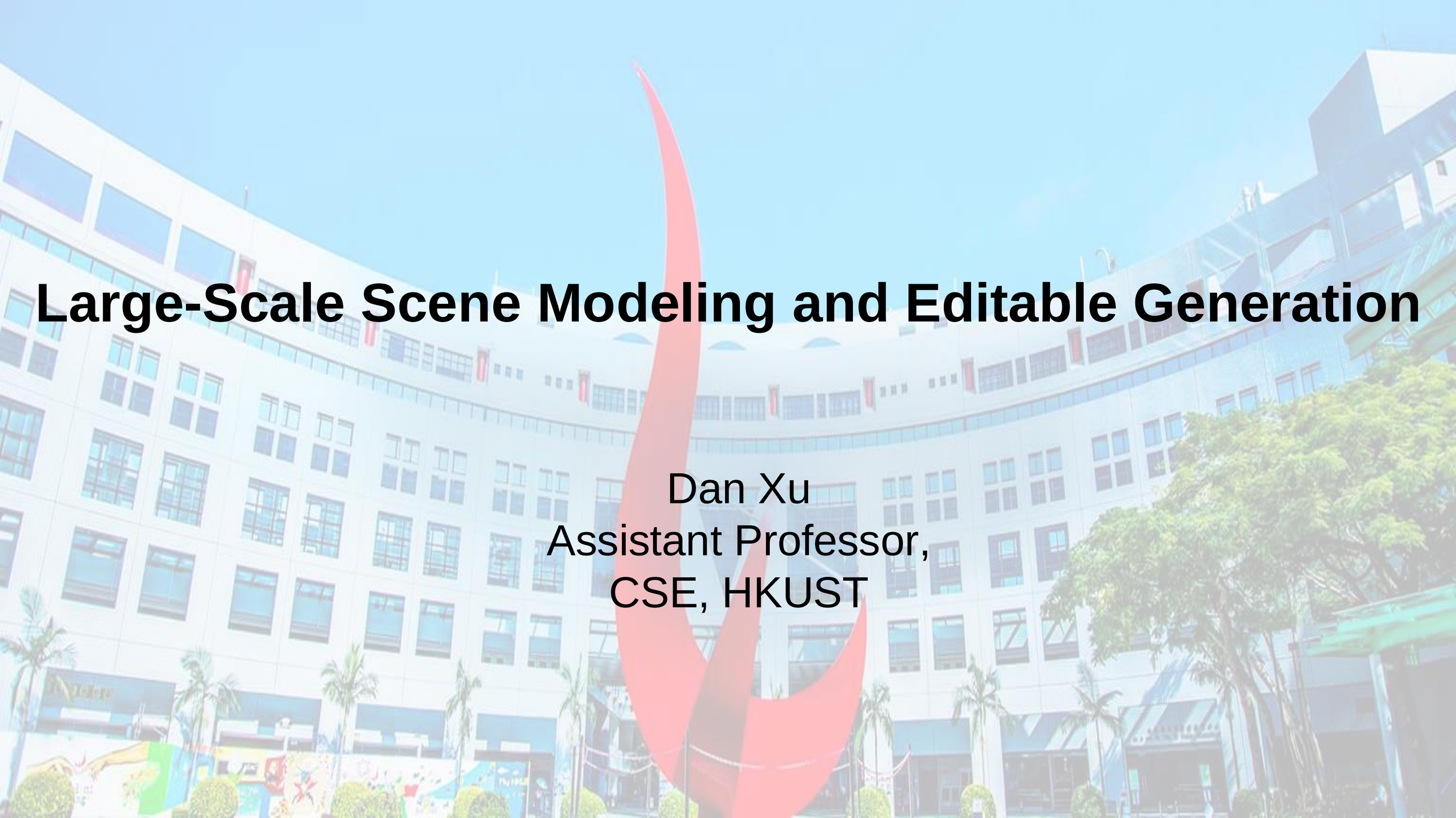


Large-Scale Scene Modeling and Editable Generation

Dan Xu
Assistant Professor,
CSE, HKUST

Research Background

THE DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
計算機科學及工程學系

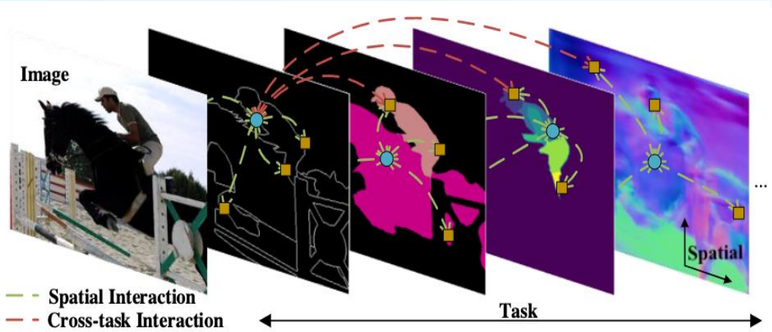
- Research Interests

Computer Vision & Deep Learning

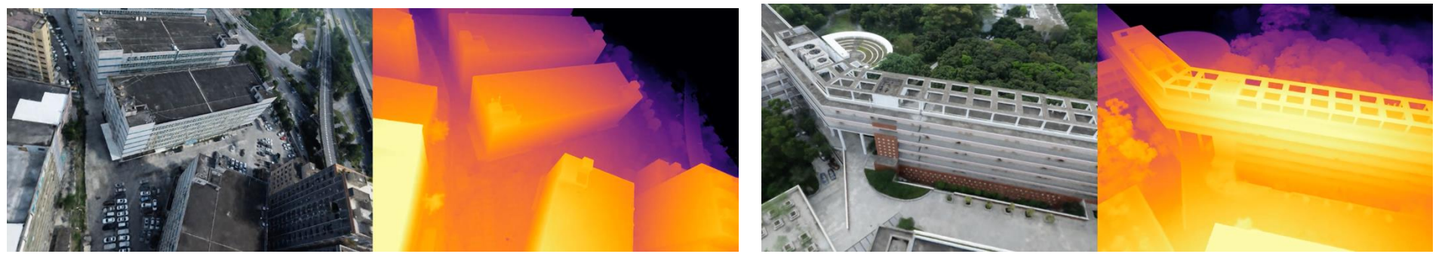
Deep Structured Feature Learning and Prediction

Deep Multi-Modal, Multi-Task Learning

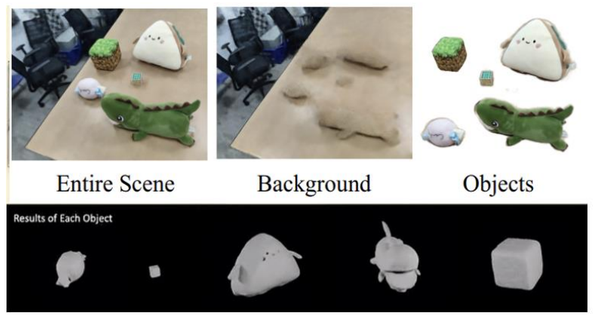
2D/3D Holistic Visual Scene Understanding and Generation



#1 Multi-Modal Multi-task Scene Perception



#2 Large-scale Scene Reconstruction



#3 Editable Scene Generation

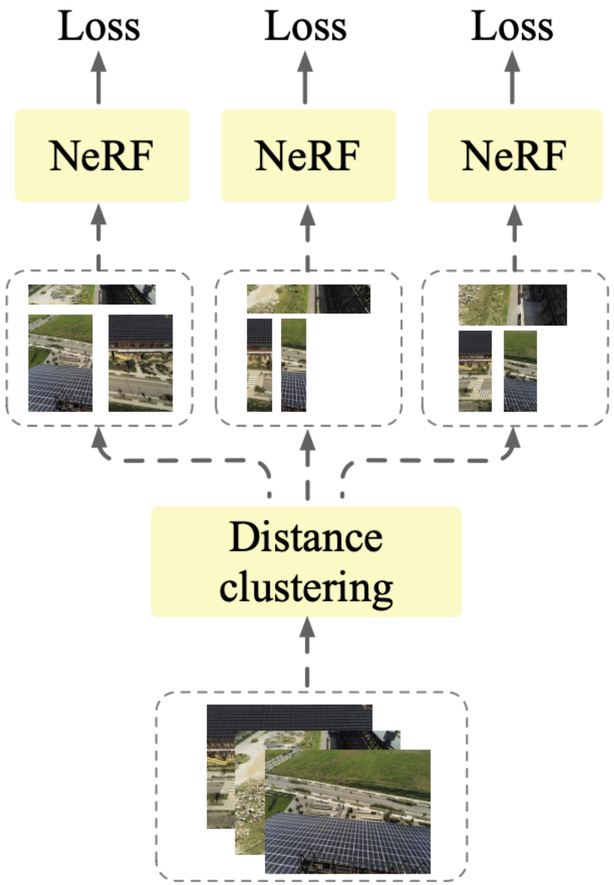
Large-scale scene modeling

- NeRF is an effective scene representation (multi-view posed images)
- Very challenging if dealing with a large-scale scene

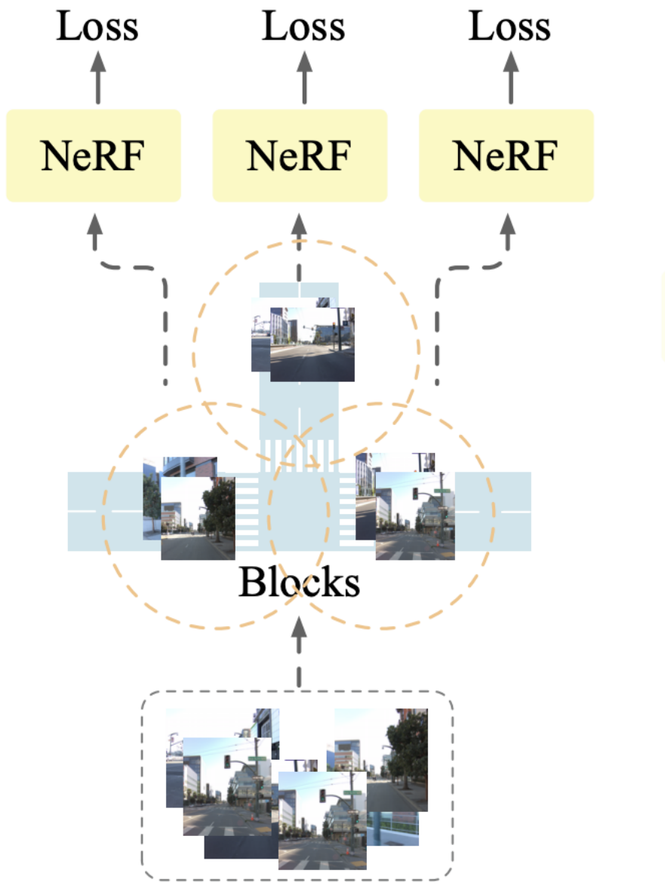
Matthew Tancik et al., BlockNeRF: Scalable Large Scene Novel View Synthesis, CVPR 2023

Large-scale scene modeling

- **Scene decomposition** is critical for efficiency and flexibility



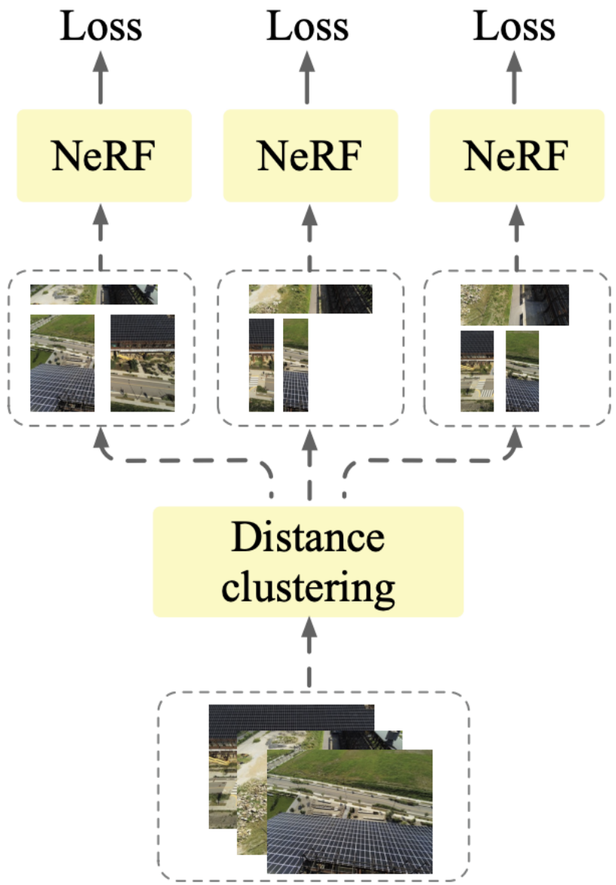
(a) Learning after distance-based decomposition (*e.g.* Mega-NeRF)



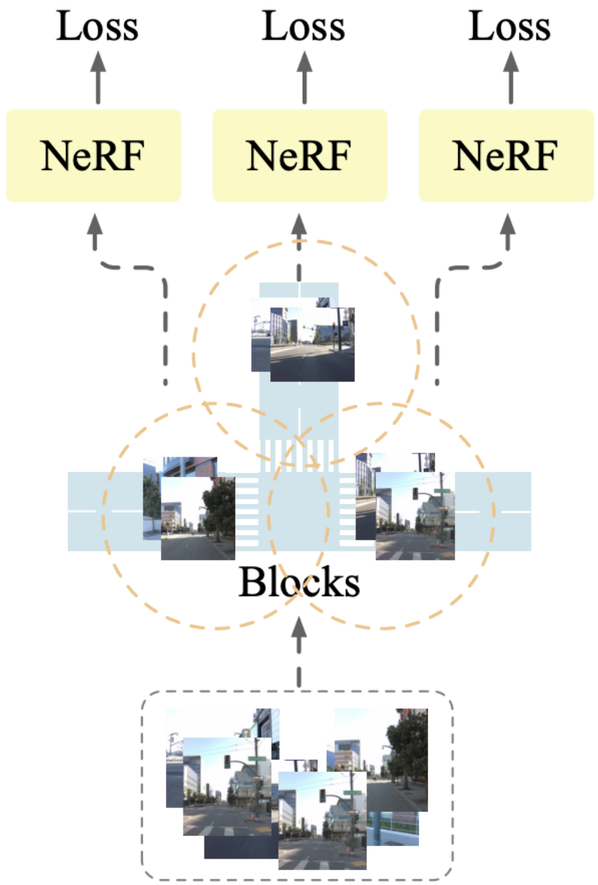
(b) Learning after physical-distribution-based decomposition (*e.g.* Block-NeRF)

Large-scale scene modeling

- **Scene decomposition** is critical for efficiency and flexibility



(a) Learning after distance-based decomposition (*e.g.* Mega-NeRF)



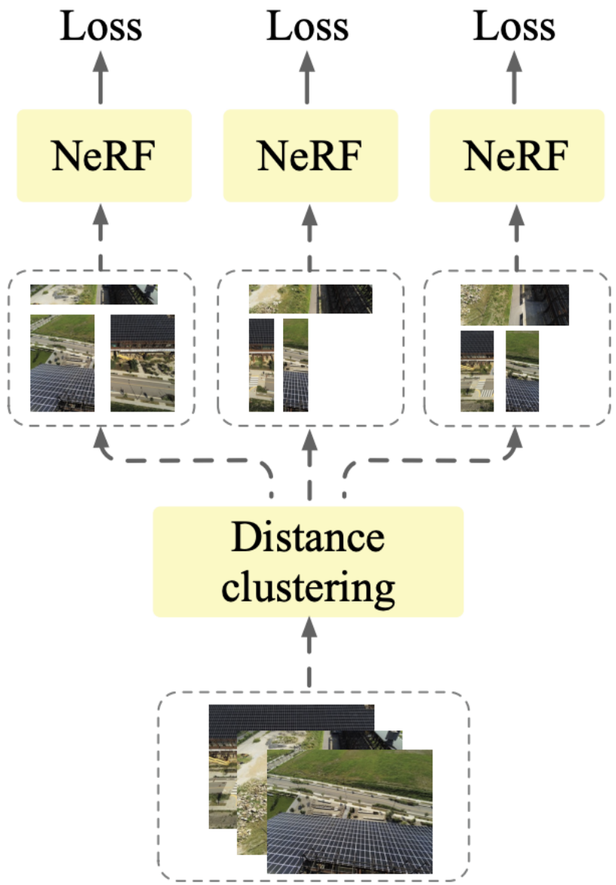
(b) Learning after physical-distribution-based decomposition (*e.g.* Block-NeRF)

Some issues:

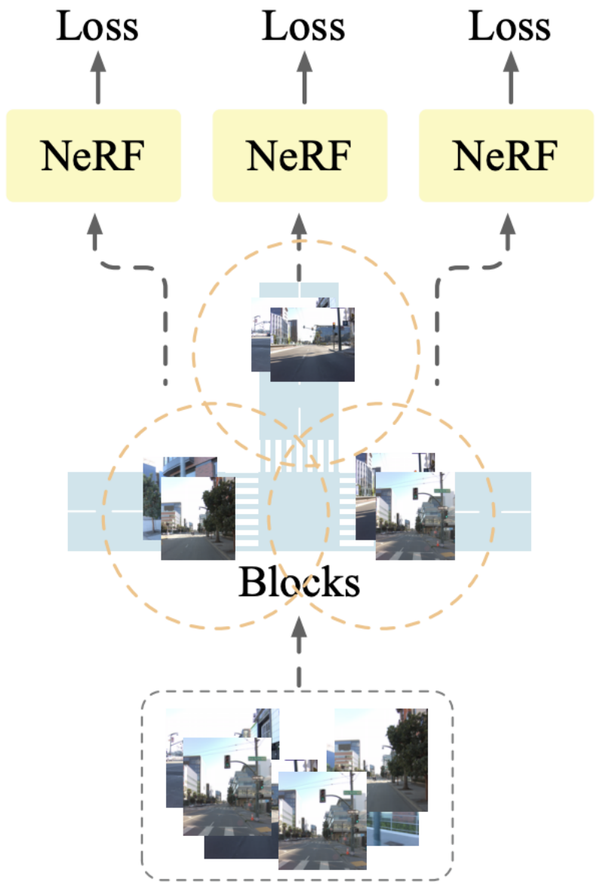
- Hand-crafted
- Need priors about the scene geometry
- Fusion of different models is sub-optimal

Large-scale scene modeling

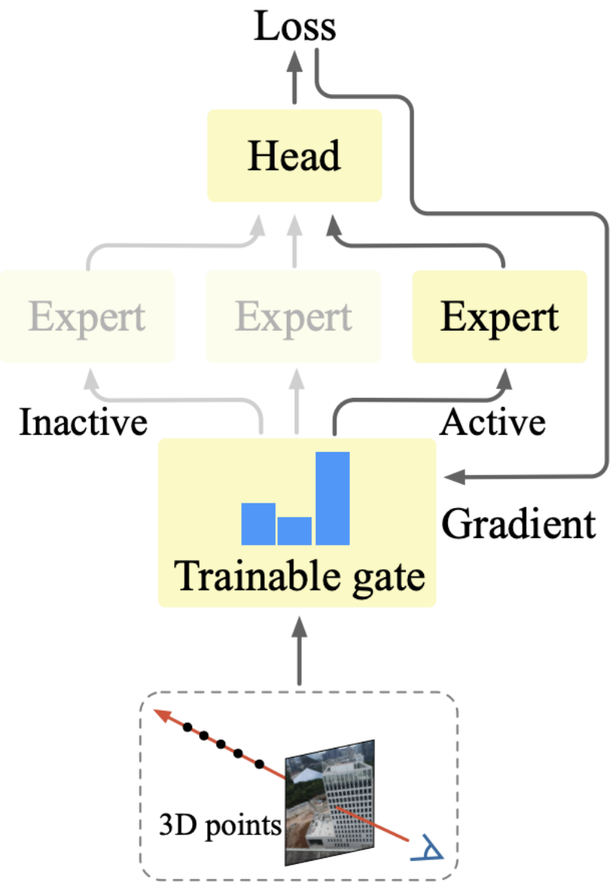
- **Scene decomposition** is critical for efficiency and flexibility



(a) Learning after distance-based decomposition (*e.g.* Mega-NeRF)



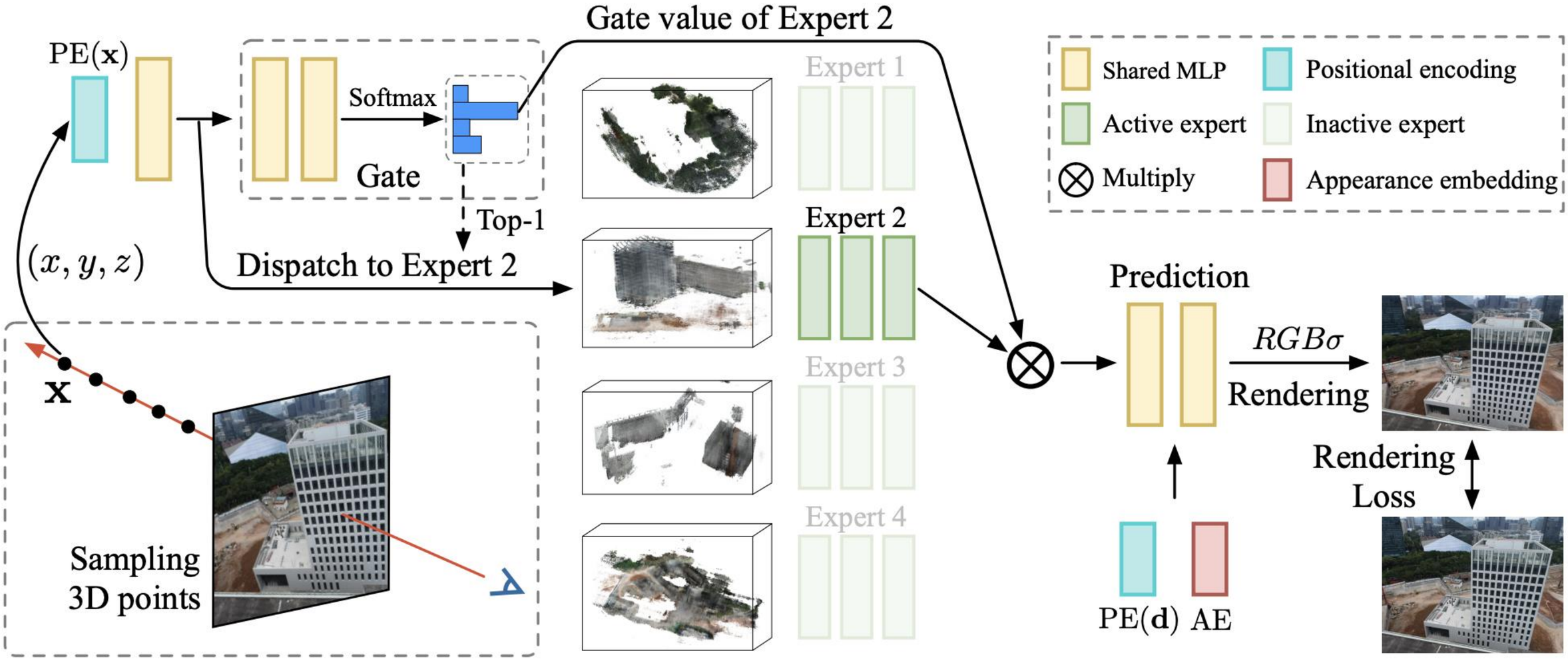
(b) Learning after physical-distribution-based decomposition (*e.g.* Block-NeRF)



(c) Learning with scene decomposition (Ours)

Large-scale scene modeling

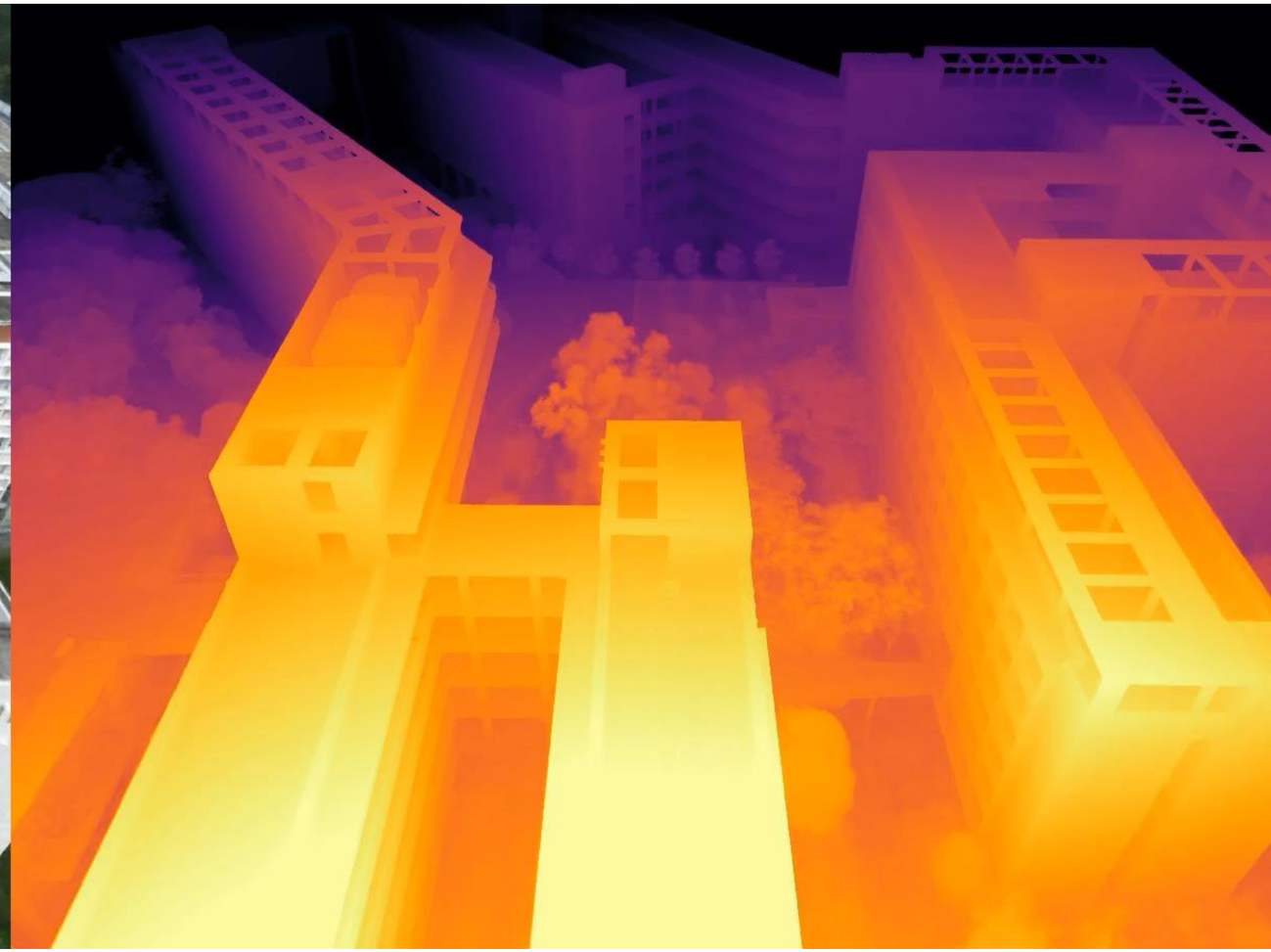
- **Framework Overview**



Zhenxing Mi, Dan Xu, 'Switch-NeRF: Learning Scene Decomposition with Mixture of Experts for Large-Scale Neural Radiance Fields', ICLR 2023

Rendering Results

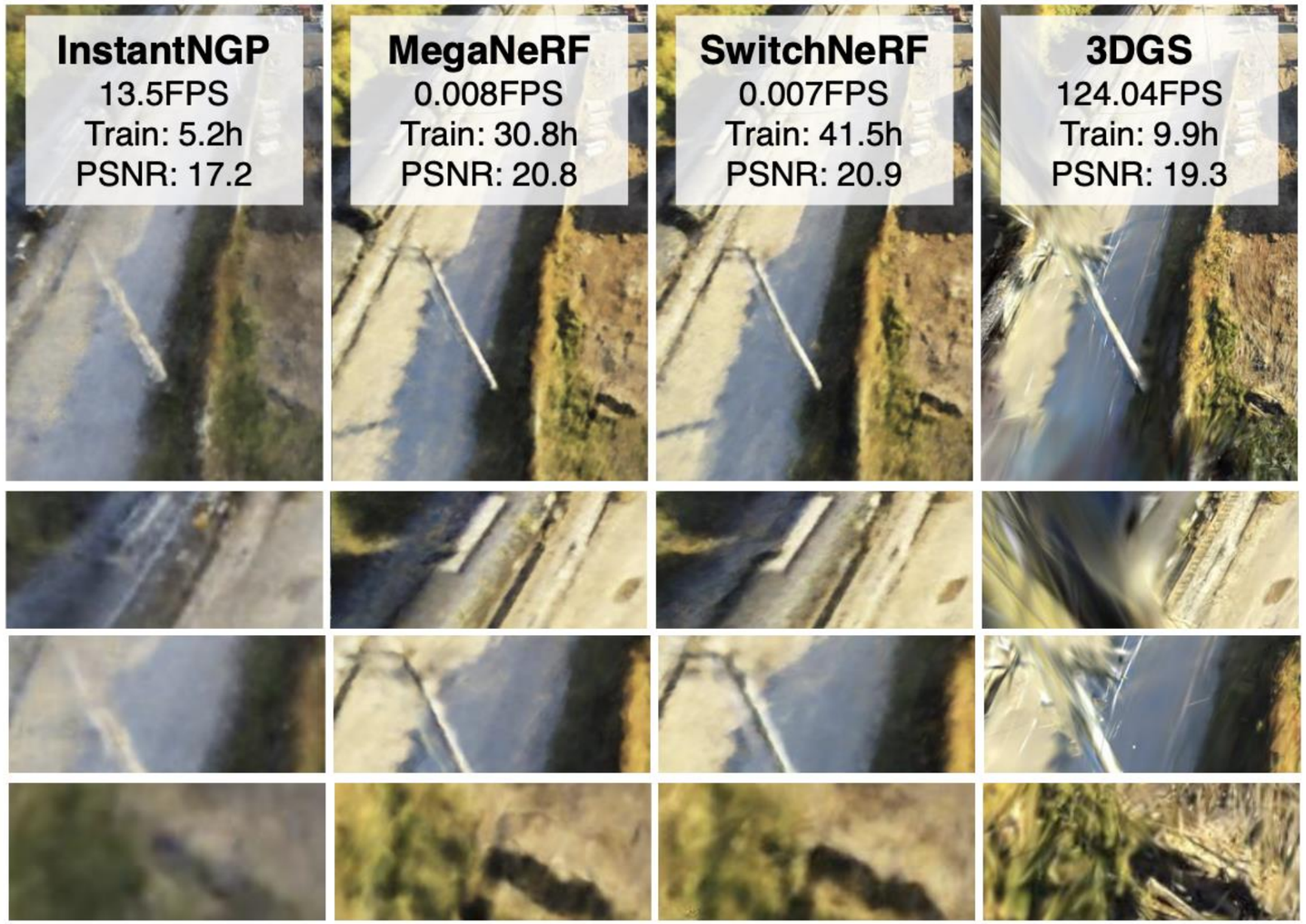
Novel view synthesis



Geometry rendering

Extend NeRF to Gaussian Splatting

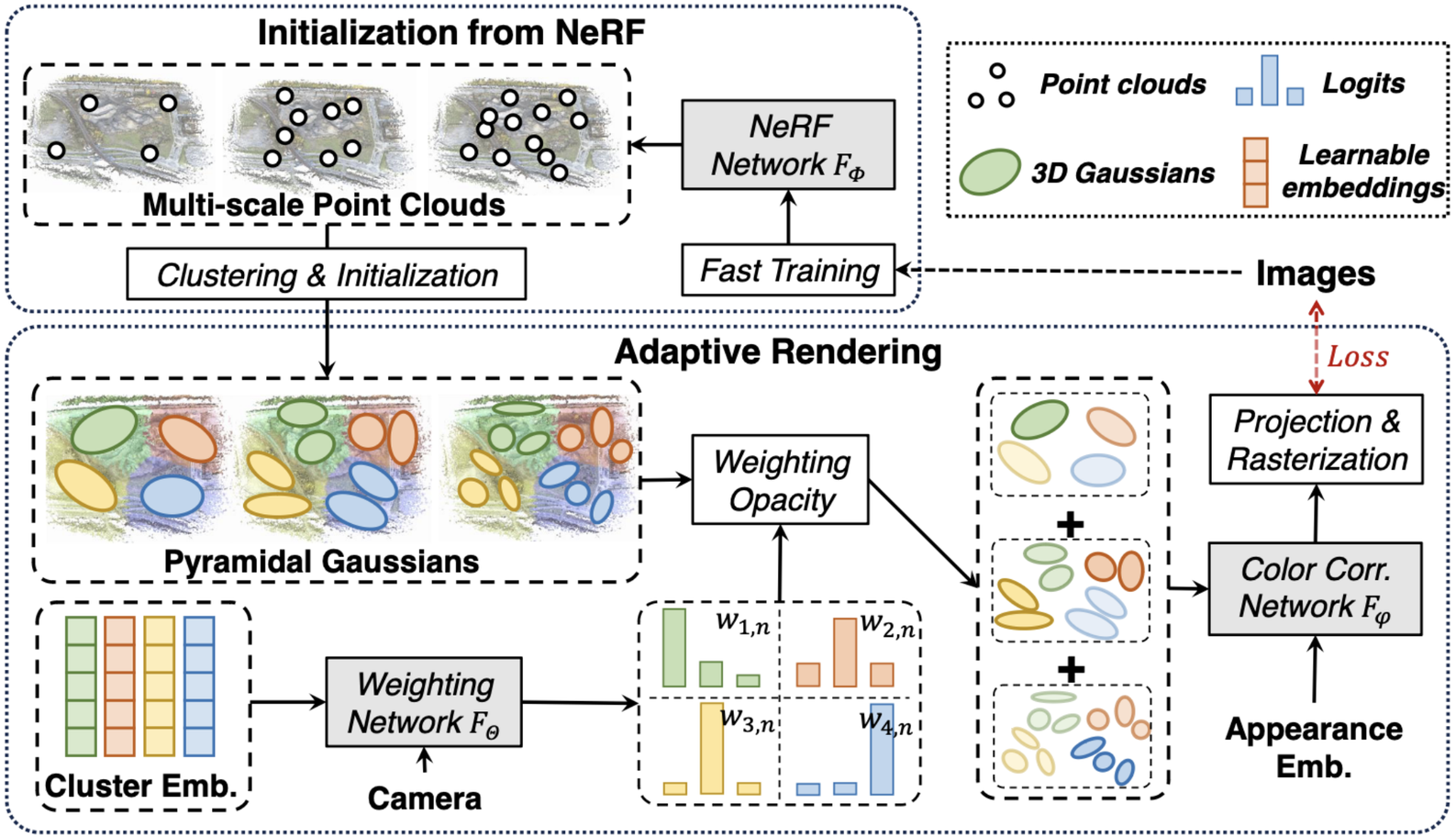
- **Some issues of NeRF for large-scale scene modeling**



- Tend to capture low-frequency element due to neural network's spectral biases, missing details
- Ray-based rendering, not able to achieve high-resolution dense rendering
- Rendering speed: significantly lower (~10 times) than Gaussian splatting techniques

Extend NeRF to Gaussian Splatting

- **Our proposal: Pyramid Gaussian Splatting**



- Replace COLMAP with with InstanceNGP (only a few minutes) for fast initialization
- Construction Multi-scale Pyramid Gaussian Blob scene representations
- Adaptive rendering based on a weighting network for automatic scale fusion

Zipeng Wang, Dan Xu, PyGS: Large-scale Scene Representation with Pyramidal 3D Gaussian Splatting, arXiv 2405.16829

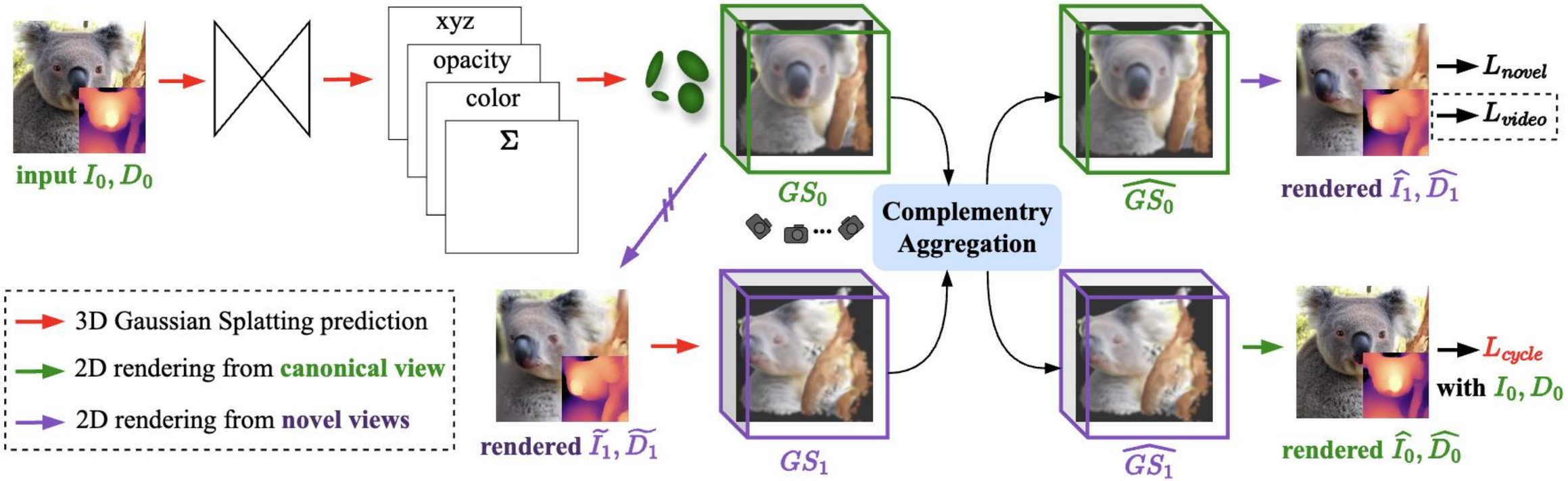
Extend NeRF to 3D Gaussian Splatting

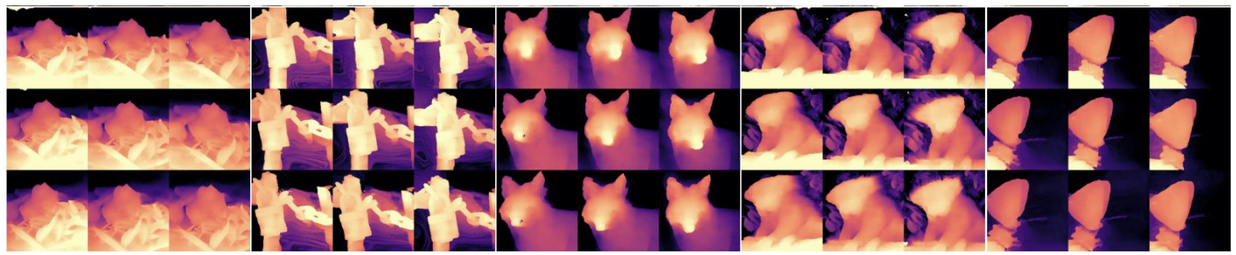
BungeeNeRF

PyGS (Ours)

Generalizable Single-view 3D Modeling

- Feedforward 3D Gaussian Modeling Framework





Yuxin Wang, Qianyi Wu, Dan Xu, 'F3D-Gaus: Feed-forward 3D-aware Generation on ImageNet with Cycle-Consistent Gaussian Splatting', arXiv 2501.06714

Generalizable Single-view 3D Modeling

Yuxin Wang, Qianyi Wu, Dan Xu, ‘F3D-Gaus: Feed-forward 3D-aware Generation on ImageNet with Cycle-Consistent Gaussian Splatting’, arXiv 2501.06714

Generalizable Single-view 3D Modeling

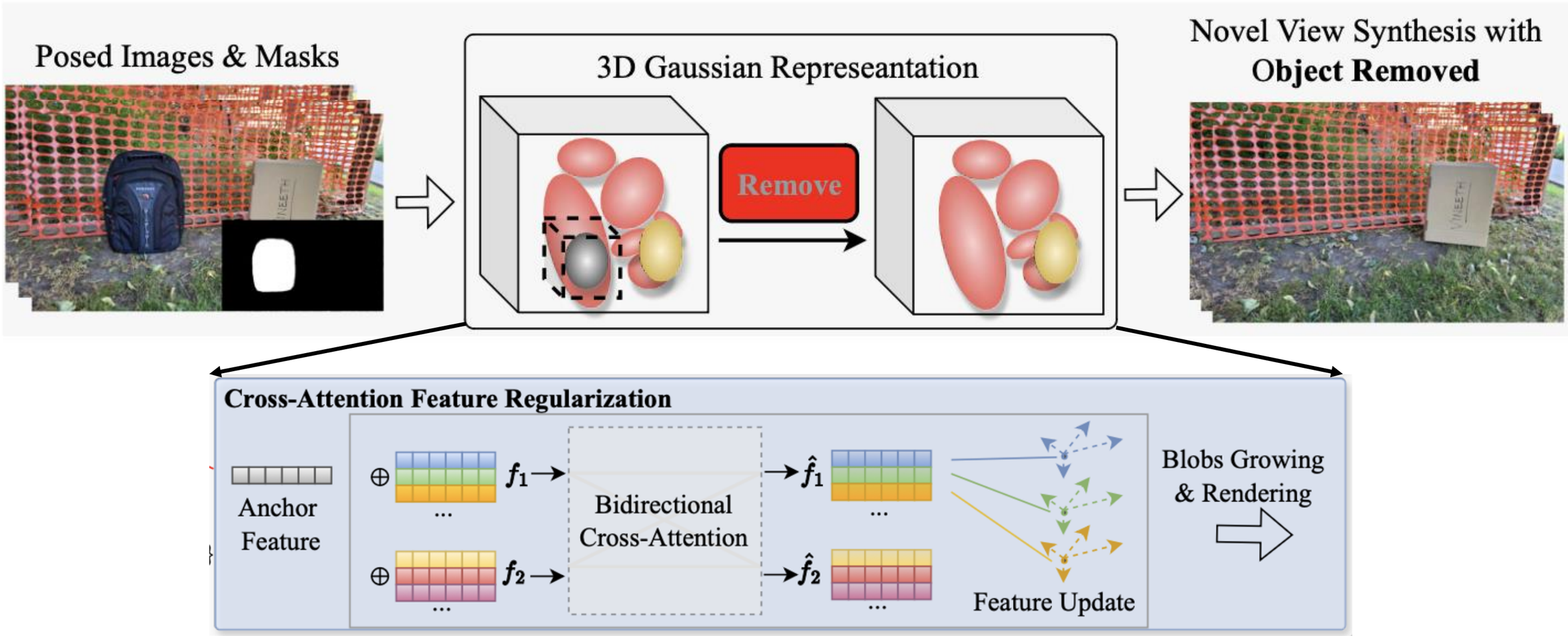





Yuxin Wang, Qianyi Wu, Dan Xu, ‘F3D-Gaus: Feed-forward 3D-aware Generation on ImageNet with Cycle-Consistent Gaussian Splatting’, arXiv 2501.06714

Explicit Editing with Gaussian Splatting

- Replace the implicit NeRF with explicit GS representations



Yuxin Wang, Qianyi Wu, Guofeng Zhang, Dan Xu, 'Learning 3D Geometry and Feature Consistent Gaussian Splatting for Object Removal', ECCV 2024

Some Qualitative Results

Yuxin Wang, Qianyi Wu, Guofeng Zhang, Dan Xu, ‘Learning 3D Geometry and Feature Consistent Gaussian Splatting for Object Removal’, ECCV 2024

Comparison with 3D Gaussian Splatting-based Methods

Scene-5

GaussianEditor[3]

GScream (Ours)

Scene-6

GaussianEditor[3]

GScream (Ours)

[3] GaussianEditor: Swift and Controllable 3D Editing with Gaussian Splatting (2024 CVPR)

Thank you!