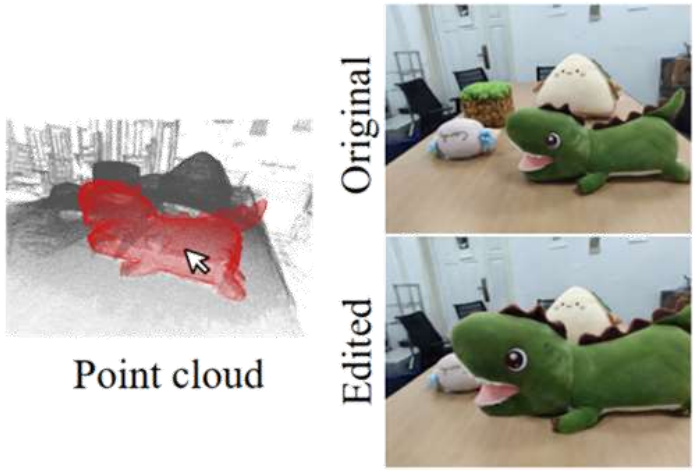
Frequency-Modulated Point Cloud Rendering with Easy Editing

Yi Zhang* · Xiaoyang Huang* · Bingbing Ni · Teng Li · Wenjun Zhang
Shanghai Jiao Tong University

Speaker: ***Xiaoyang Huang***

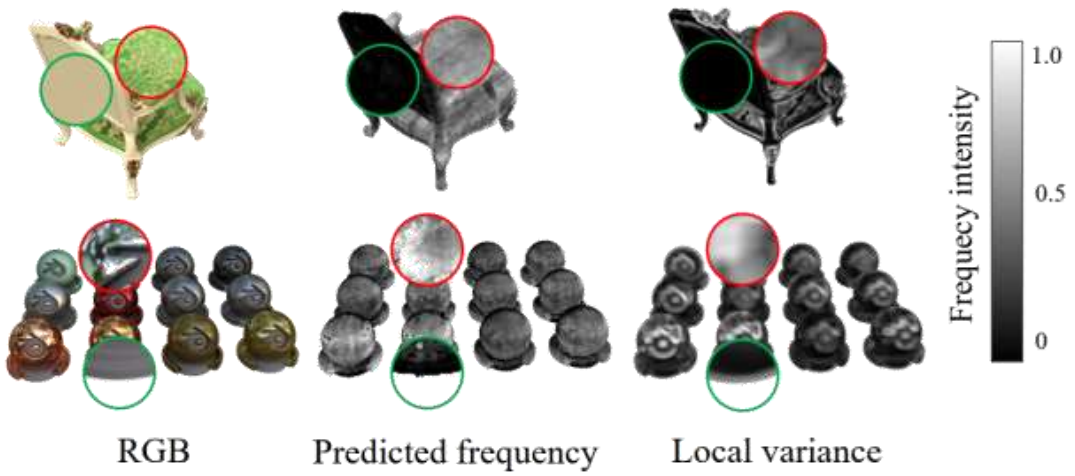
Preview

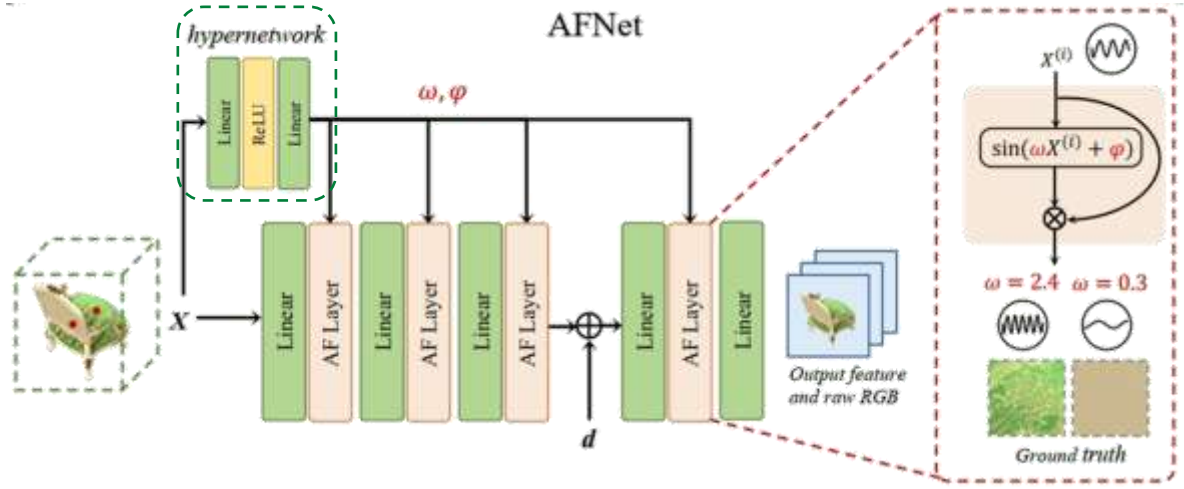
Real-time high-fidelity rendering



Object-level editing

Scene composition


RGB
Predicted frequency
Local variance
Frequecy intensity
1.0
0.5
0

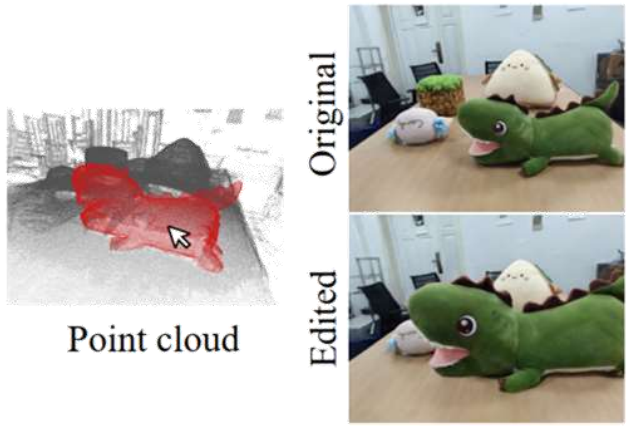

AFNet
hypernetwork
Linear
ReLU
Linear
ω, φ
X
Linear
AF Layer
Linear
AF Layer
Linear
AF Layer
d
Linear
AF Layer
Linear
Output feature and raw RGB
sin(ωX^(i) + φ)
ω = 2.4
ω = 0.3
Ground truth

Motivation

Why point clouds rendering?

- Advantage
  - Explicit geometry enables single sampling and benefits ***acceleration***.
  - Explicit geometry benefits user selection in ***object/scene editing/composition***.

Real-time high-fidelity rendering

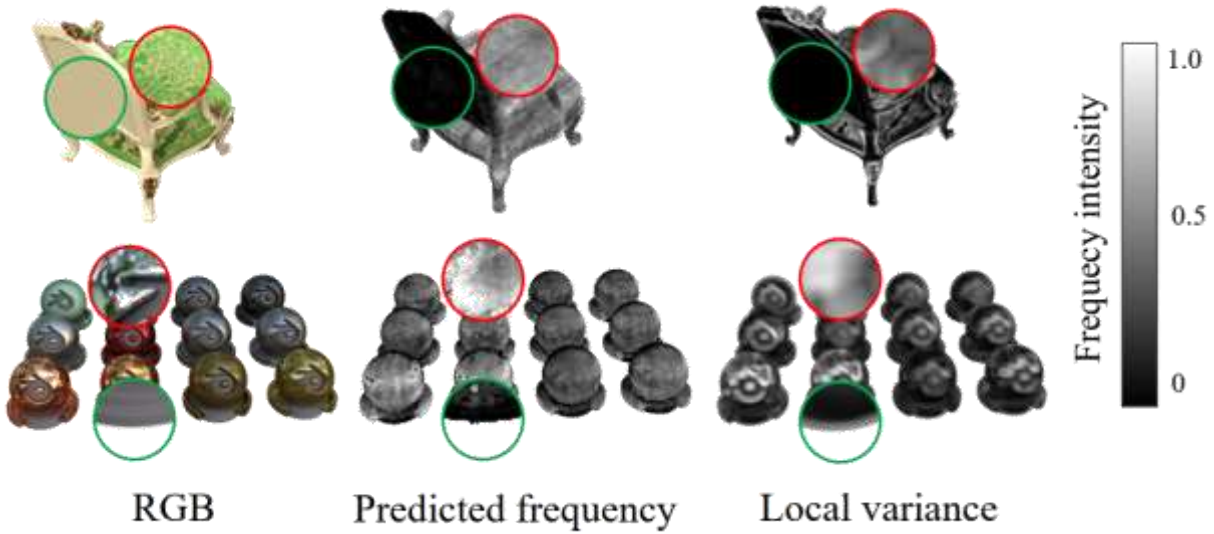


Object-level editing

Why Frequency-Modulated?

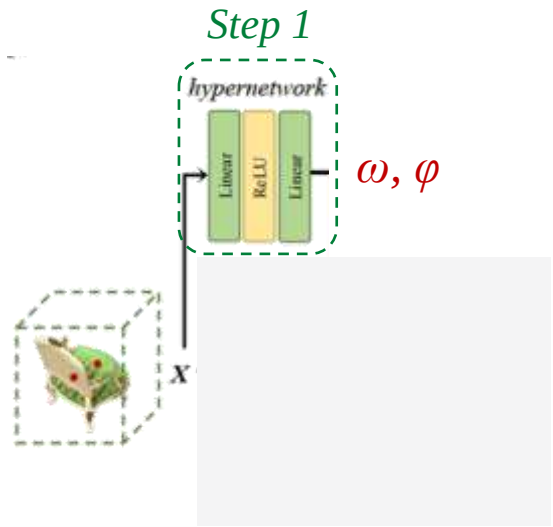
The texture frequencies of a 3D scene are region-dependent.

Weak-textured and strong-textured regions should be processed by neural networks adaptively.



Method: Frequency Modulation

Step 1: Encode texture frequency ω, φ from the query coordinate.

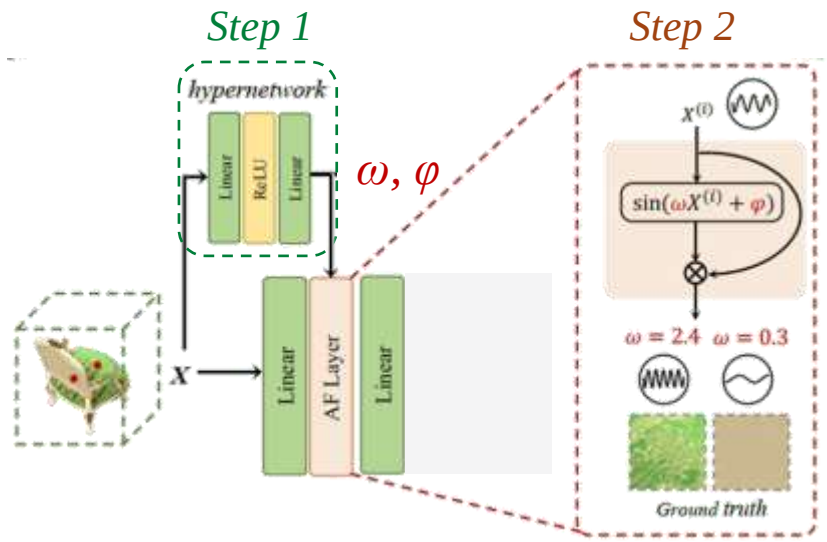


AF Layer

Method: Frequency Modulation

Step 1: Encode texture frequency ω, φ from the query coordinate.

Step 2: Inject frequency using a sine function: $\mathbf{Y}^{(i)} = \mathbf{X}^{(i)} \circ \sin(\omega^{(i)} \circ \mathbf{X}^{(i)} + \phi^{(i)})$.

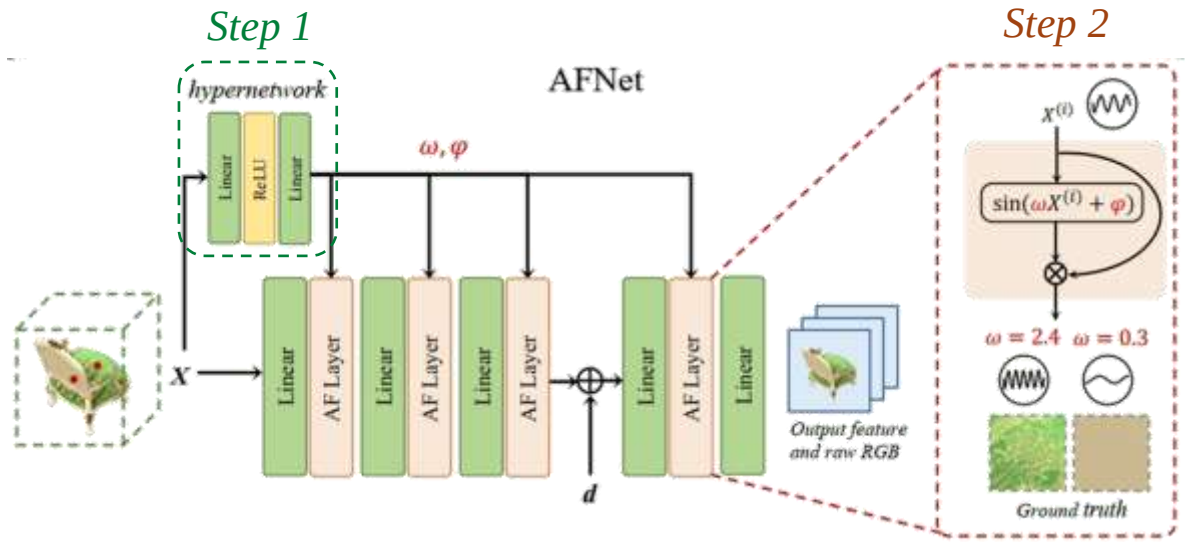


Method: Frequency Modulation

Step 1: Encode texture frequency ω, φ from the query coordinate..

Step 2: Inject frequency using a sine function: $\mathbf{Y}^{(i)} = \mathbf{X}^{(i)} \circ \sin(\omega^{(i)} \circ \mathbf{X}^{(i)} + \phi^{(i)})$.

AFNet: hypernetwork + stacks of {AF layers and linear layers}



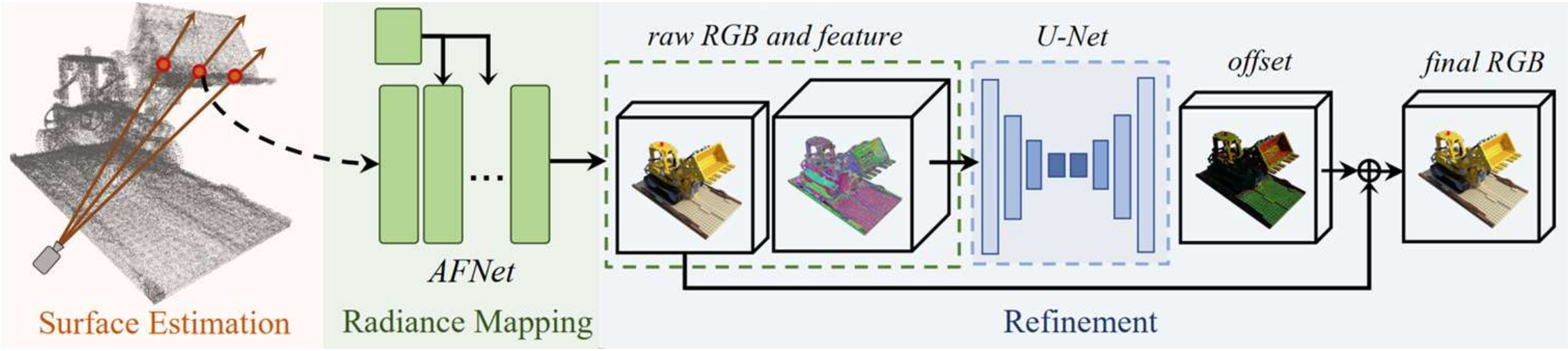
AF Layer

Method: Pipeline

We build the pipeline upon our prior work:

Boosting Point Clouds Rendering via Radiance Mapping. Xiaoyang Huang*, Yi Zhang*, Bingbing Ni, et al. *AAAI 2023.*

Consisted of three steps:



Method: Editing

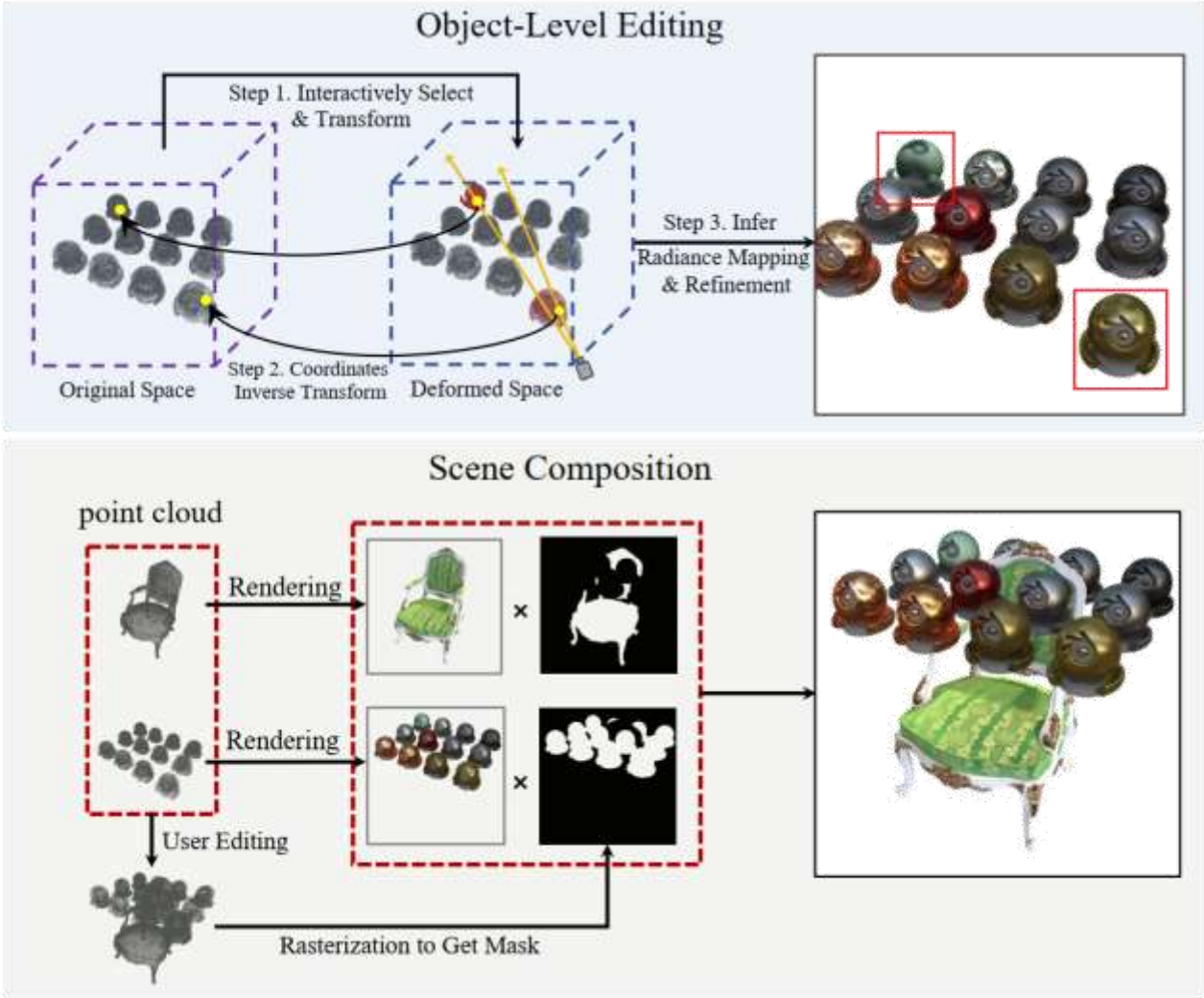
- Object-level Editing

  Coordinate inverse transformation.

  Translation, scaling, rotation.

- Scene Composition

  No need for cross-scene training.

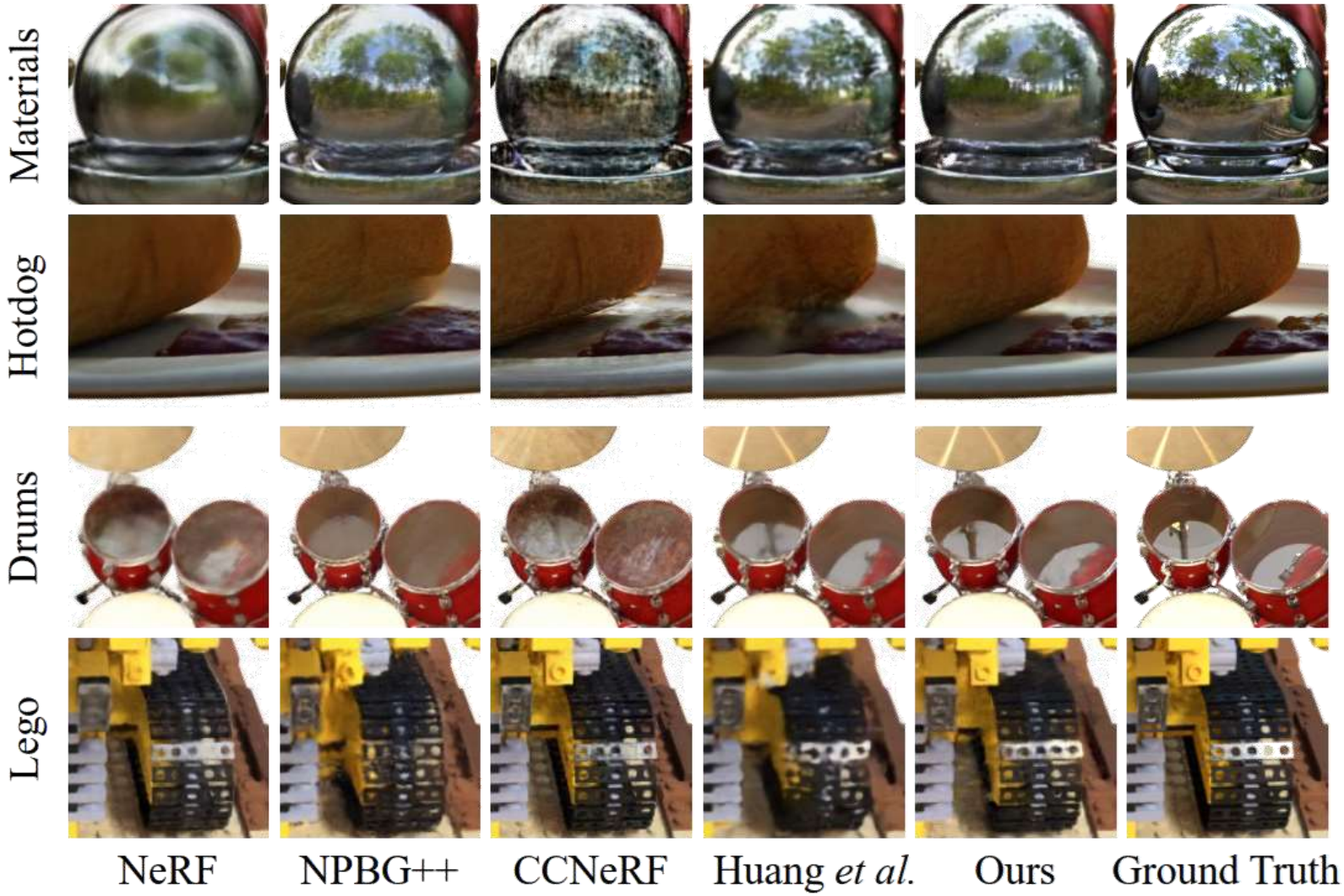


Experiments: Benchmark Results

			Editing Ability		NeRF-Synthetic			Tanks&Temples		
Method	Size(MB)	FPS	Object	Scene	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
NeRF [21]	5.0	0.023	✗	✗	31.01	0.947	0.081	25.78	0.864	0.198
NSVF [19]	16.0	0.815	✗	✓	31.75	0.953	0.047	28.40	0.900	0.153
Object-NeRF [52]	121.2	0.1	✓	✗	31.19	0.949	0.079	25.96	0.866	0.194
PlenOctrees [55]	1976.3	168	✗	✓	31.71	0.958	0.053	27.99	0.917	0.131
Point-NeRF [49]	20.0	0.125	✗	✗	33.00	0.978	0.055	29.61	0.954	0.115
Plenoxels [8]	778.1	15	✗	✓	31.71	0.958	0.049	27.43	0.906	0.142
TensoRF [3]	71.8	1.15	✗	✓	33.14	0.963	0.056	28.56	0.920	0.118
Instant-NGP [22]	63.3	60	✗	✗	33.18	0.963	0.050	28.78	0.925	0.113
CCNeRF-CP [44]	4.4	1.05	✗	✓	30.55	0.935	0.076	27.01	0.879	0.180
CCNeRF-HY-S [44]	68.9	1.05	✗	✓	31.22	0.947	0.074	27.53	0.901	0.177
NPBG [1]	44.2	33.4	✓	✓	28.10	0.923	0.077	25.97	0.889	0.137
NPBG++ [33]	28.6	35.4	✓	✓	28.12	0.928	0.076	26.04	0.892	0.130
Huang *et al.* [13]	18.5	39.1	✓	✓	28.96	0.932	0.061	26.35	0.893	0.130
Ours	11.8	39.3	✓	✓	31.24	0.950	0.049	27.79	0.902	0.125

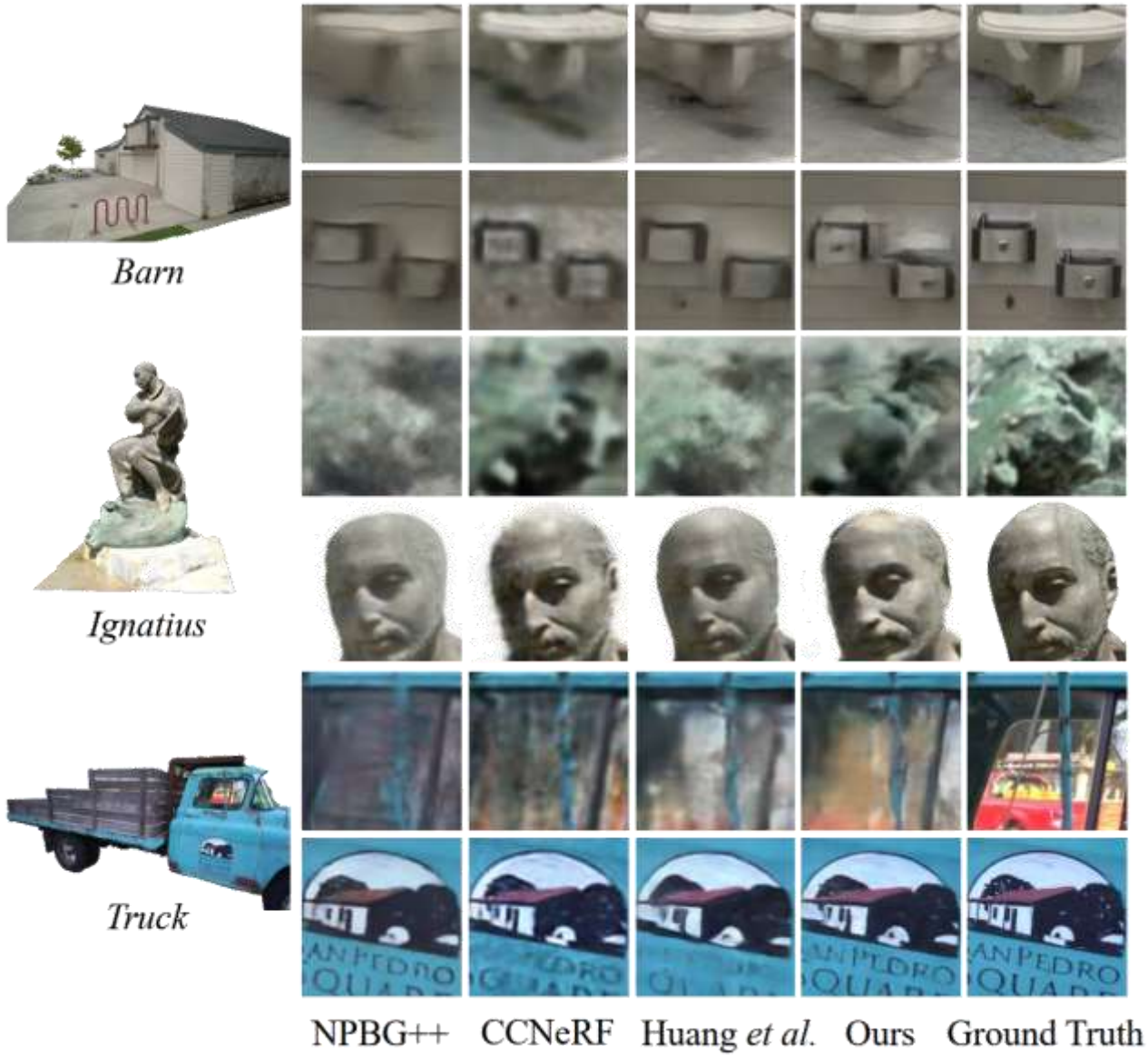
Experiments: Static scenes

- NeRF-Synthetic

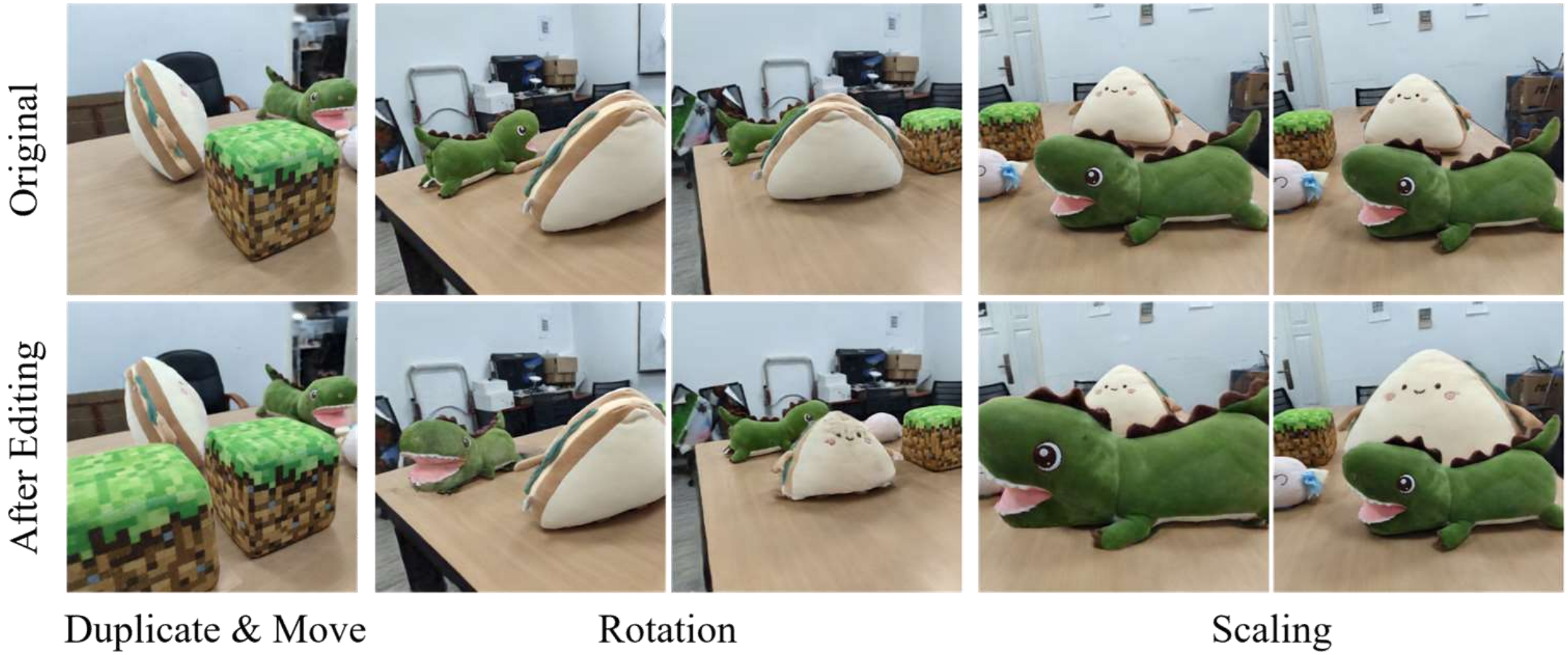


Experiments: Static scenes

- Tanks and Temples



Experiments: Object Editing



Experiments: Scene Composition

Conclusion

- We develop a novel point cloud rendering pipeline which enables:
  - high fidelity reconstruction
  - real-time rendering
  - user-friendly editing
- Experiments on major benchmarks demonstrate the proposed method outperforms existing point cloud rendering methods and achieves the state-of-the-art.

Check out the code for this study

https://github.com/yizhangphd/FreqPCR

Thanks for Listening

WeChat

Yi Zhang *Xiaoyang Huang*

饮 水 思 源 爱 国 荣 校