SA-ConvONet: Sign-Agnostic Optimization of Convolutional Occupancy Networks

Jiapeng Tang[1,4] Jiabao Lei [1] Dan Xu[2] Feiying Ma[4] Kui Jia[1] Lei Zhang[3,4]

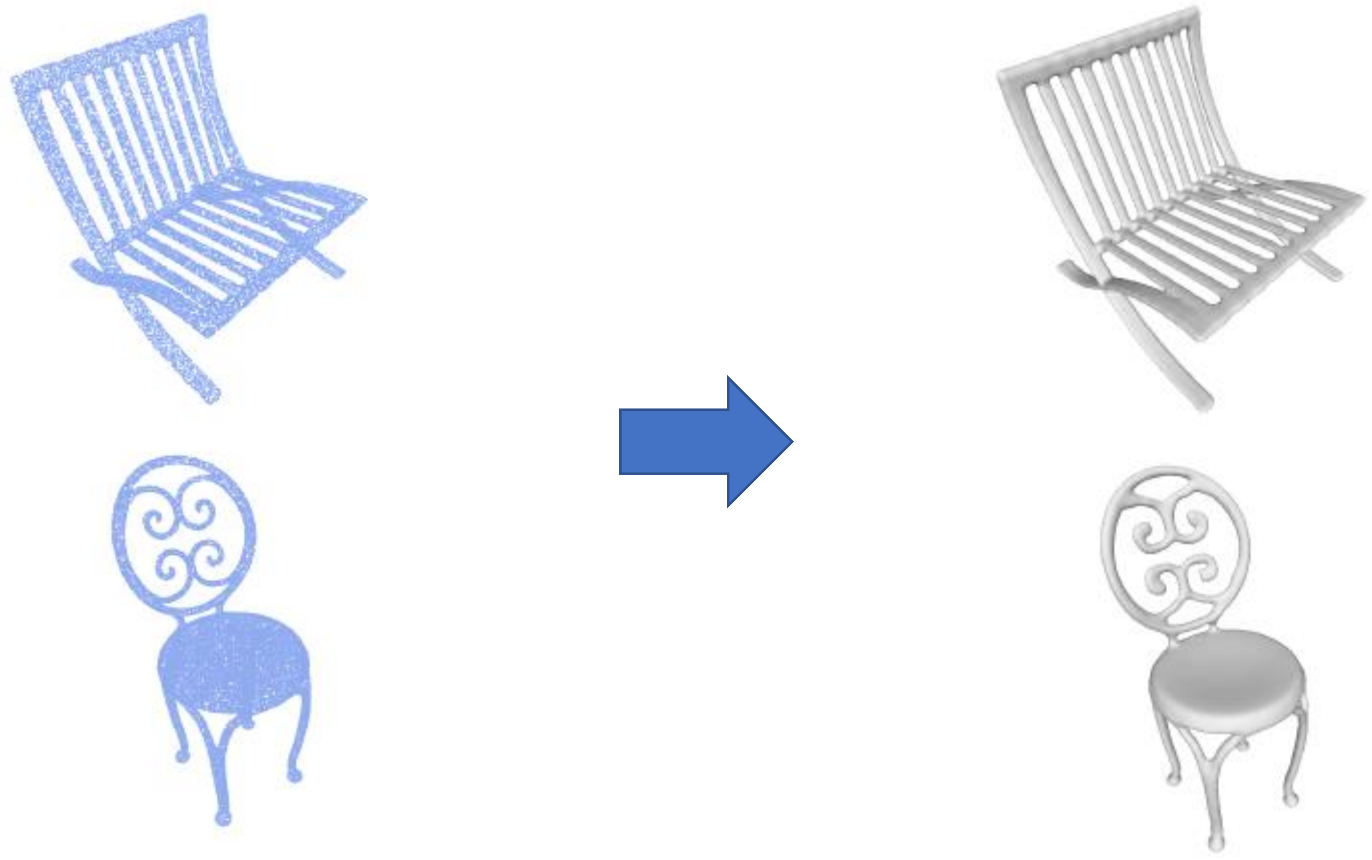
[1] South China University of Technology

[2] The Hong Kong University of Sciences and Technology

[3] The Hong Kong Polytechnic University

[4] DAMO Academy, Alibaba Group

Task: Surface Reconstruction

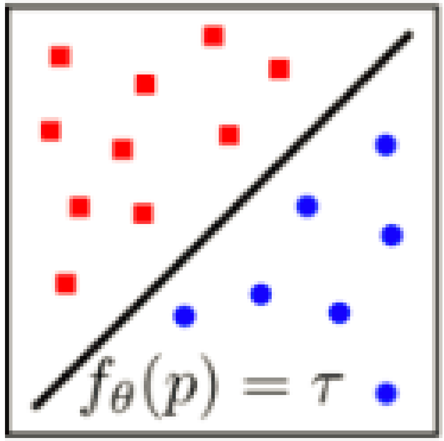


Un-oriented Point clouds Surface Meshes

Related Works

Neural Implicit Representation

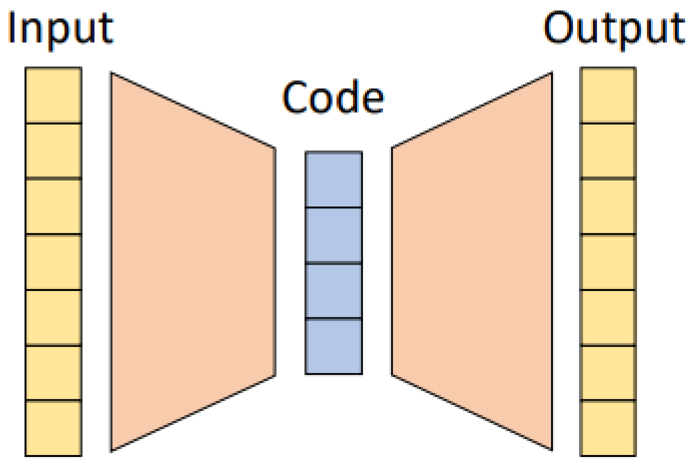
- ***represent a 3D shape as the continuous decision boundary of a binary classifier.***



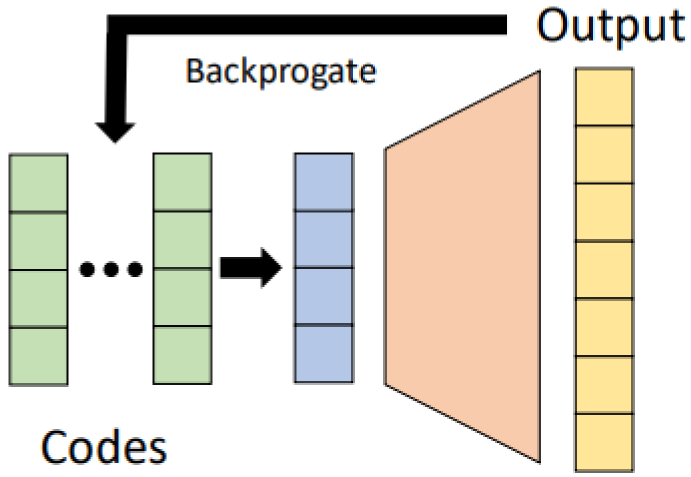
- ***surface reconstruction with infinite resolution and arbitrary topology***

Mescheder, L., et al. Occupancy networks. CVPR 2019.

Improve the **generality to novel shapes**



single forward pass

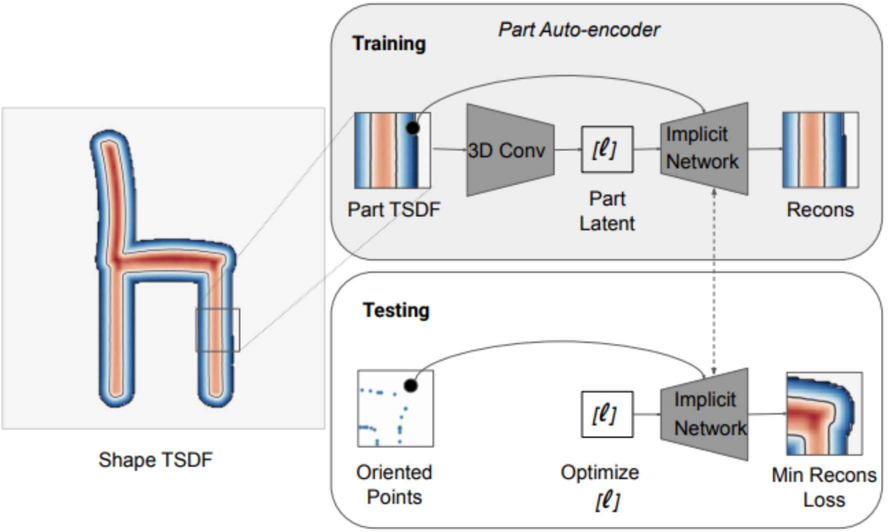


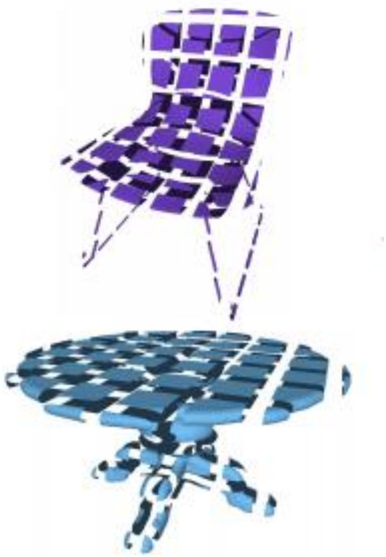
test-time optimization

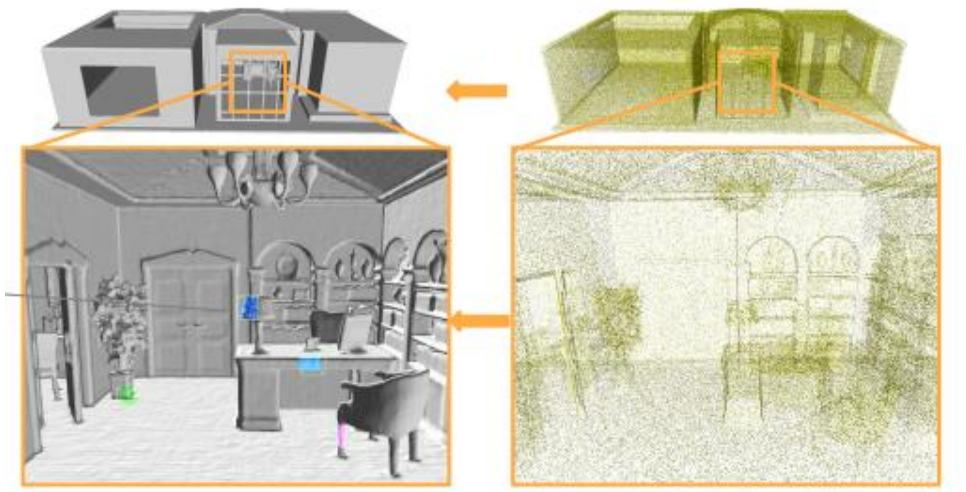
- ***Further optimize network parameters during inference to find a better solution***

Park, J., et al. DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation. CVPR2019.

Improve the scalability to large-scale scenes





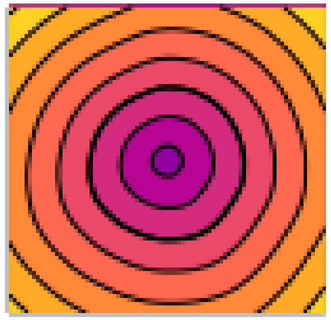


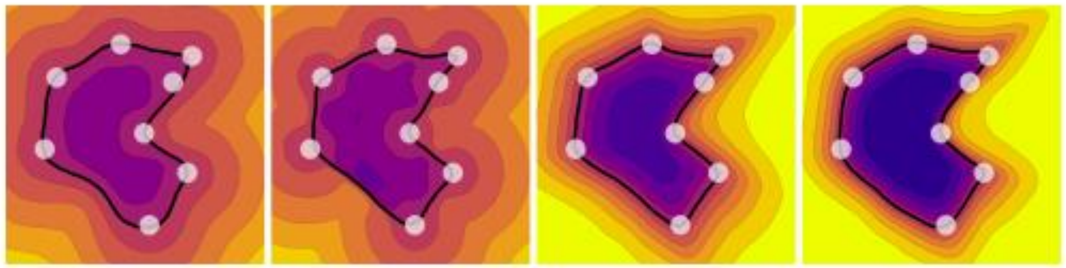
Part auto-encoder | Training object parts | Test: scalable to large scenes

- ***Pros: local shape modeling for 3D scenes***
- ***Cons: require accurate oriented normals to enforce global consistency***

Jiang, C., et al. Local implicit grid representations for 3d scenes. CVPR2020.

Improve the applicability to real-world scans





initialize the SDF decoder to represent a signed field

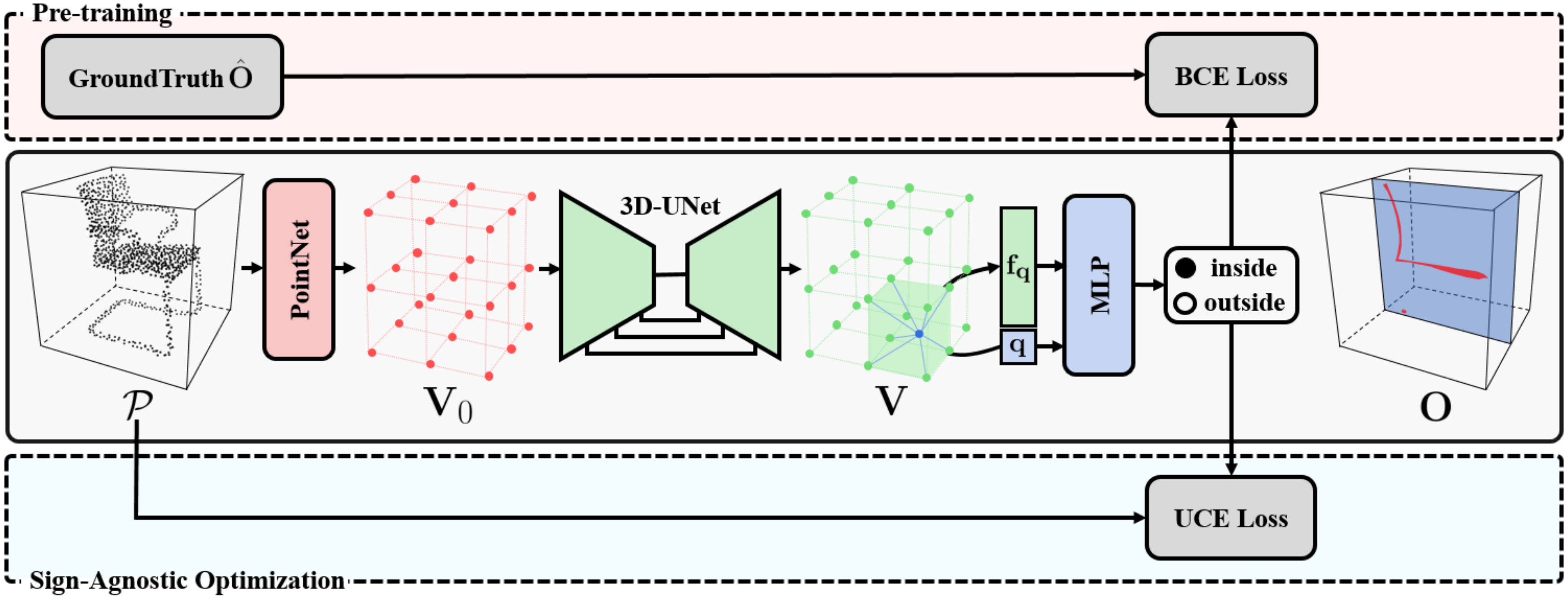
learn SDF by unsigned distance loss

- ***Pros: not require oriented normals***
- ***Cons: struggle to recover fine-grained scene surfaces***

Atzmon, M., et al. Sal: Sign agnostic learning of shapes from raw data. CVPR2020.

Approach

Sign-Agnostic Optimization of Convolutional Occupancy Networks

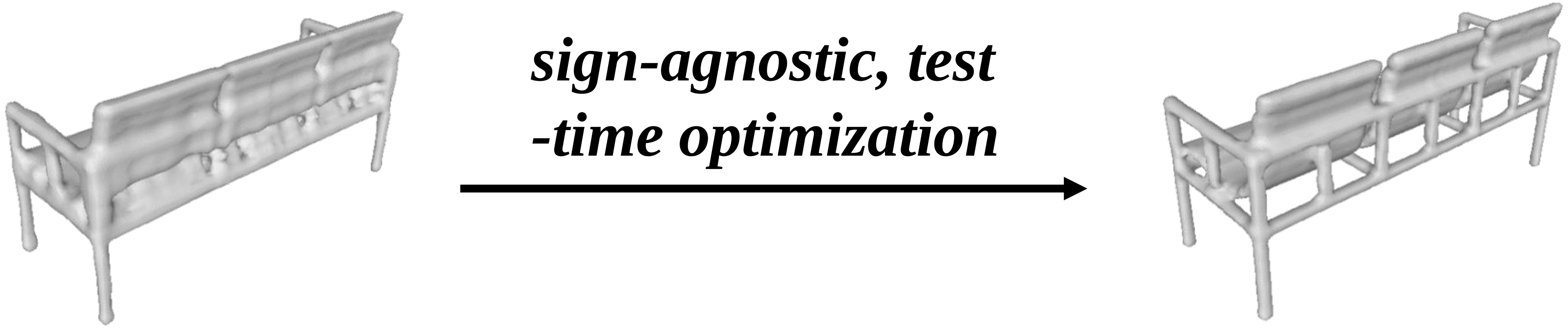


- ***Middle:*** ***local implicit fields*** conditioned on ***convolutional features*** from a 3D U-Net.
- ***Top:*** ***network pre-training on 3D datasets*** by binary cross-entropy (BCE) loss.
- ***Bottom:*** ***sign-agnostic, test-time optimization*** via unsigned cross entropy (UCE) loss.

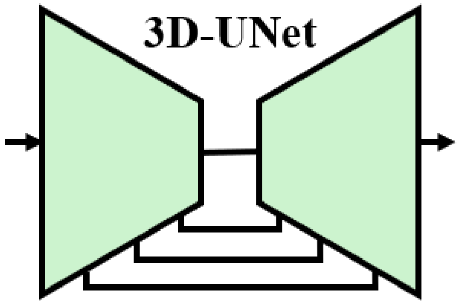
Peng, S., et al. Convolutional occupancy networks. ECCV 2020.

Motivations

- *Characteristic 1: Pre-trained occupancy field prediction networks provide signed fields as initialization for the test-time optimization.*



- *Characteristic 2: 3D U-Net aggregates both local and global shape features*



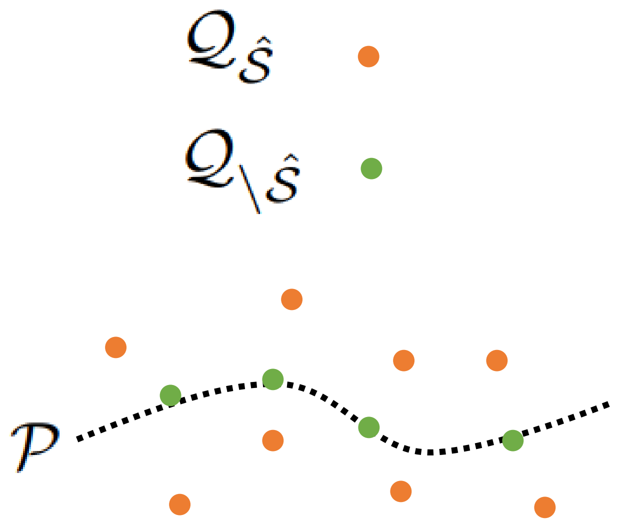
  - *local shape features: preserve scene surface details.*
  - *global shape priors: enforce global consistency between local fields.*

Unsigned Cross Entropy

$$\mathcal{L}_{uce} = \sum_{\mathbf{q} \in \mathcal{Q}} \mathrm{BCE}\left(\mathbf{O}^{\dagger}(\mathbf{q}), \hat{\mathbf{O}}^{\dagger}(\mathbf{q})\right)$$

pred $\mathbf{O}^{\dagger}(\mathbf{q}) = \mathrm{sigmoid}\left(\boxed{|g(\mathbf{q}, \mathbf{f}_{\mathbf{q}})|}\right) \in [0.5, 1)$

target $\hat{\mathbf{O}}^{\dagger}(\mathbf{q}) = \begin{cases} 0.5, & \text{for } \mathbf{q} \in \mathcal{Q}_{\hat{\mathcal{S}}} \\ 1.0, & \text{for } \mathbf{q} \in \mathcal{Q}_{\setminus \hat{\mathcal{S}}} \end{cases}$



$\mathcal{Q}_{\hat{\mathcal{S}}}$: a point set obtained from the ***observed surface.***

$\mathcal{Q}_{\setminus \hat{\mathcal{S}}}$: a point set sampled from ***non-surface volume.***

Work Condition Summary

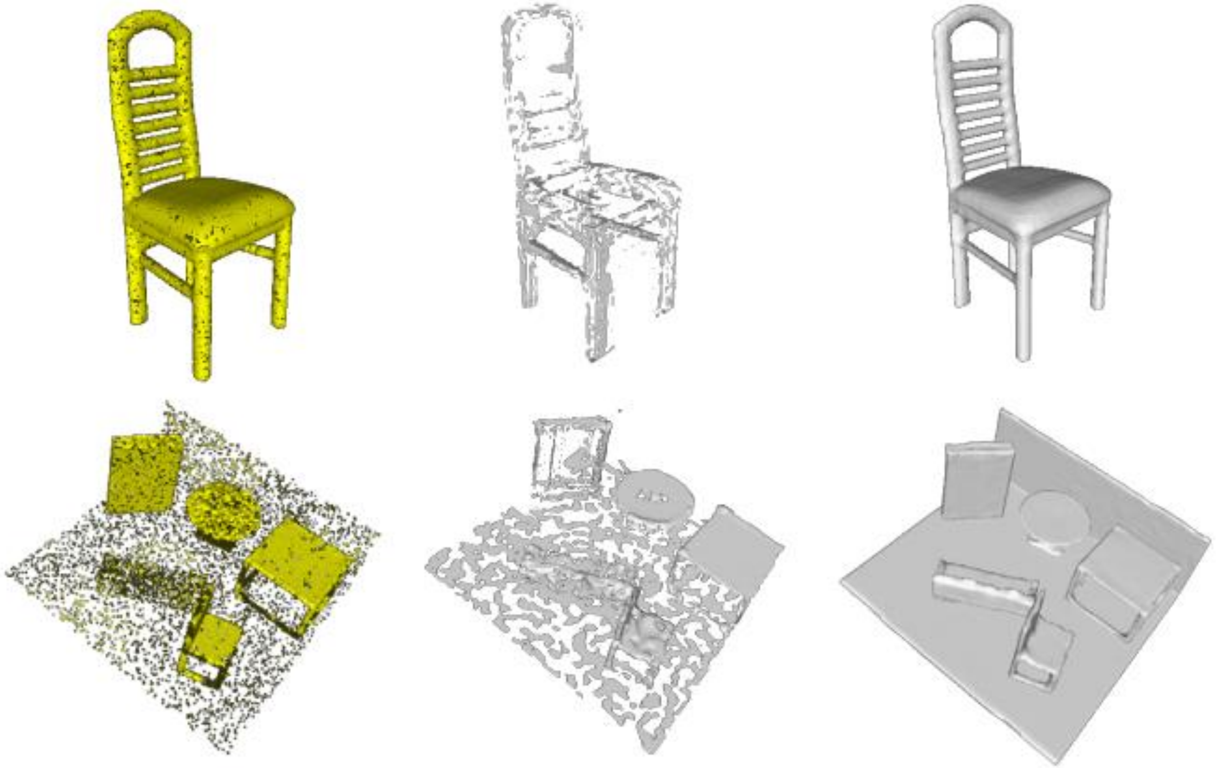
Methods	Without normals	Optimization of network parameters	Local geometry modeling
SPSR [26]	×	✓	✓
ONet [30]	✓	×	×
SAL [2]	✓	×	×
IGR [16]	✓	✓	×
CONet [33]	✓	×	✓
LIG [23]	×	✓	✓
Ours	✓	✓	✓

Our method is the first to maximize the three reconstruction objectives in a unified framework: ***scale well to large scenes, generalize well to novel shapes, and robust to real-world scans.***

Ablation Studies

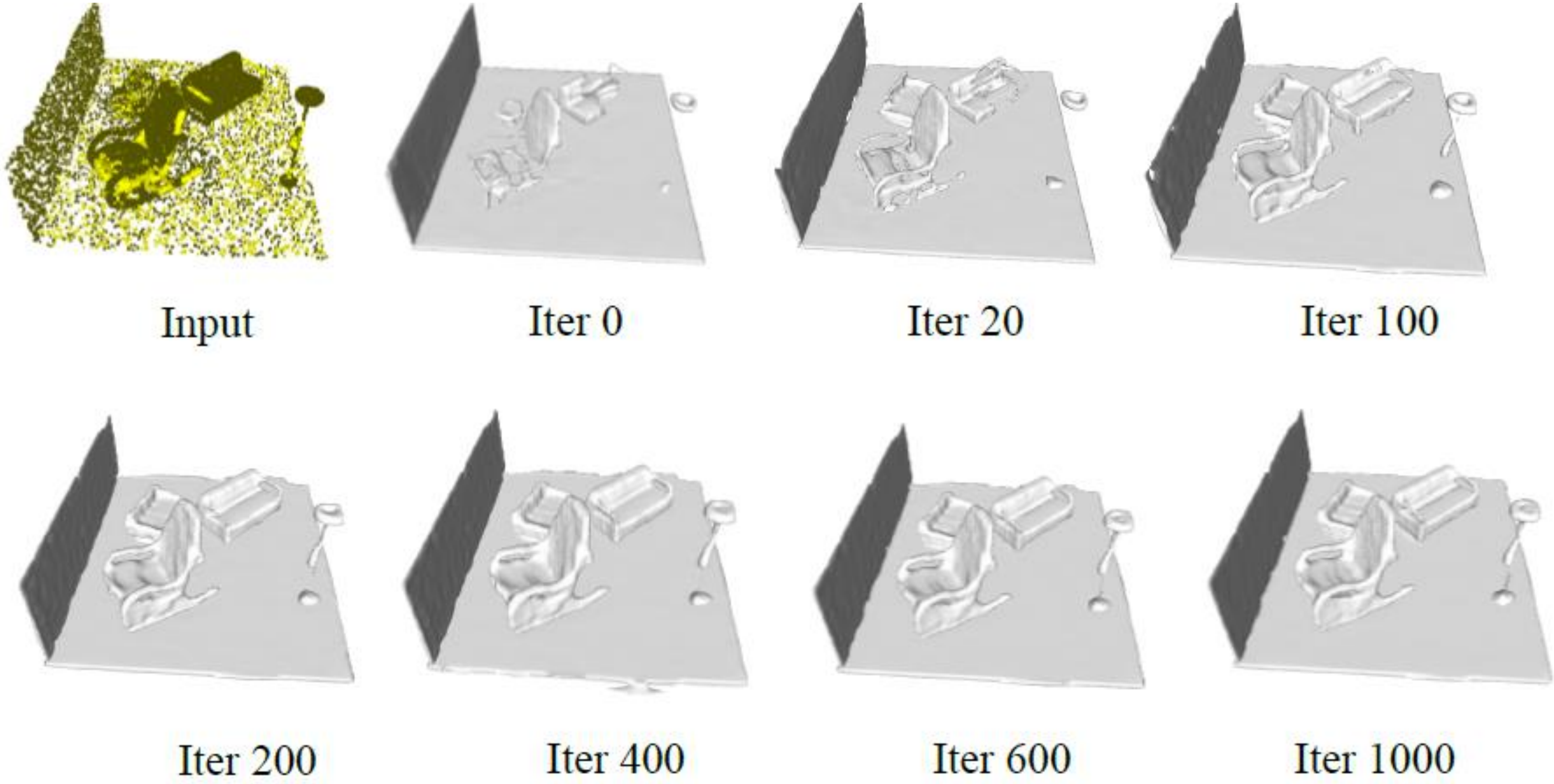
Effect of network pre-training

- ***Without pre-trained shape priors :*** ***fail to reconstruct reasonable geometries***



(a) Input (b) w/o pretraining (c) Ours

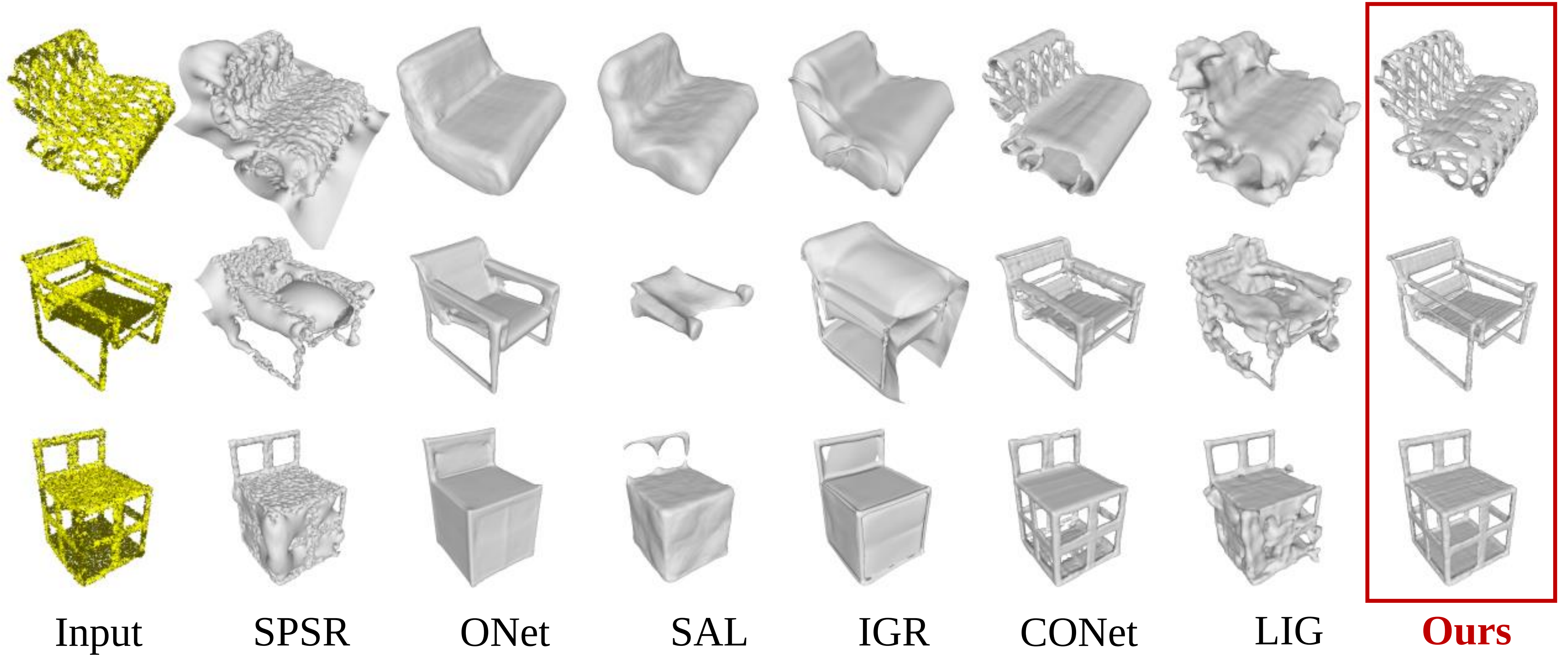
Sensitivity to the iteration number of test-time optimization



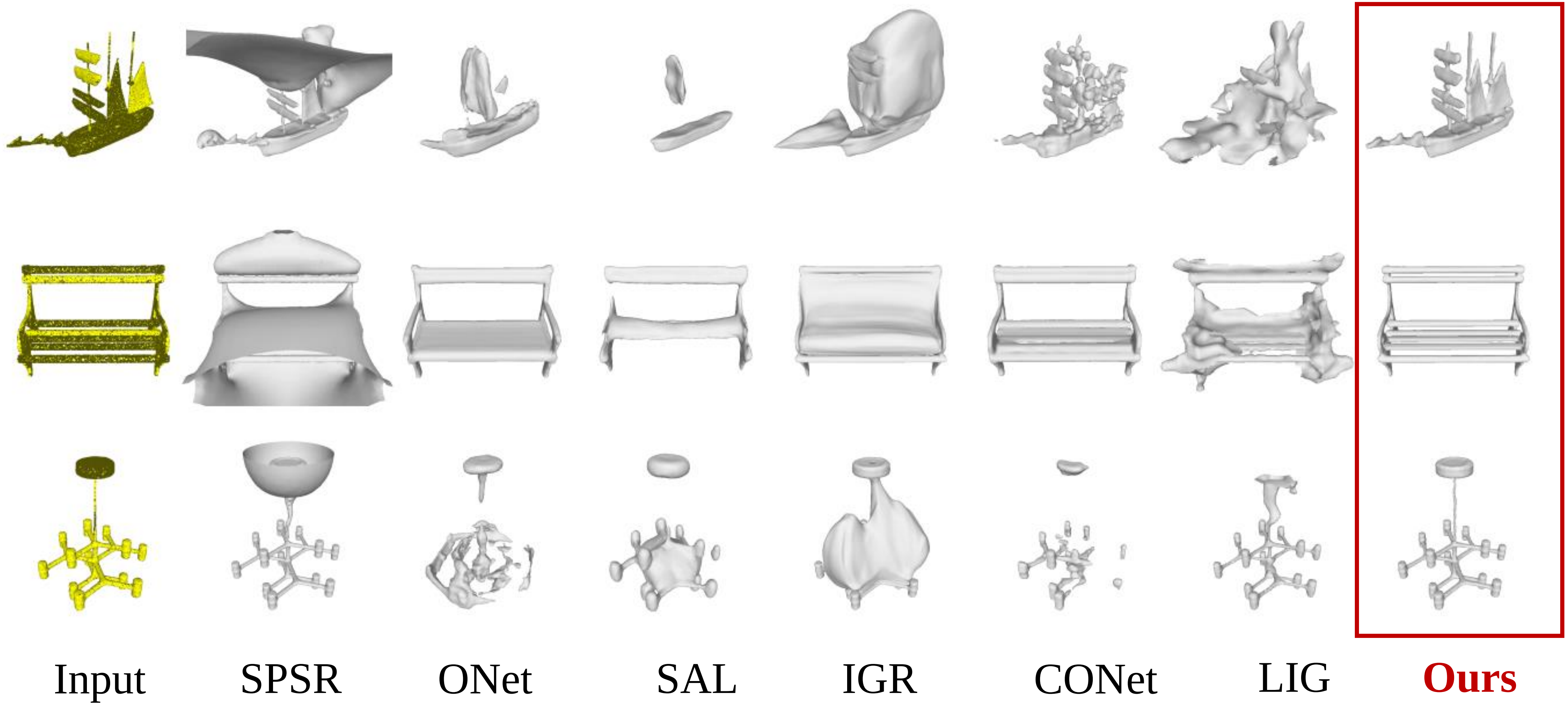
after about 600 iterations, the results become stable.

Object-level Reconstruction

ShapeNet-chair

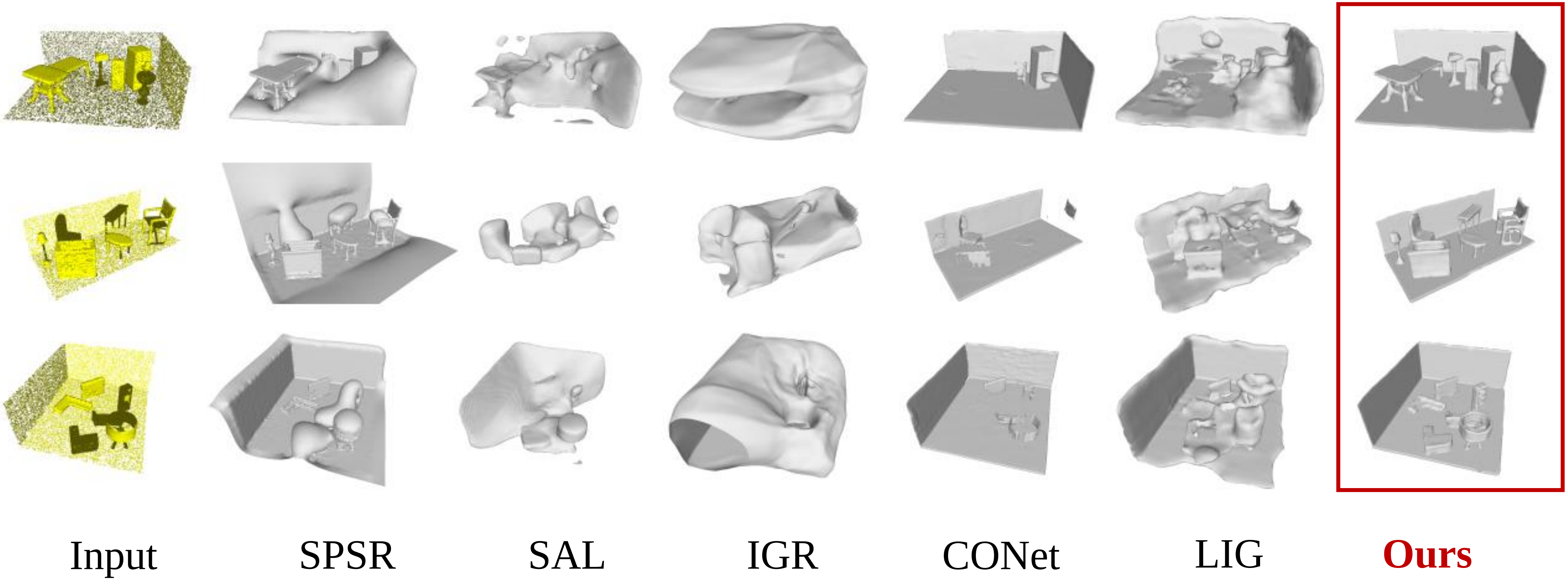


Novel categories generalization



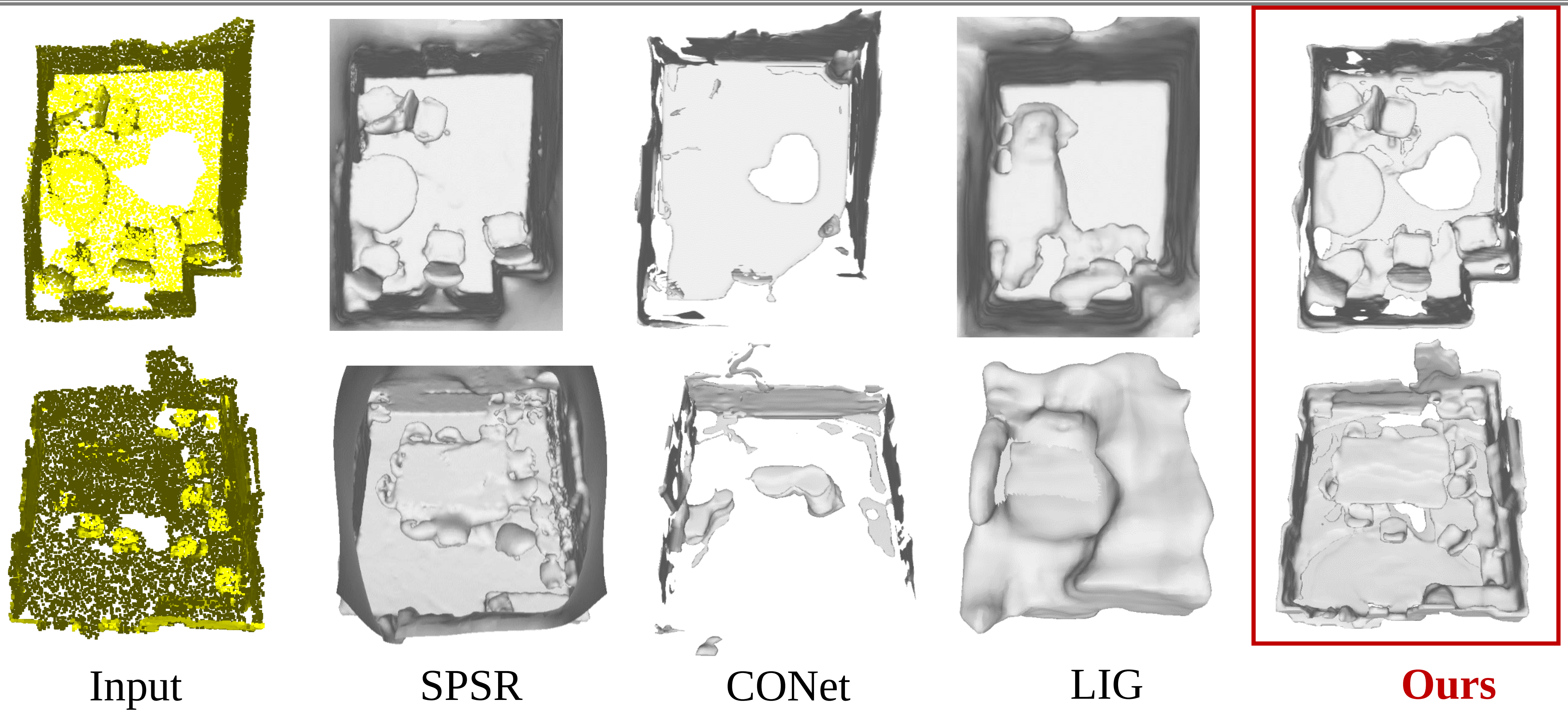
Scene-level Reconstruction

Synthetic indoor rooms

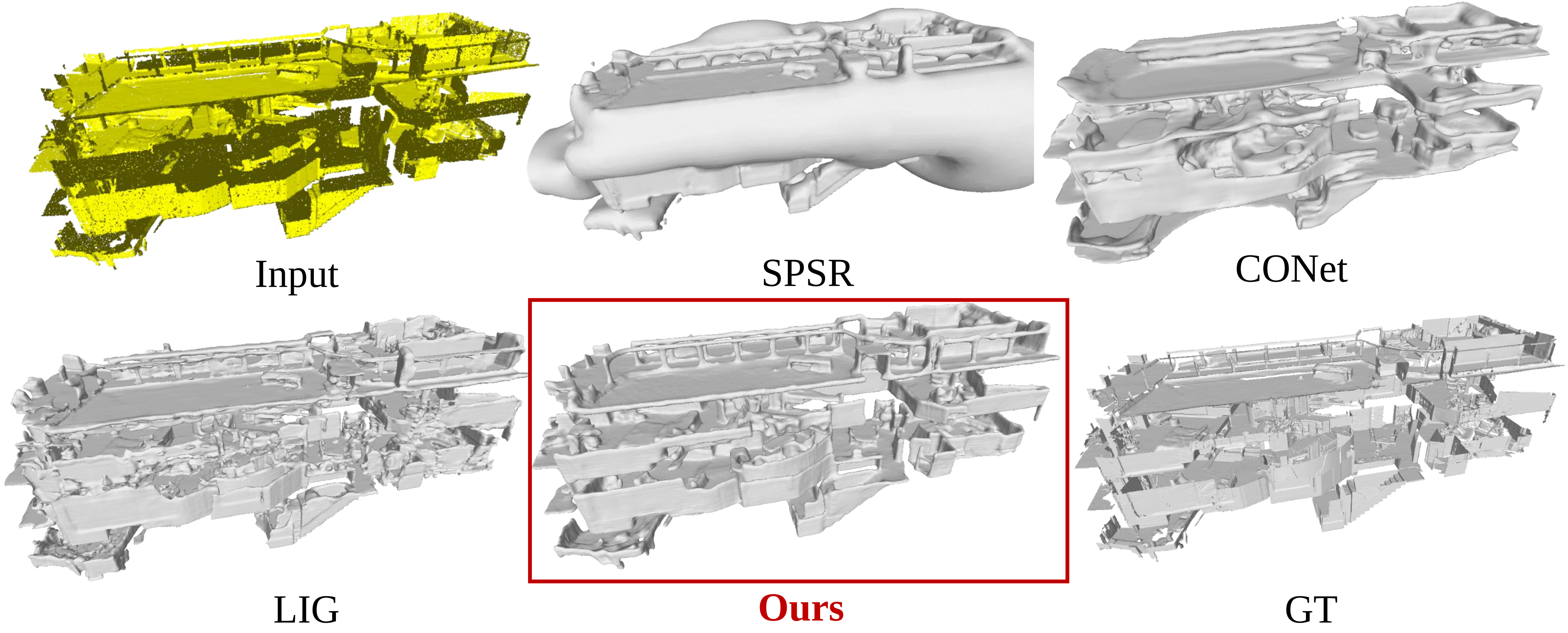


Real-world Scenes

ScanNet



Matterport3D



2021 ICCV OCTOBER 11-17 VIRTUAL

THANK YOU!

The code is available at

https://github.com/tangjiapeng/SA-ConvNet

华南理工大学 SOUTH CHINA UNIVERSITY OF TECHNOLOGY 1934

香港科技大學 THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

达摩院 Alibaba DAMO Academy

THE HONG KONG POLYTECHNIC UNIVERSITY 香港理工大學