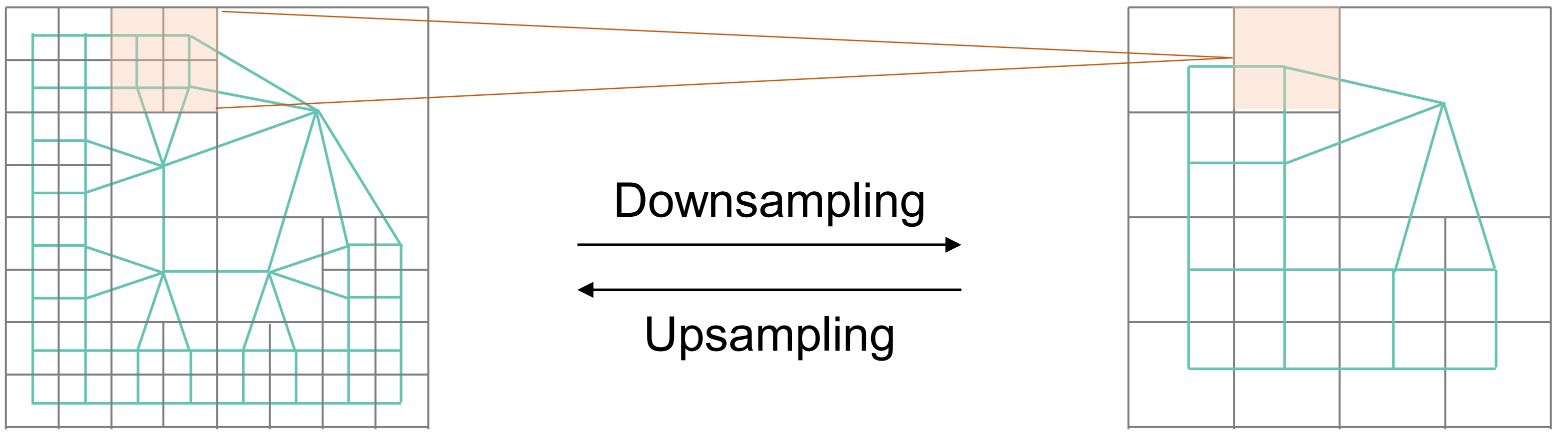
- ✅ 5 times faster than previous graph convolutions



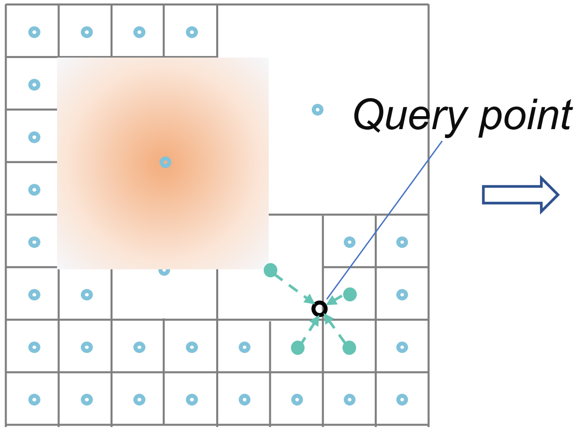


$$F_i = \sum_{j \in N_i} W_{I(\Delta p_{ij})} \times F_j$$

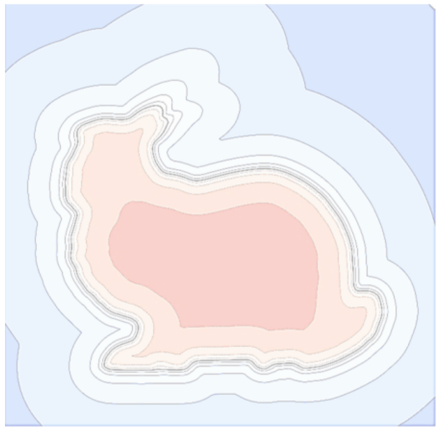
Graph Downsampling and Upsampling



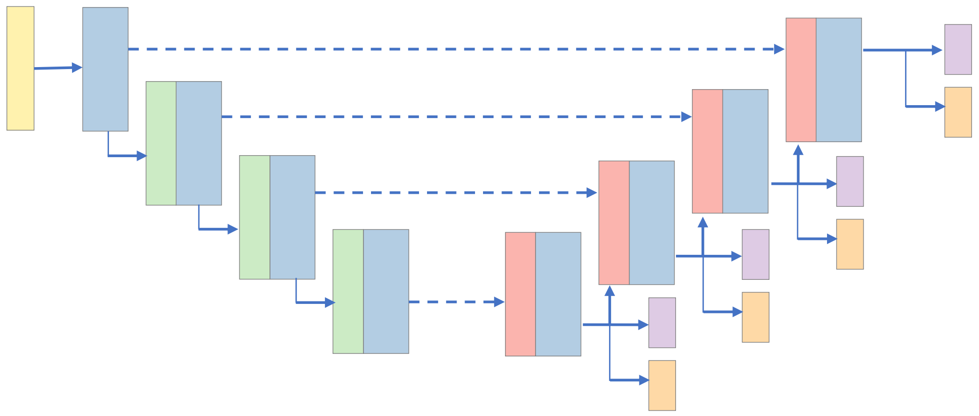
Neural Multi-level Partition of Unity (MPU)



$$\frac{\sum weight(x) \times F(\ \ ,x)}{\sum weight(x)}$$



Network and Loss Functions



- *Input*
- *PredictionModule* ⟶ Octree Splitting Loss
- *NeuralMPU* ⟶ Regression Loss
- *Upsample(c)*
- *Downsample(c)*
- *Resblock(n, c)*

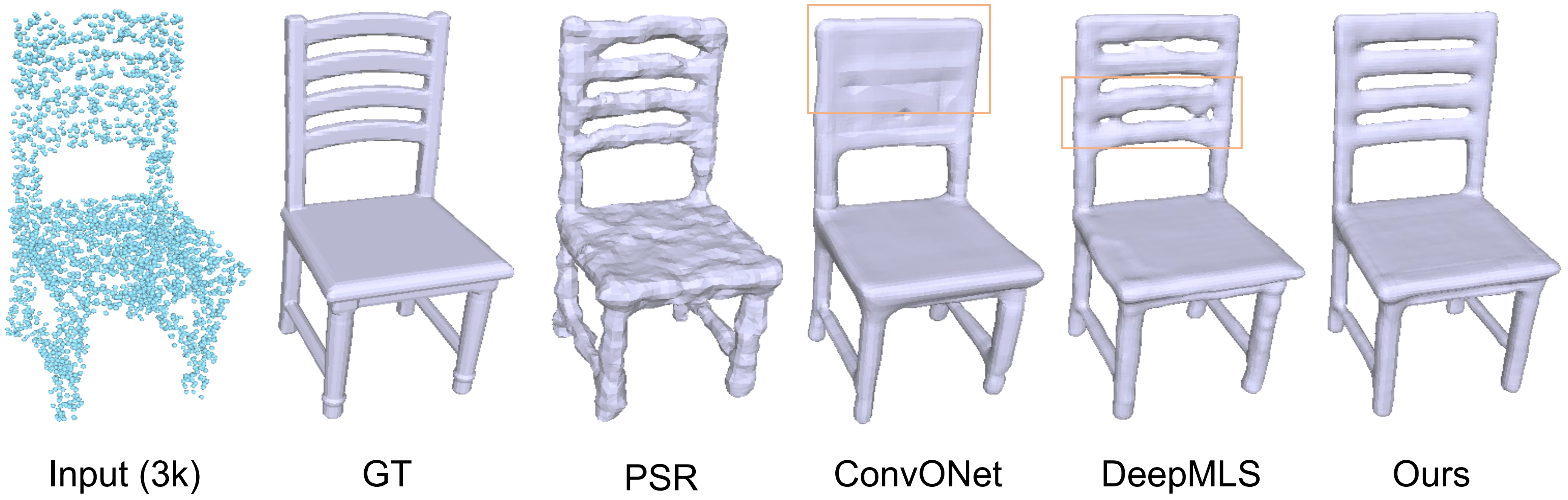
$$\mathcal{L}_{octree} = \sum_d \frac{1}{N_d} \sum_{o \in O_d} \mathrm{CrossEntropy}(o, o_{gt})$$

$$\mathcal{L}_{regress} = \sum_d \frac{1}{N_{\mathcal{P}}} \sum_{x \in \mathcal{P}} \left(\lambda_v \| F(x) - G(x) \|_2^2 + \| \nabla F(x) - \nabla G(x) \|_2^2 \right)$$

$$\mathcal{L}_{grad} = \sum_d \left\{ \frac{1}{N_{\mathcal{S}}} \sum_{x \in \mathcal{S}} \left(\lambda_v \| F(x) \|_2^2 + \| \nabla F(x) - \mathcal{N}(x) \|_2^2 \right) + \frac{1}{N_{\mathcal{Q}}} \sum_{x \in \mathcal{Q}} \lambda_g \| \nabla F(x) \|_2^2 \right\}$$

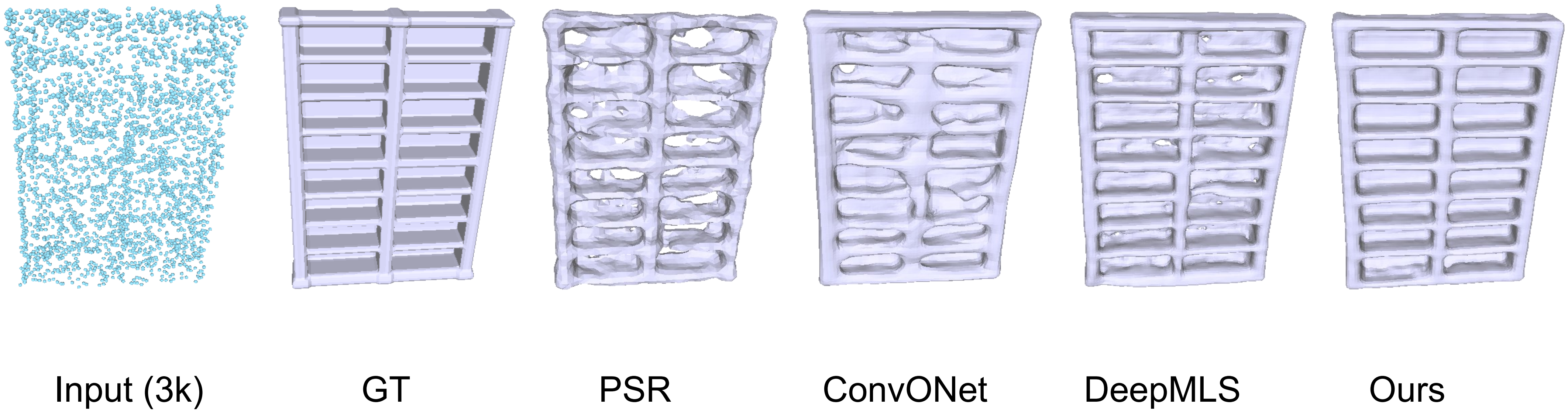
Shape Reconstruction from Point Clouds

- Dataset: 13 categories from ShapeNet

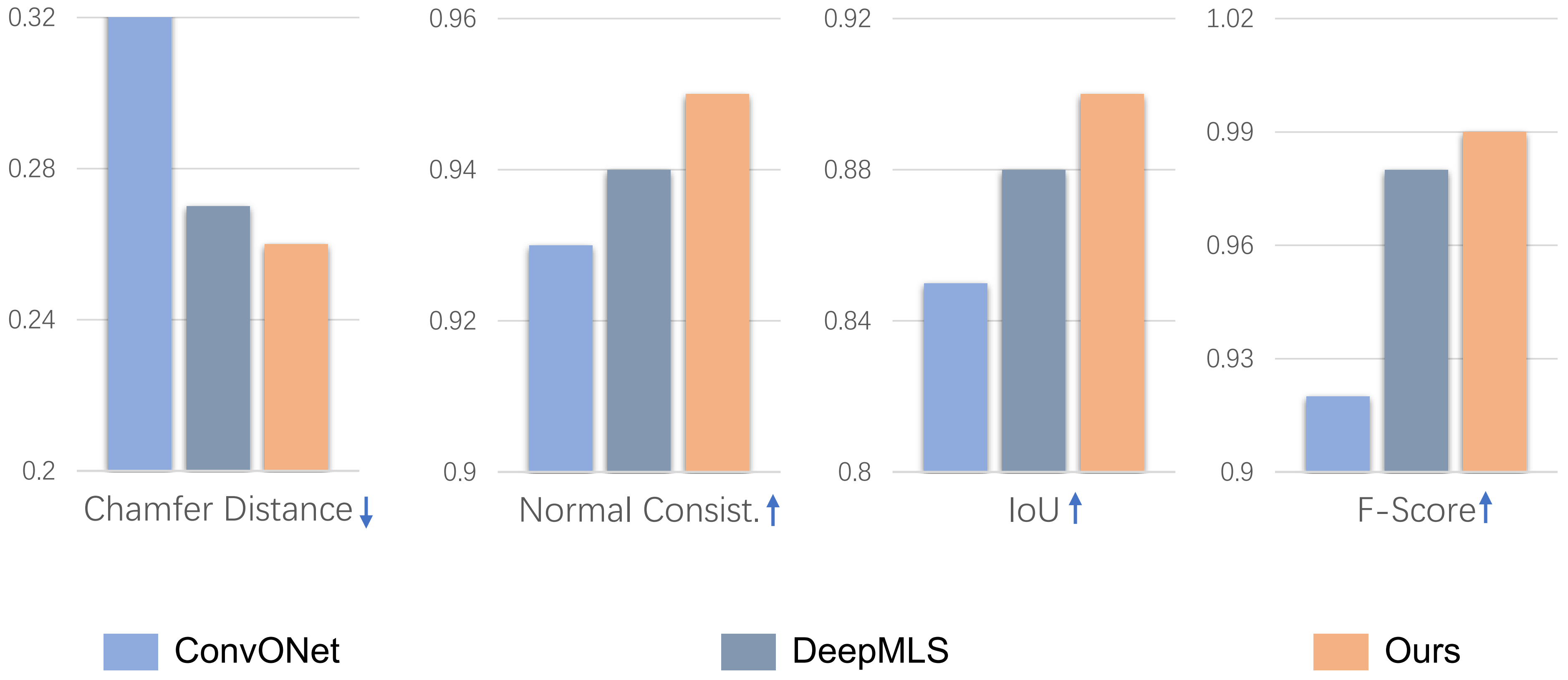


Shape Reconstruction from Point Clouds

- 13 categories from ShapeNet



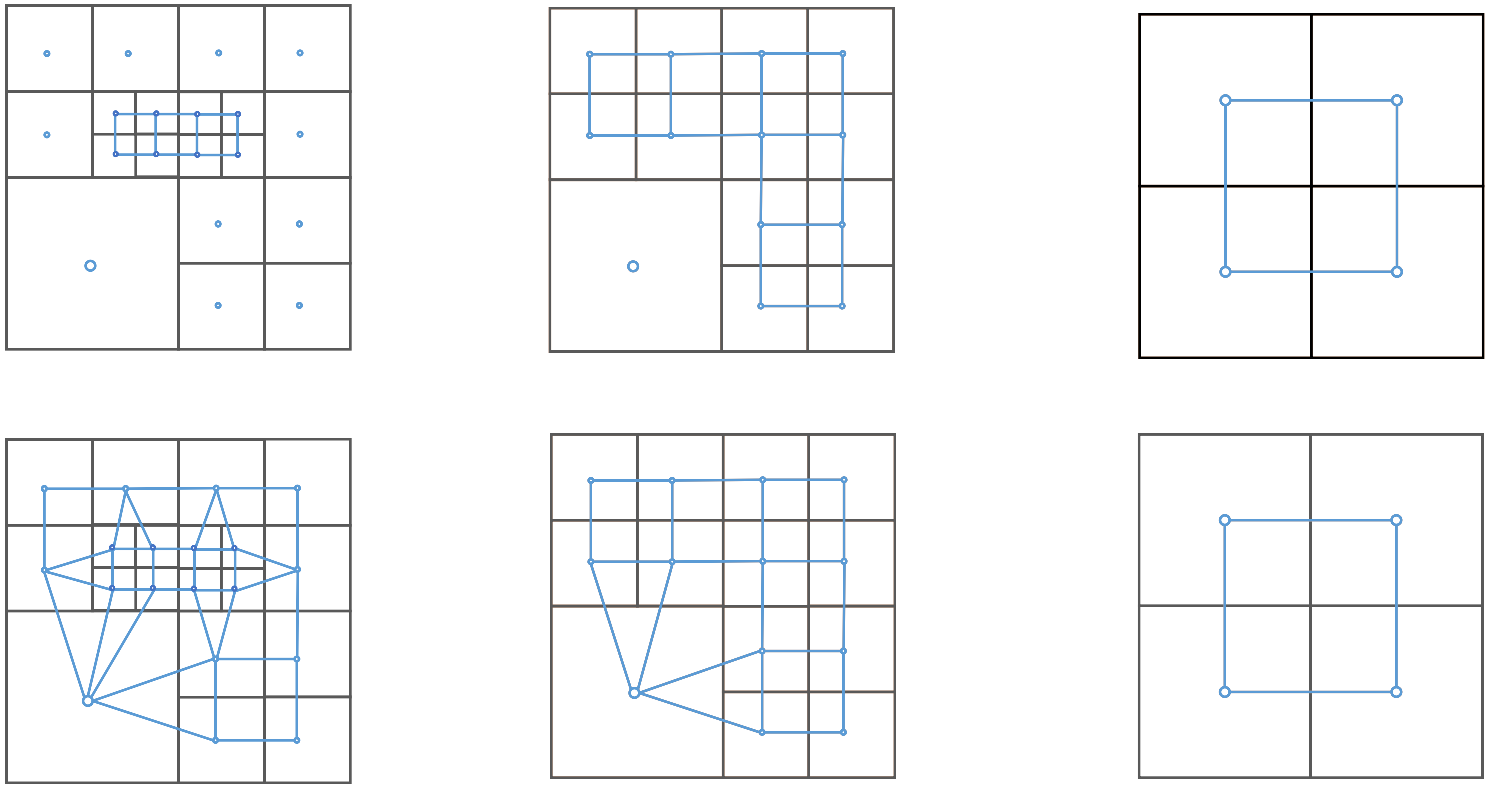
Shape Reconstruction from Point Clouds



Effectiveness: Multiscale Features in Convolutions

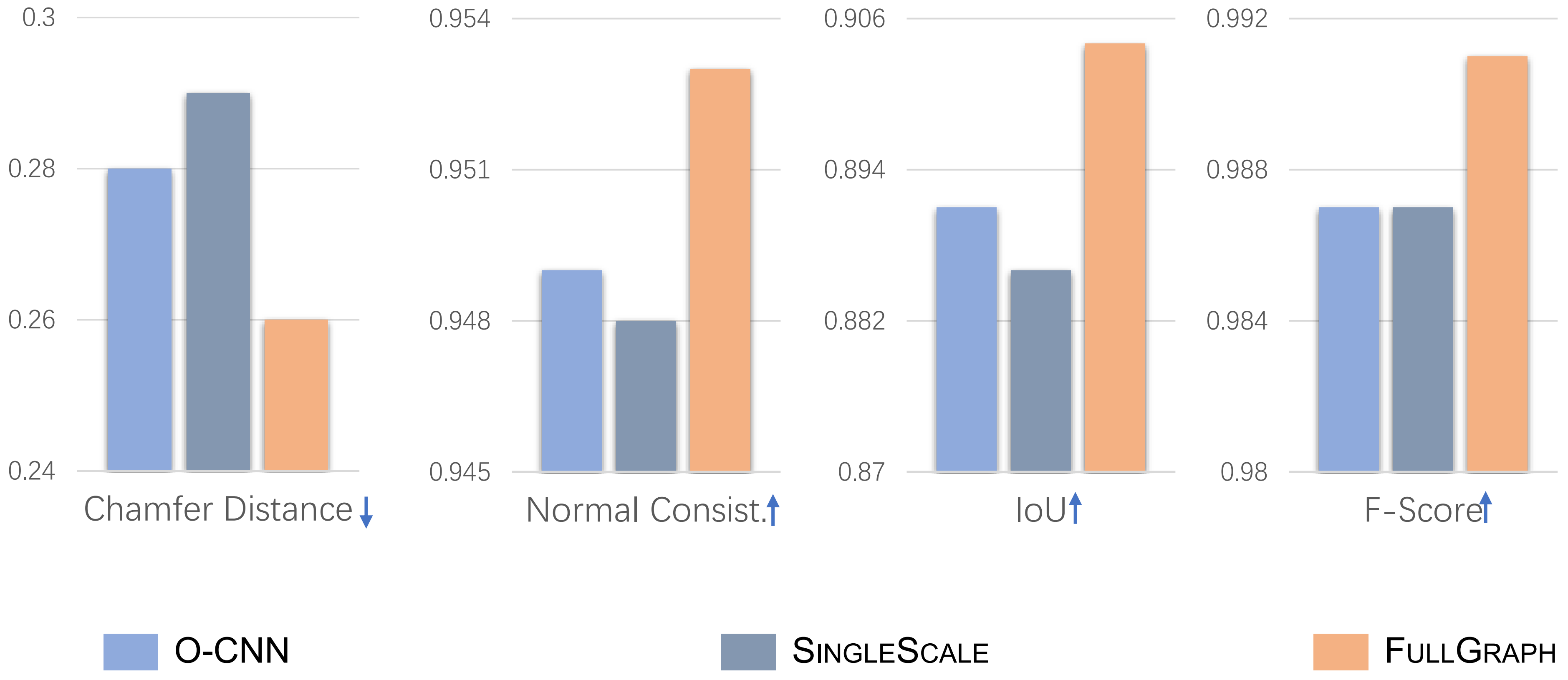
O-CNN

Dual
Octree GNNs

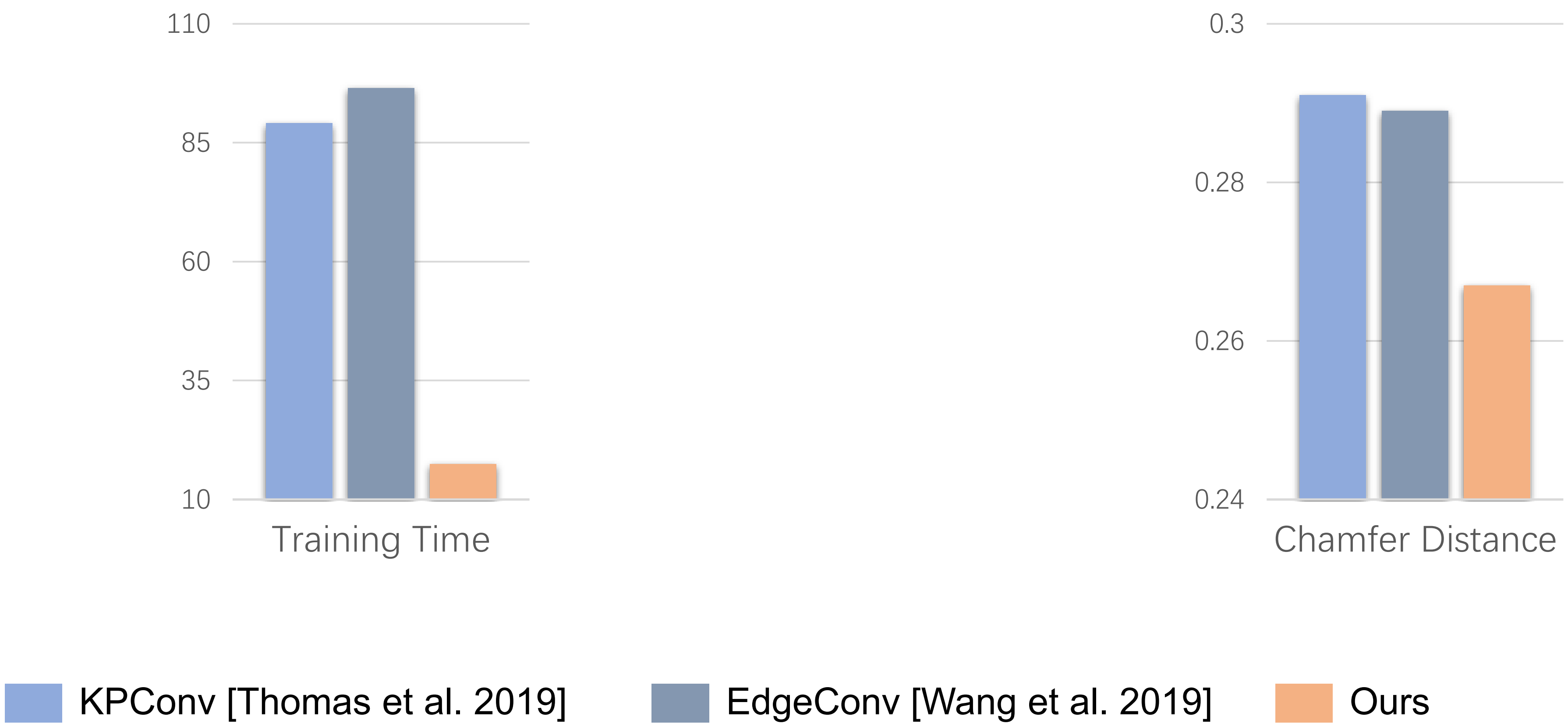


O-CNN [Wang et at. 2017] SIGGRAPH

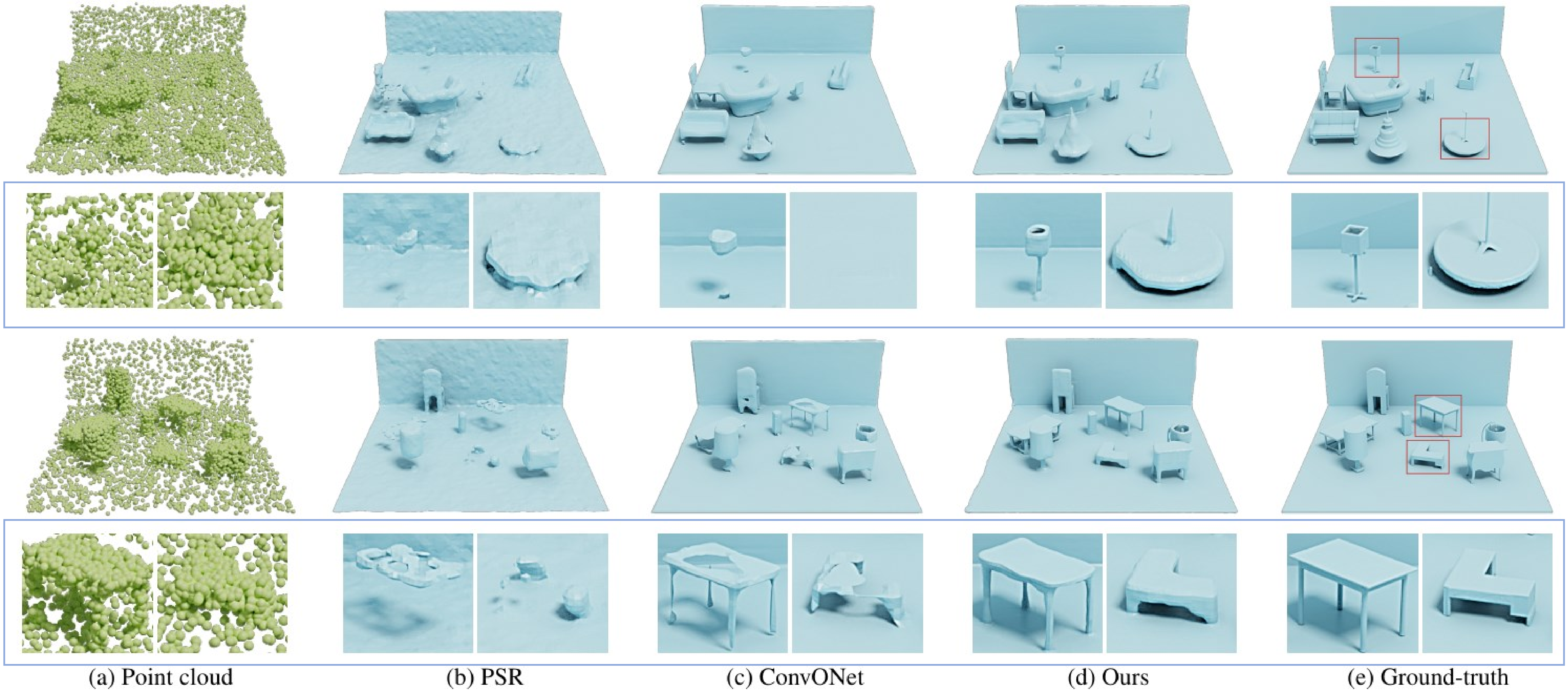
Effectiveness: Multiscale Features in Convolutions



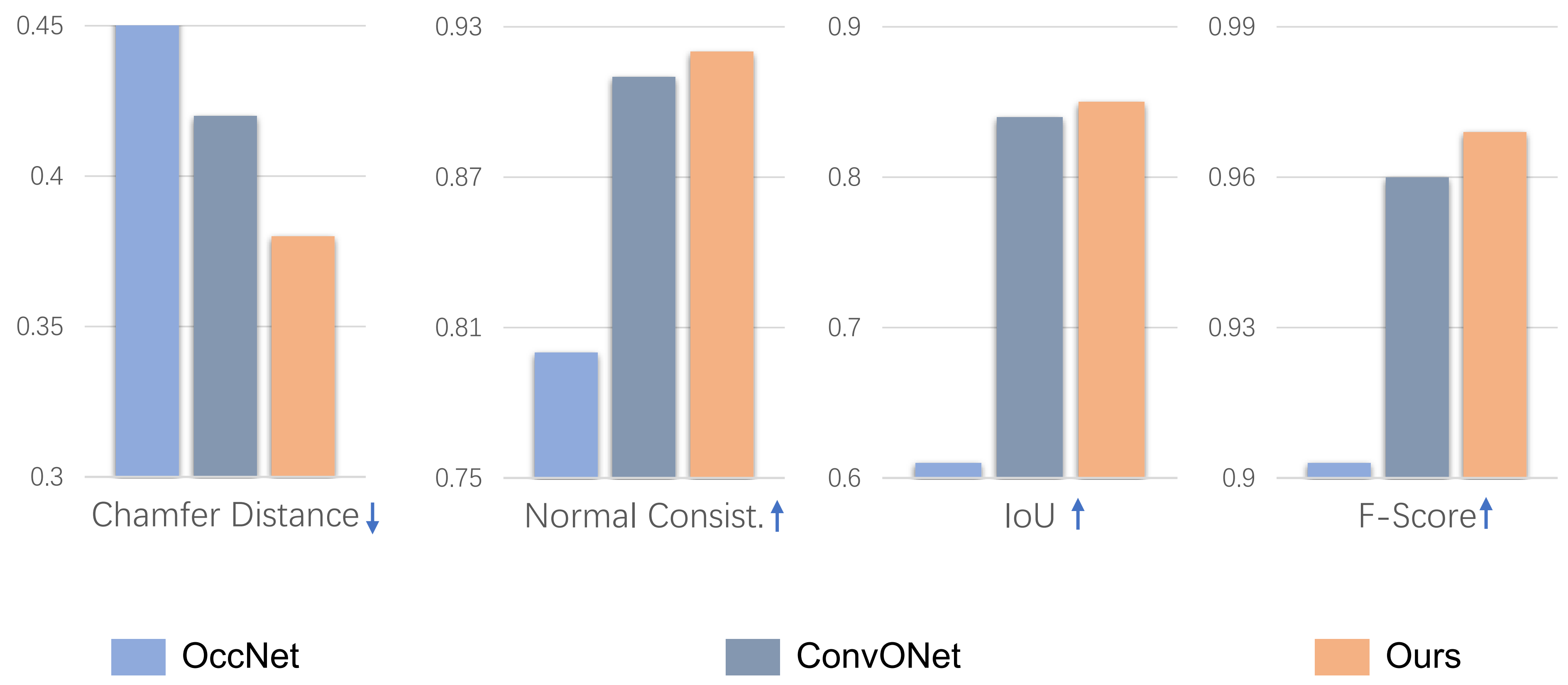
Efficiency: Faster than Other Graph Convolutions



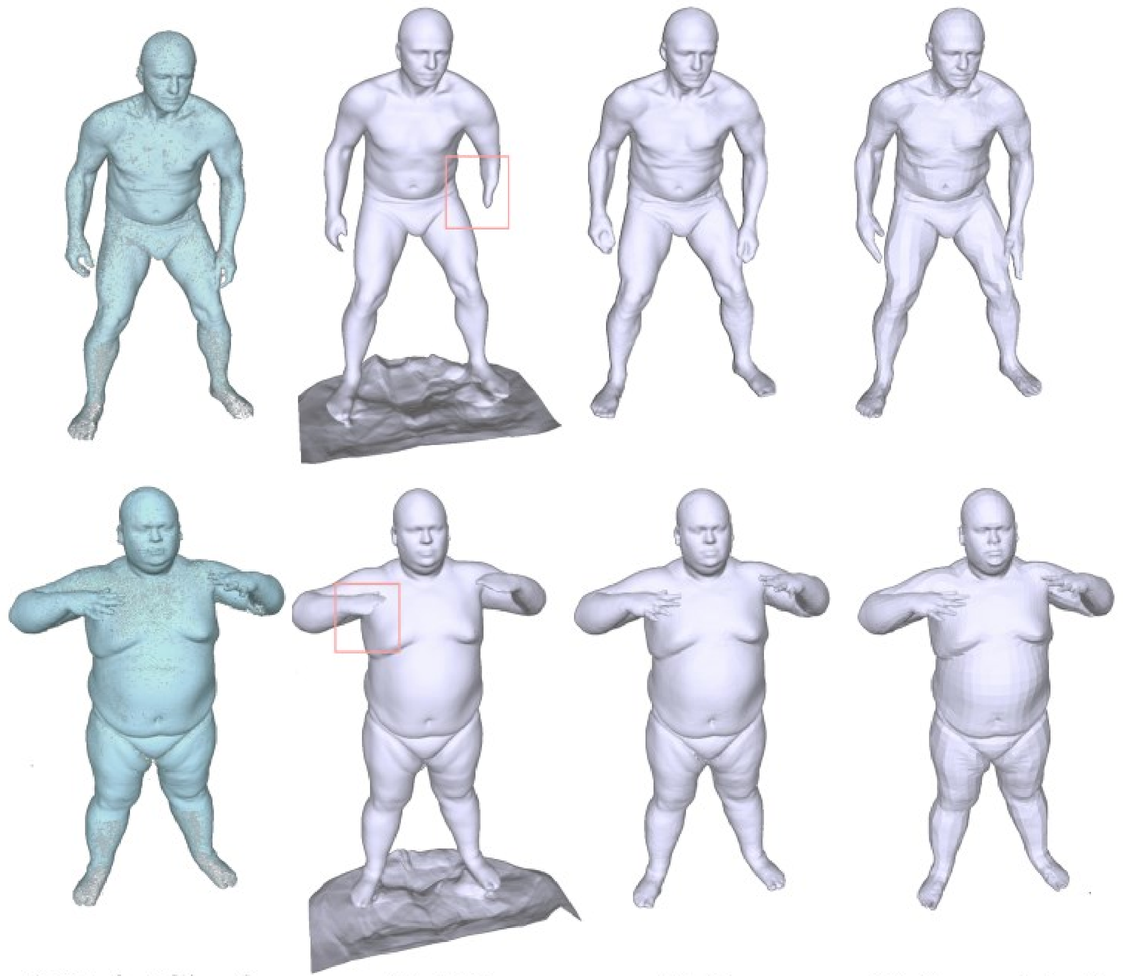
Scene Reconstruction from Point Clouds



Scene Reconstruction from Point Clouds

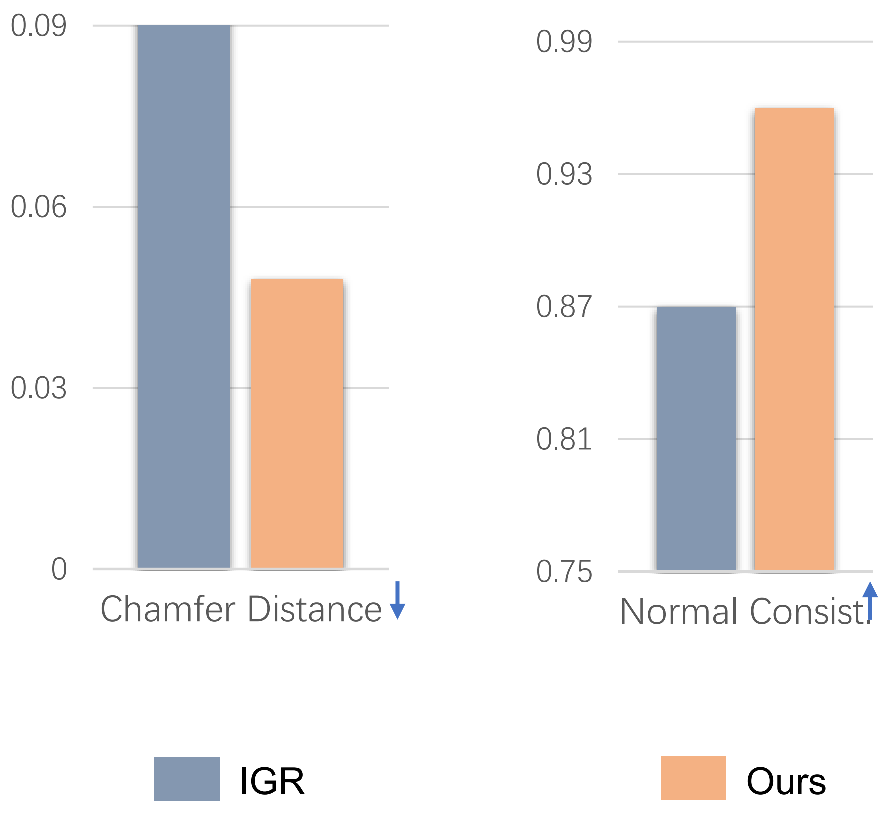


Surface Reconstruction from Point Clouds



(a) Point Cloud (b) IGR (c) Ours (d) Ground-truth

- ✔ 390 TIMES FASTER!

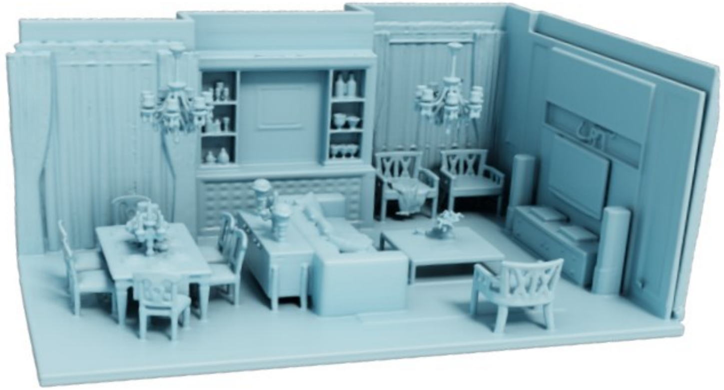


Generalization Ability

- The network trained on human bodies can be applied to general shapes

Summary

- Learning to predict adaptive volumetric fields
  - Dual octree graph networks
  - Neural MPU
  - Produce high-quality continuous surfaces
- General for more applications
  - Superior performances for shape autoencoding
  - Promising performance for shape analysis
- Future work
  - Shape generation, like Diffusion Models



Reconstruct a complex scene in 478ms

SCAN ME

Code and data are available online