NeRF的基础及后续扩展

彭思达
浙江大学

任务：Image-based rendering

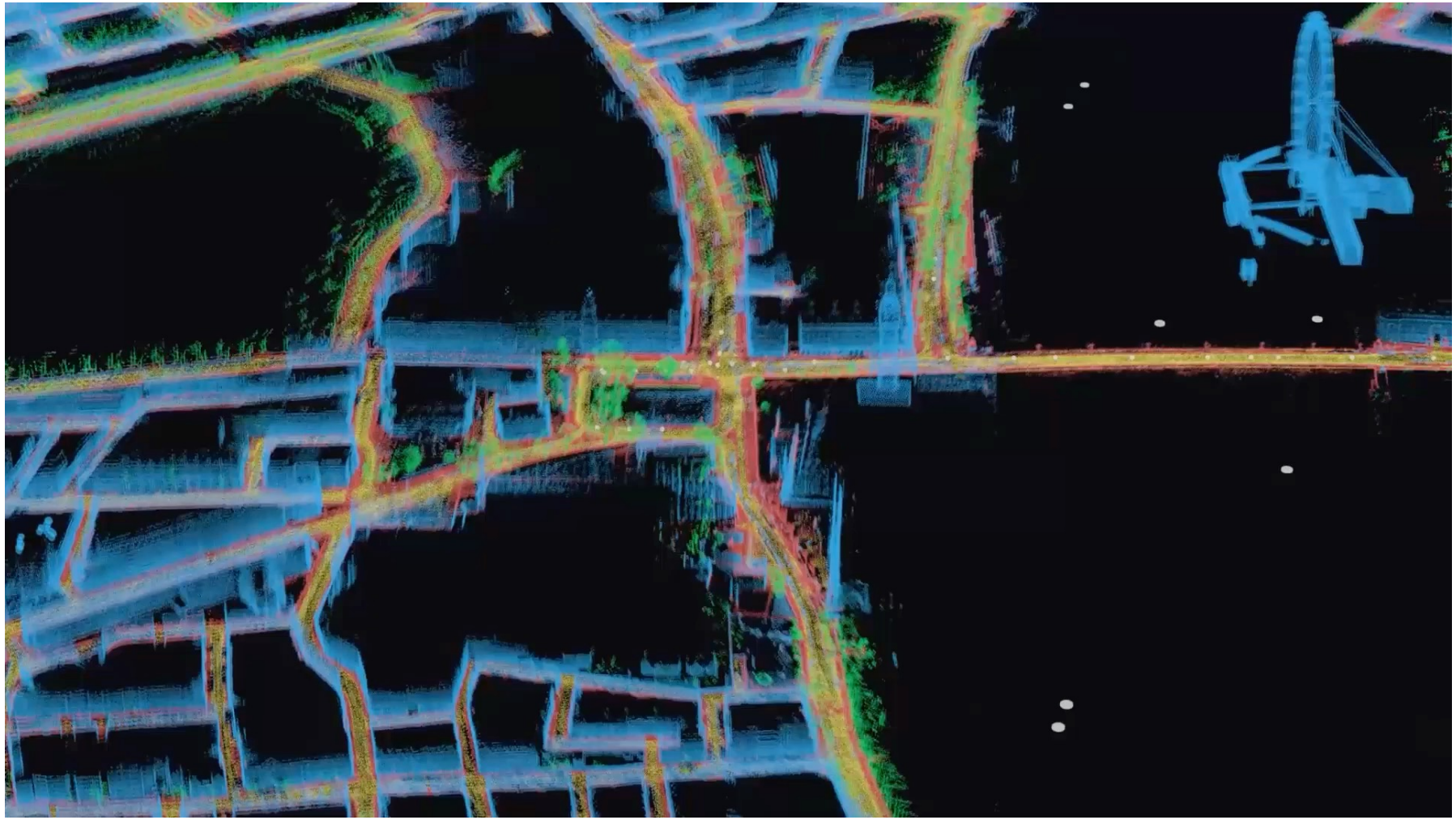
Input images

3D model

Novel views

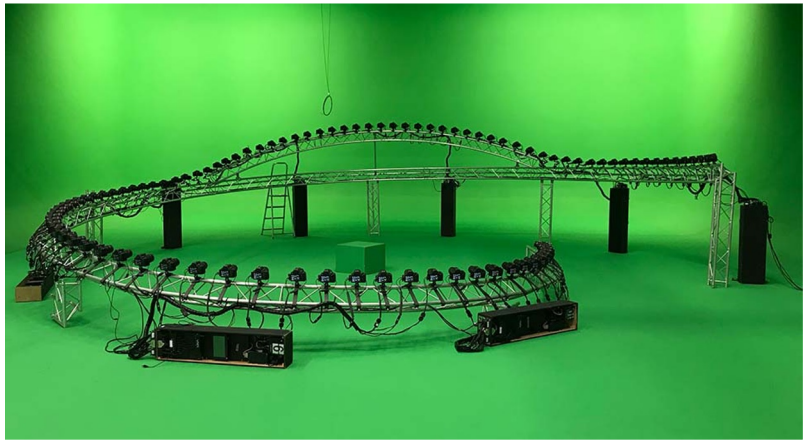
NeRF in the wild: Neural radiance fields for unconstrained photo collections. CVPR. 2021.

Applications: (1) VR城市漫游



Google Immersive View

Applications: (2) 影视特效



Film “Matrix”

Applications: (3) 沉浸式远程观赛

https://www.youtube.com/clip/Ugkx200I8Vsn3ciOS4I38PSXI_gOqevhiN35

Applications: (4) Immersive memory

https://www.youtube.com/watch?v=TX9qSaGXFyg

Applications: (5) Simulator

https://www.youtube.com/watch?v=RVFIDEuNtt0

Image-based rendering的基本流程

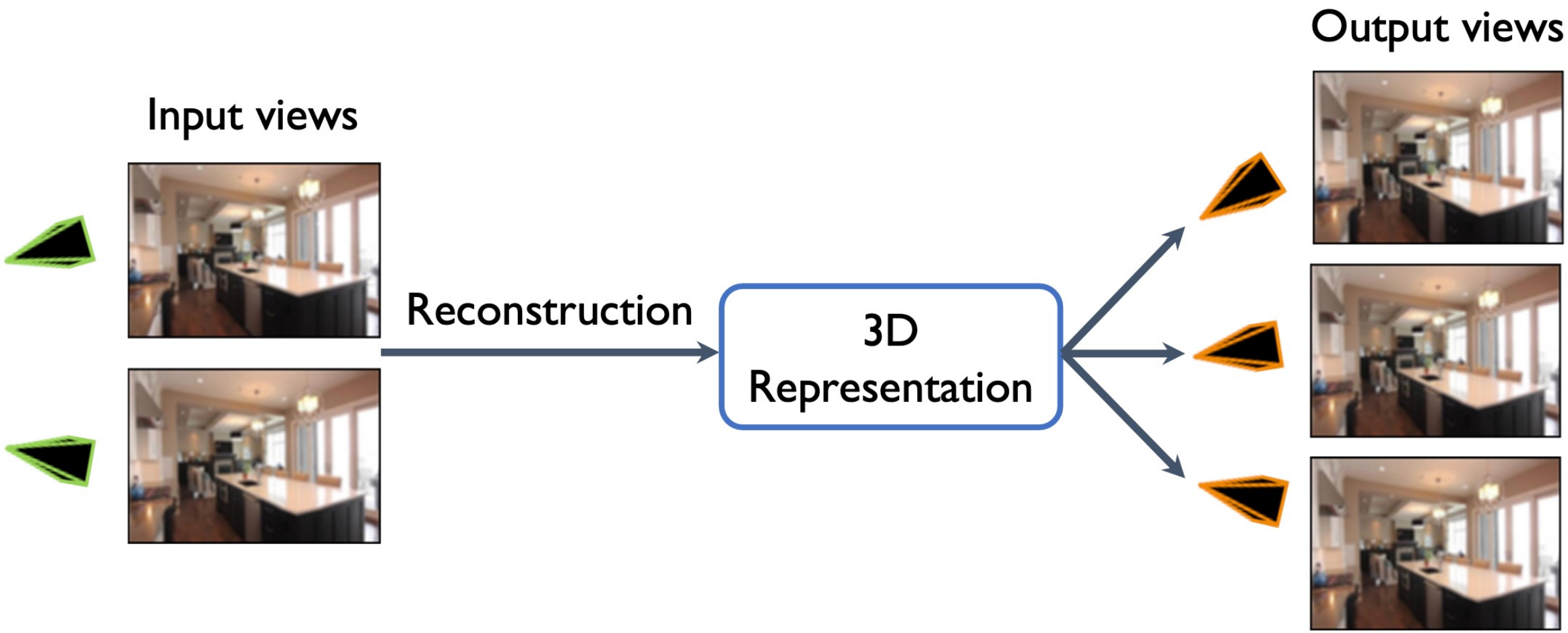
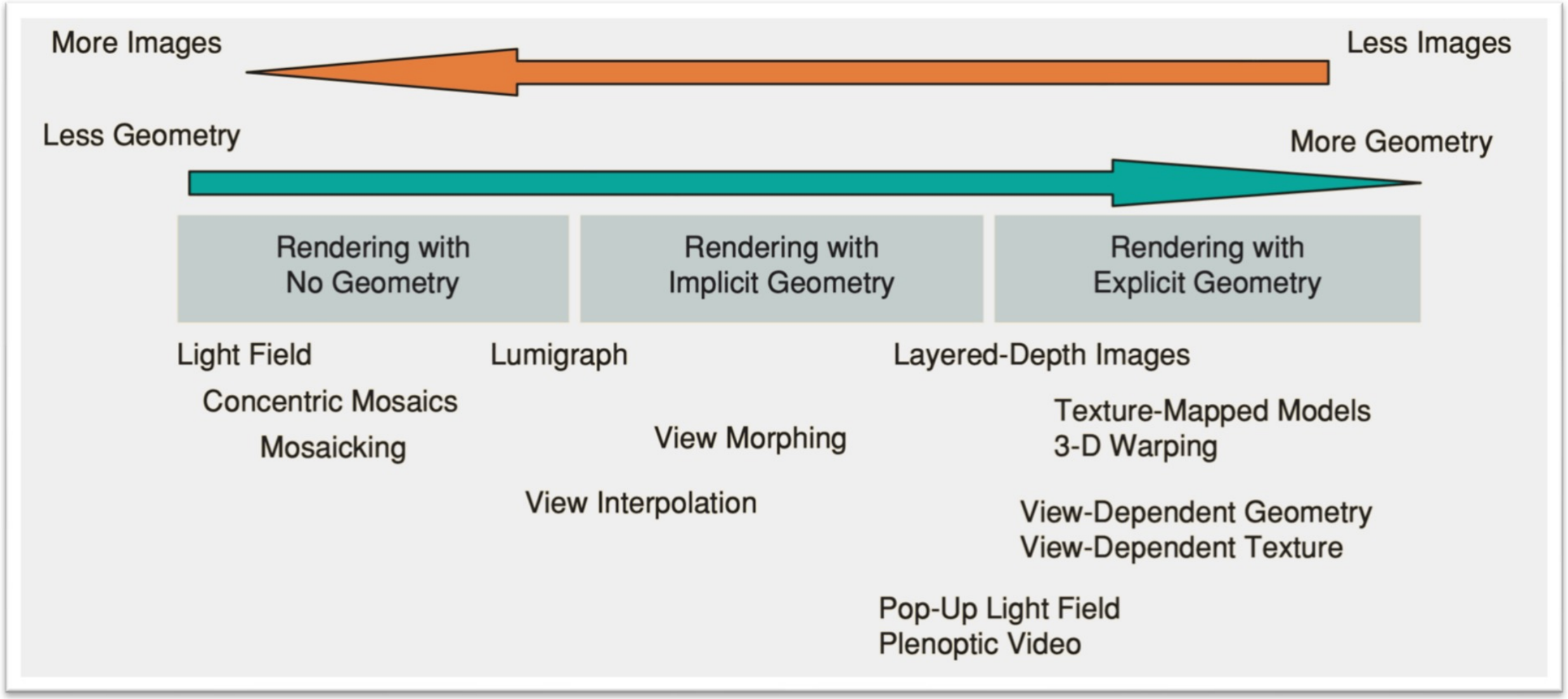


Image credit: Noah Snavely

传统方法



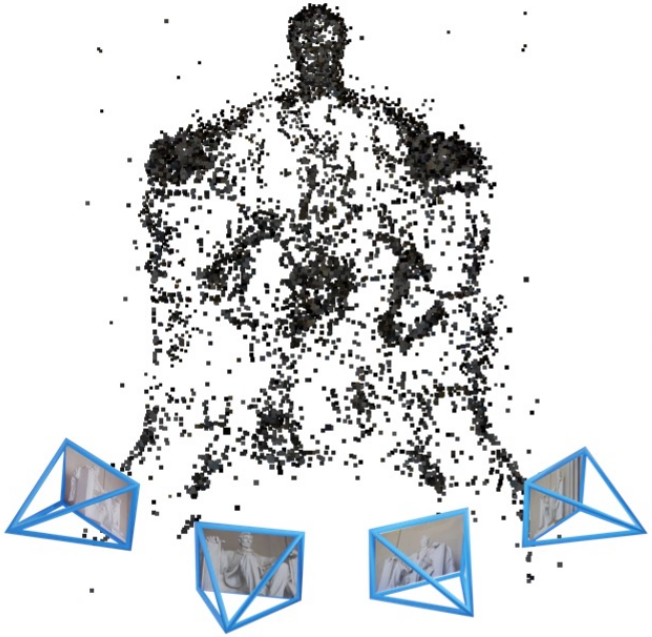
A Review of Image-based Rendering Techniques, Visual Communications and Image Processing 2000.

常用的三维表示1：Surface-based representations

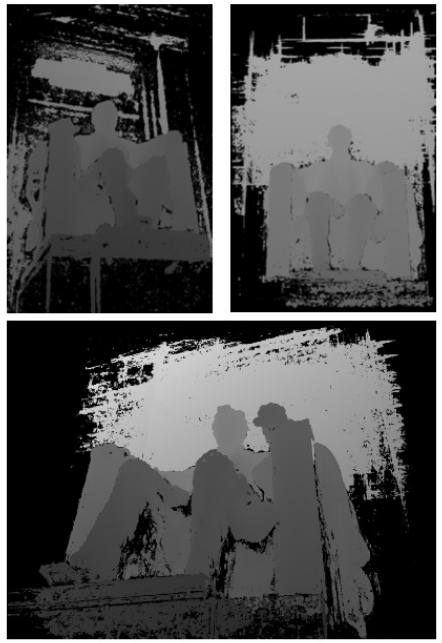
- 带纹理的三维网格

常用的三维表示1：Surface-based representations

- 重建流程：

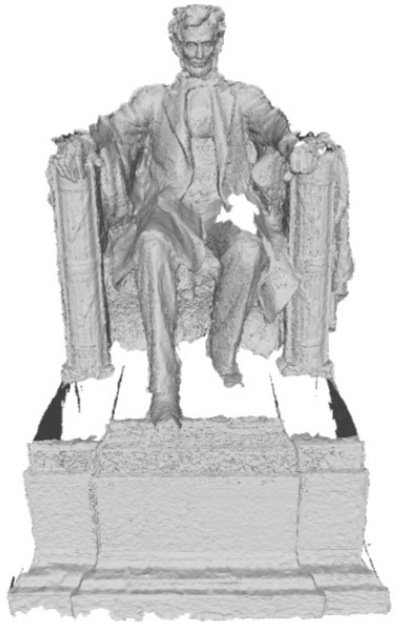


SfM



MVS

Fusion



Mesh extraction

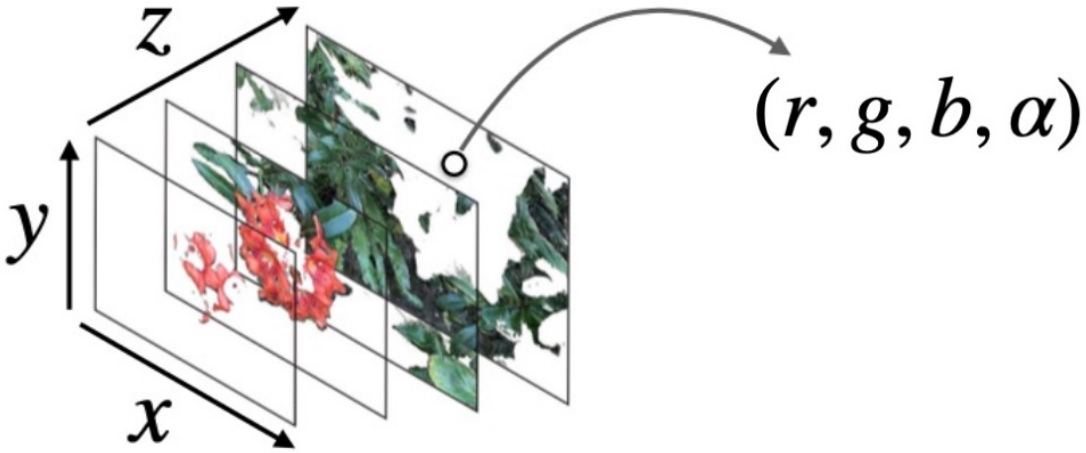
Surface-based representations面临的挑战

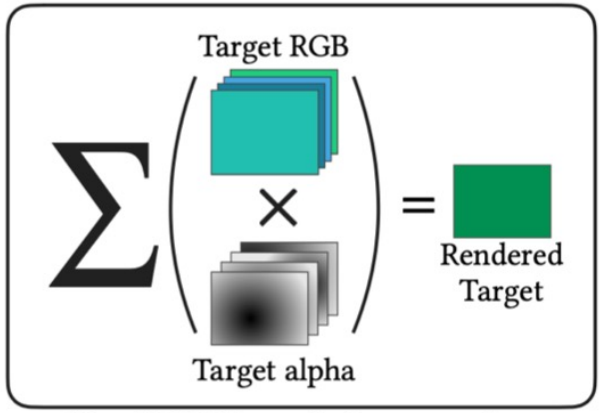
两大挑战：

- 难以重建高质量表面几何。
- 难以表示非常复杂的三维场景。

常用的三维表示2：Volume-based representations

- Multi-Plane image (MPI)





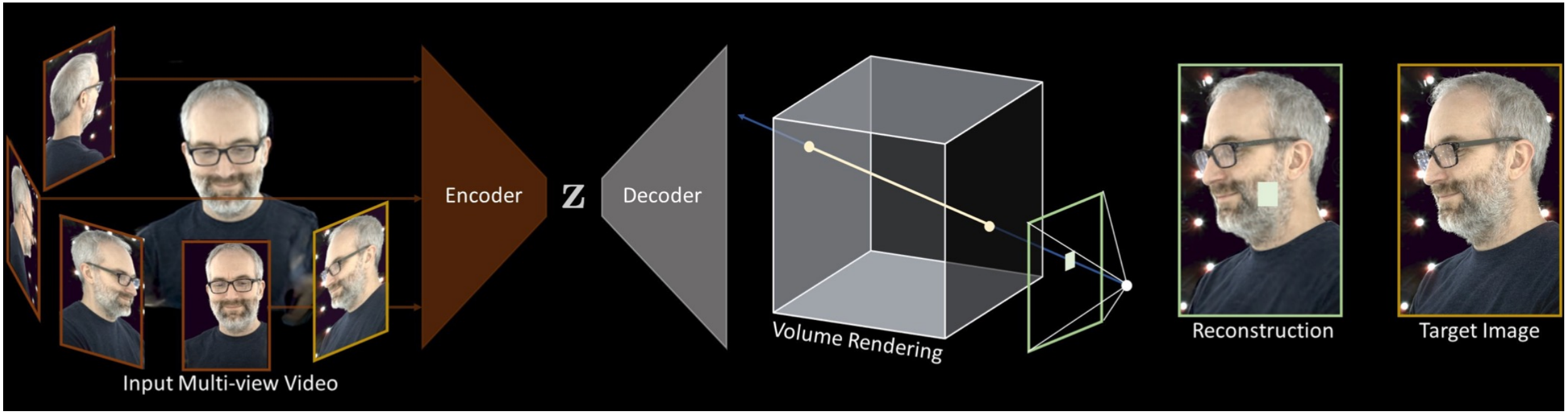
Blend RGBA renderings together to render final output image

Multi-Plane image的效果

https://augmentedperception.github.io/deepview/

常用的三维表示2：Volume-based representations

- RGB-alpha volume



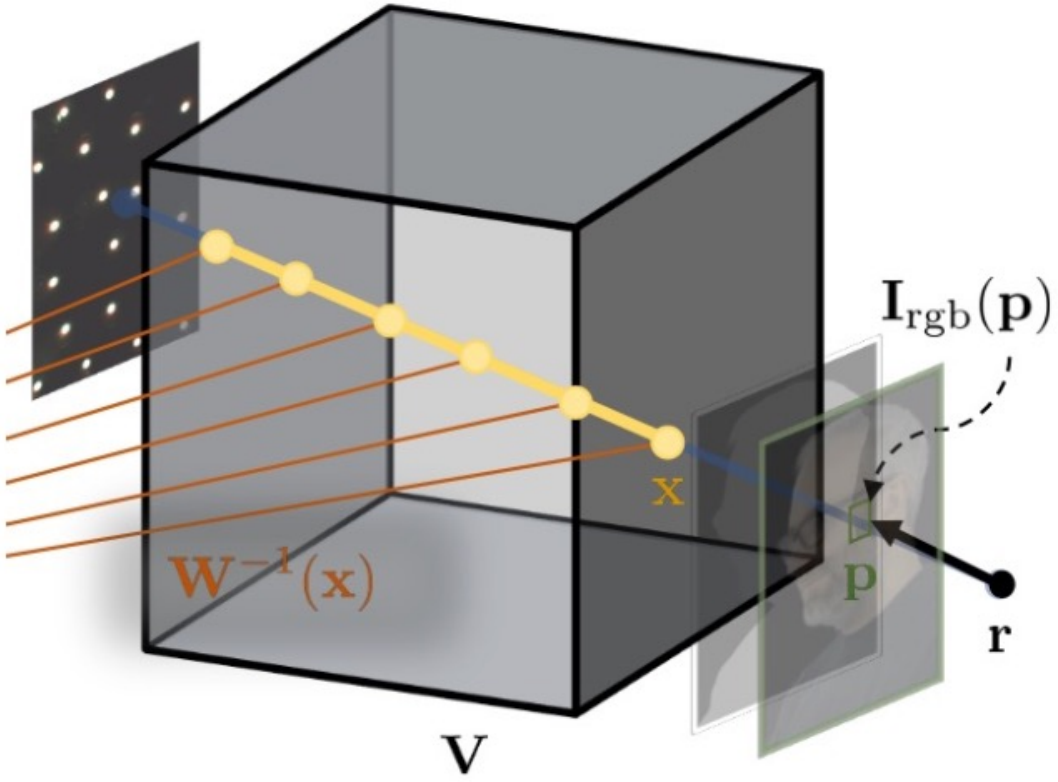
Learning Dynamic Renderable Volumes from Images. SIGGRAPH 2019.

RGB-alpha volume的效果

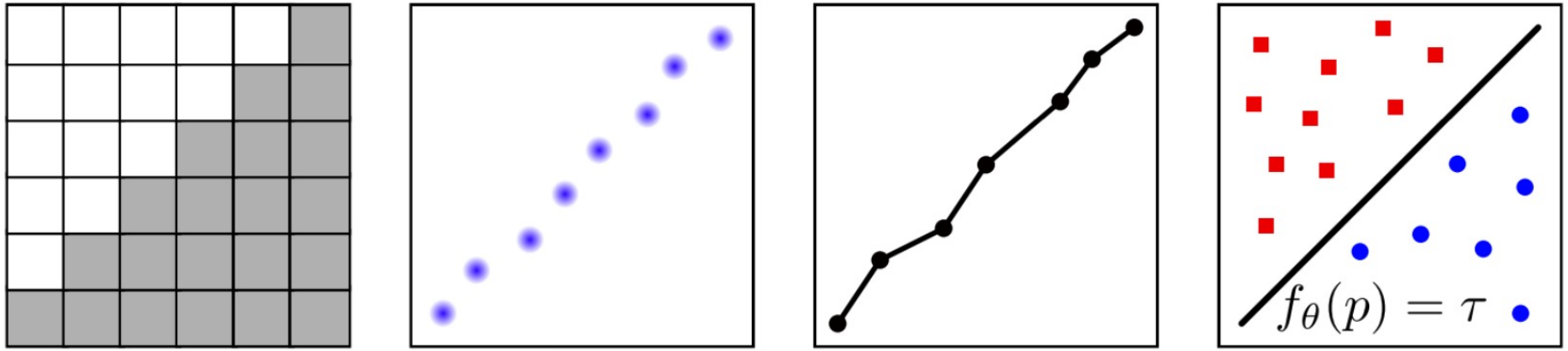
Learning Dynamic Renderable Volumes from Images. SIGGRAPH 2019.

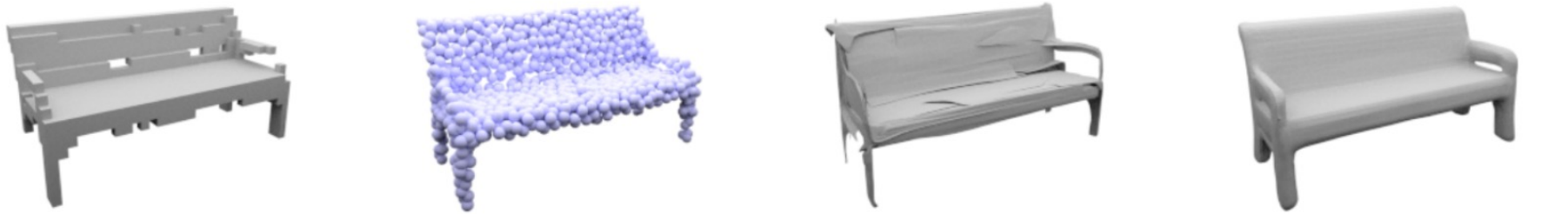
Volume-based representations的优缺点

- 优点
  - 可以表示非常复杂的场景
  - 渲染速度快
- 缺点
  - 容易占据大量显存，无法表示高分辨率场景



常用的三维表示3：Implicit representations





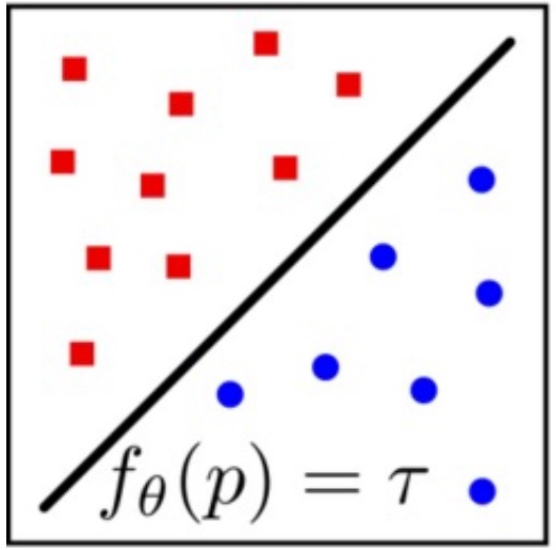
Volume | Point cloud | Mesh | Implicit function

Explicit & discrete | **Implicit & continous**

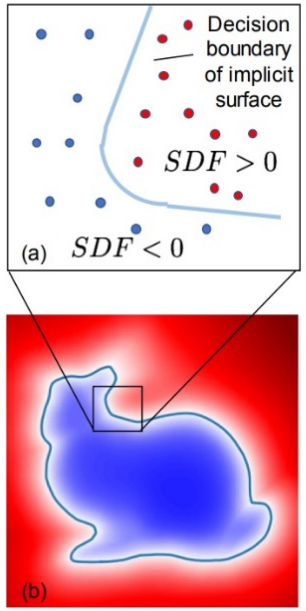
Occupancy Networks: Learning 3D Reconstruction in Function Space, CVPR 2019.

常用的三维表示3：Implicit representations

Occpupacy



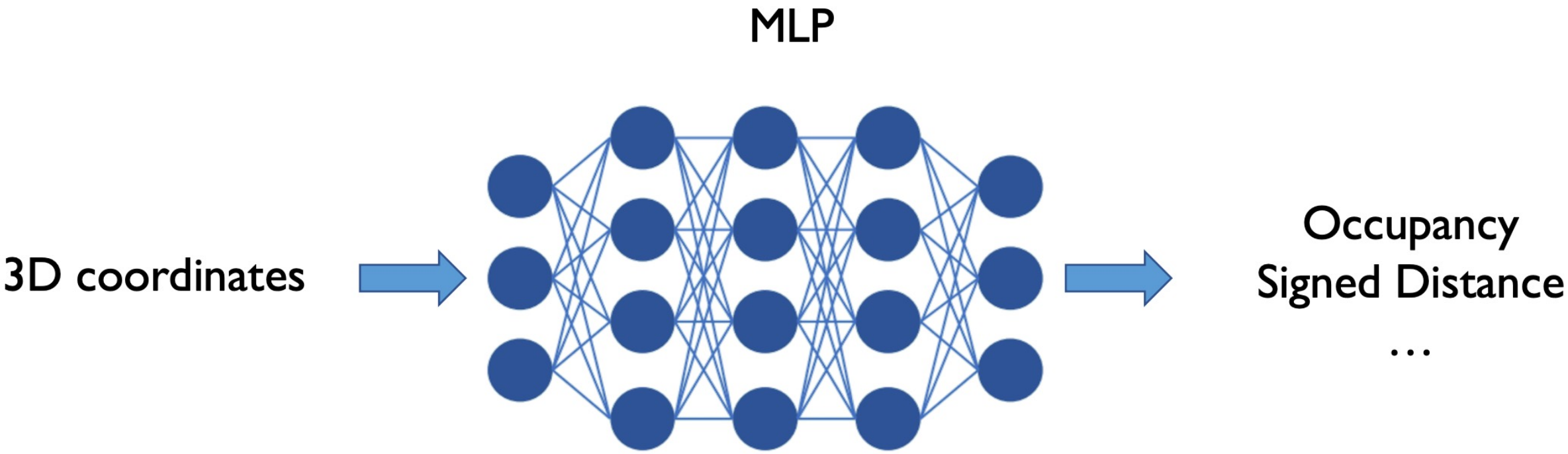
Signed distance function (SDF)



Occupancy Networks: Learning 3D Reconstruction in Function Space, CVPR 2019.

DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation, CVPR 2019.

Implicit neural representations

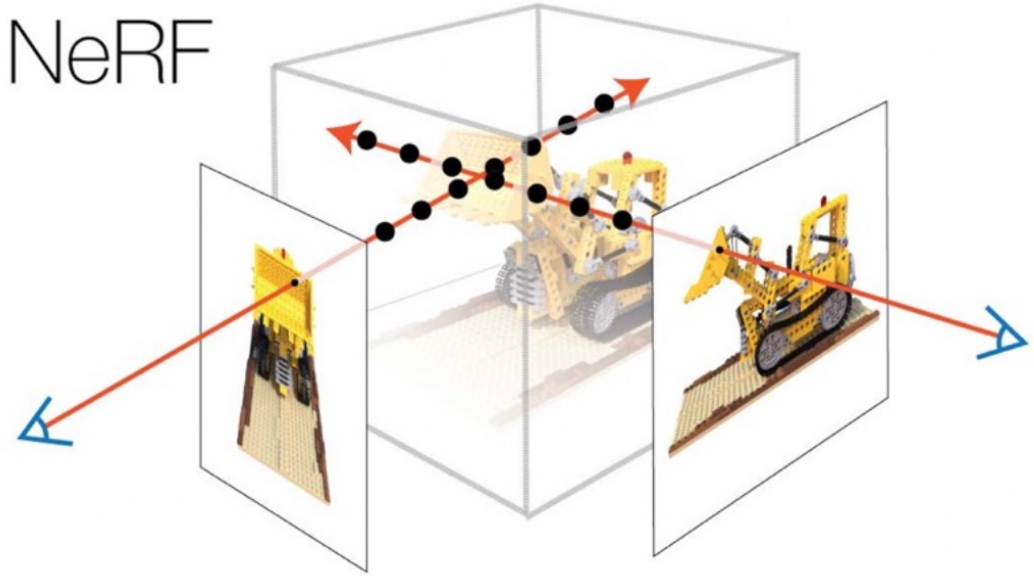


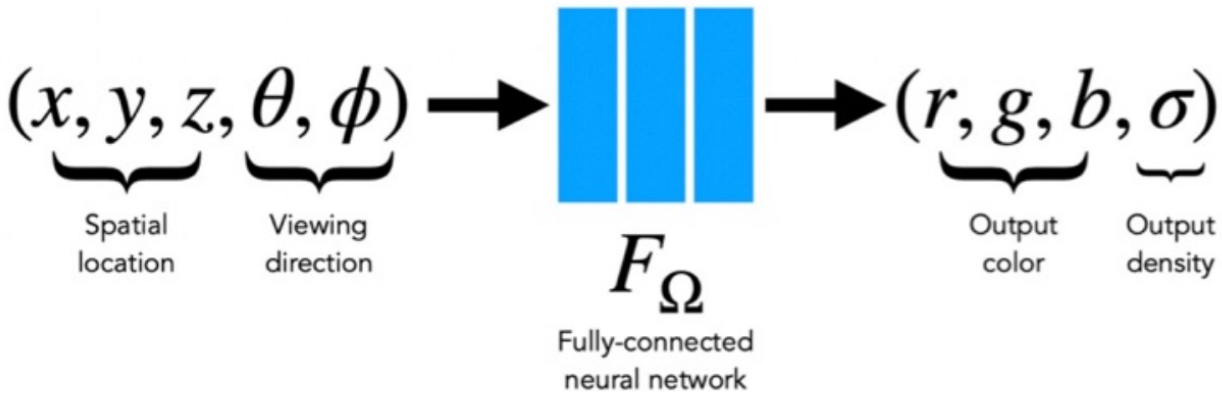
Occupancy Networks: Learning 3D Reconstruction in Function Space, CVPR 2019.

DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation, CVPR 2019.

Neural Radiance fields (NeRF)

- 将场景表示为连续的体素密度场与颜色场。

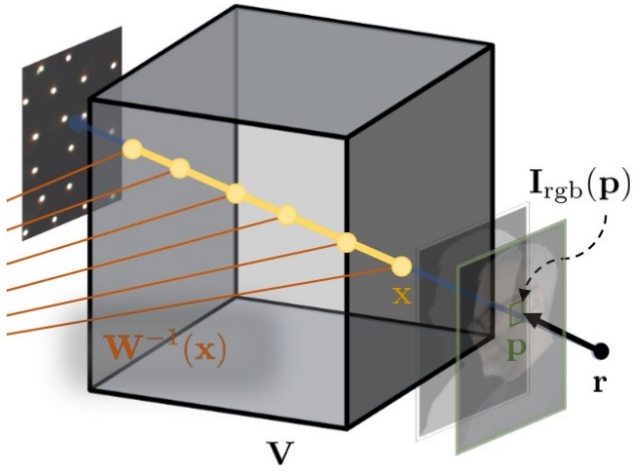




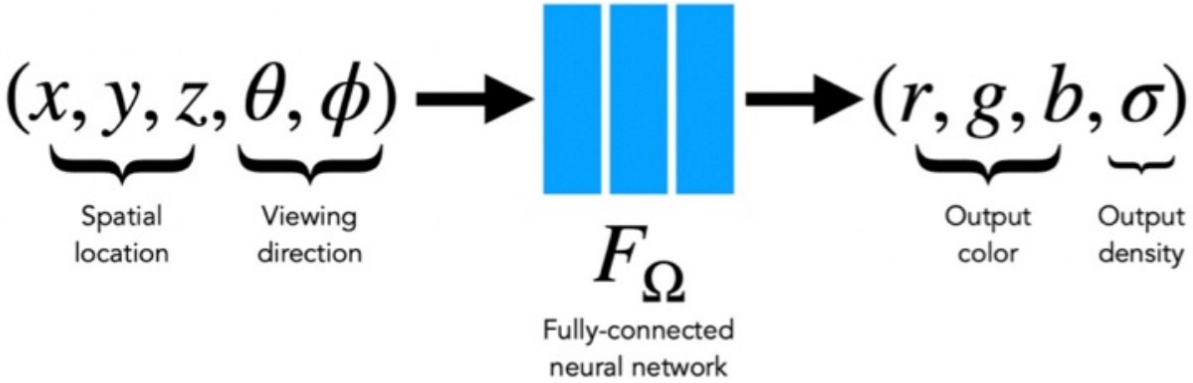
NeRF: Representing scenes as neural radiance fields for view synthesis, ECCV 2020.

Neural Radiance fields (NeRF)

- 离散表示与连续表示的对比



Discrete RGB-α volume



Continuous RGB-α field

Neural Volumes: Learning Dynamic Renderable Volumes from Images, SIGGRAPH 2019.

NeRF: Representing scenes as neural radiance fields for view synthesis, ECCV 2020.

NeRF的渲染流程：Volume rendering

Rendering model for ray r(t) = o + td:

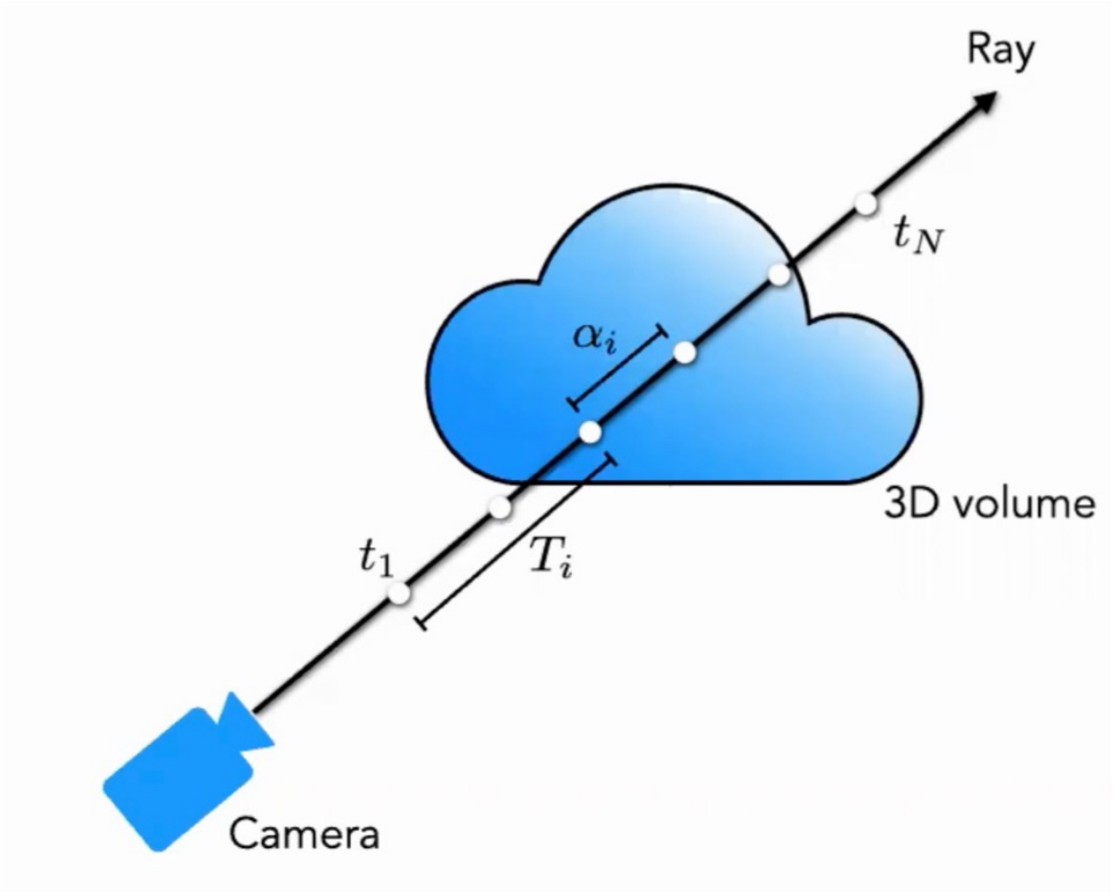
$$C \approx \sum_{i=1}^{N} T_i \alpha_i c_i$$

weights

colors

How much light is blocked earlier along ray:

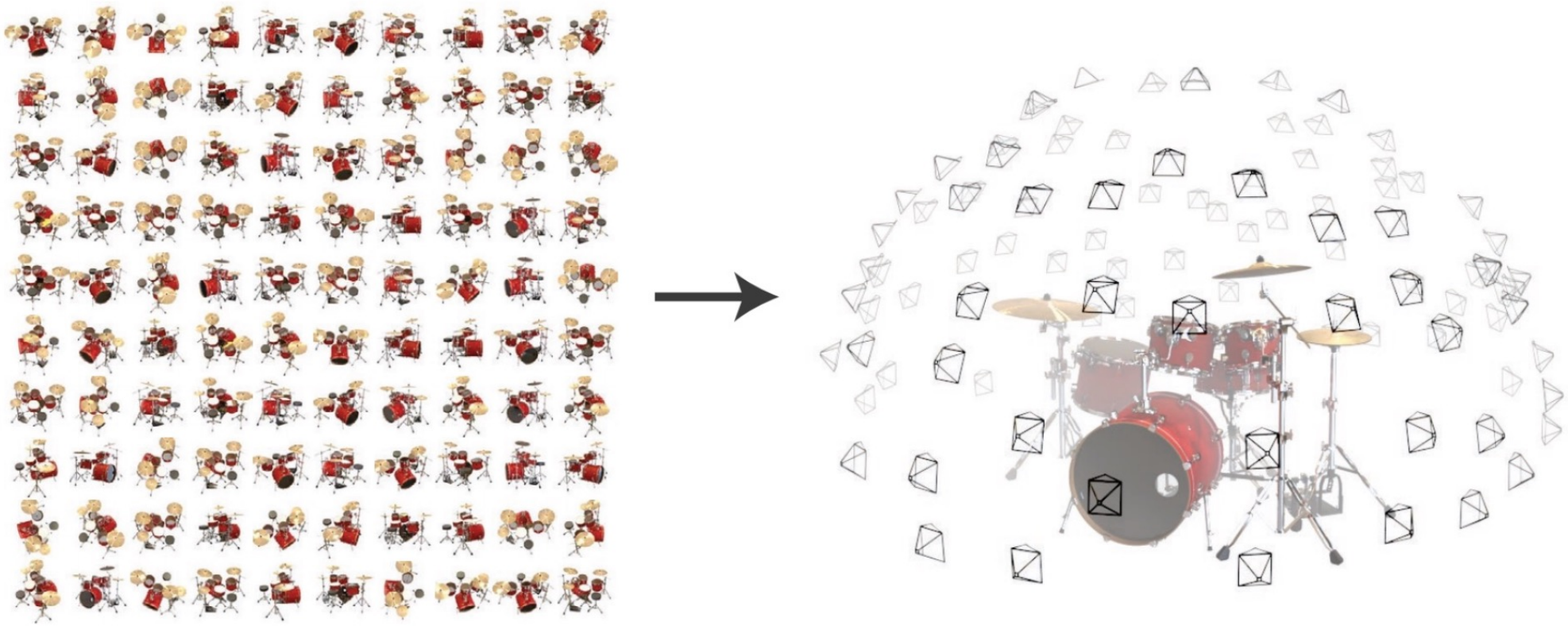
$$T_i = \prod_{j=1}^{i-1} (1 - \alpha_j)$$



NeRF: Representing scenes as neural radiance fields for view synthesis, ECCV 2020.

NeRF的训练

- 基于多视角图片优化NeRF

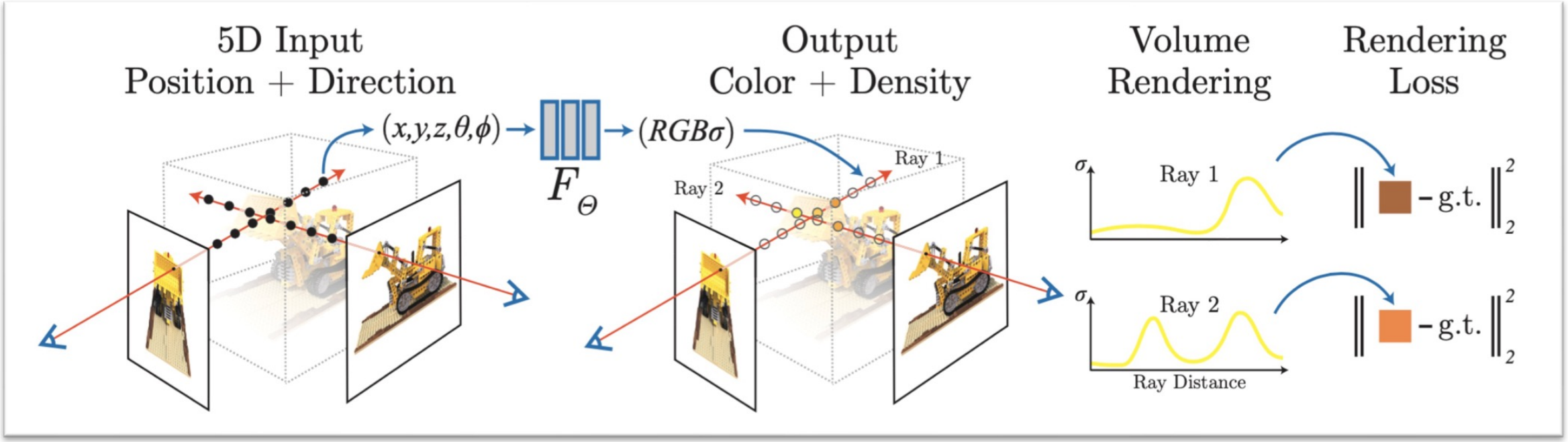


Input multi-view images

Optimizing NeRF

NeRF: Representing scenes as neural radiance fields for view synthesis, ECCV 2020.

NeRF的训练

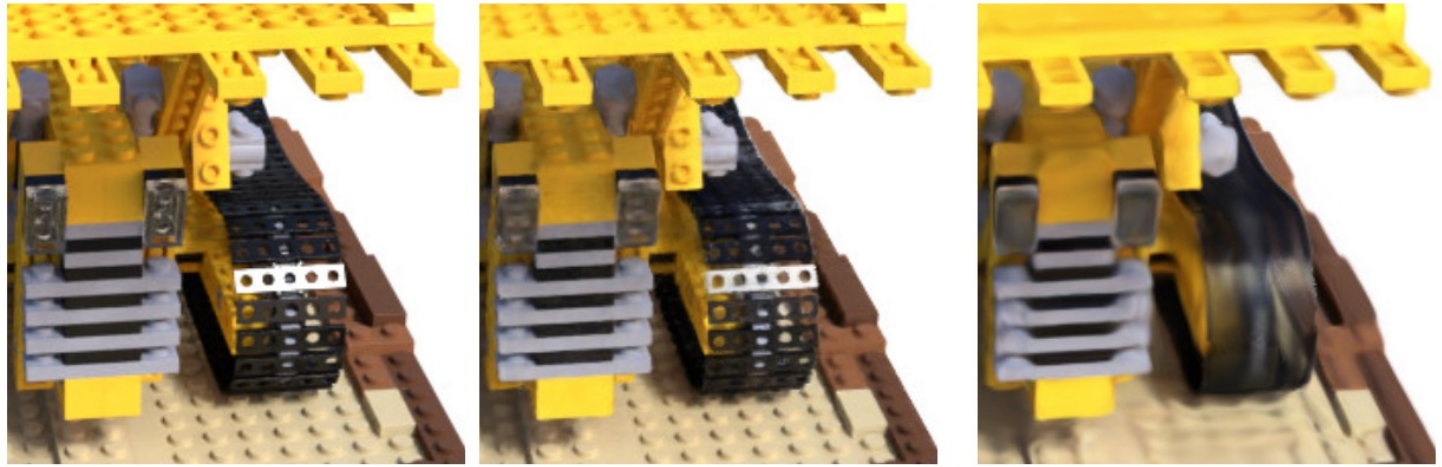


NeRF: Representing scenes as neural radiance fields for view synthesis, ECCV 2020.

NeRF的关键：Positional encoding

- 将输入坐标映射为高维向量

$$\gamma(p) = \left(\sin\left(2^0\pi p\right), \cos\left(2^0\pi p\right), \cdots, \sin\left(2^{L-1}\pi p\right), \cos\left(2^{L-1}\pi p\right)\right)$$



Ground Truth　　Complete Model　　No Positional Encoding

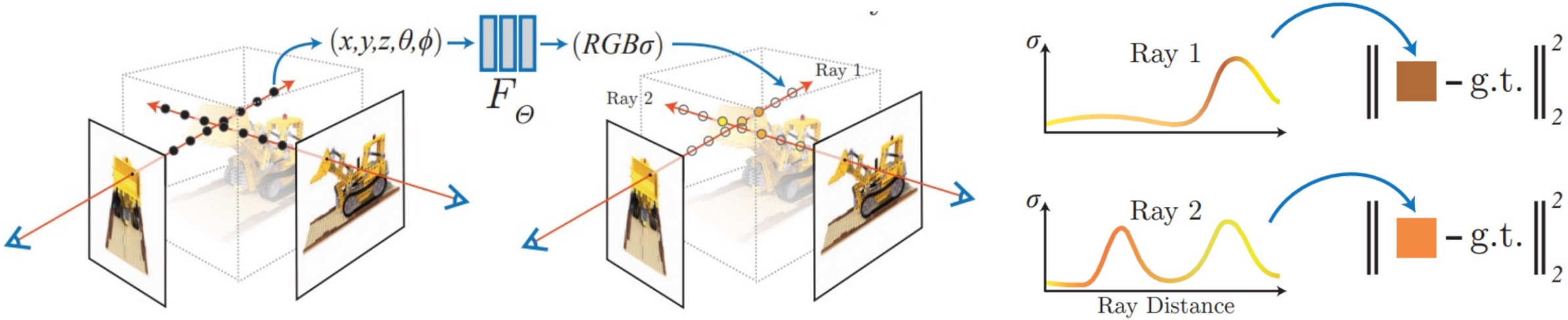
NeRF: Representing scenes as neural radiance fields for view synthesis, ECCV 2020.

NeRF的效果

NeRF: Representing scenes as neural radiance fields for view synthesis, ECCV 2020.

为什么NeRF效果更好

- MLP network可以表示连续高分辨率场景

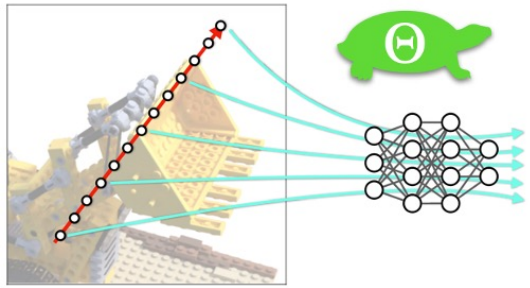


NeRF: Representing scenes as neural radiance fields for view synthesis, ECCV 2020.

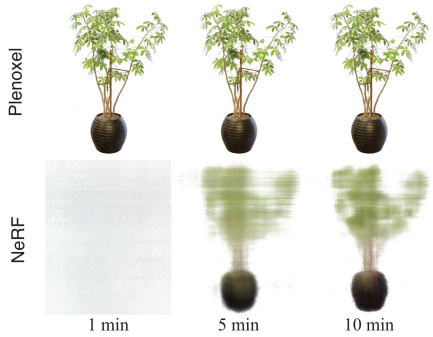
NeRF面临的问题

NeRF在静态场景建模上遇到的问题

渲染速度慢

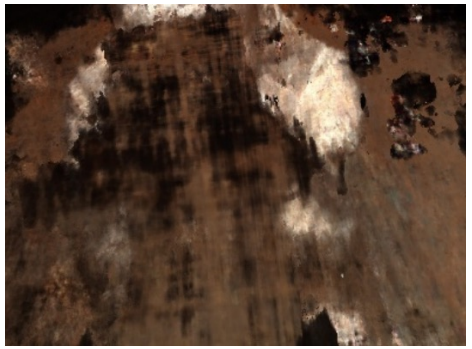


训练速度慢



场景建模能力弱

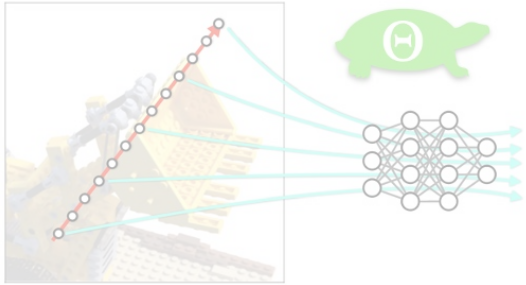
模型鲁棒性差



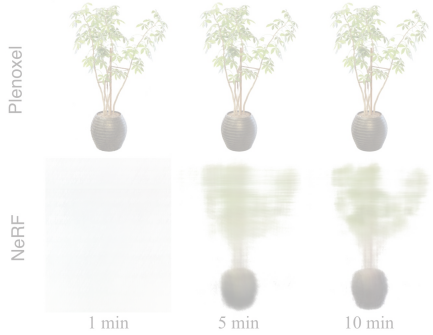
NeRF面临的问题

NeRF在静态场景建模上遇到的问题

渲染速度慢

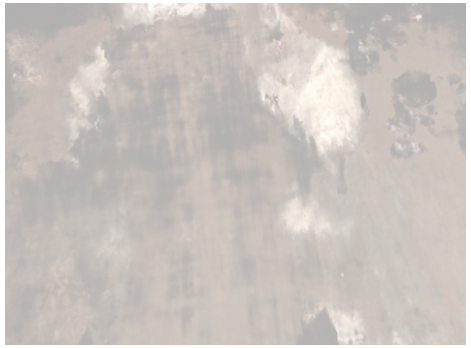


训练速度慢



场景建模能力弱

模型鲁棒性差



NeRF在其他类型场景建模上遇到的问题

几何重建质量低

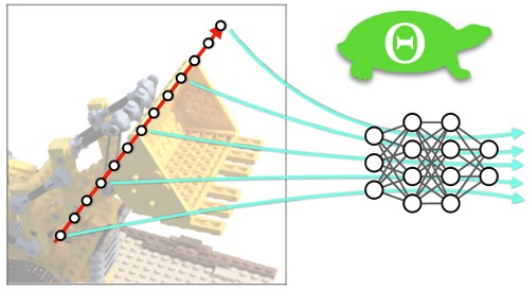
无法建模多模态信号

无法建模动态场景

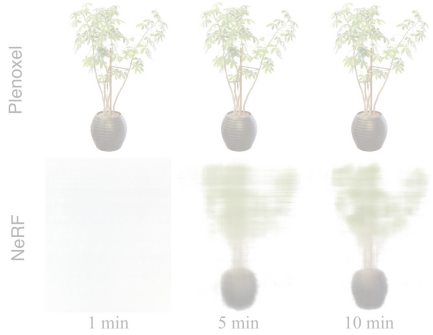
NeRF面临的问题1：渲染速度慢

NeRF在静态场景建模上遇到的问题

渲染速度慢

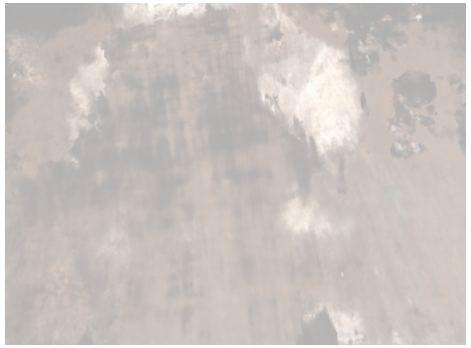


训练速度慢



场景建模能力弱

模型鲁棒性差



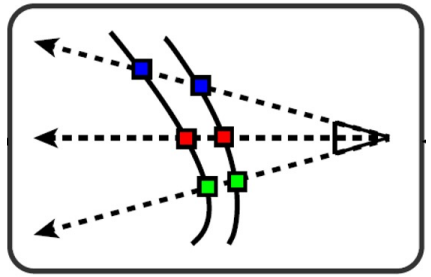
NeRF在其他类型场景建模上遇到的问题

几何重建质量低

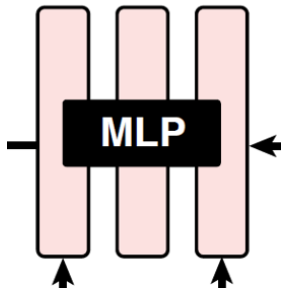
无法建模多模态信号

无法建模动态场景

NeRF渲染成本分析



*

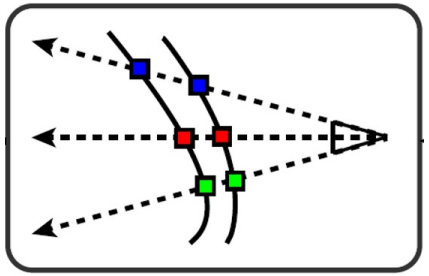


=

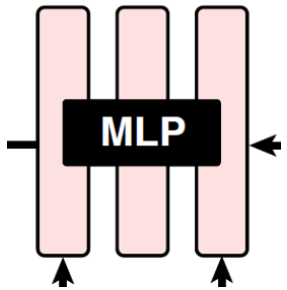
采样点数量 * 逐点计算成本 = 渲染成本

降低NeRF渲染成本的思路

1. 降低采样点的数量：NSVF、AdaNeRF、ENeRF
2. 降低逐点计算成本：SNeRG、PlenOctree、KiloNeRF、ObjectNeRF



*

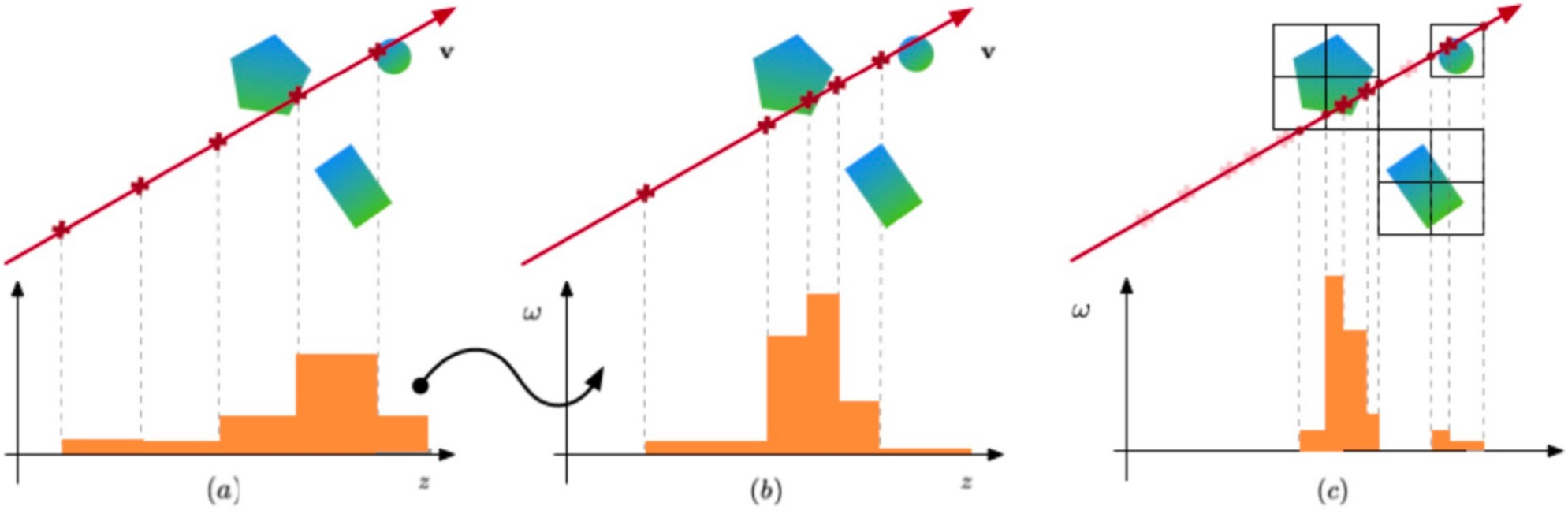


=

采样点数量 * 逐点计算成本 = 渲染成本

降低采样点的数量：NSVF

- 建立Sparse voxels来跳过empty space的采样

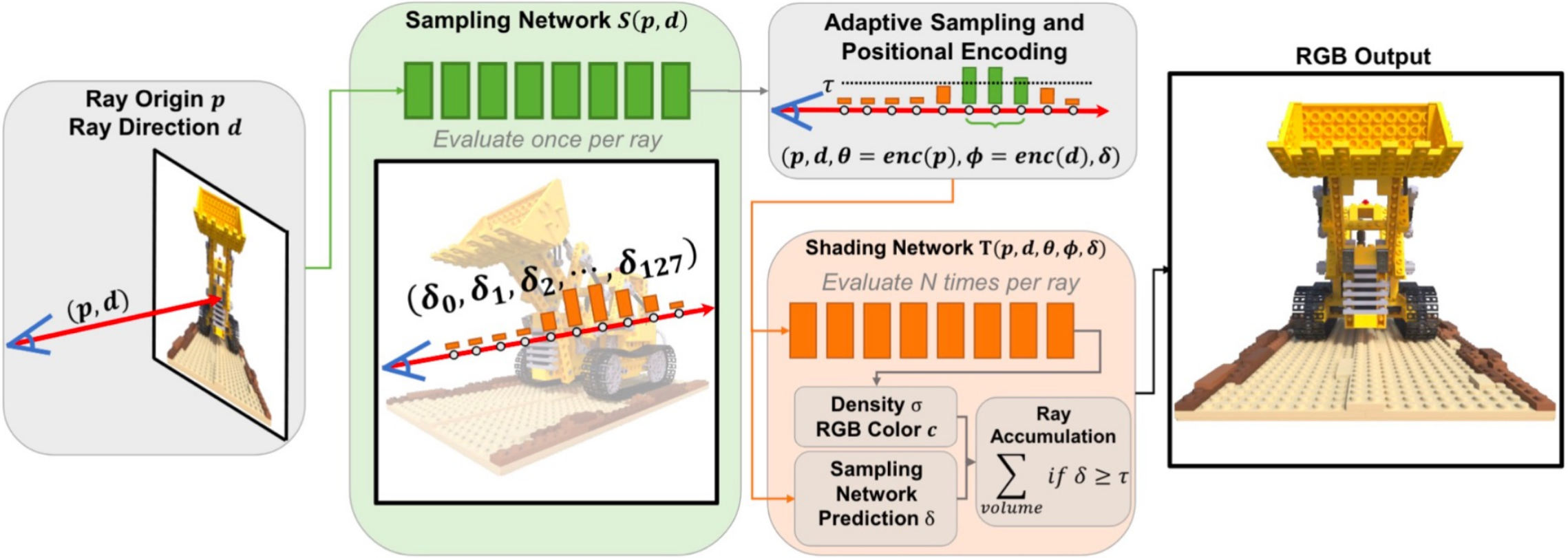


(a)uniform sampling (b) importance sampling (c) ESS sampling

Neural Sparse Voxel Fields. NeurIPS 2020.

降低采样点的数量：AdaNeRF

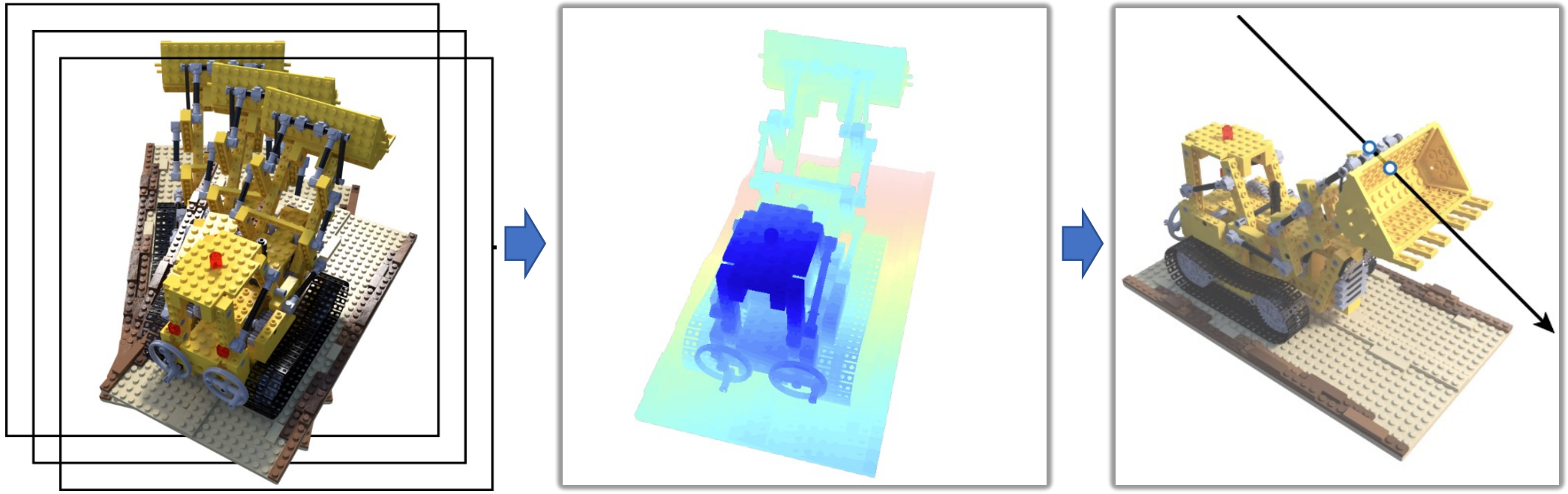
- 训练一个Sampling network，一次性预测一条线上所有点的weights



AdaNeRF: Adaptive Sampling for Real-time Rendering of Neural Radiance Fields. ECCV 2022.

降低采样点的数量：ENeRF

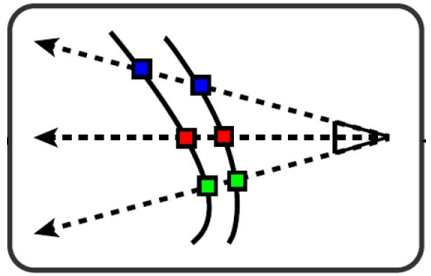
- 使用多视图几何算法预测场景表面深度，在表面做采样



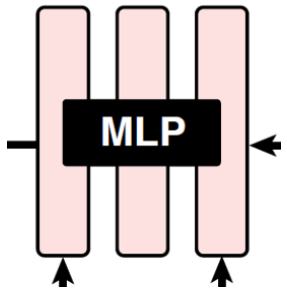
Efficient Neural Radiance Fields for Interactive Free-viewpoint Video. SIGGRAPH Asia 2022.

降低NeRF渲染成本的思路

1. 降低采样点的数量：NSVF、AdaNeRF、ENeRF
2. 降低逐点计算成本：SNeRG、PlenOctree、KiloNeRF、ObjectNeRF



*

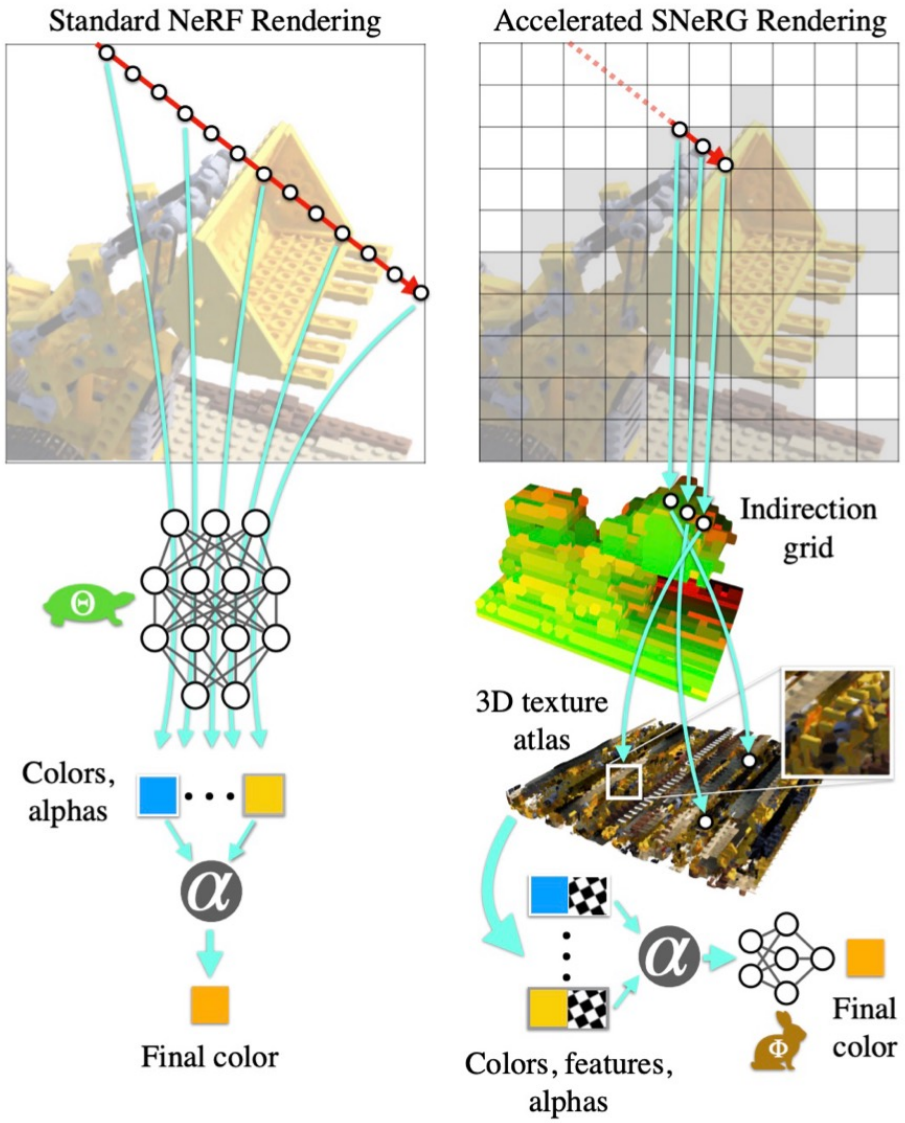


=

采样点数量 * 逐点计算成本 = 渲染成本

降低逐点计算成本的数量：SNeRG

- 四方面的加速：
  1. 使用cached sparse 3D grid快速定位表面。
  2. 将颜色缓存到3D texture atlas。
  3. 通过小网络增加view-dependent effect。



Baking neural radiance fields for real-time view synthesis. ICCV, 2021.

降低逐点计算成本的数量：SNeRG

- 使用sparsity loss来减小sparse voxels的显存占用。

(a) Frame rendered by our method.

(b) Cross-section of (a).

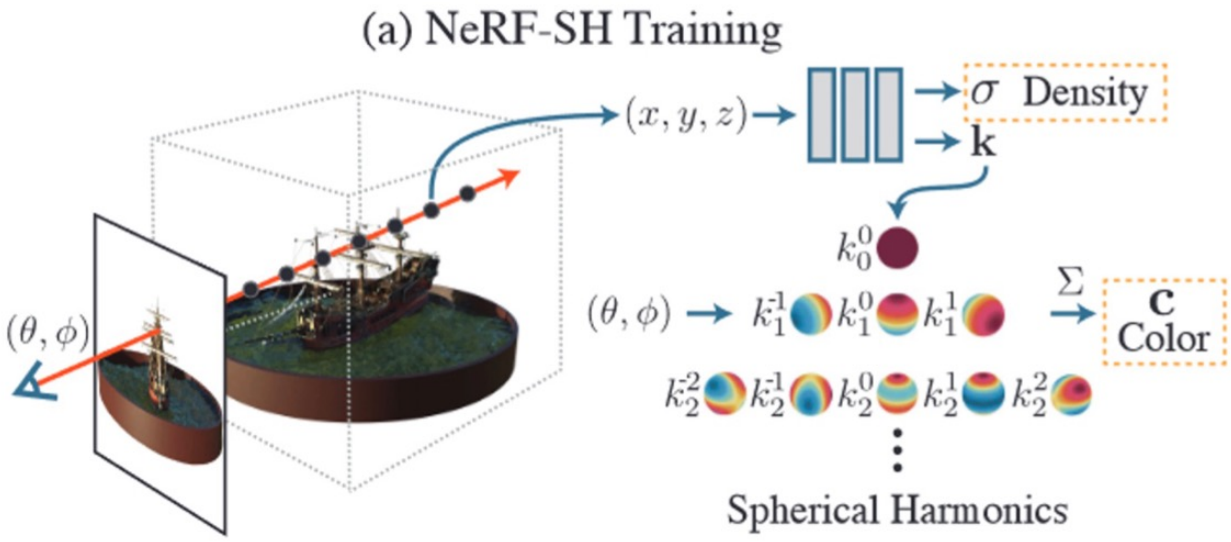
(c) Trained without $\mathcal{L}_s$.

(d) No $\mathcal{L}_s$, no visibility culling.

Baking neural radiance fields for real-time view synthesis. ICCV, 2021.

降低逐点计算成本的数量：PlenOctree

- 使用spherical harmonics model建模颜色。
- 使用octree structure让volume rendering跳过empty space。



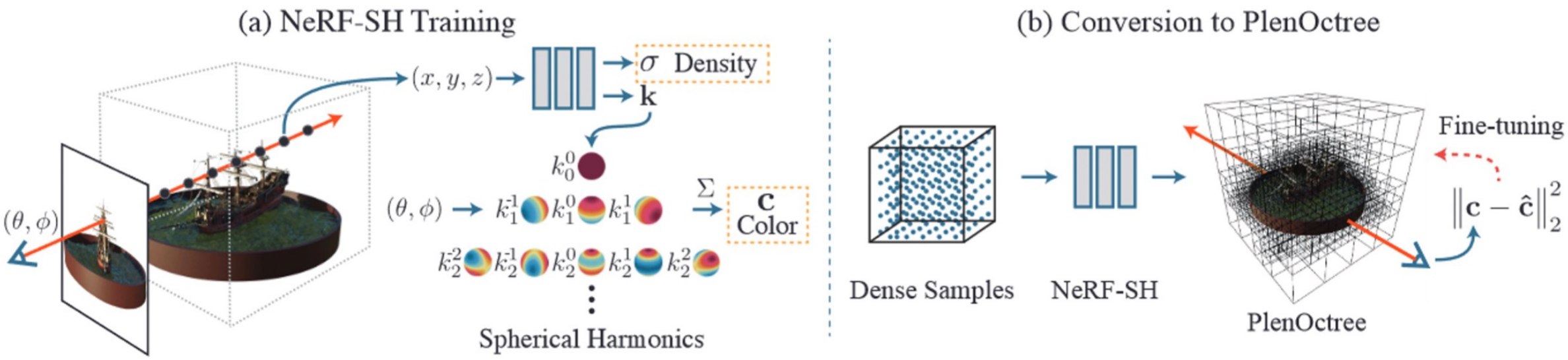
$$c(\mathbf{d}; \mathbf{k}) = S\left(\sum_{\ell=0}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} k_\ell^m Y_\ell^m(\mathbf{d})\right)$$

Y是基函数，输入视角方向，输出颜色

Plenoctrees for real-time rendering of neural radiance fields. ICCV, 2021.

降低逐点计算成本的数量：PlenOctree

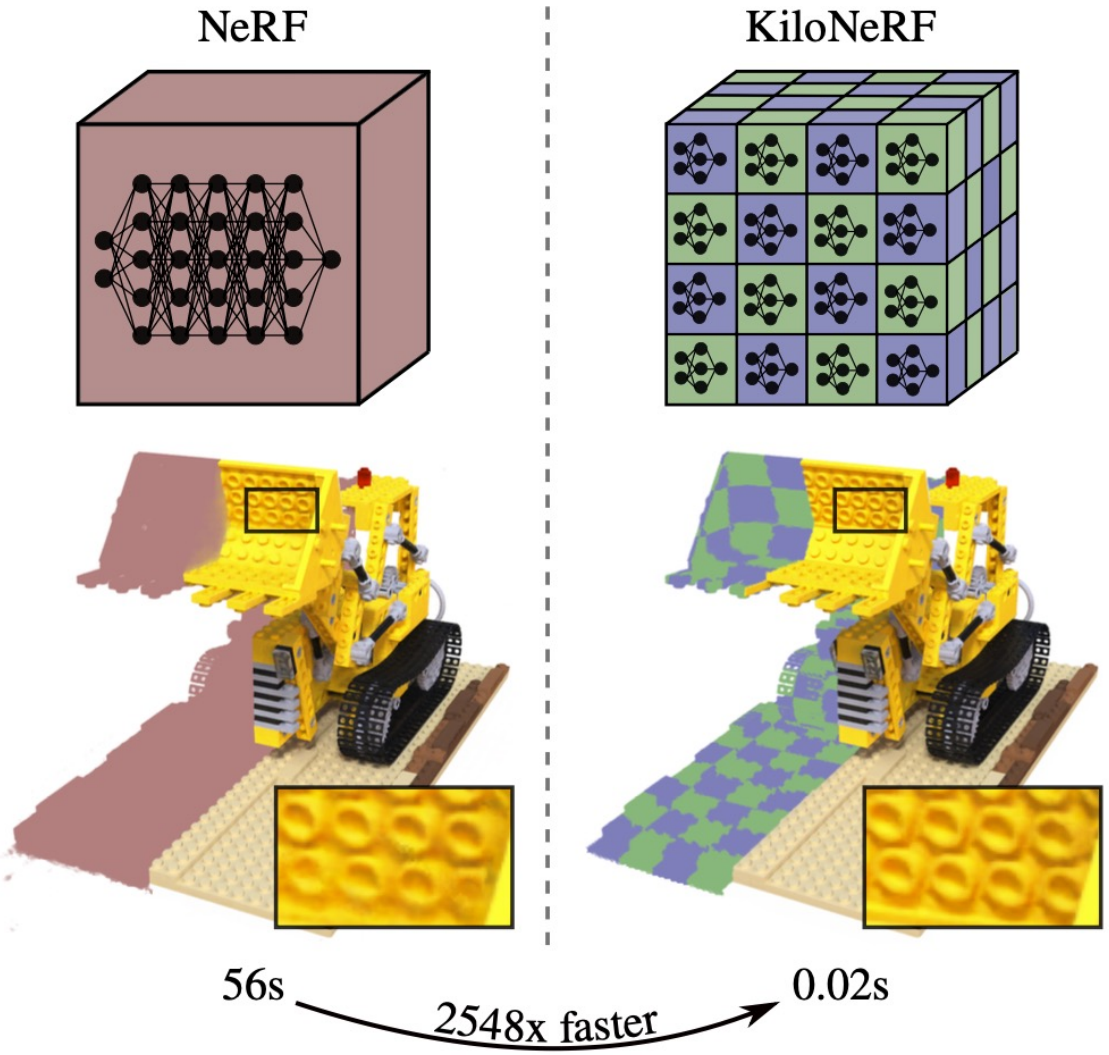
- 使用spherical harmonics model建模颜色。
- 使用octree structure让volume rendering跳过empty space。



Plenoctrees for real-time rendering of neural radiance fields. ICCV, 2021.

降低逐点计算成本的数量：KiloNeRF

- 两方面的加速：
  1. 使用一组小型MLP网络表示场景。
  2. 使用cached sparse 3D grid快速定位表面。



KiloNeRF: Speeding up Neural Radiance Fields with Thousands of Tiny MLPs. ICCV, 2021.

降低逐点计算成本的数量：KiloNeRF

- 两方面的加速：
  1. 使用一组小型MLP网络表示场景。
  2. 使用cached sparse 3D grid快速定位表面。

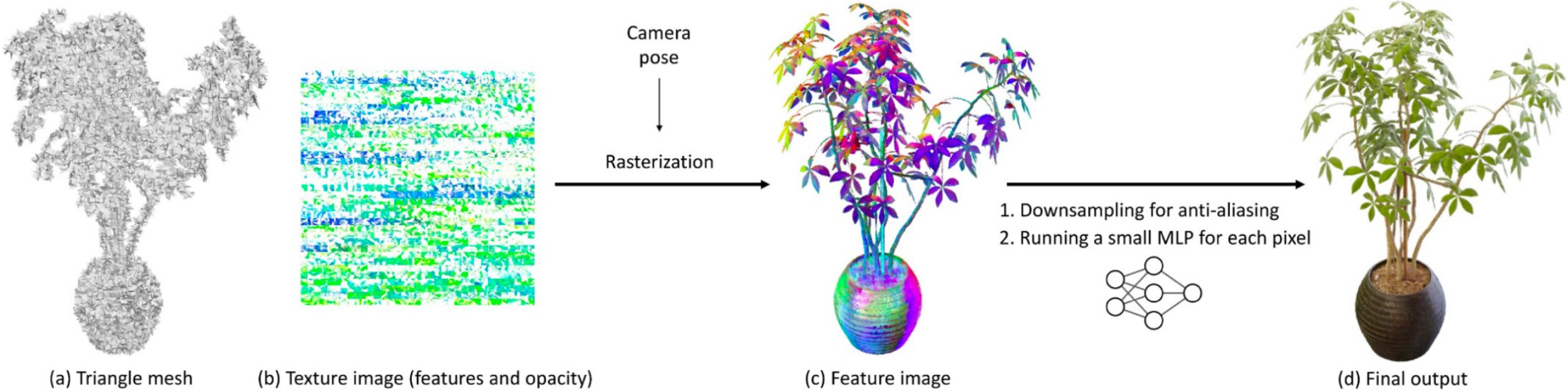
Method	Render time ↓	Speedup ↑
NeRF	56185 ms	–
NeRF + ESS + ERT	788 ms	71
KiloNeRF	**22** ms	**2548**

Table 2: **Speedup Breakdown.** The original NeRF model combined with KiloNeRF's implementation of ESS and ERT is compared against the full KiloNeRF technique.

KiloNeRF: Speeding up Neural Radiance Fields with Thousands of Tiny MLPs. ICCV, 2021.

降低逐点计算成本的数量：MobileNeRF

- 基于mesh rasterization的渲染，支持GPU硬件加速。
- 在Image space做颜色预测。



KiloNeRF: Speeding up Neural Radiance Fields with Thousands of Tiny MLPs. ICCV, 2021.

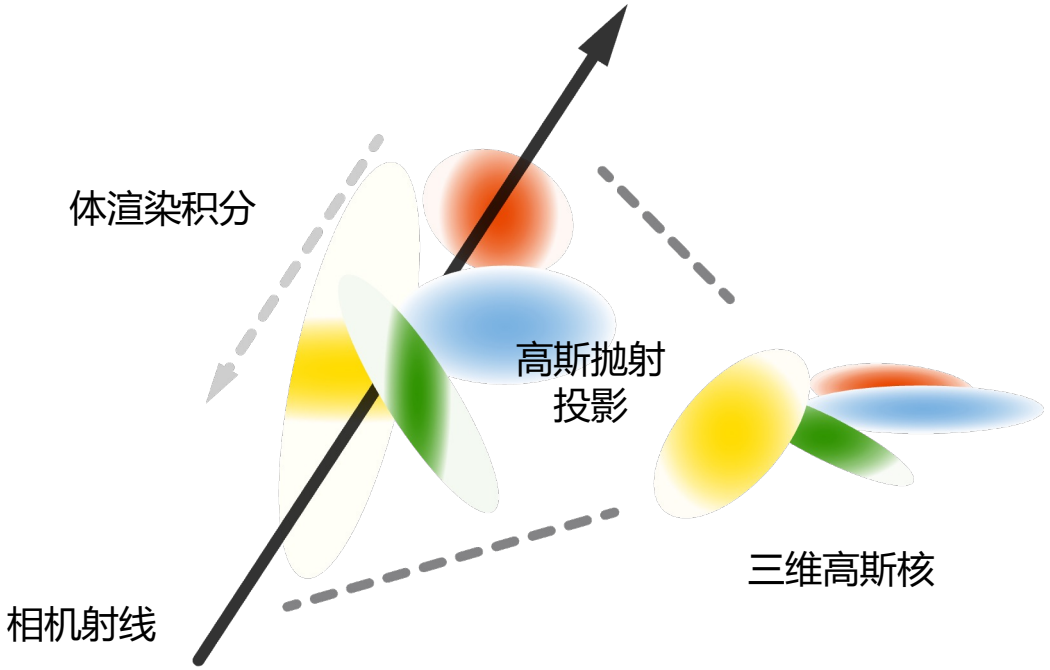
降低采样点数&逐点计算成本：3DGS

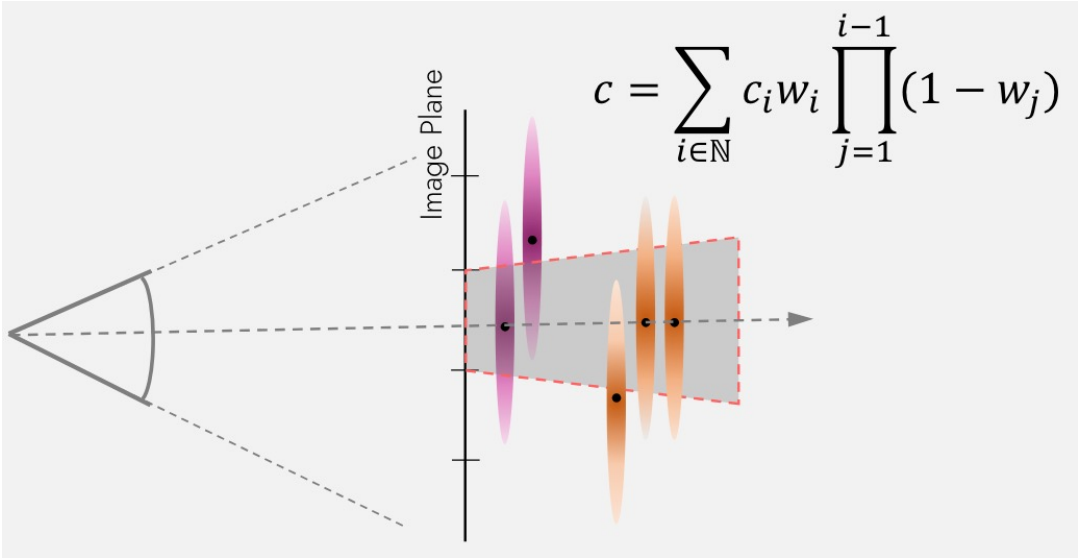
- 基于一组三维高斯表示场景，每个三维高斯包含position, rotation, opacity和Spherical harmonics系数。

$$p = \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sigma_x & \sigma_{xy} & \sigma_{xz} \\ & \sigma_y & \sigma_{yz} \\ & & \sigma_z \end{bmatrix} \quad o \quad SH$$

spherical harmonics

covariance matrix





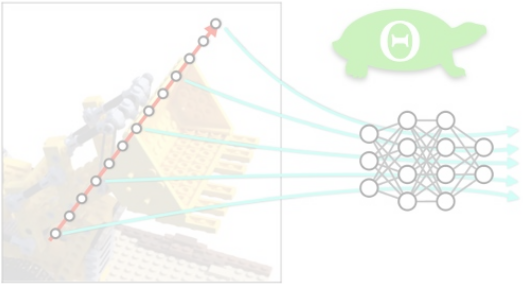
3D Gaussian Splatting for Real-Time Radiance Field Rendering. SIGGRAPH 2023.

https://3dgstutorial.github.io/3dv_part1.pdf

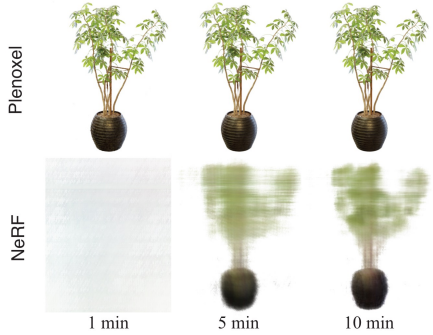
NeRF面临的问题2：训练速度慢

NeRF在静态场景建模上遇到的问题

渲染速度慢

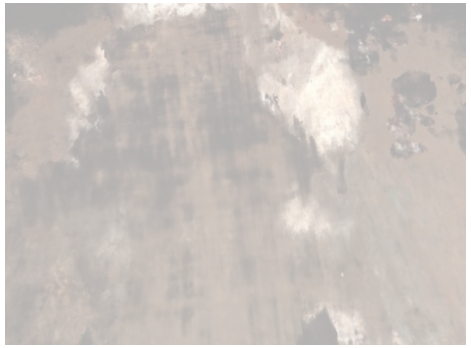


训练速度慢



场景建模能力弱

模型鲁棒性差



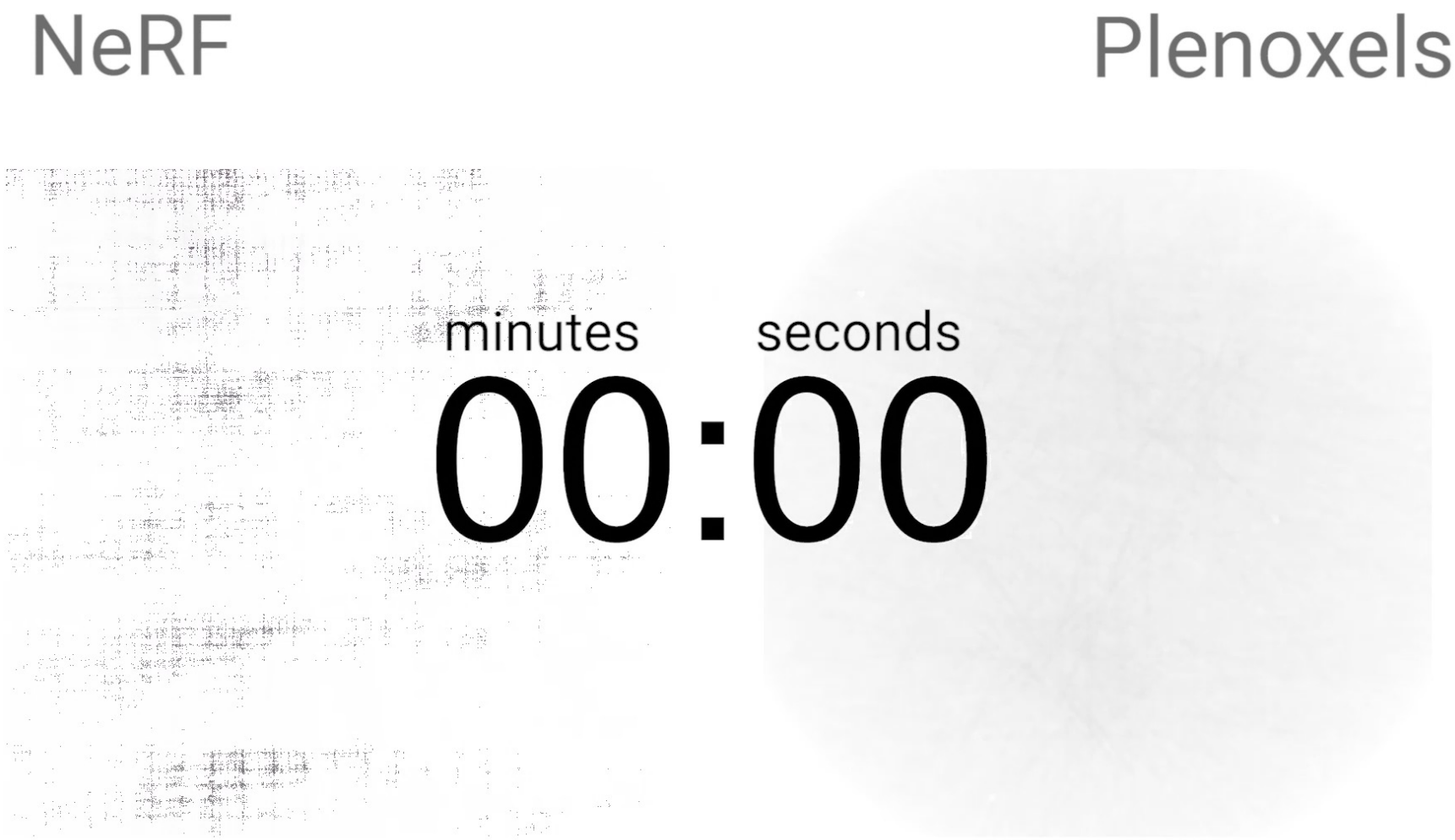
NeRF在其他类型场景建模上遇到的问题

几何重建质量低

无法建模多模态信号

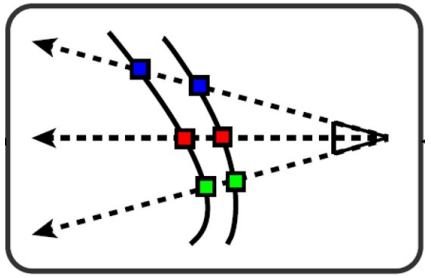
无法建模动态场景

NeRF训练速度慢

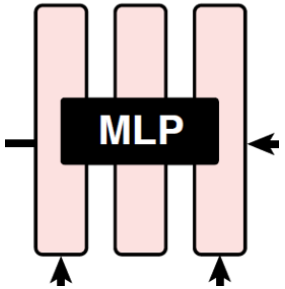


Plenoxels: Radiance Fields without Neural Networks. CVPR 2022.

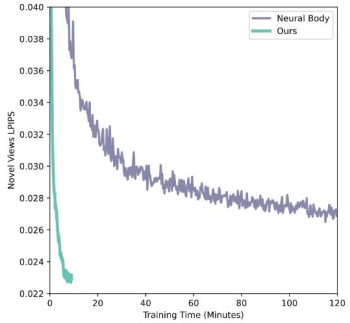
NeRF训练成本分析



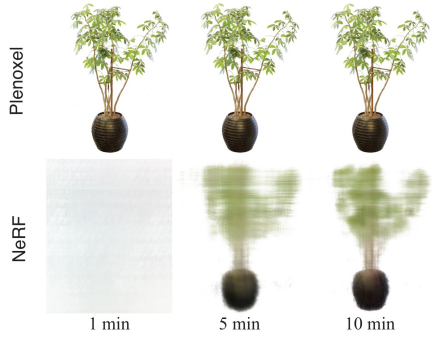
*



*



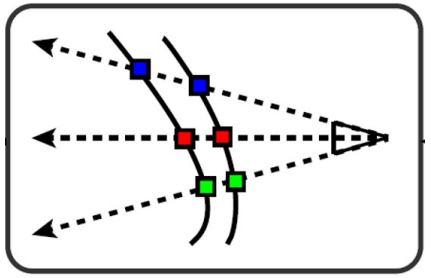
=



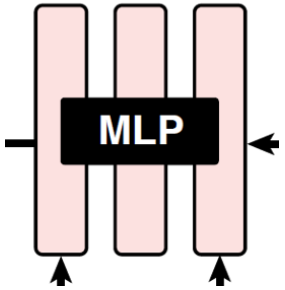
三维点数量 * 逐点训练成本 * 迭代次数 = 训练成本

降低NeRF训练成本的思路

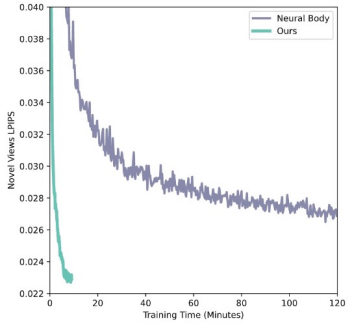
1. 降低三维点的数量：Plenoxel
2. 降低逐点训练成本：Instant NGP、TensoRF
3. 减少迭代次数：IBRNet



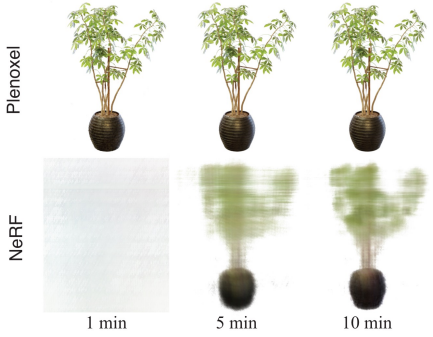
*



*



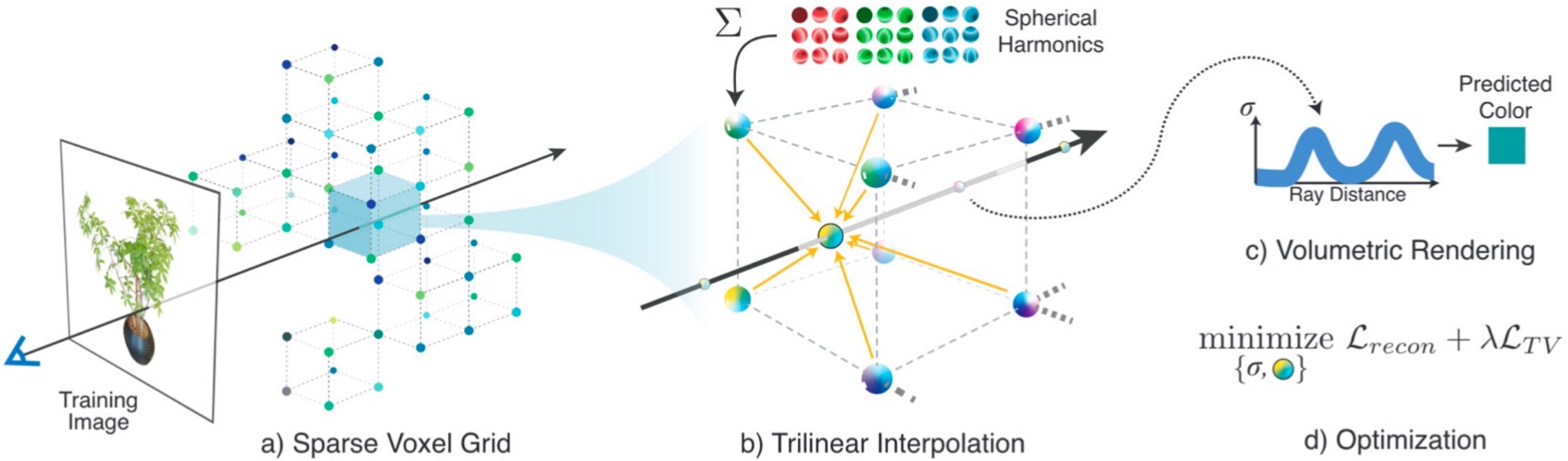
=



三维点数量 * 逐点训练成本 * 迭代次数 = 训练成本

降低三维点的数量&逐点训练成本&迭代次数：Plenoxels

- 使用sparse 3D voxel grids降低三维点
- 使用显式的opacity model和color model降低逐点训练成本&迭代次数。

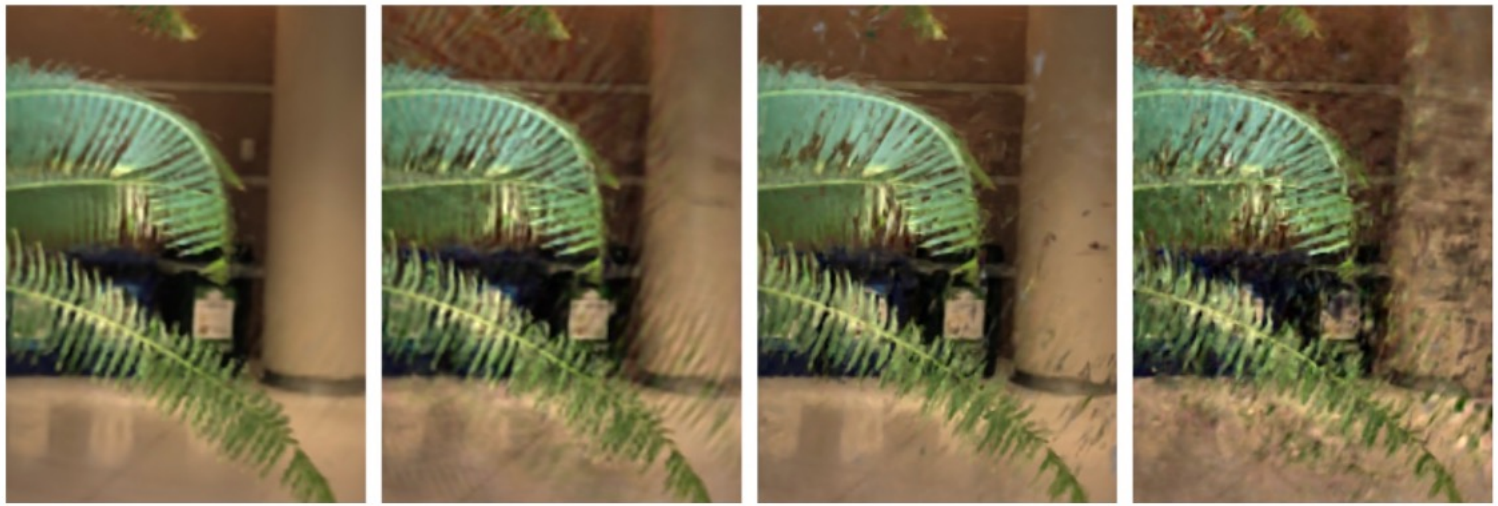


Plenoxels: Radiance Fields without Neural Networks. CVPR 2022.

Plenoxels需要regularization technique

- Total variation loss:

$$\mathcal{L}_{TV} = \frac{1}{|\mathcal{V}|} \sum_{\substack{\mathbf{v} \in \mathcal{V} \\ d \in [D]}} \sqrt{\Delta_x^2(\mathbf{v}, d) + \Delta_y^2(\mathbf{v}, d) + \Delta_z^2(\mathbf{v}, d)}$$

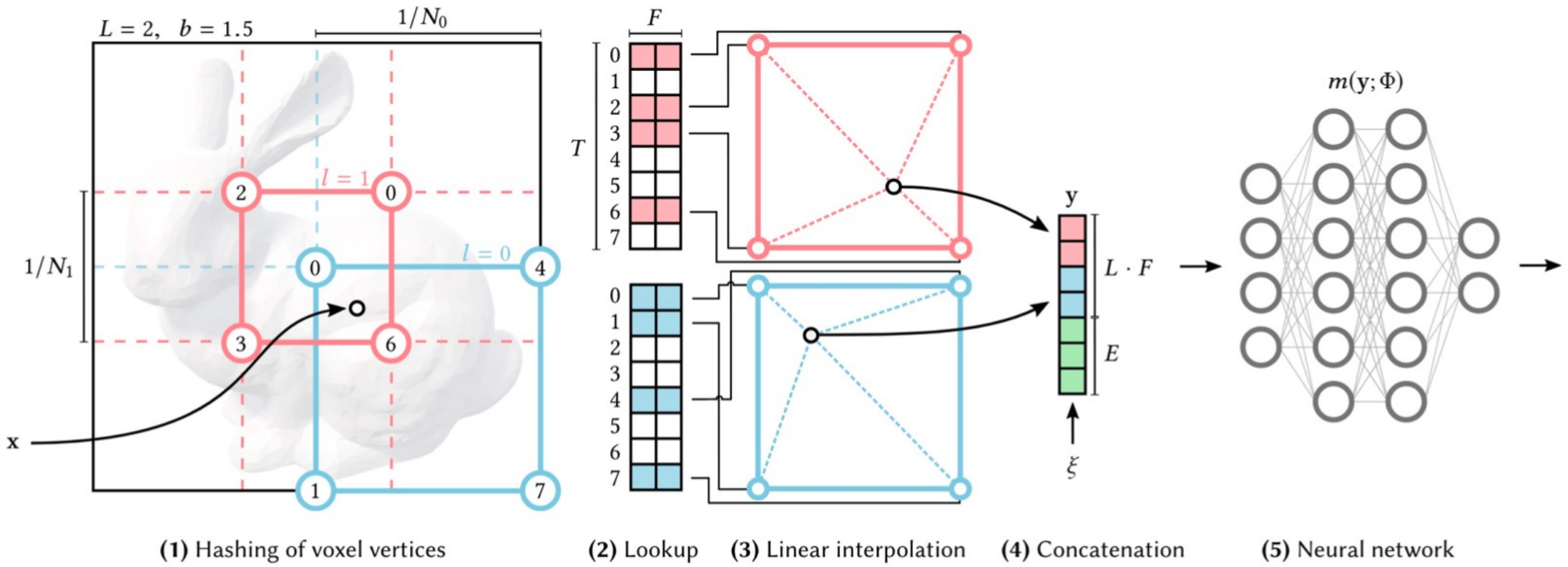


Full | No SH TV | No σ TV | No TV

Plenoxels: Radiance Fields without Neural Networks. CVPR 2022.

降低逐点训练成本&迭代次数：Instant NGP

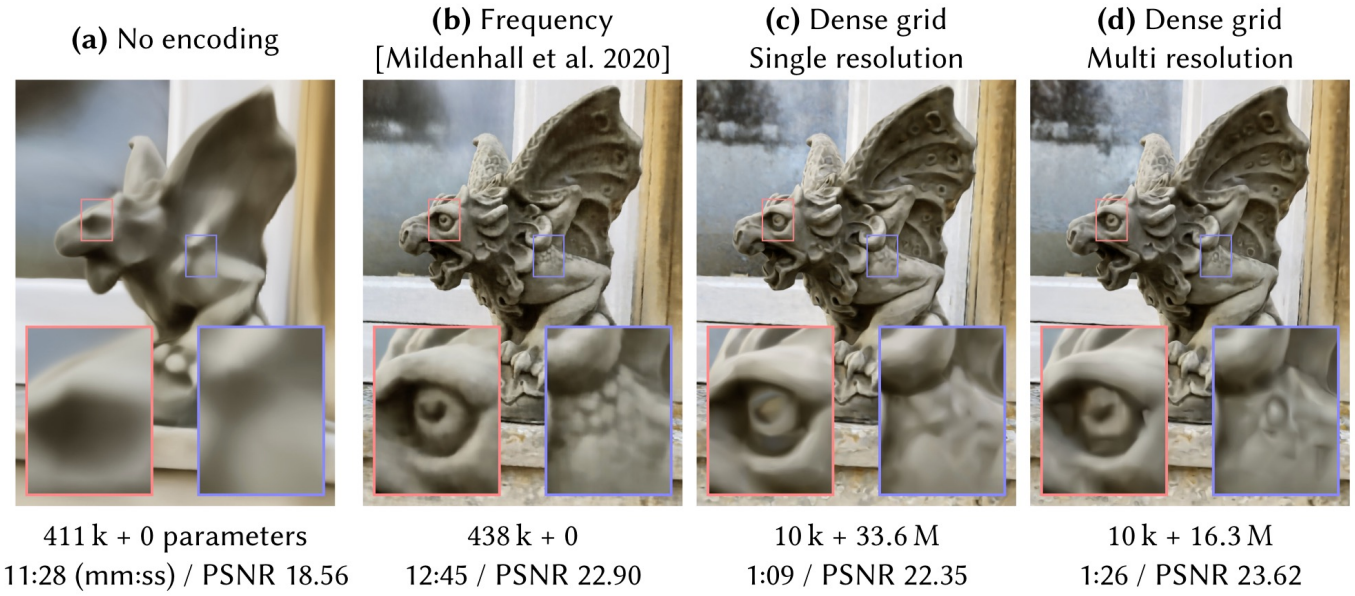
- 使用hash table中的feature vector和一个小型MLP网络表示场景。

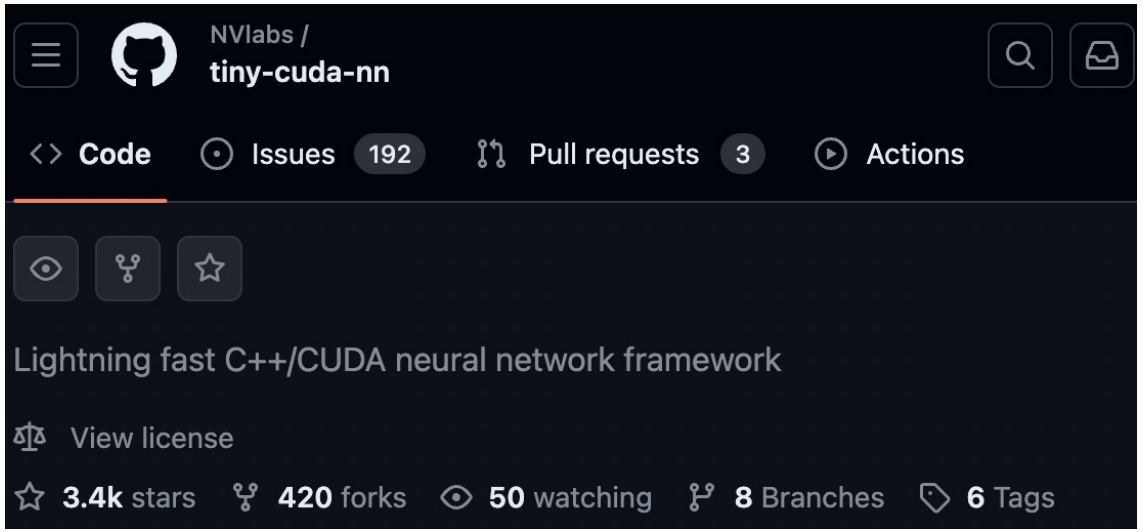


Instant Neural Graphics Primitives with a Multiresolution Hash Encoding. SIGGRAPH 2022.

Instant NGP为什么快

- 使用hash table中的feature vector和一个小型MLP网络表示场景。





Instant Neural Graphics Primitives with a Multiresolution Hash Encoding. SIGGRAPH 2022.

Instant NGP怎么降低model size

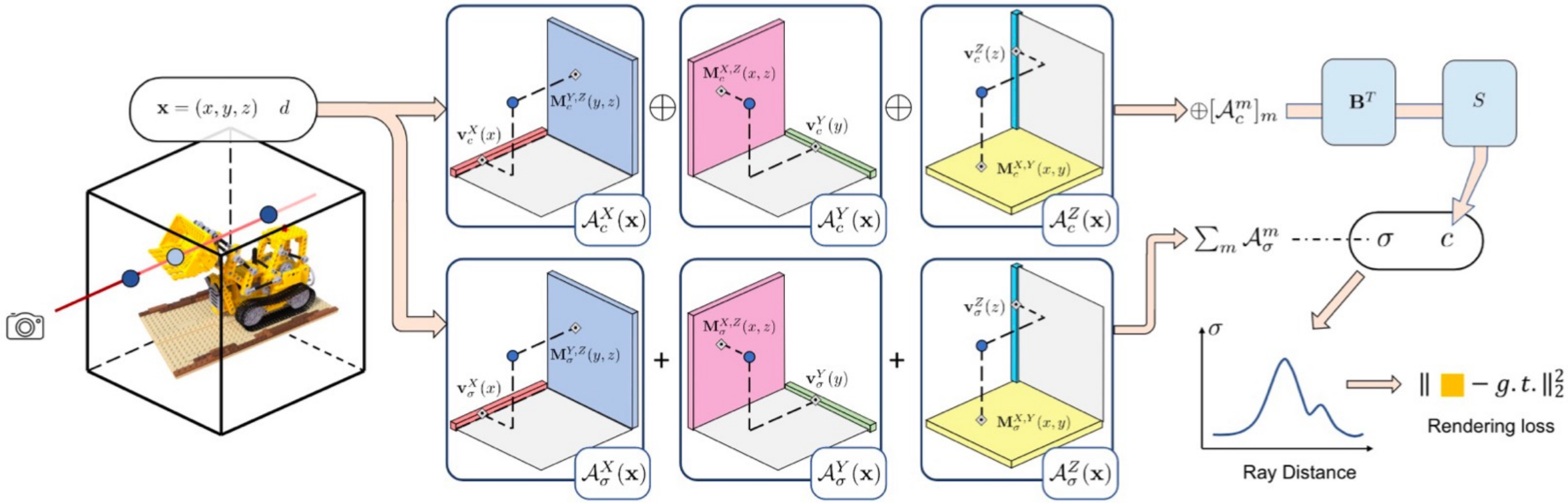
- 两大关键：(1) Multi-resolution。(2) Hash table。

(a) No encoding	(b) Frequency [Mildenhall et al. 2020]	(c) Dense grid Single resolution	(d) Dense grid Multi resolution	(e) Hash table (ours) $T = 2^{14}$
411 k + 0 parameters	438 k + 0	10 k + 33.6 M	10 k + 16.3 M	10 k + 494 k
11:28 (mm:ss) / PSNR 18.56	12:45 / PSNR 22.90	1:09 / PSNR 22.35	1:26 / PSNR 23.62	1:48 / PSNR 22.61

Instant Neural Graphics Primitives with a Multiresolution Hash Encoding. SIGGRAPH 2022.

降低逐点训练成本&迭代次数：TensoRF

- 将三维场景解耦成一组平面和一组向量



TensoRF: Tensorial Radiance Fields. ECCV 2022.

TensoRF的效果

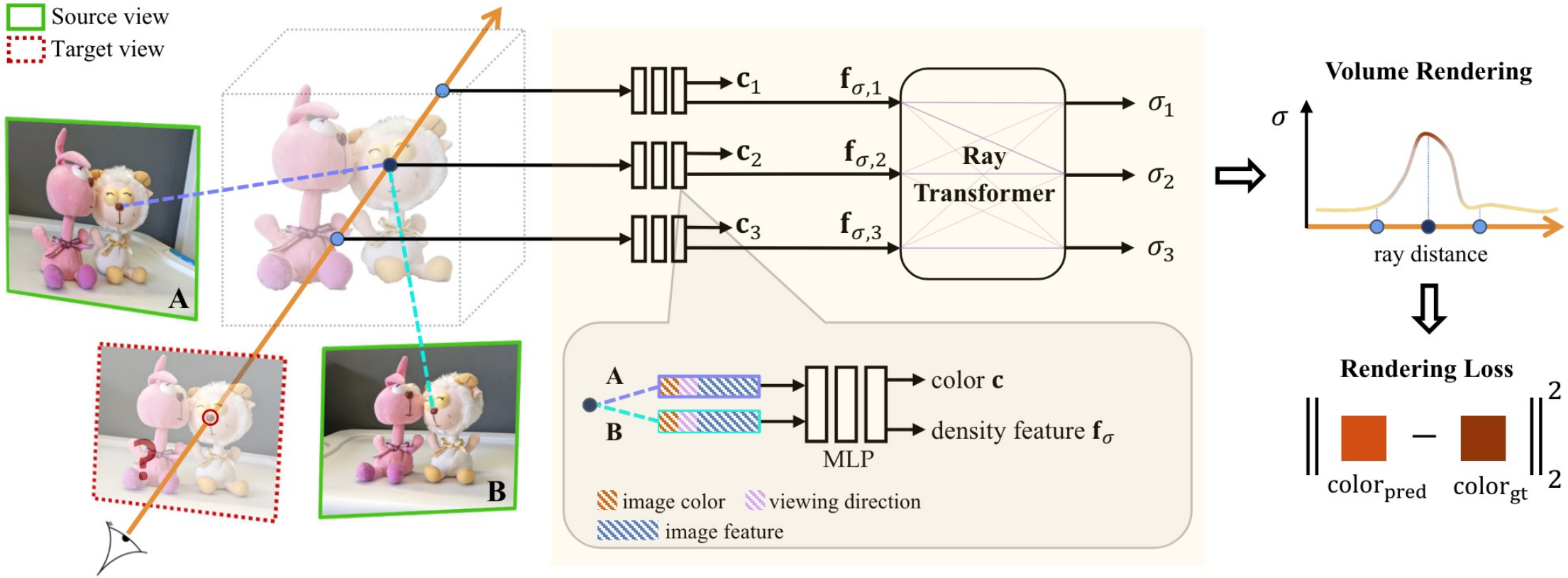
- Synthetic–NeRF dataset上的比较

Method	Device	Training Time	Memory (MB)	PSNR	Initial LR
NeRF	NVIDIA V100	1-2 days	6MB	31.01	5e-4
DVGO	NVIDIA RTX2080 Ti	15 mins	612.1	31.95	0.1
Plenoxel	NVIDIA Titan RTX	11 mins	778.1	31.71	30
Instant NGP	NVIDIA RTX 3090	**5 mins**	**50.4**	**33.18**	0.01
Instant NGP's NeRF	NVIDIA RTX 3090	**5 mins**	–	30.06	–
TensoRF	NVIDIA V100	17 mins	71.8	**33.14**	0.02

TensoRF: Tensorial Radiance Fields. ECCV 2022.

降低迭代次数：IBRNet

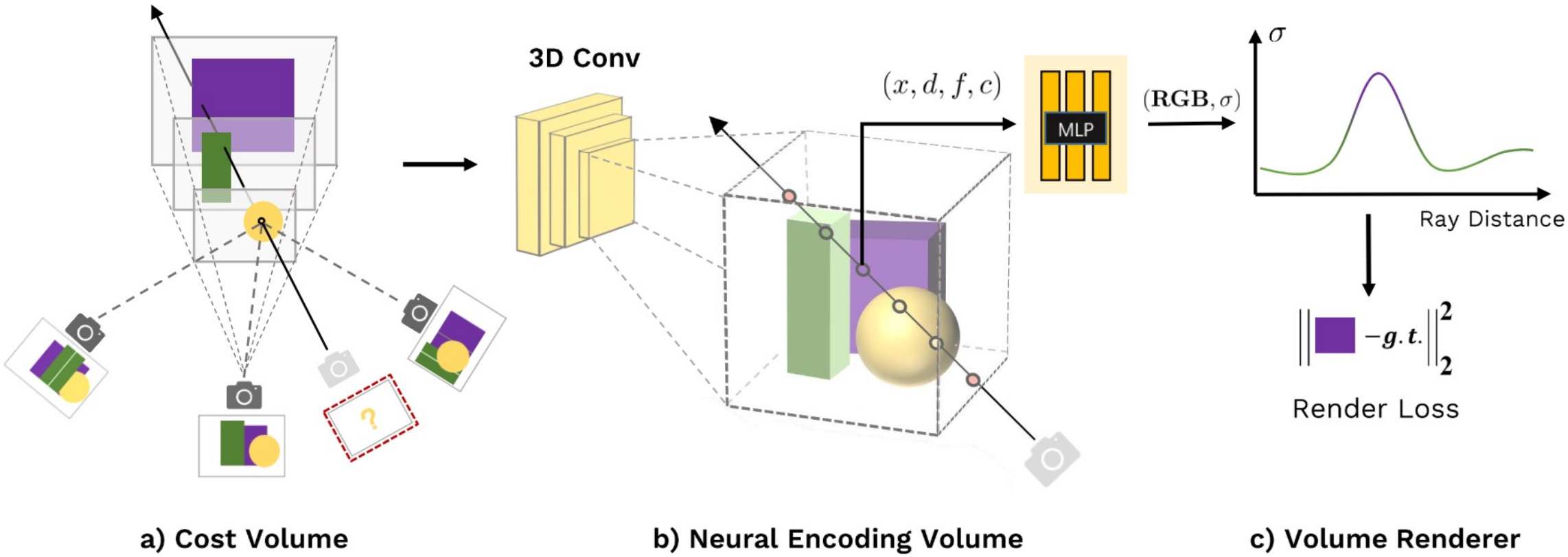
- 使用feedforward model学习data prior



IBRNet: Learning Multi-View Image-Based Rendering. CVPR 2021.

降低迭代次数：MVSNeRF

- 使用feedforward model学习data prior



MVSNeRF: Fast Generalizable Radiance Field Reconstruction from Multi-View Stereo. ICCV 2021.

降低迭代次数：MVSNeRF

- feedforward model的data prior可以加速收敛

Method	Settings	Synthetic Data (NeRF [31])			Real Data (DTU [20] / Forward-Facing [33])		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
PixelNeRF [54]	No per-scene optimization	7.39	0.658	0.411	19.31/11.24	0.789/0.486	0.382/0.671
IBRNet [46]		22.44	0.874	0.195	26.04/21.79	0.917/0.786	0.190/0.279
Ours		**23.62**	**0.897**	**0.176**	**26.63/21.93**	**0.931/0.795**	**0.168/0.252**
$\text{NeRF}_{10.2h}$ [34]	Per-scene optimization	**30.63**	**0.962**	**0.093**	27.01/**25.97**	0.902/0.870	0.263/ 0.236
$\text{IBRNet}_{ft-1.0h}$ [46]		25.62	0.939	0.110	**31.35**/24.88	**0.956**/0.861	**0.131/0.189**
$\text{Ours}_{ft-15min}$		27.07	0.931	0.168	28.50/25.45	0.933/**0.877**	0.179/0.192

MVSNeRF: Fast Generalizable Radiance Field Reconstruction from Multi-View Stereo. ICCV 2021.

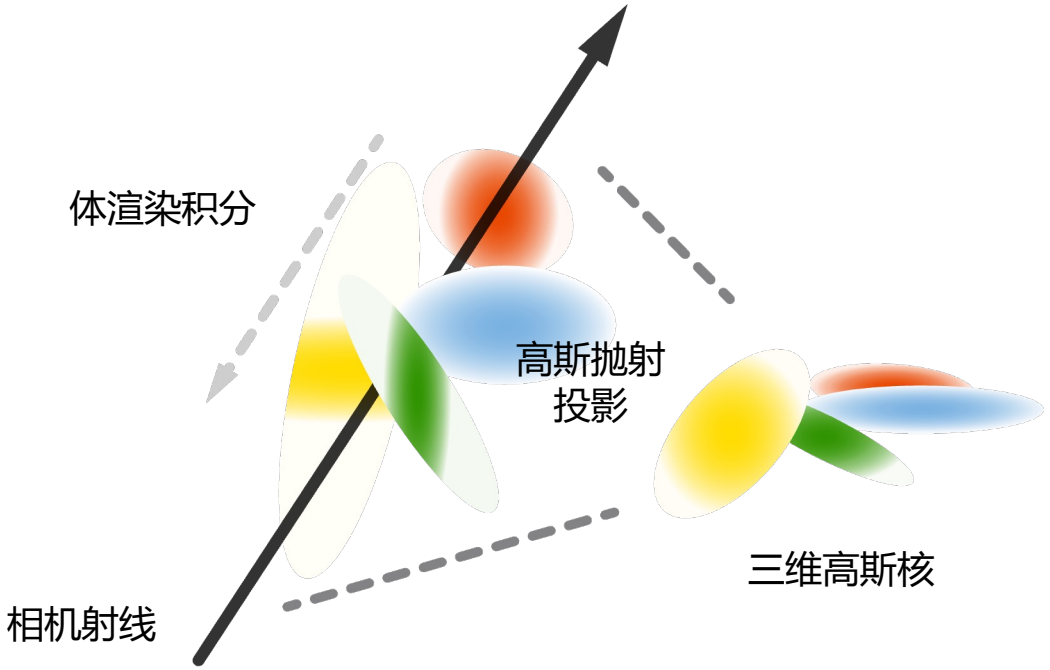
降低三维点的数量&逐点训练成本&迭代次数：3DGS

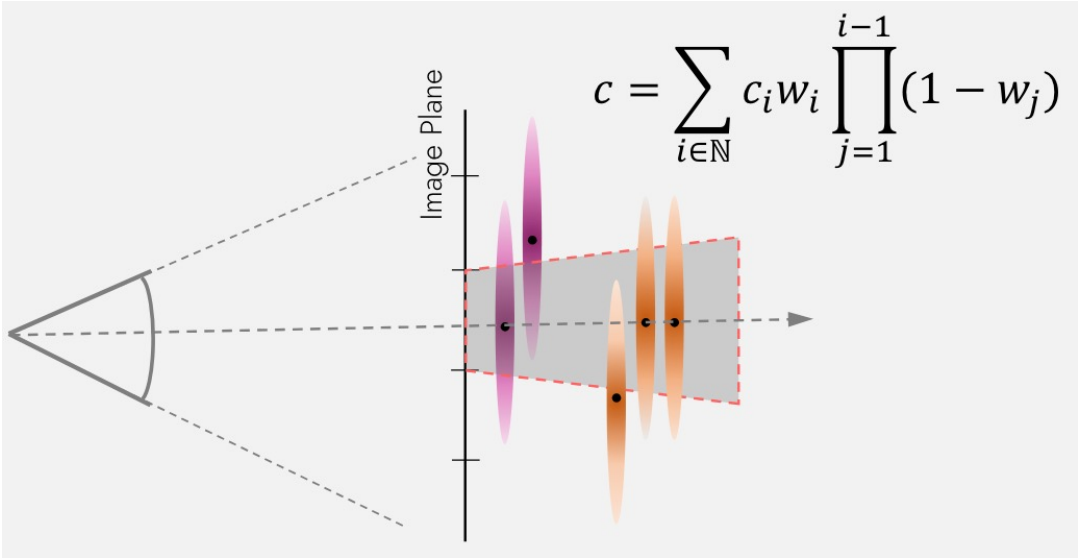
- 基于一组三维高斯表示场景，每个三维高斯包含position, rotation, opacity和Spherical harmonics系数。

$$p = \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sigma_x & \sigma_{xy} & \sigma_{xz} \\ & \sigma_y & \sigma_{yz} \\ & & \sigma_z \end{bmatrix} \quad o \quad SH$$

spherical harmonics

covariance matrix





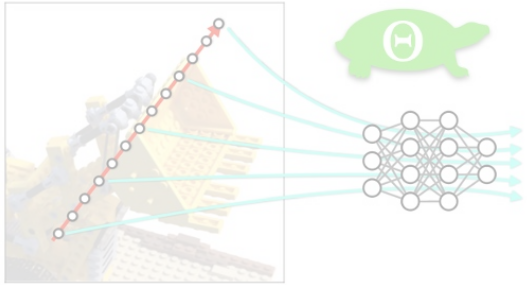
3D Gaussian Splatting for Real-Time Radiance Field Rendering. SIGGRAPH 2023.

https://3dgstutorial.github.io/3dv_part1.pdf

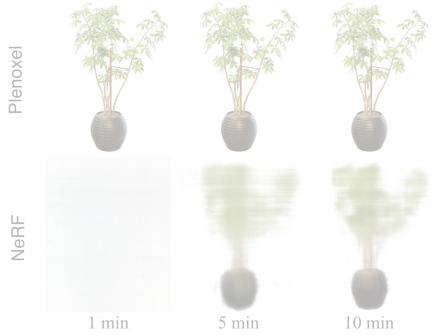
NeRF面临的问题3：场景建模能力弱

NeRF在静态场景建模上遇到的问题

渲染速度慢

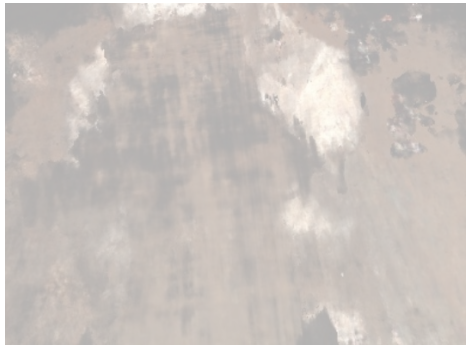


训练速度慢



场景建模能力弱

模型鲁棒性差



NeRF在其他类型场景建模上遇到的问题

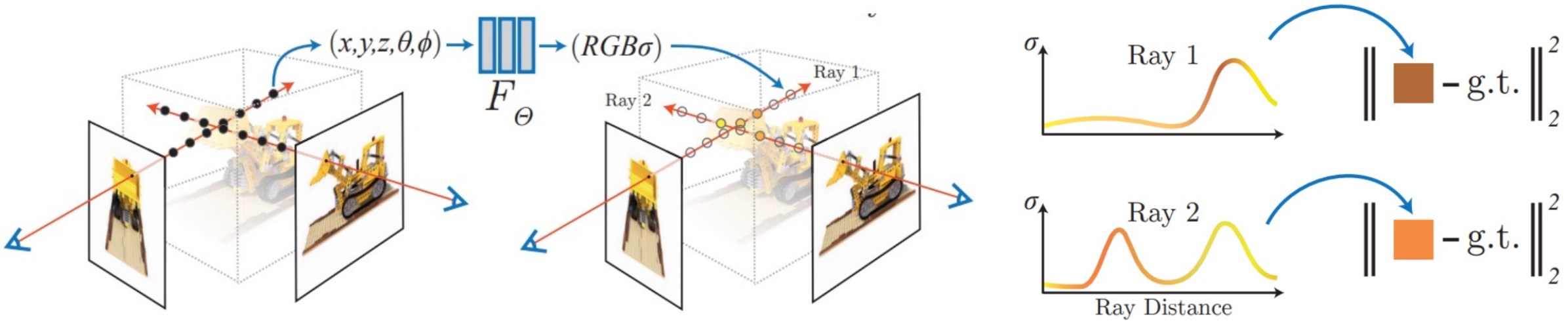
几何重建质量低

无法建模多模态信号

无法建模动态场景

NeRF无法建模大场景的原因

- NeRF的MLP network capability有限
- NeRF沿着camera ray采样三维点，因此无法处理unbounded scene

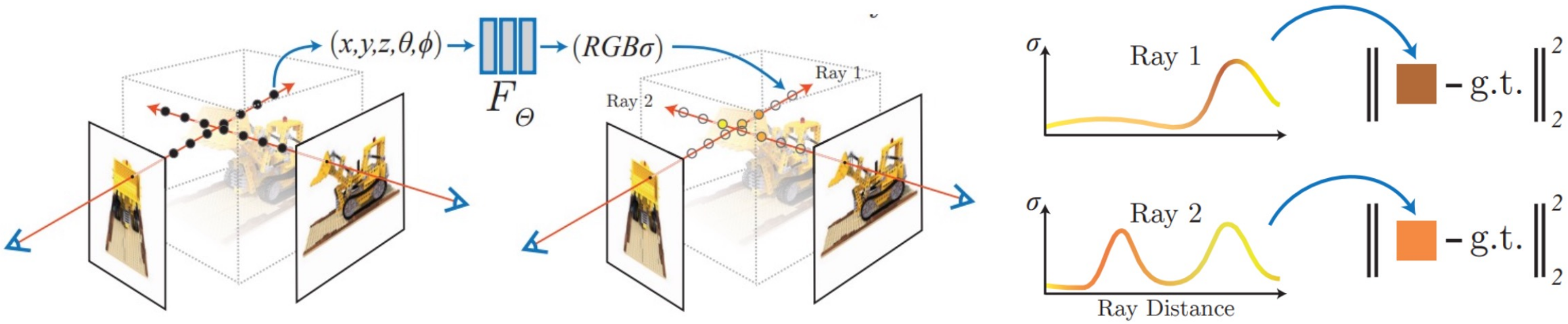


NeRF: Representing scenes as neural radiance fields for view synthesis, ECCV 2020.

提升NeRF大场景建模能力的思路

- NeRF的MLP network capability有限
- NeRF沿着camera ray采样三维点，因此无法处理unbounded scene

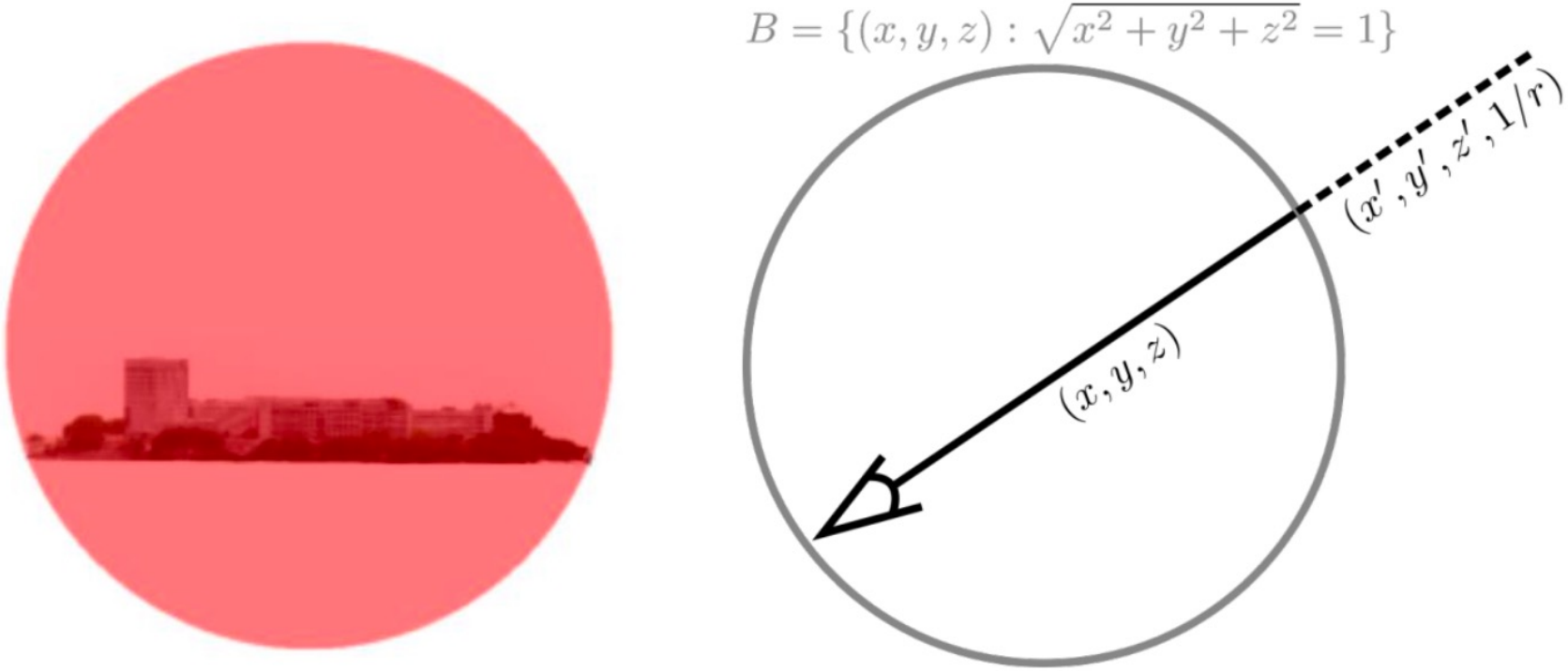
- 解决思路：(1) 设计新的场景表示；(2) 提升模型表达能力。



NeRF: Representing scenes as neural radiance fields for view synthesis, ECCV 2020.

设计场景表示：NeRF++

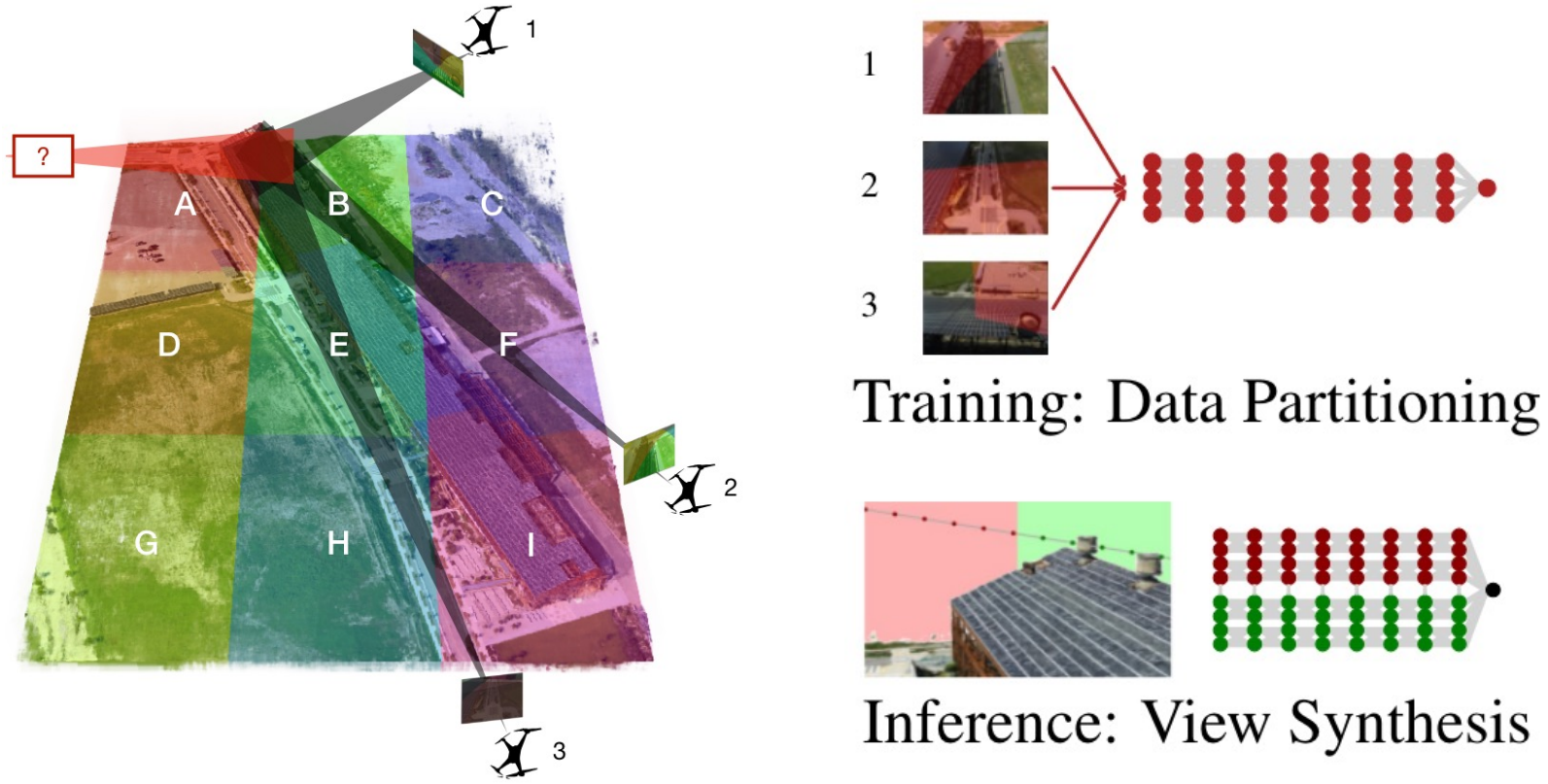
- NeRF++设定了一个半径范围，在半径内基于均匀三维点建模场景，在半径外基于一组球面建模场景



Nerf++: Analyzing and improving neural radiance fields. arXiv 2020.

提升模型表达能力：MegaNeRF

- 将大场景划分为一组区域，每个场景用一个MLP network表示



Mega-NeRF: Scalable Construction of Large-Scale NeRFs for Virtual Fly-Throughs. CVPR 2022.

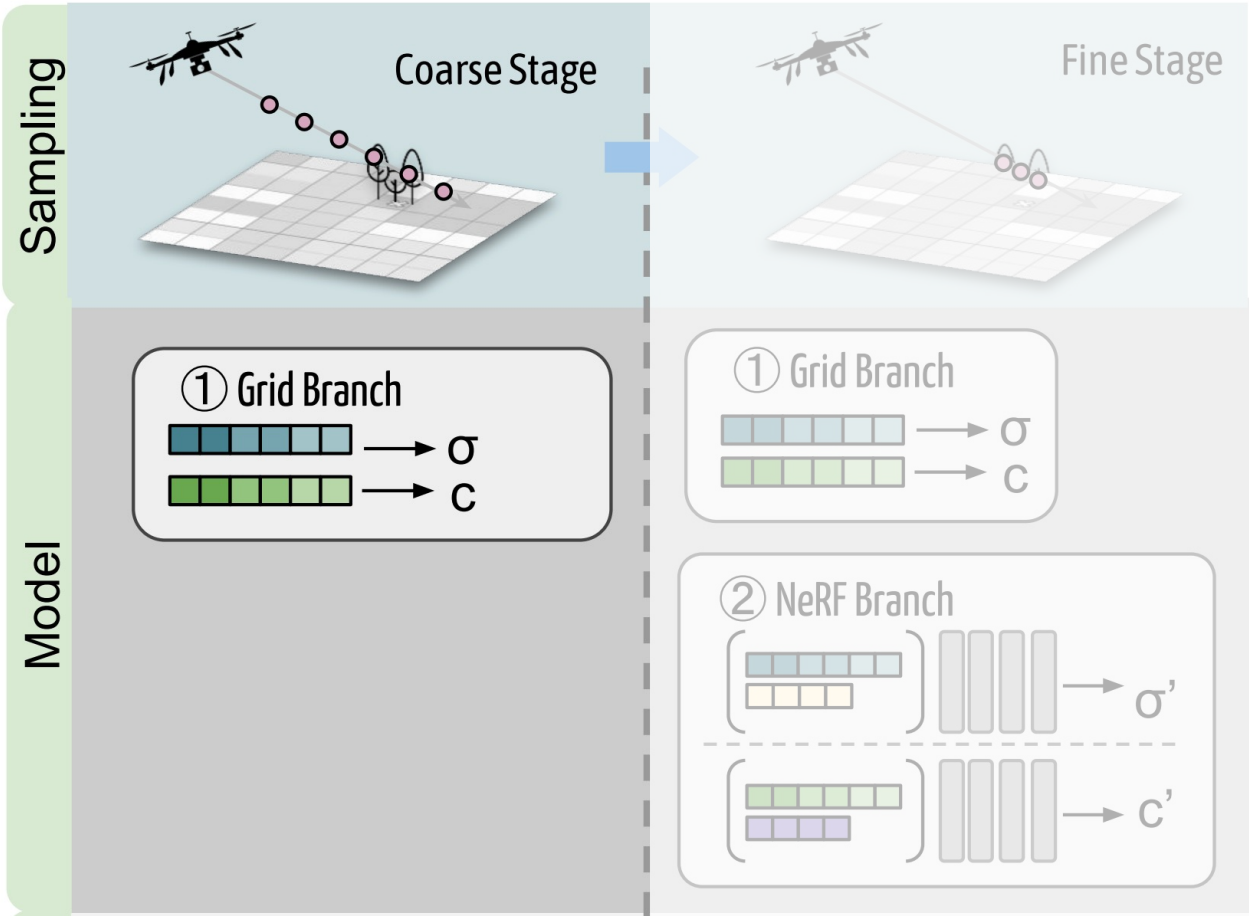
提升模型表达能力：BlockNeRF

- 将大场景划分为一组区域，每个场景用一个MLP network表示

Block-NeRF: Scalable Large Scene Neural View Synthesis. CVPR 2022.

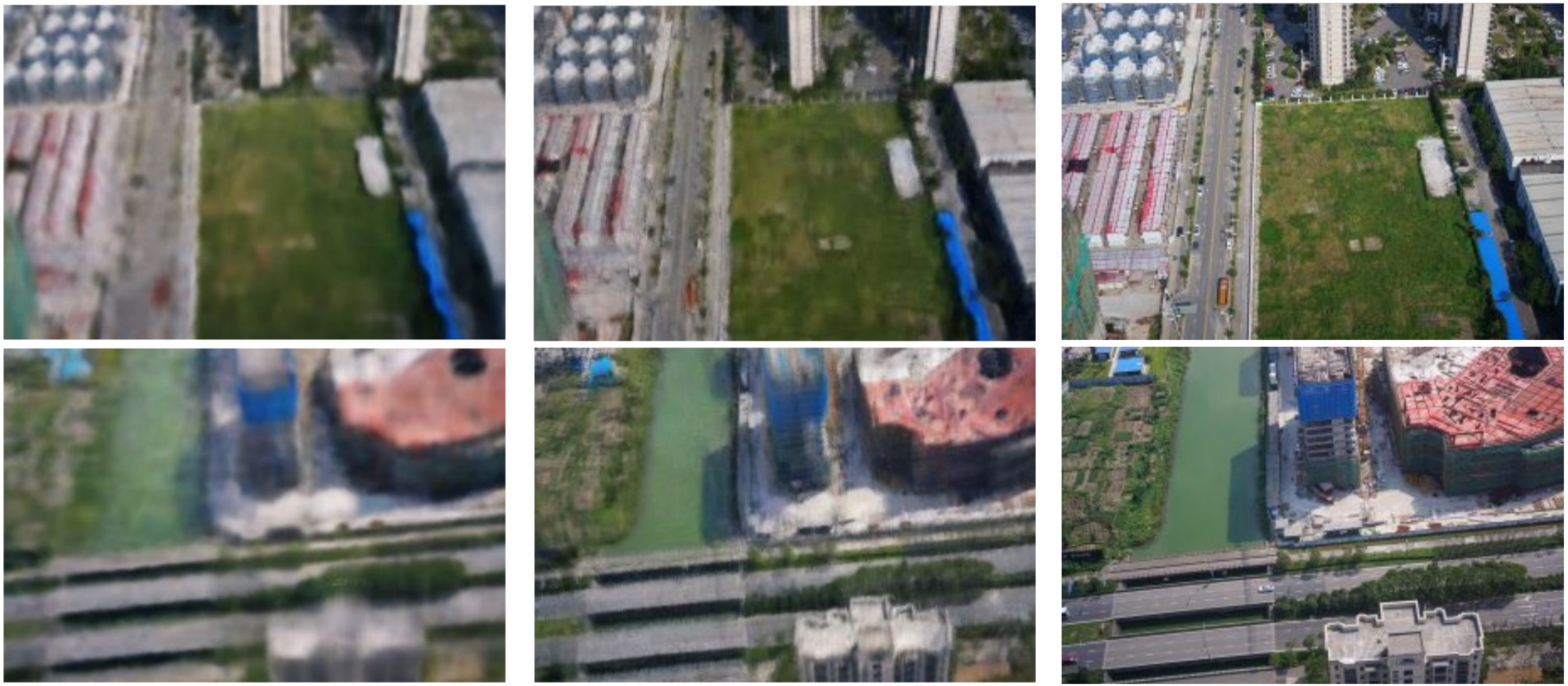
提升模型表达能力：Grid-NeRF

- 使用平面网格特征与小型MLP网络表示三维场景



Grid-guided Neural Radiance Fields for Large Urban Scenes. CVPR 2023.

基于网格特征的方法存在的问题

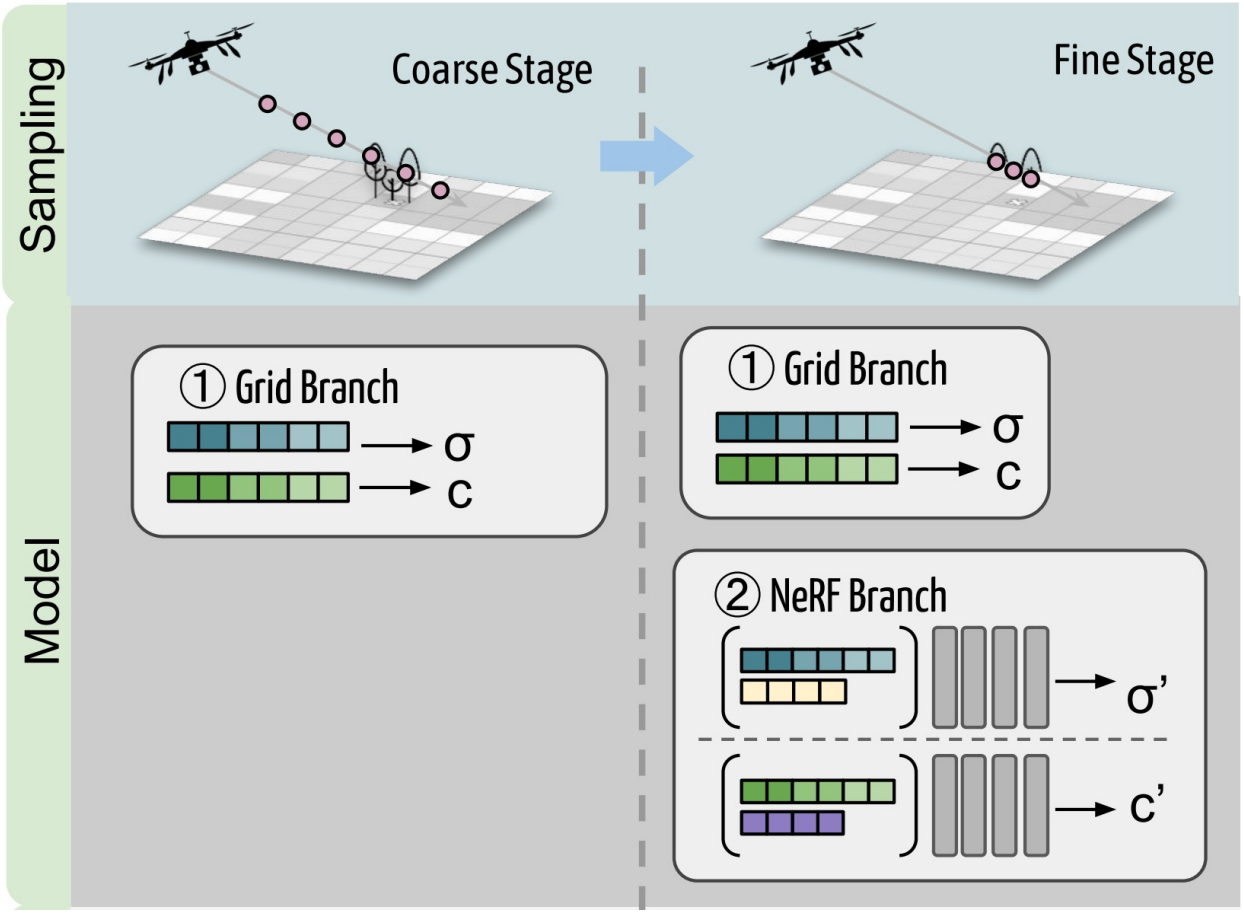


NeRF　　Plane-based method　　Ground truth

Grid-guided Neural Radiance Fields for Large Urban Scenes. CVPR 2023.

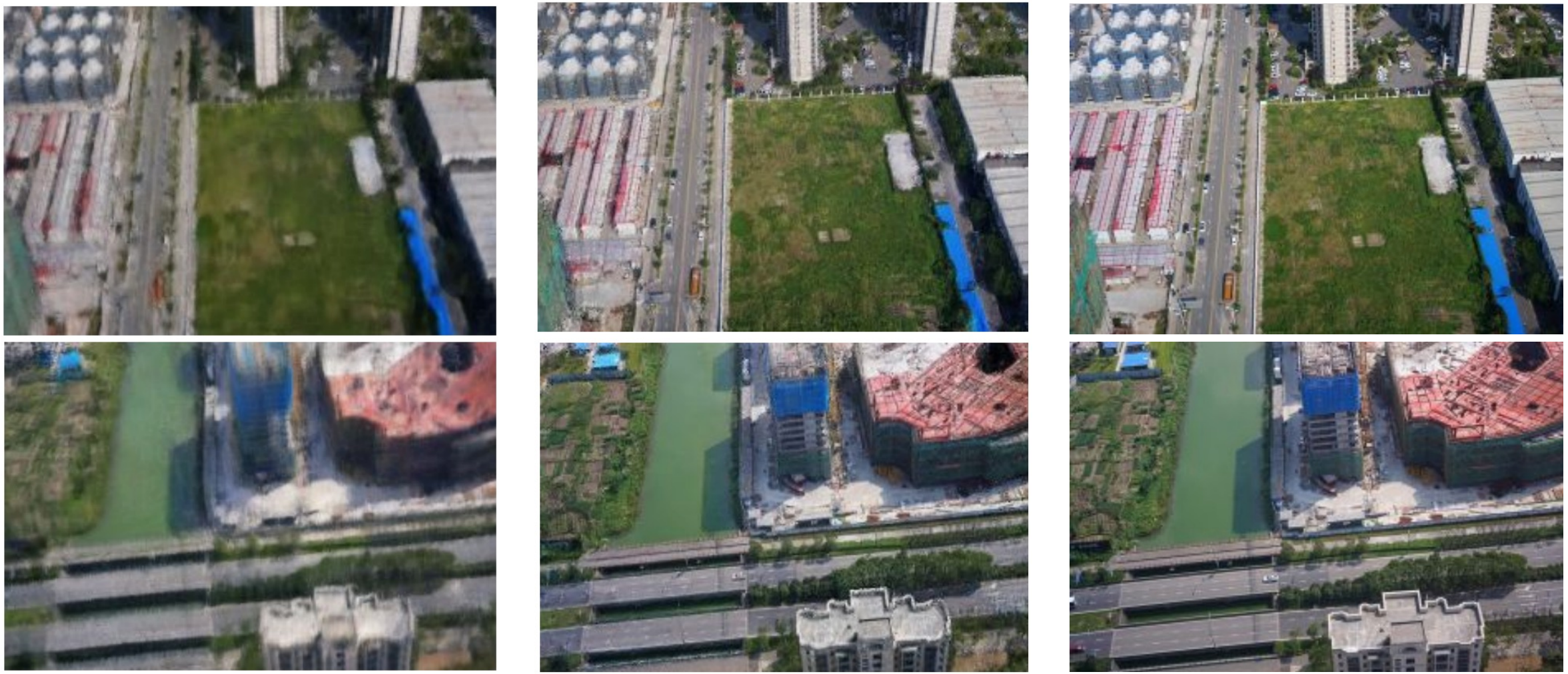
提升模型表达能力：Grid-NeRF

- 利用NeRF branch来引导场景的优化过程



Grid-guided Neural Radiance Fields for Large Urban Scenes. CVPR 2023.

提升模型表达能力：Grid-NeRF



Plane-based method　　Full method　　Ground truth

Grid-guided Neural Radiance Fields for Large Urban Scenes. CVPR 2023.

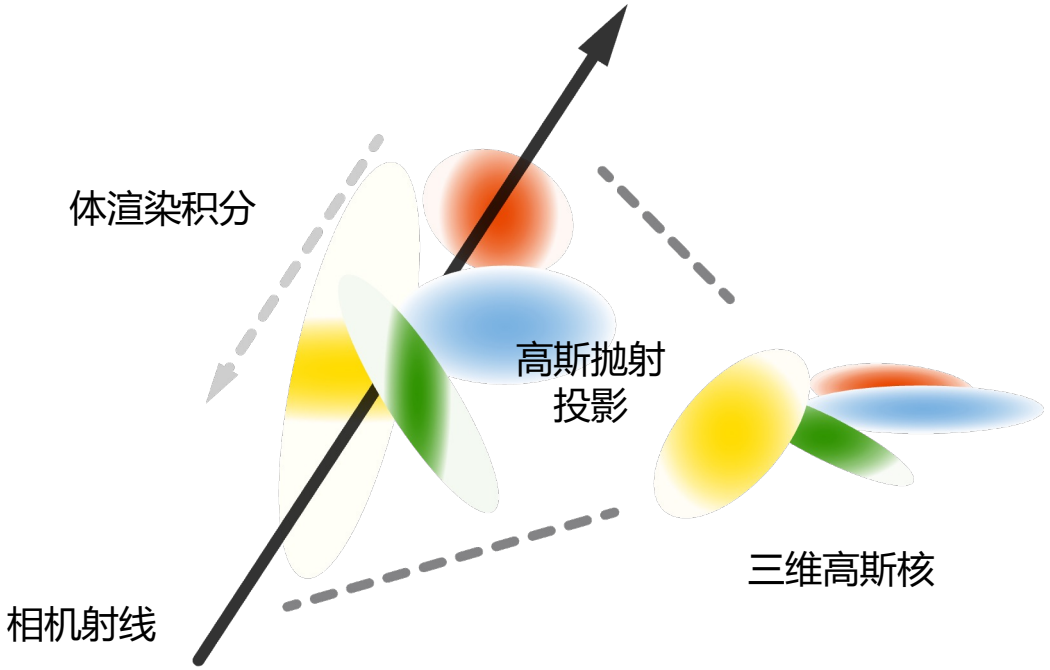
设计场景表示&提升模型表达能力：3DGS

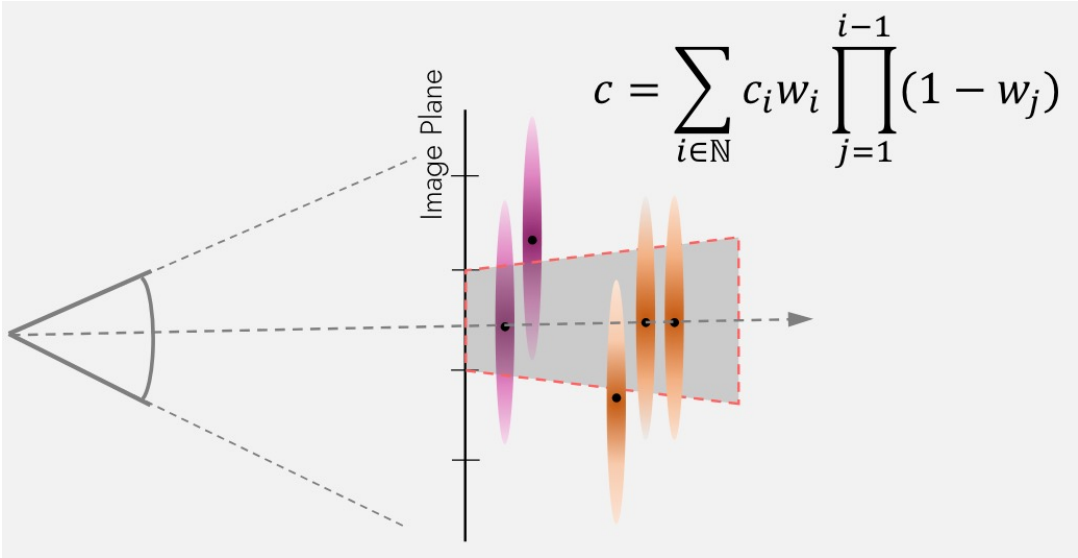
- 基于一组三维高斯表示场景，每个三维高斯包含position, rotation, opacity和Spherical harmonics系数。

$$p = \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sigma_x & \sigma_{xy} & \sigma_{xz} \\ & \sigma_y & \sigma_{yz} \\ & & \sigma_z \end{bmatrix} \quad o \quad SH$$

spherical harmonics

covariance matrix





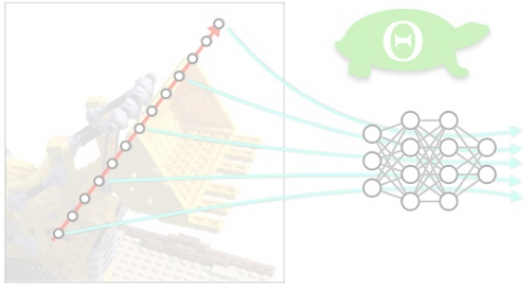
3D Gaussian Splatting for Real-Time Radiance Field Rendering. SIGGRAPH 2023.

https://3dgstutorial.github.io/3dv_part1.pdf

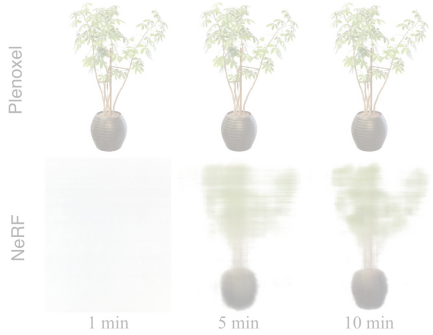
NeRF面临的问题4：模型鲁棒性差

NeRF在静态场景建模上遇到的问题

渲染速度慢

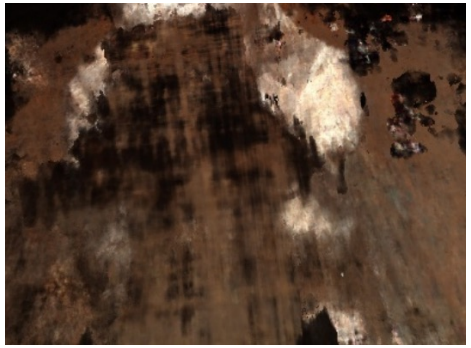


训练速度慢



场景建模能力弱

模型鲁棒性差



NeRF在其他类型场景建模上遇到的问题

几何重建质量低

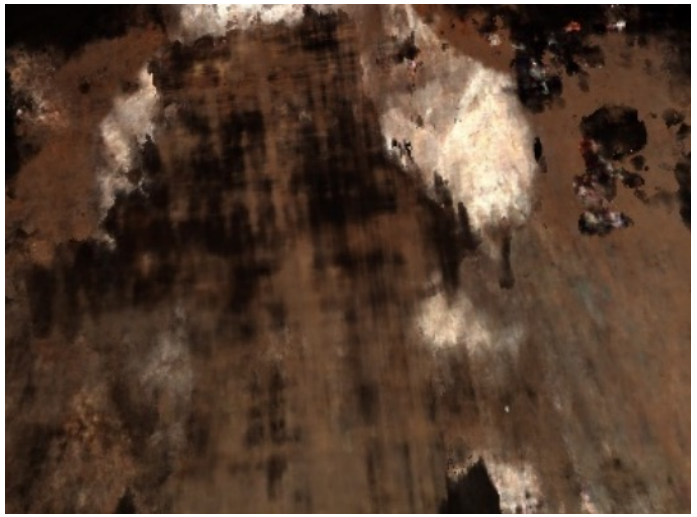
无法建模多模态信号

无法建模动态场景

NeRF模型不鲁棒的体现

相机位姿不准确

输入视角稀疏

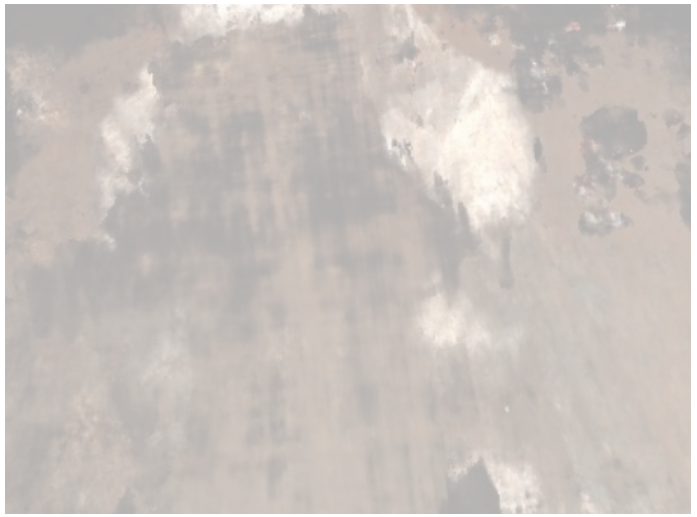


光照变化

NeRF模型不鲁棒1：对相机位姿精度敏感

相机位姿不准确

输入视角稀疏



光照变化

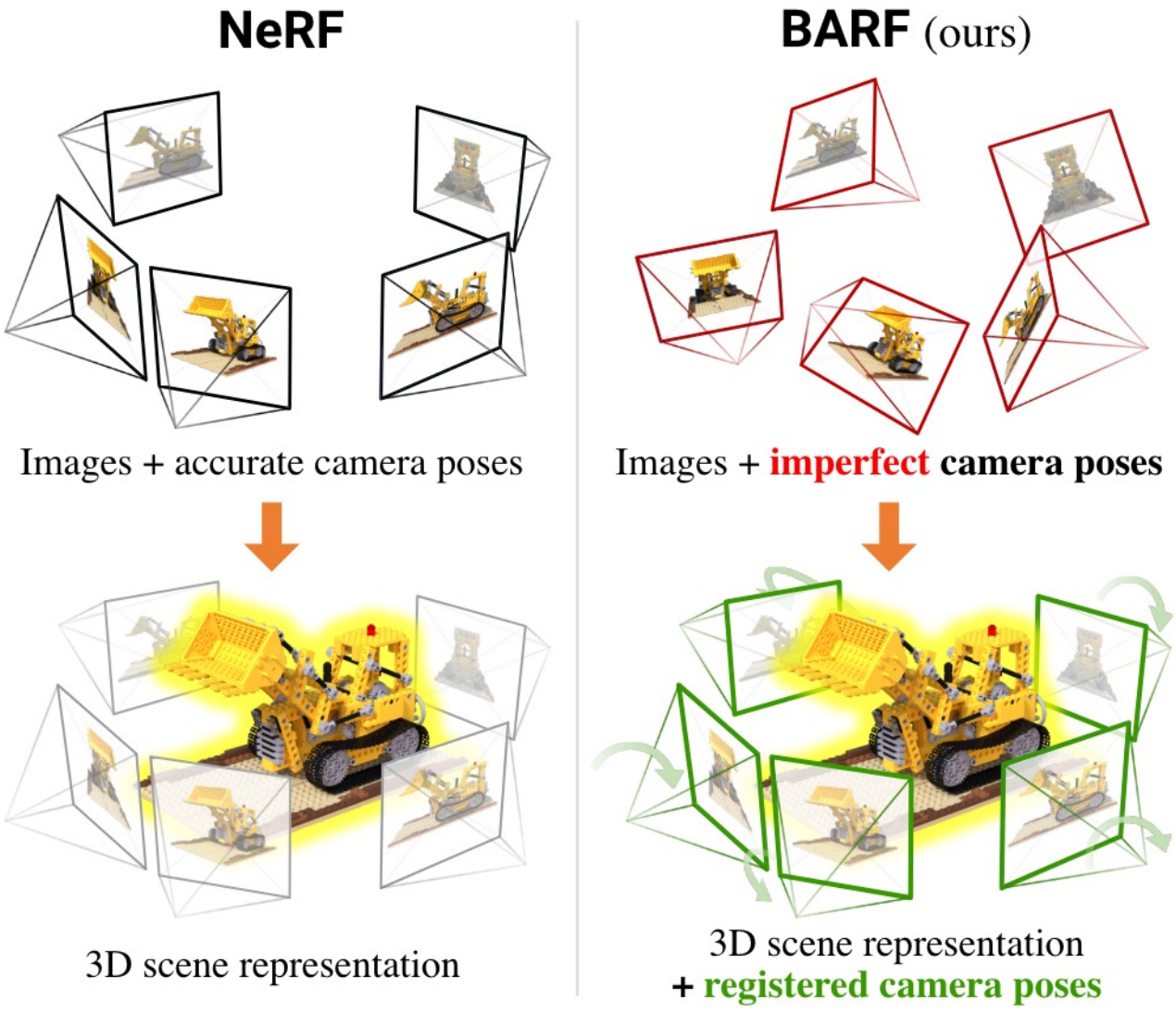
BARF

- 给定带误差的camera poses与相应的多视角图片，重建NeRF。

$$\min_{\mathbf{p}_1,\dots,\mathbf{p}_M,\boldsymbol{\Theta}} \sum_{i=1}^{\overset{\text{frames}}{M}} \sum_{\mathbf{u}} \Big\| \underbrace{\hat{\mathcal{I}}(\mathbf{u};\mathbf{p}_i,\boldsymbol{\Theta})}_{\substack{\text{RGB}\\ \text{(rendered)}}} - \underbrace{\mathcal{I}_i(\mathbf{u})}_{\text{RGB}} \Big\|_2^2$$

camera pose

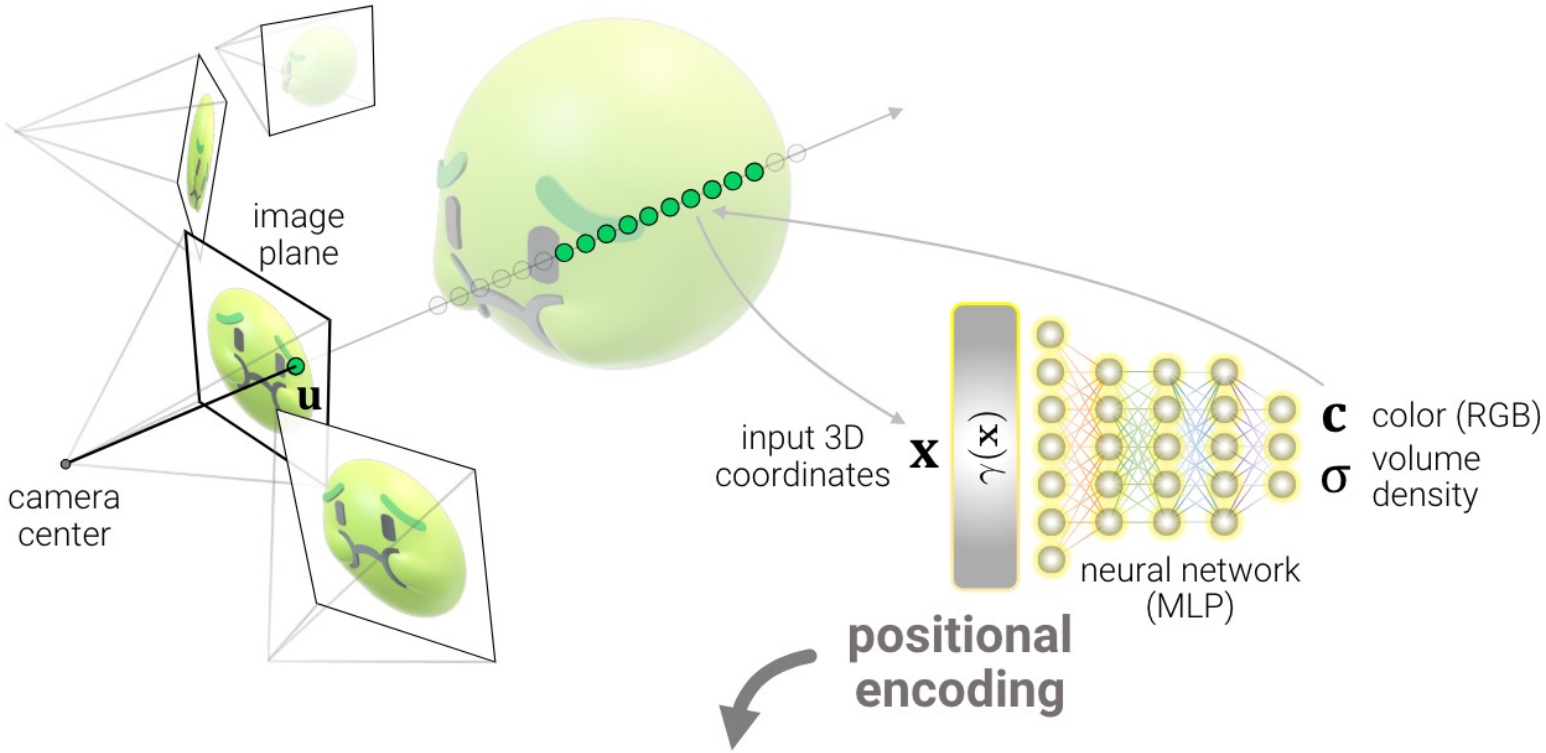
network params.



BARF: Bundle-Adjusting Neural Radiance Fields. ICCV 2021.

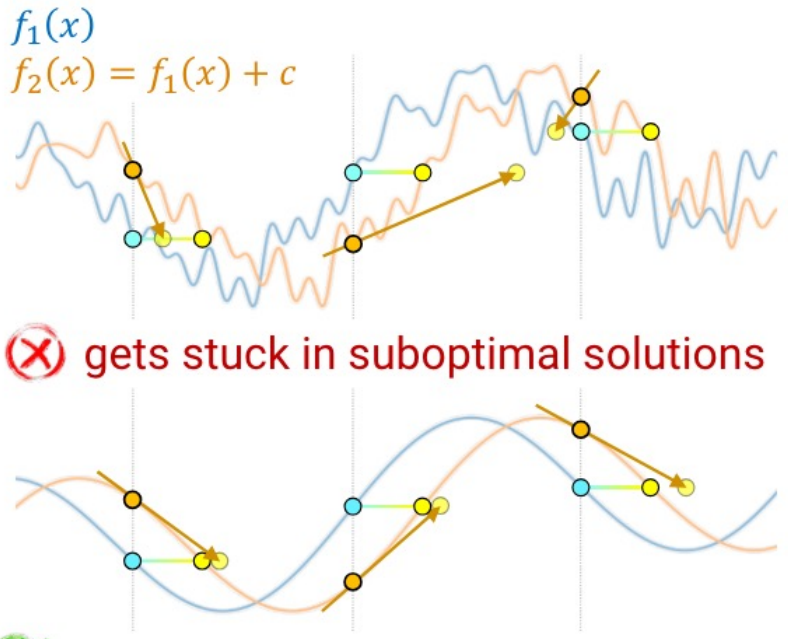
https://chenhsuanlin.bitbucket.io/bundle-adjusting-NeRF/poster.pdf

Positional encoding对Pose优化带来影响



encourages representation with high frequency

detrimental to gradient-based **registration**!!



gets stuck in suboptimal solutions

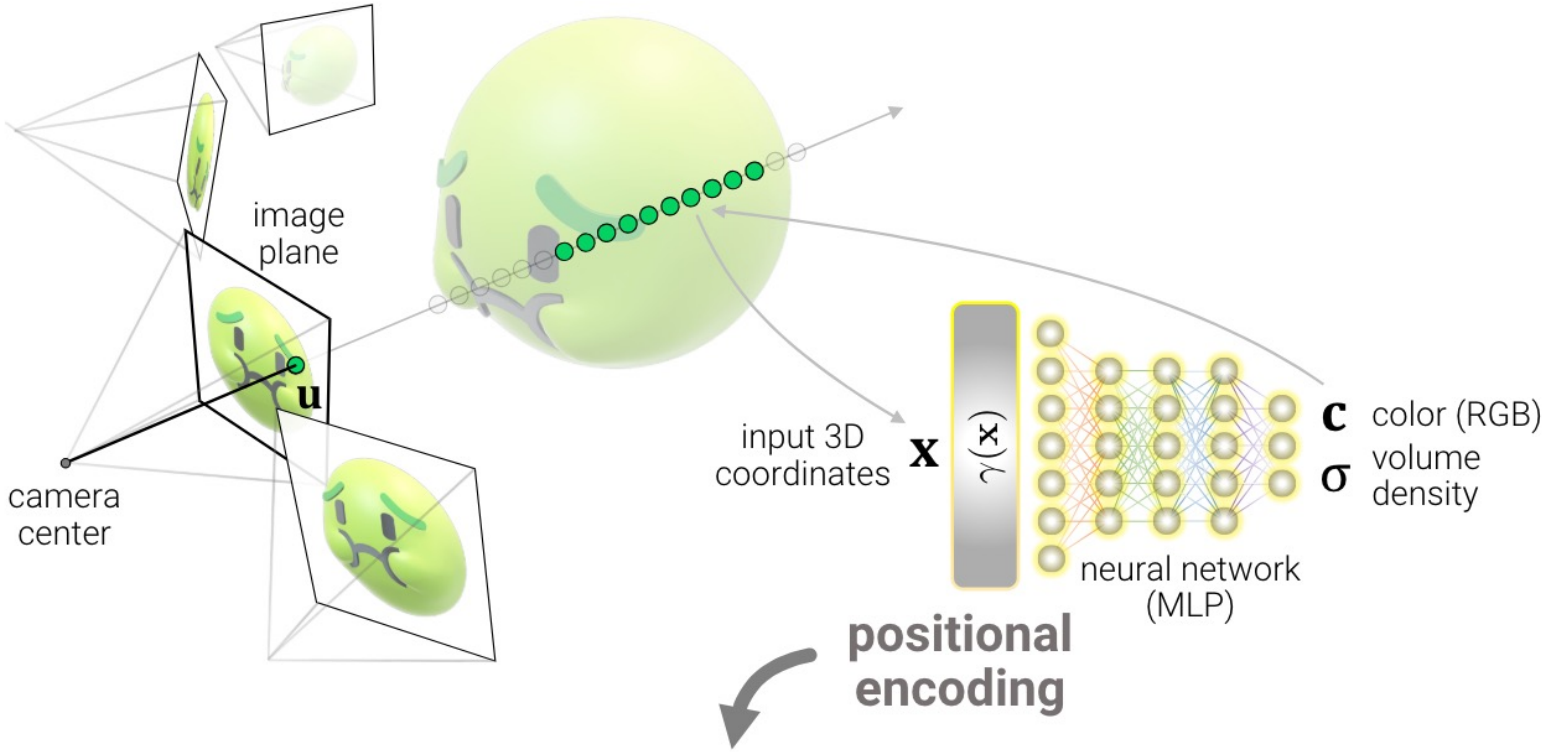
smooth signals ➔ coherent updates

BARF: Bundle-Adjusting Neural Radiance Fields. ICCV 2021.

https://chenhsuanlin.bitbucket.io/bundle-adjusting-NeRF/poster.pdf

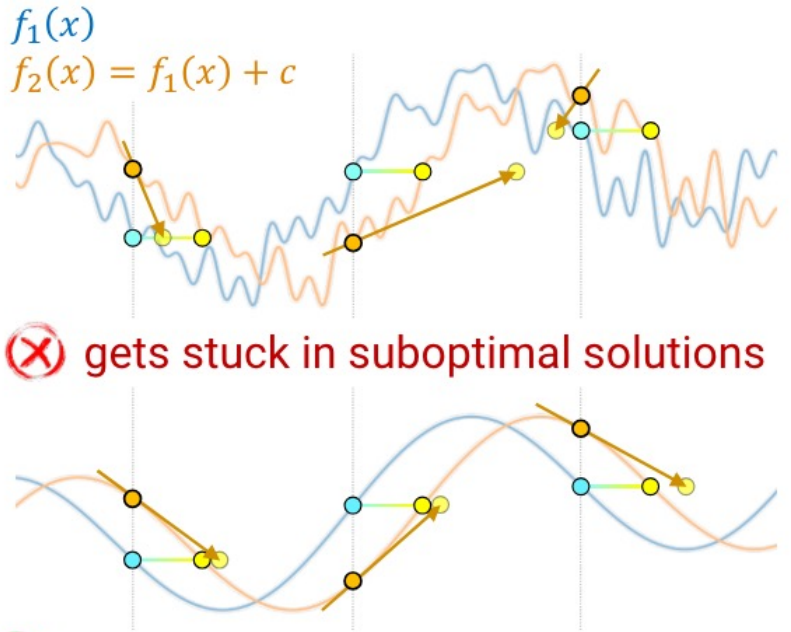
BARF的解决方法

- Coarse–to–fine refinement：逐渐增加positional encoding的高频部分。



encourages representation with high frequency

detrimental to gradient-based **registration**!!



gets stuck in suboptimal solutions

smooth signals ➔ coherent updates

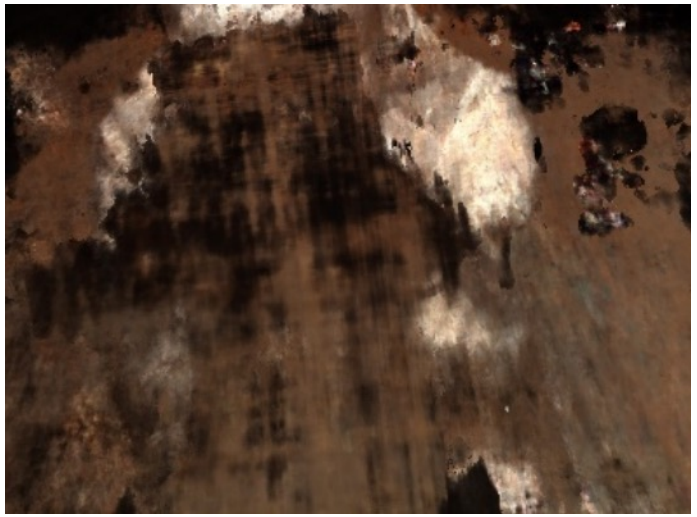
BARF: Bundle–Adjusting Neural Radiance Fields. ICCV 2021.

https://chenhsuanlin.bitbucket.io/bundle–adjusting–NeRF/poster.pdf

NeRF模型不鲁棒2：要求稠密输入视角

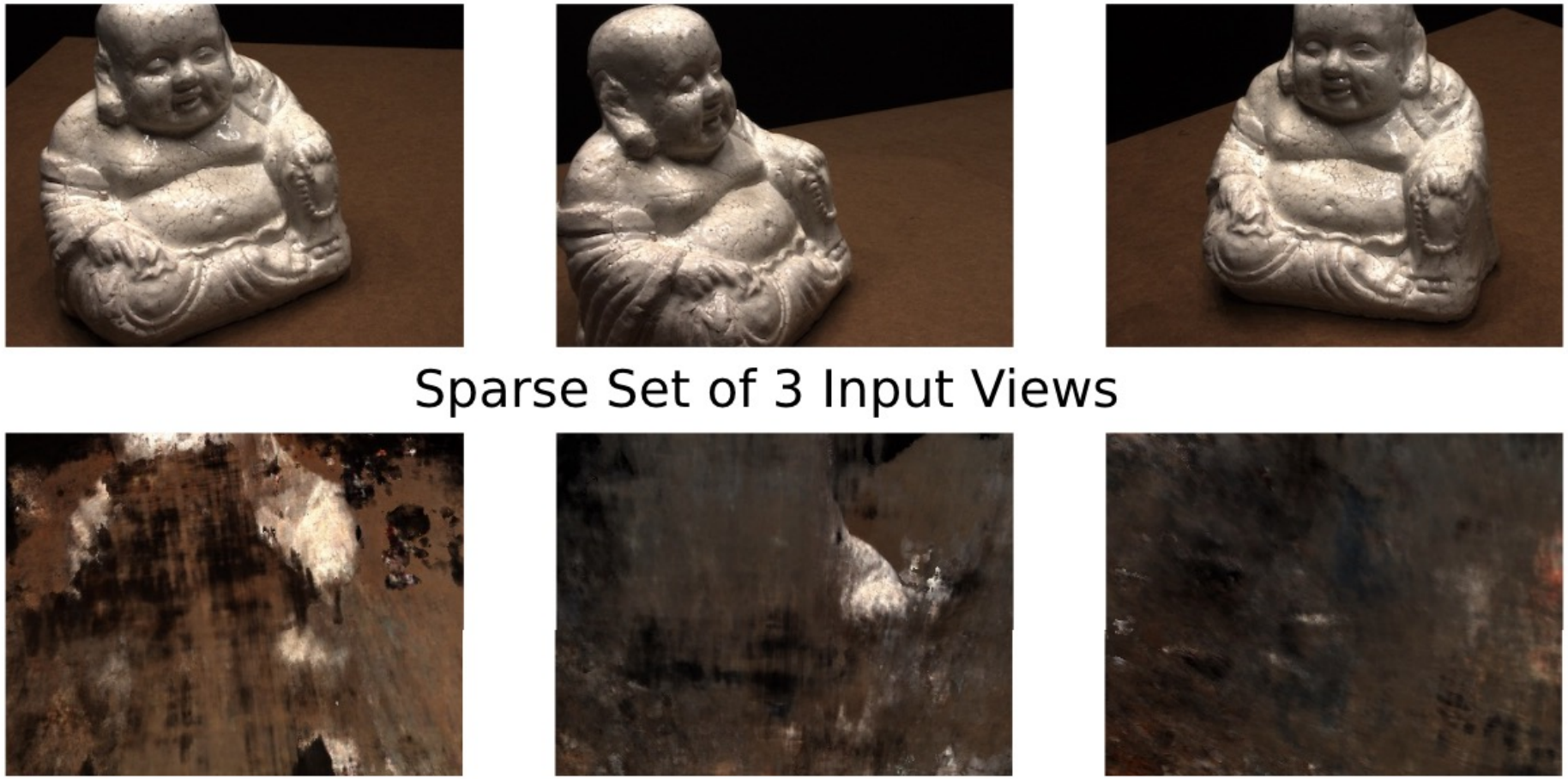
相机位姿不准确

输入视角稀疏



光照变化

NeRF在稀疏视角下表现不佳

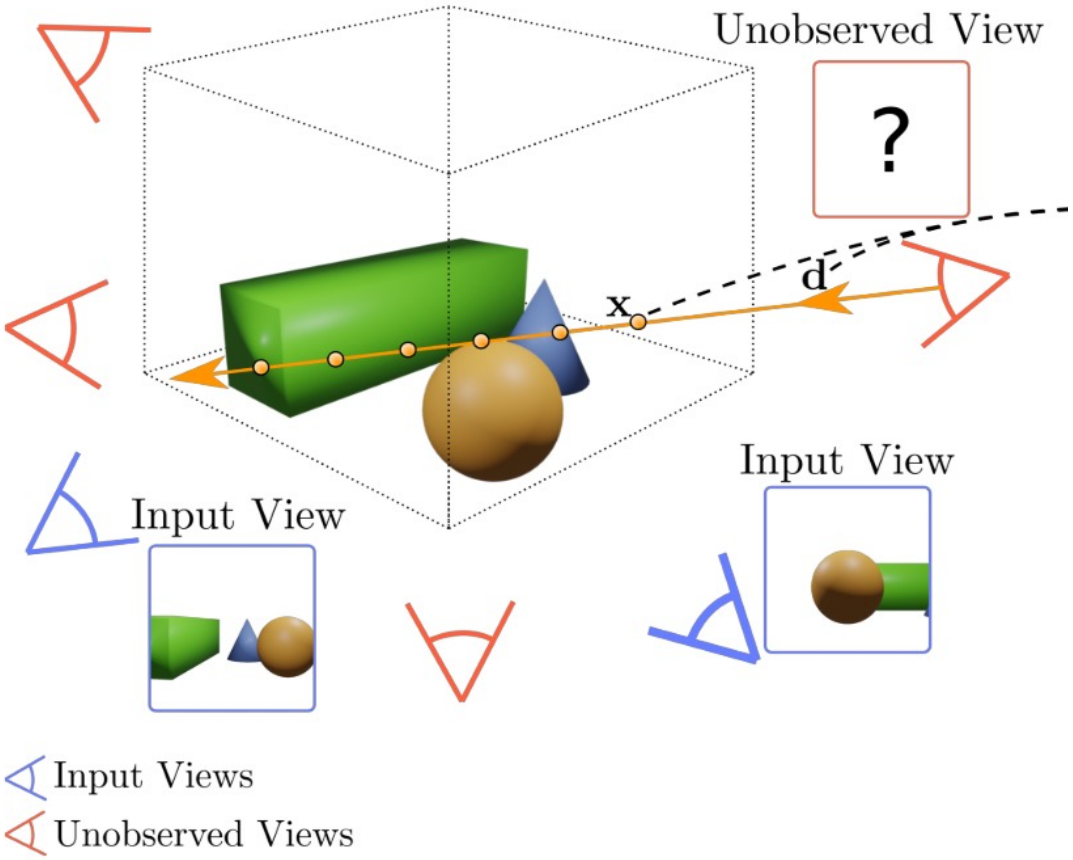


RegNeRF: Regularizing Neural Radiance Fields for View Synthesis from Sparse Inputs. CVPR 2022.

file:///Users/pengsida/Downloads/regnerf_poster.pdf

提升稀疏视角重建效果的思路

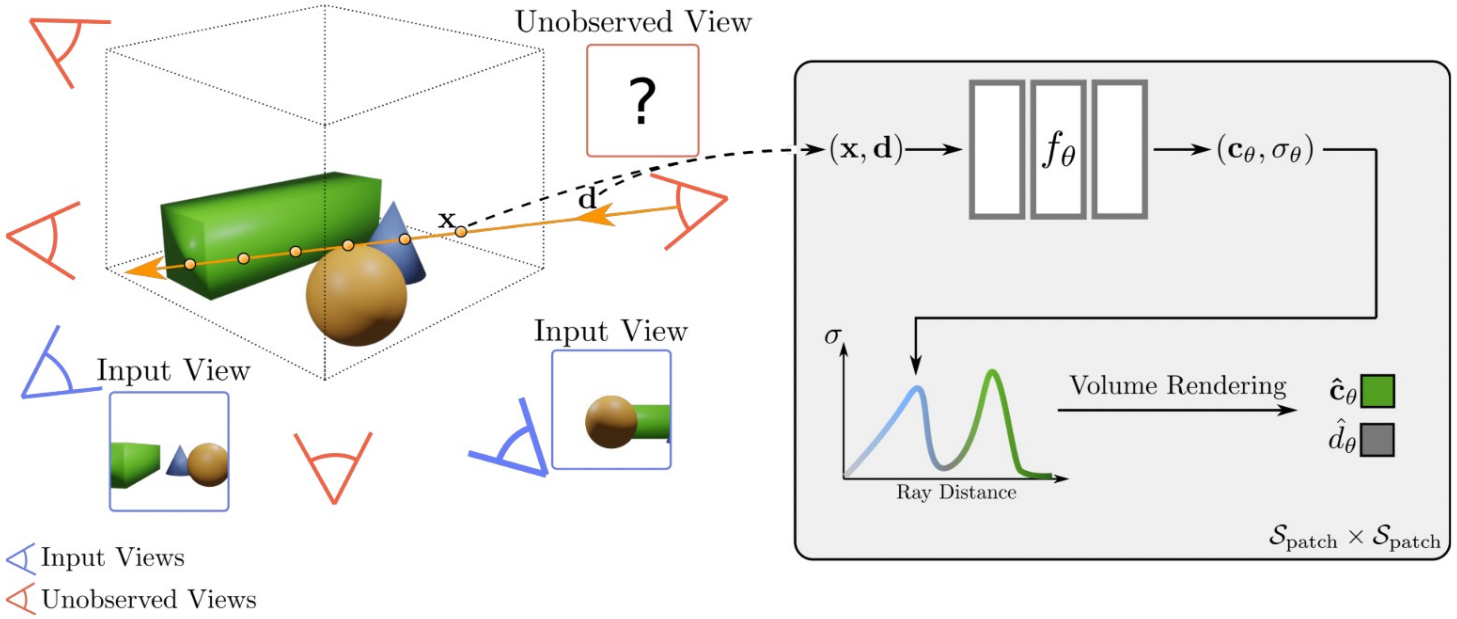
- 使用先验约束未观测到的视角
  1. Smoothness Prior: RegNeRF
  2. Depth Prior: SinNeRF
  3. Generative Prior: ReconFusion
- 使用Feedforward model直接预测
  1. Transformer：LRM、GRM
  2. Diffusion model：iNVS



RegNeRF: Regularizing Neural Radiance Fields for View Synthesis from Sparse Inputs. CVPR 2022.

使用先验约束未观测到的视角： RegNeRF

- (1) Normalizing flow prior; (2) Smoothness prior



1. Appearance Regularization:

$$\hat{P} = \{\hat{\mathbf{c}}_\theta(\mathbf{r}_{ij})\}_{ij}$$

$$\longrightarrow \mathcal{L}_{\text{NLL}}(\theta) = -\log p_Z\left(\phi\left(\hat{P}\right)\right)$$

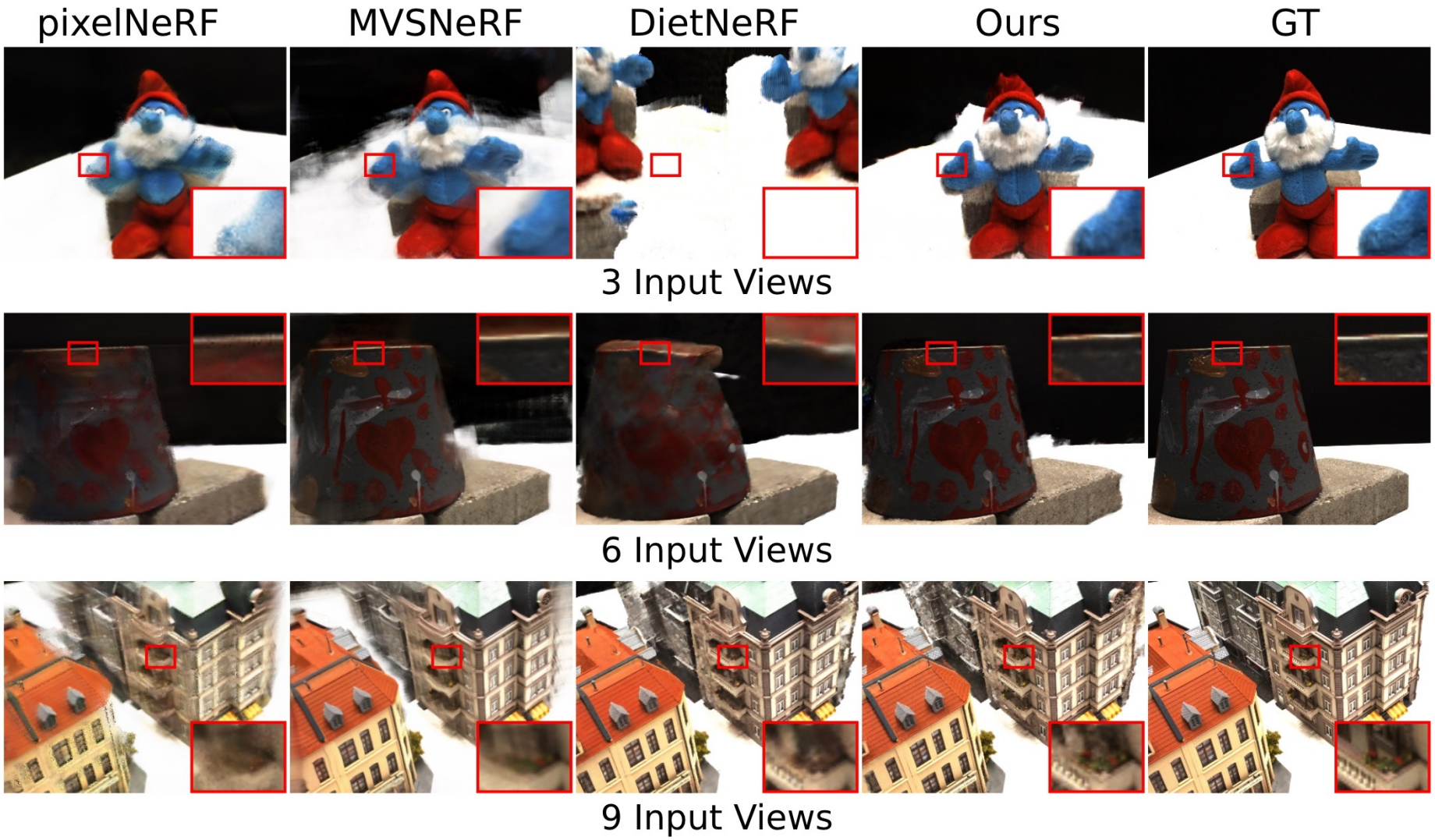
2. Geometry Regularization:

$$\left\{\hat{d}_\theta(\mathbf{r}_{ij})\right\}_{ij}$$

$$\longrightarrow \mathcal{L}_{\text{DS}}(\theta) = \sum_{i,j} \left(\hat{d}_\theta(\mathbf{r}_{ij}) - \hat{d}_\theta(\mathbf{r}_{i+1j})\right)^2 + \left(\hat{d}_\theta(\mathbf{r}_{ij}) - \hat{d}_\theta(\mathbf{r}_{ij+1})\right)^2$$

RegNeRF: Regularizing Neural Radiance Fields for View Synthesis from Sparse Inputs. CVPR 2022.

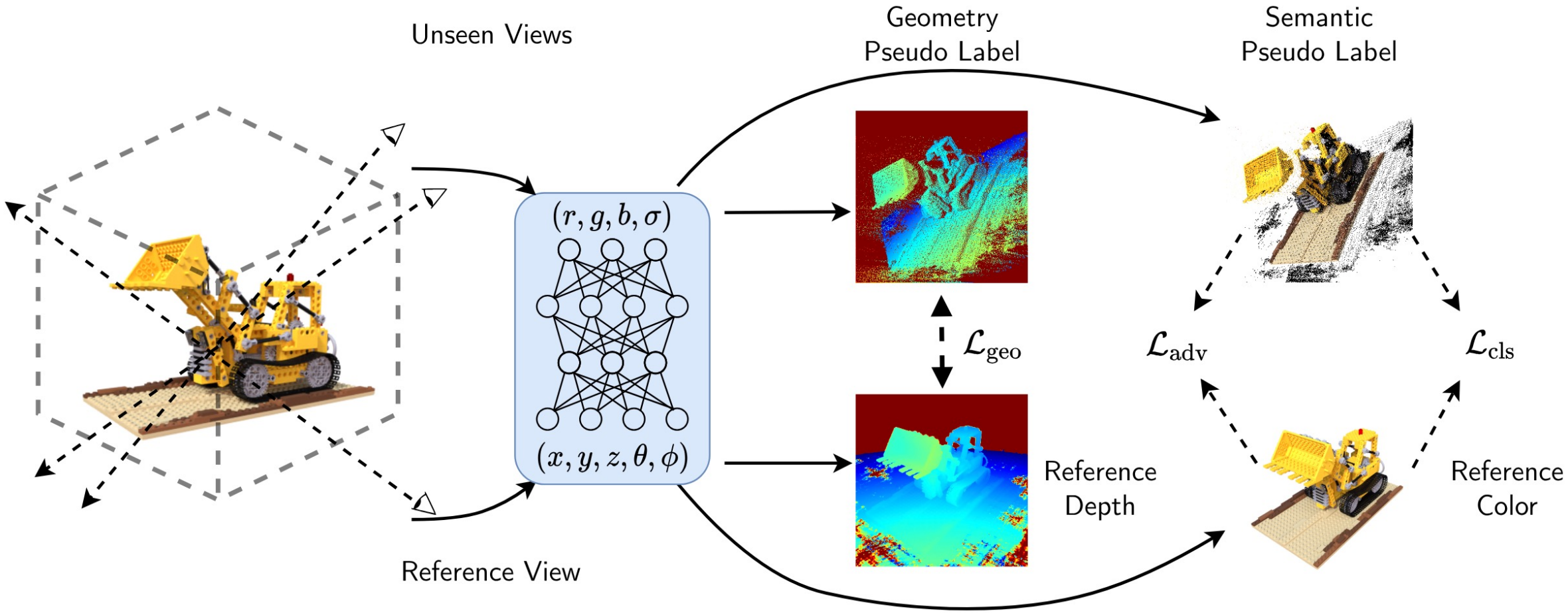
DTU数据集上的对比



RegNeRF: Regularizing Neural Radiance Fields for View Synthesis from Sparse Inputs. CVPR 2022.

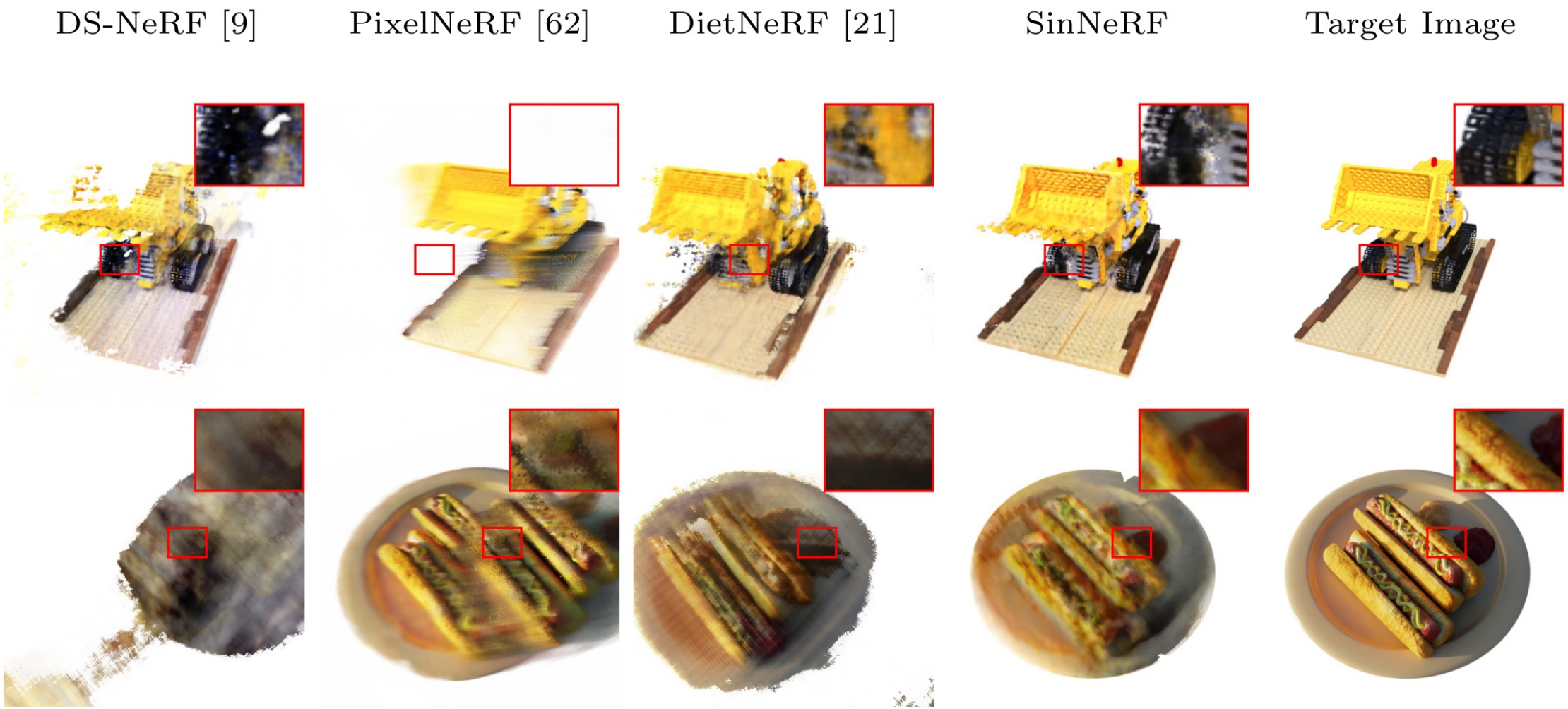
使用先验约束未观测到的视角：SinNeRF

- 利用Depth map约束未观测到的视角：(1) Adversarial loss; (2) Perceptual loss



SinNeRF: Training Neural Radiance Fields on Complex Scenes from a Single Image. ECCV 2022.

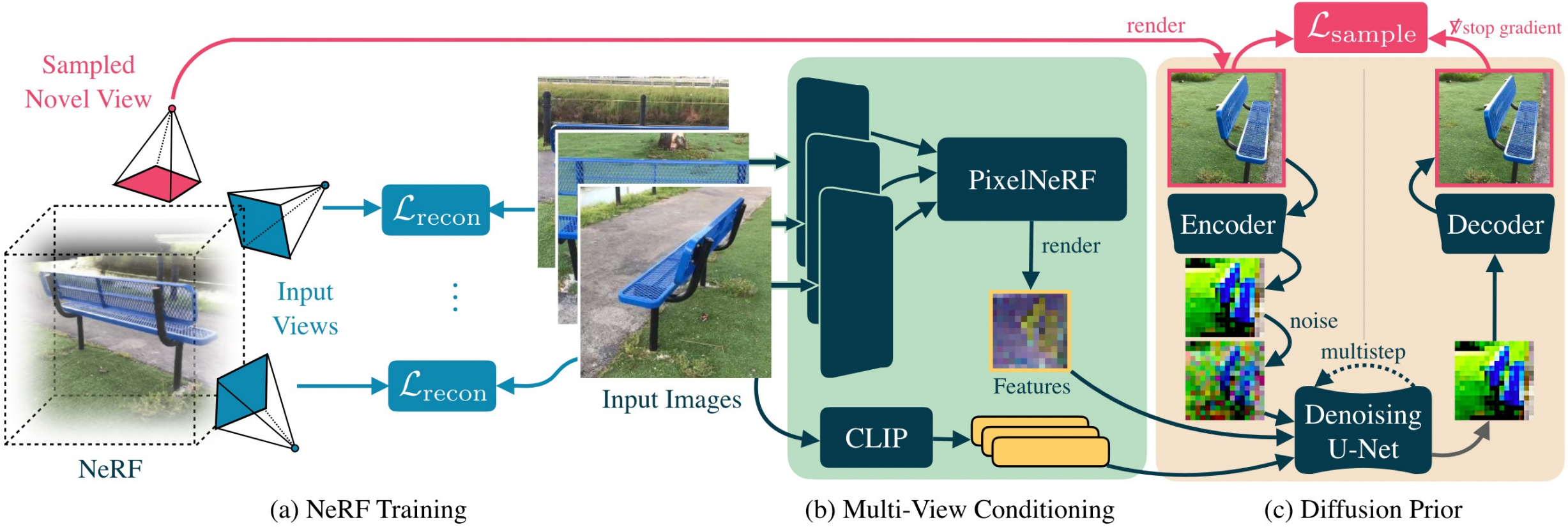
Synthetic-NeRF数据集上的对比



SinNeRF: Training Neural Radiance Fields on Complex Scenes from a Single Image. ECCV 2022.

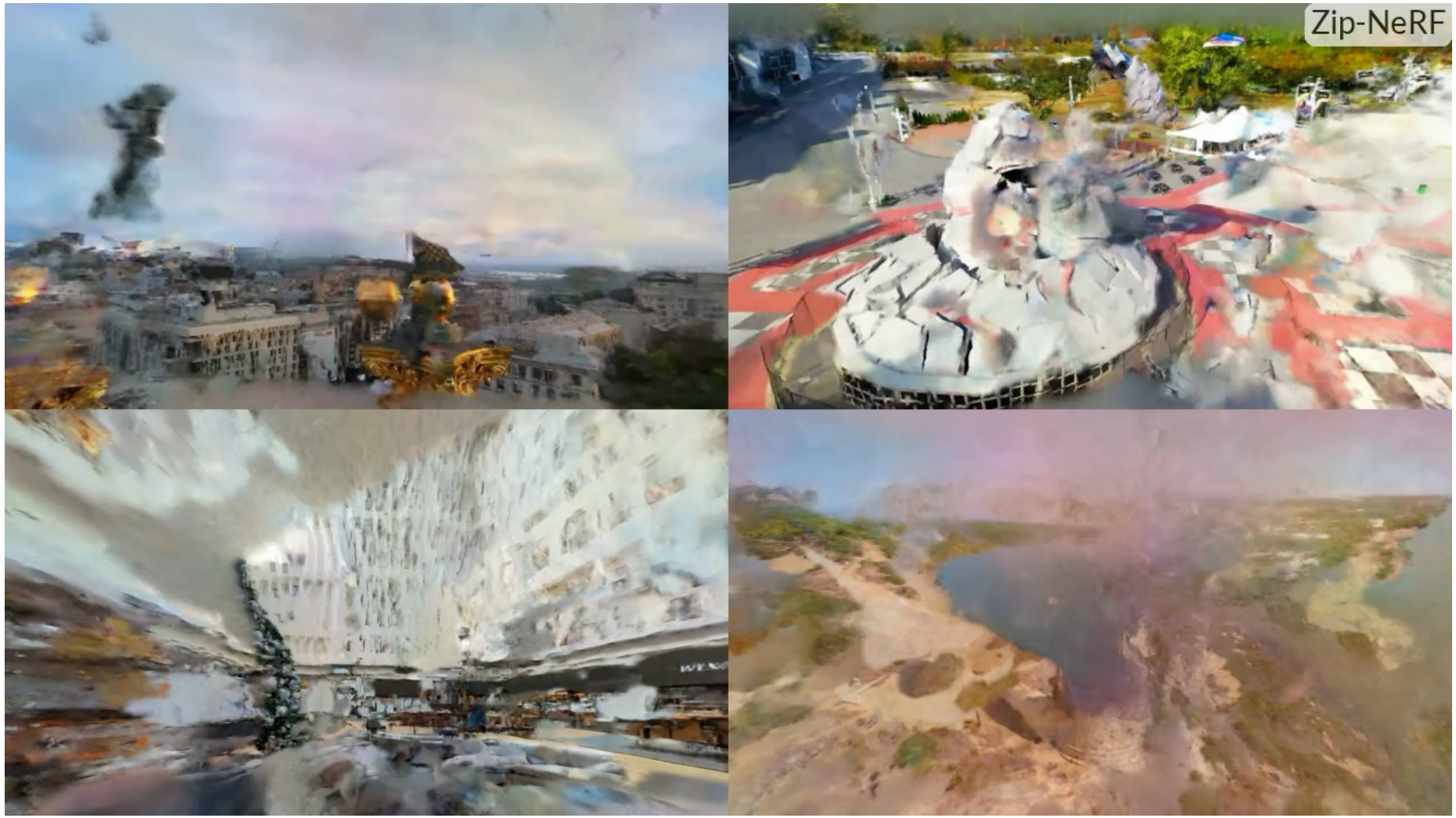
使用先验约束未观测到的视角： ReconFusion

- 使用Diffusion Prior约束未观测到的视角 $\mathcal{L}_{\text{sample}}(\psi) = \mathbb{E}_{\pi,t}\left[w(t)\left(\|x - \hat{x}_\pi\|_1 + \mathcal{L}_{\text{p}}(x, \hat{x}_\pi)\right)\right]$



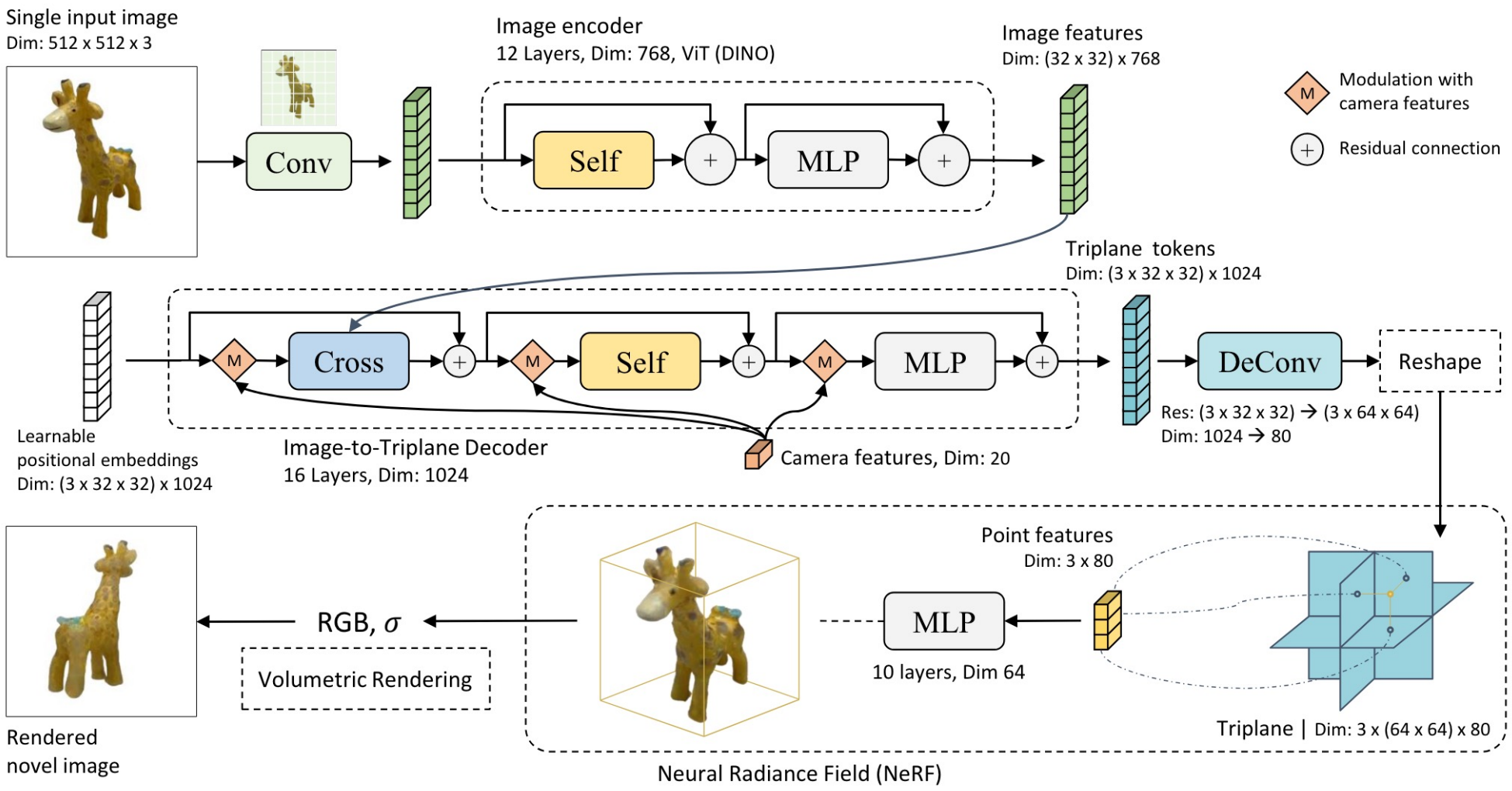
ReconFusion: 3D Reconstruction with Diffusion Priors. arXiv 2023.

ReconFusion的结果



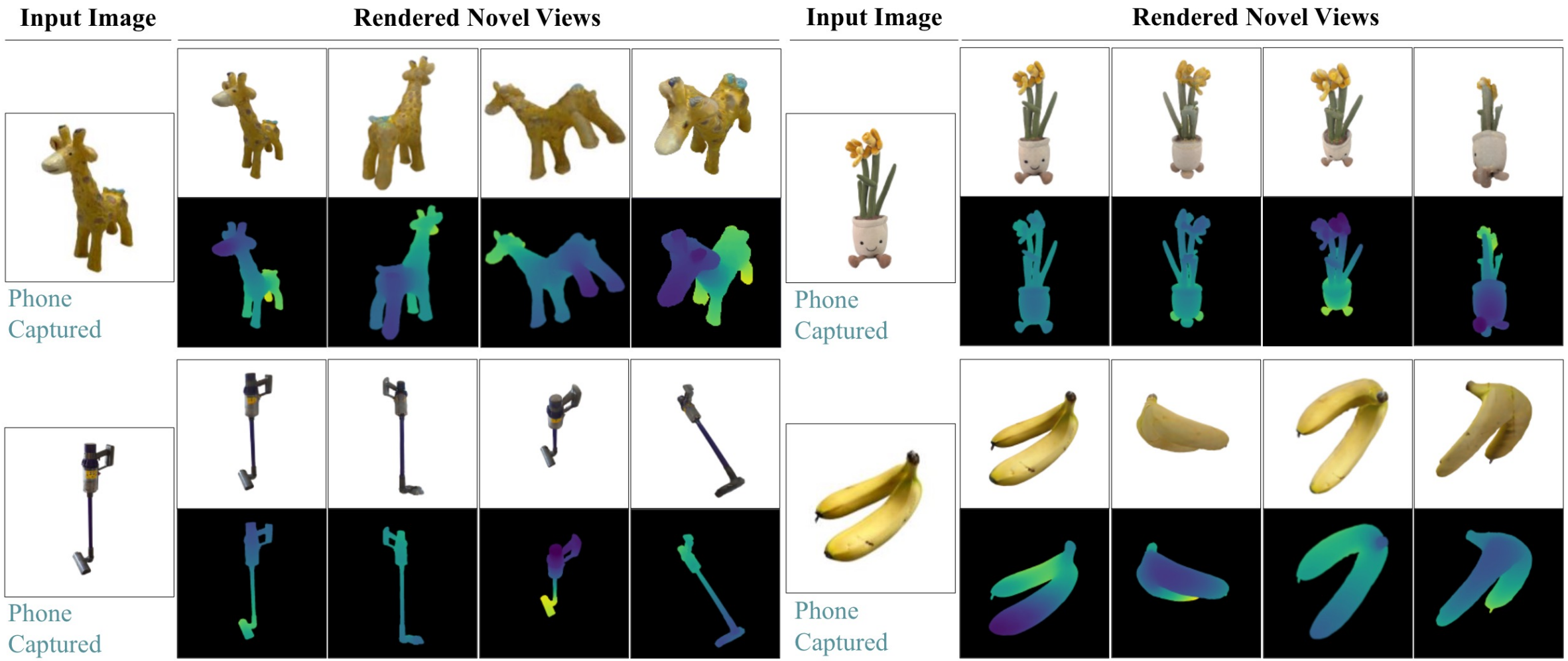
ReconFusion: 3D Reconstruction with Diffusion Priors. arXiv 2023.

使用Feedforward model直接预测：LRM



LRM: Large Reconstruction Model for Single Image to 3D. ICLR 2024.

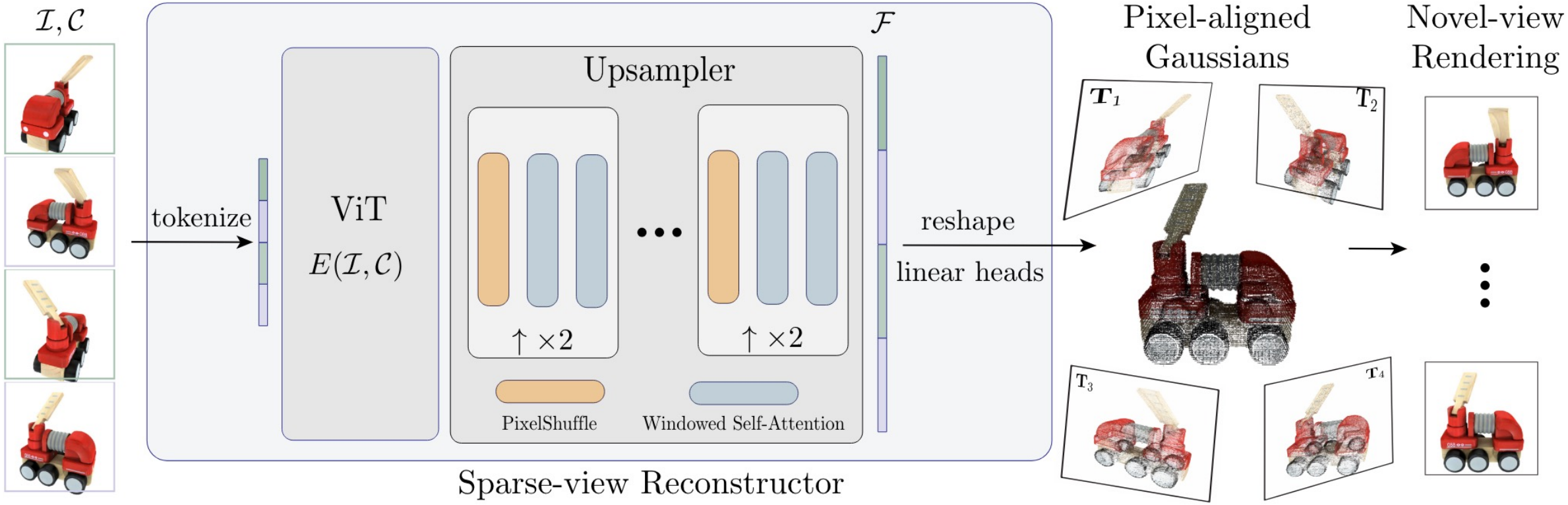
LRM的重建结果



LRM: Large Reconstruction Model for Single Image to 3D. ICLR 2024.

使用Feedforward model直接预测： GRM

- 将Gaussians maps作为网络输出结果
- 支持多视角的输入



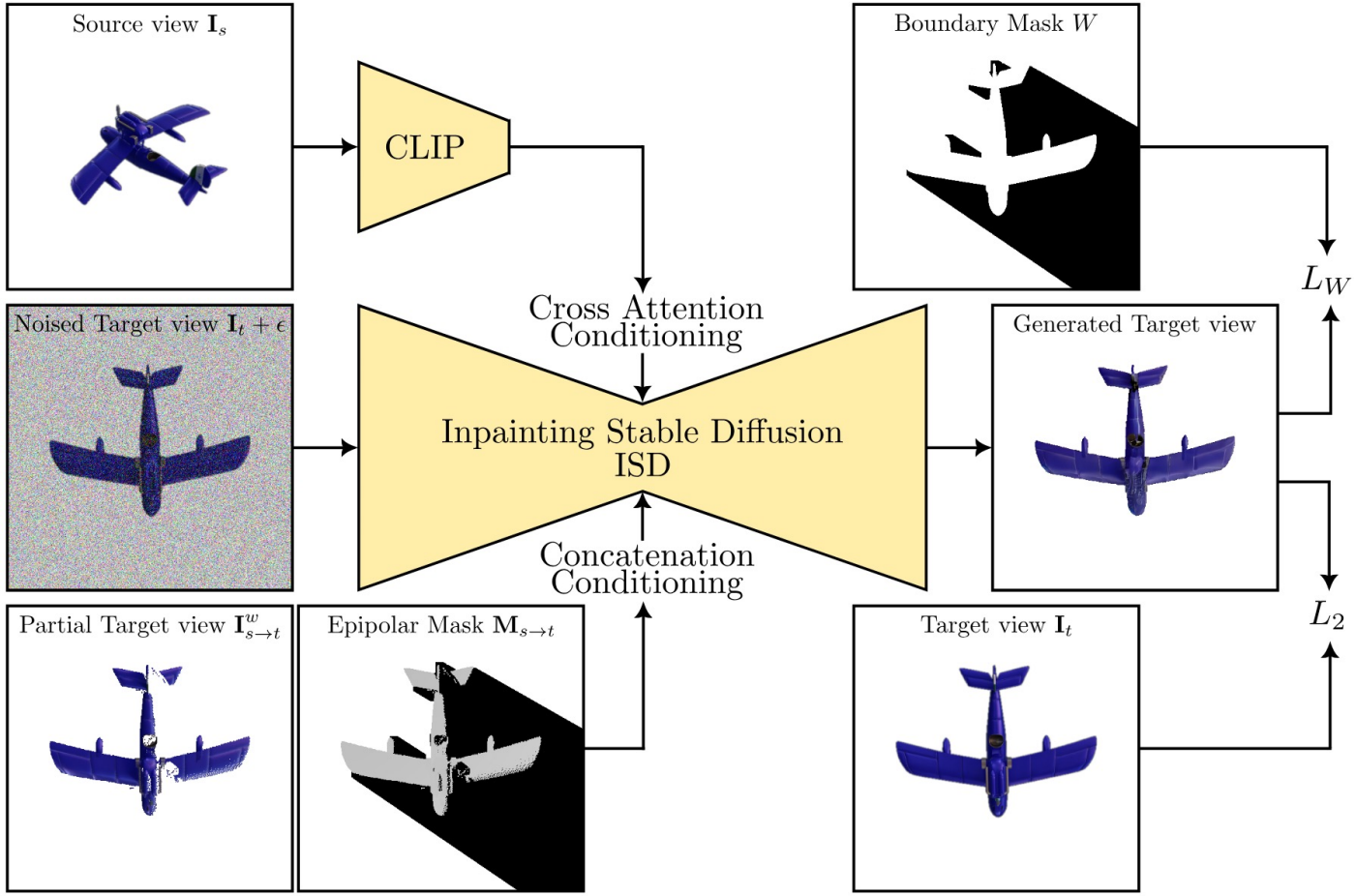
GRM: Large Gaussian Reconstruction Model for Efficient 3D Reconstruction and Generation. arXiv 2024.

GRM的重建结果

GRM: Large Gaussian Reconstruction Model for Efficient 3D Reconstruction and Generation. arXiv 2024.

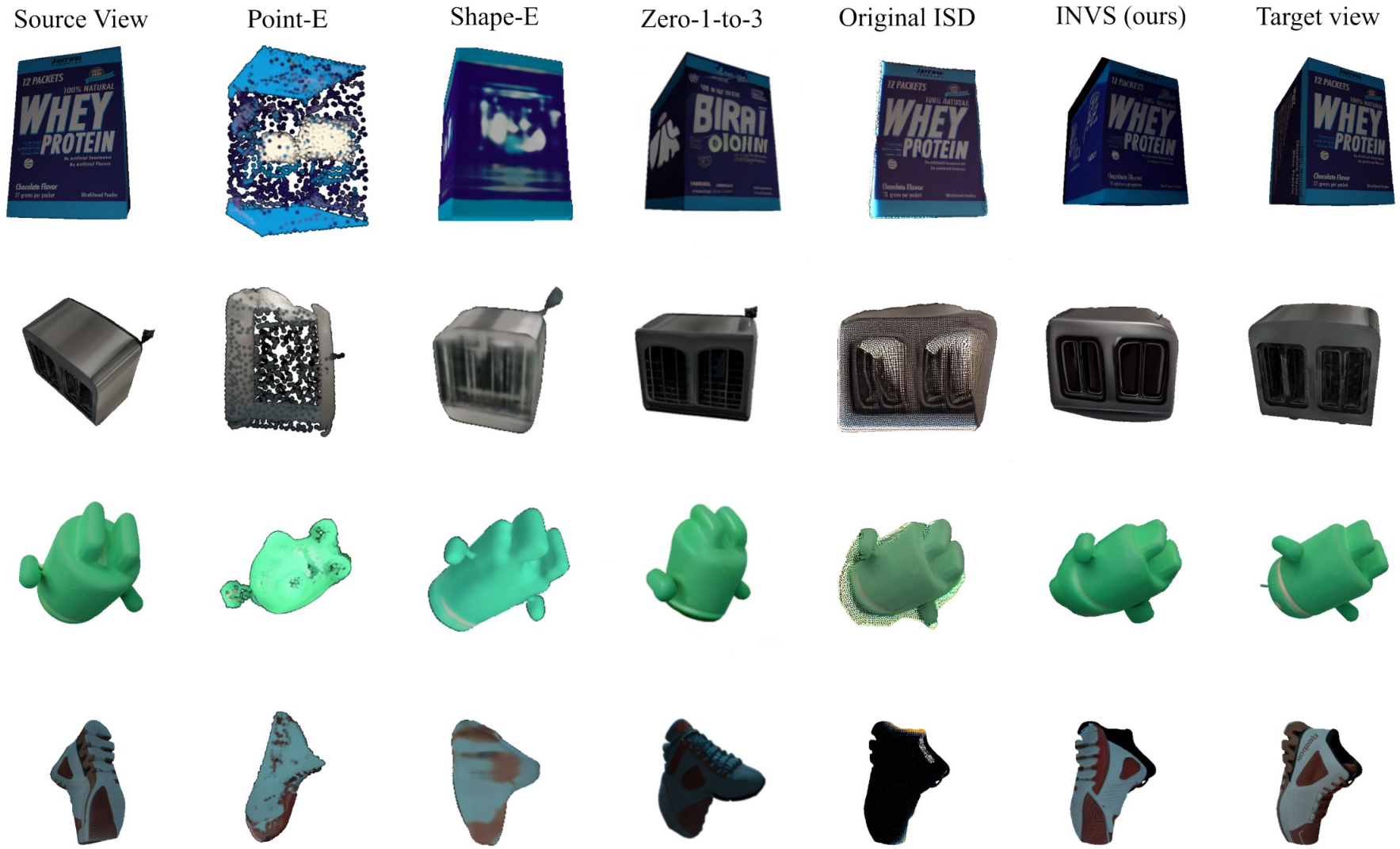
使用Feedforward model直接预测：iNVS

- 具体做法：
  - 将source view warp到novel view。
  - 使用Diffusion model补全novel view image。



Repurposing Diffusion Inpainters for Novel View Synthesis. SIGGRAPH 2023.

iNVS的重建结果

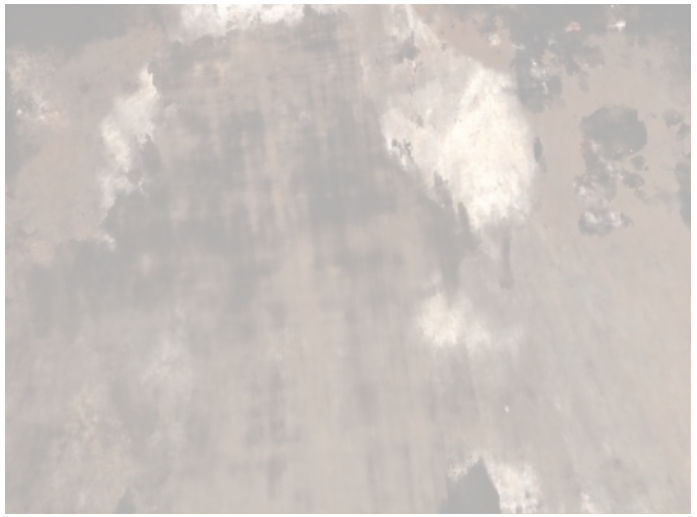


Repurposing Diffusion Inpainters for Novel View Synthesis. SIGGRAPH 2023.

NeRF模型不鲁棒3：要求输入图片光照一致

相机位姿不准确

输入视角稀疏



光照变化

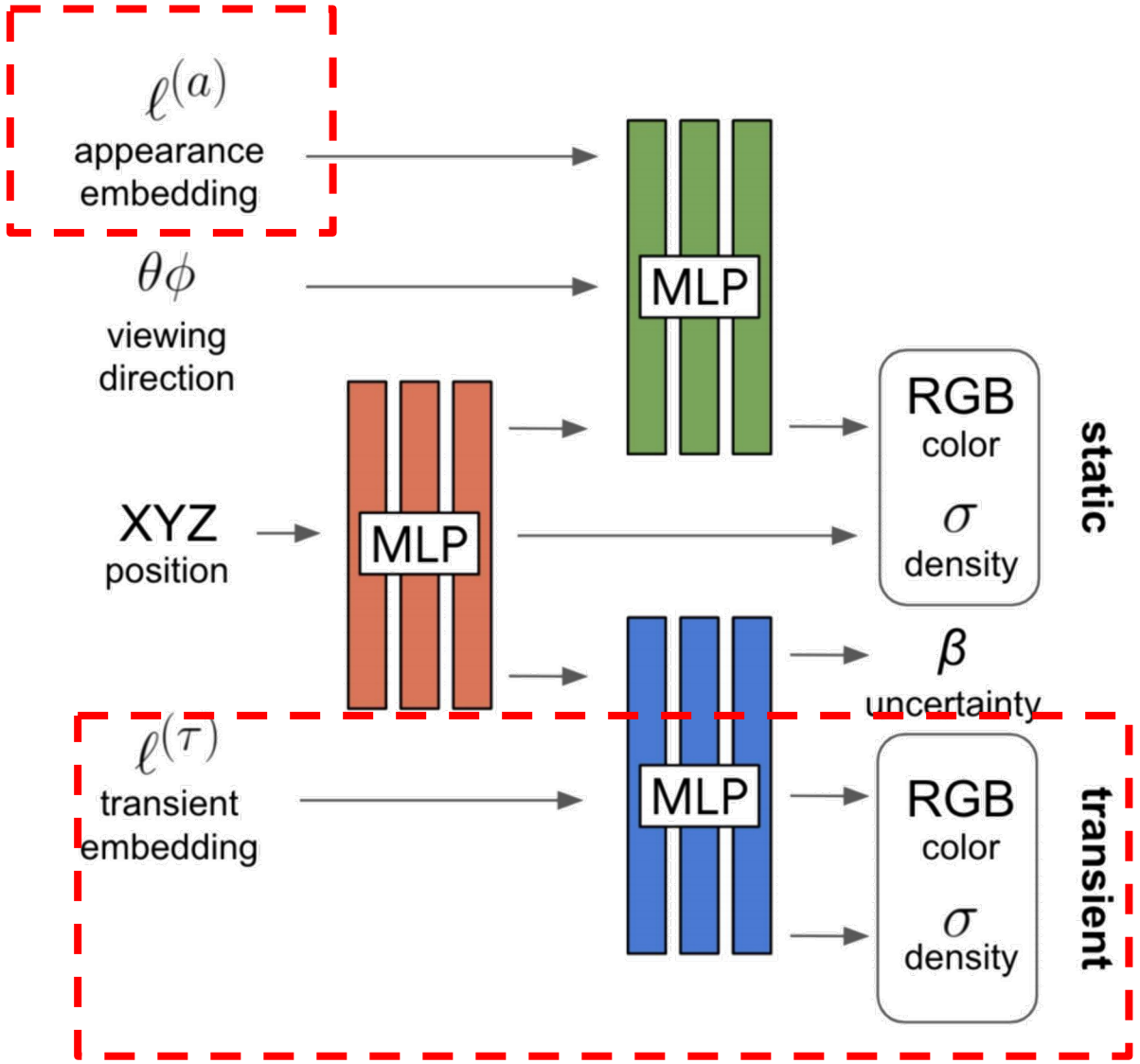
NeRF in the Wild

- 互联网图片存在光照不一致的情况

NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections. CVPR 2021.

NeRF in the Wild

- 给NeRF增加了两个部分：
  1. Appearance embedding，使得可以表达不同光照。
  2. Transient embedding，使得可以建模变化的几何。



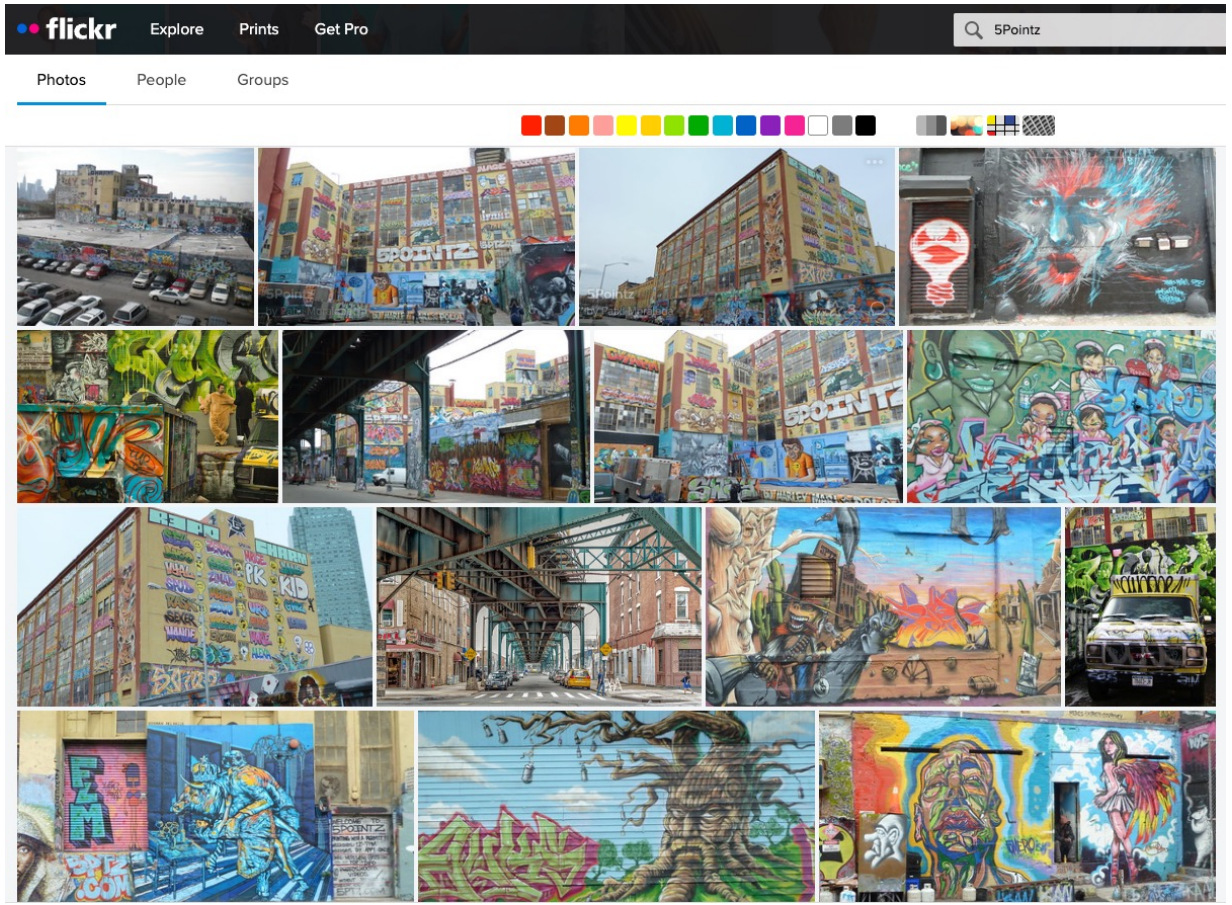
NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections. CVPR 2021.

NeRF in the Wild

NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections. CVPR 2021.

NeRF-W无法处理外观发生重大变化的场景

- 互联网图片存在图片内容变化的情况



Times Square, December, 2011

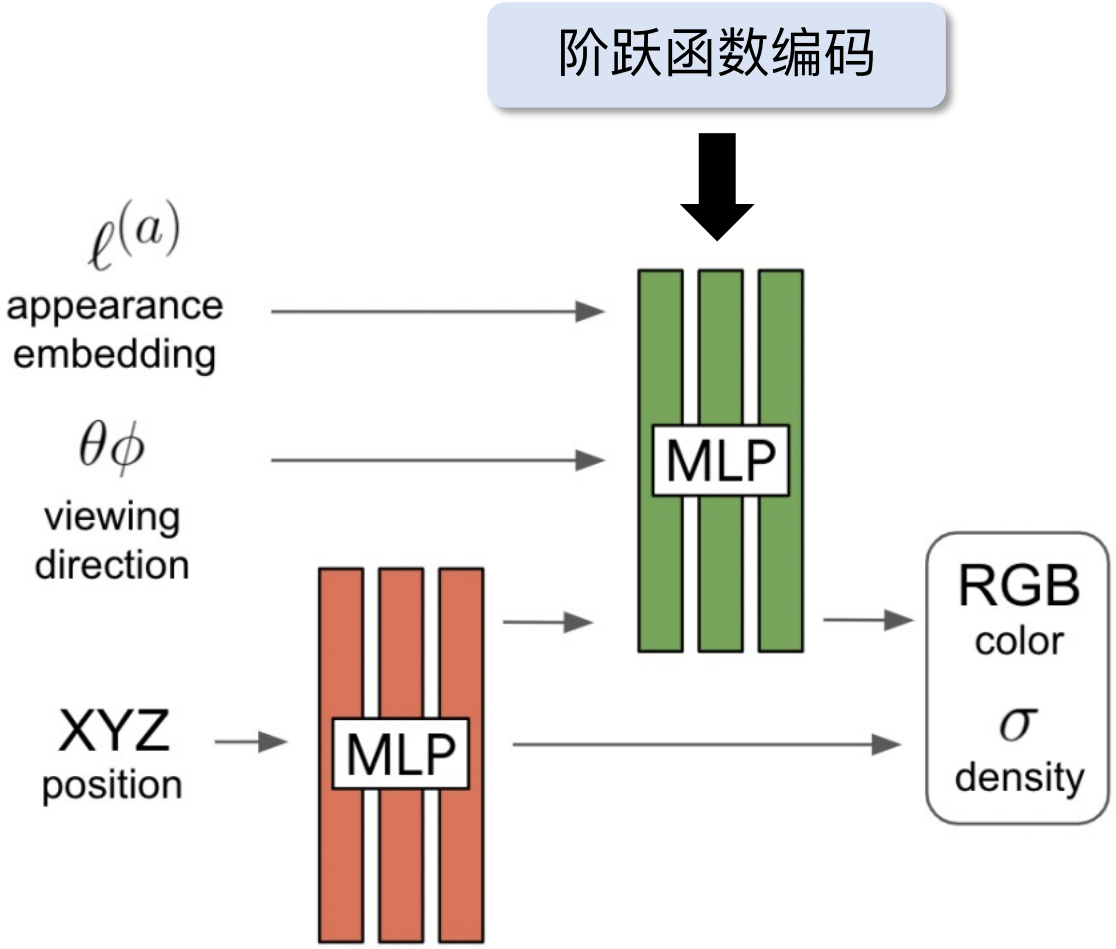
Times Square, February, 2011

Neural Scene Chronology. CVPR 2023.

NeRF-W无法处理外观发生重大变化的场景

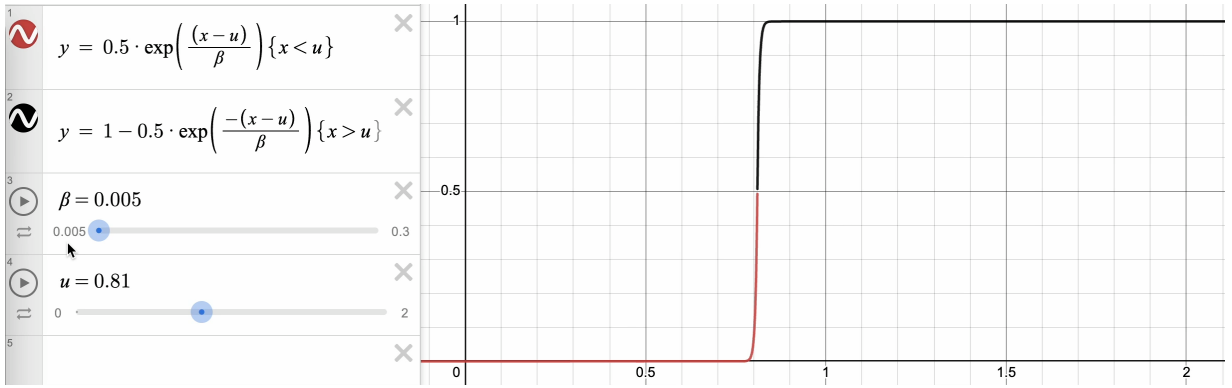
Neural Scene Chronology. CVPR 2023.

Neural Scene Chronology



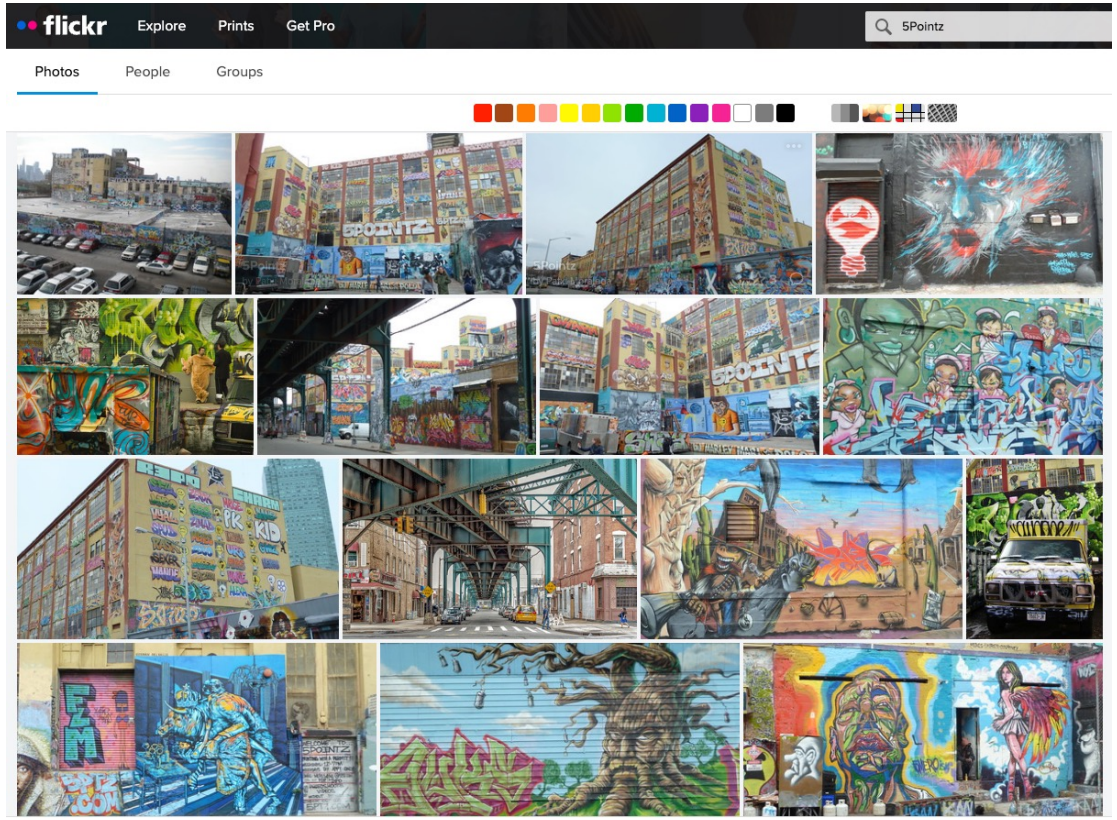
- 阶跃函数编码:

$$\bar{h}(t) = \begin{cases} \frac{1}{2}\exp(\frac{t-u}{\beta}) & \text{if } t \leq u \\ 1 - \frac{1}{2}\exp(\frac{-(t-u)}{\beta}) & \text{if } t > u \end{cases}$$



Neural Scene Chronology. CVPR 2023.

Neural Scene Chronology

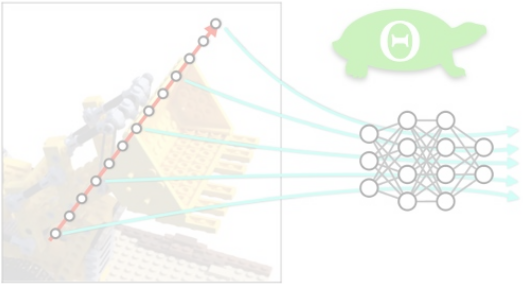


Neural Scene Chronology. CVPR 2023.

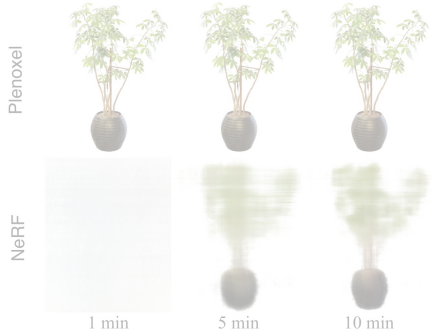
NeRF面临的问题5：几何重建质量低

NeRF在静态场景建模上遇到的问题

渲染速度慢

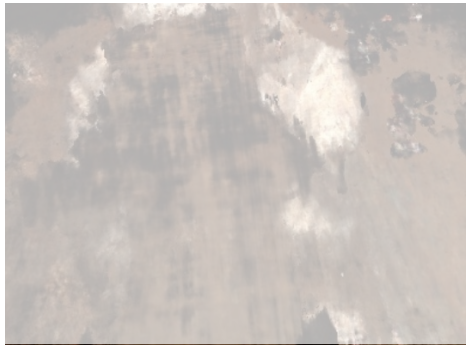


训练速度慢



场景建模能力弱

模型鲁棒性差



NeRF在其他类型场景建模上遇到的问题

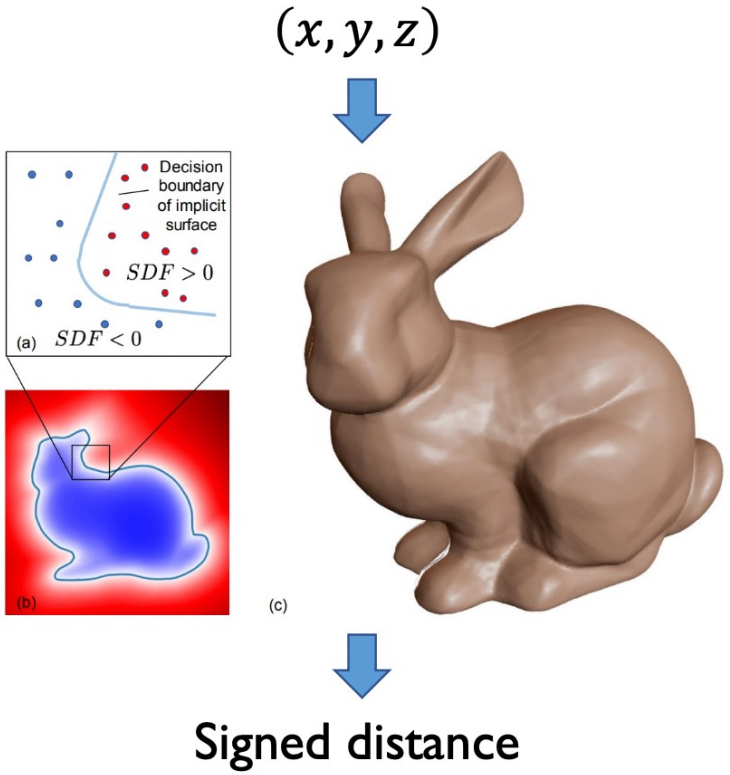
几何重建质量低

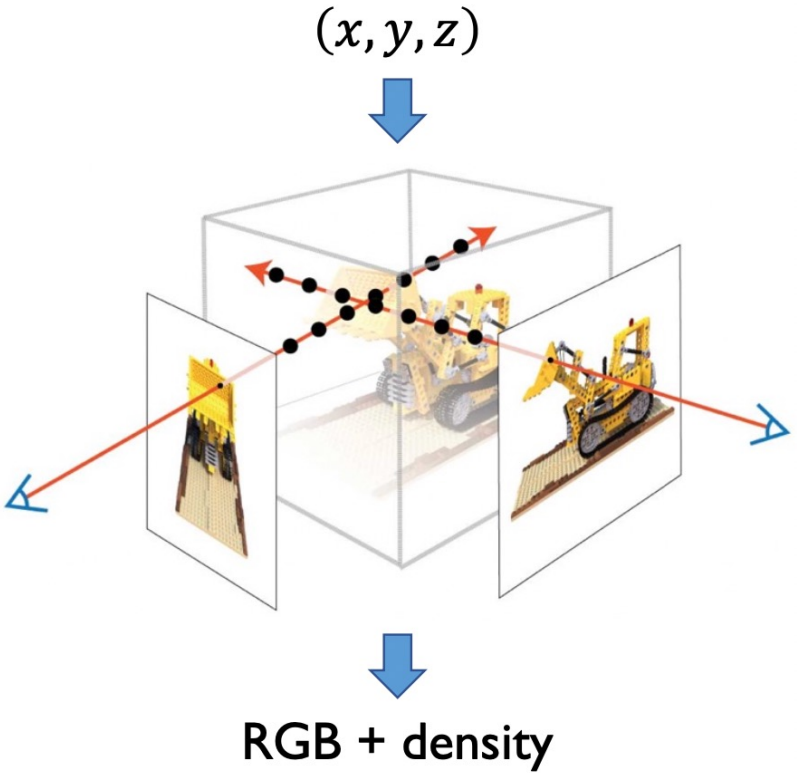
无法建模多模态信号

无法建模动态场景

NeRF几何重建质量差的原因

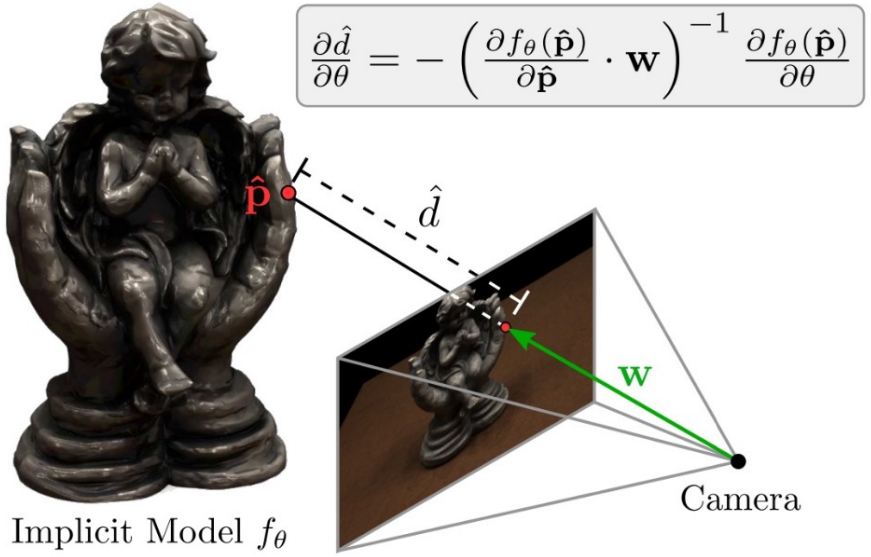
- NeRF缺少对表面几何的定义



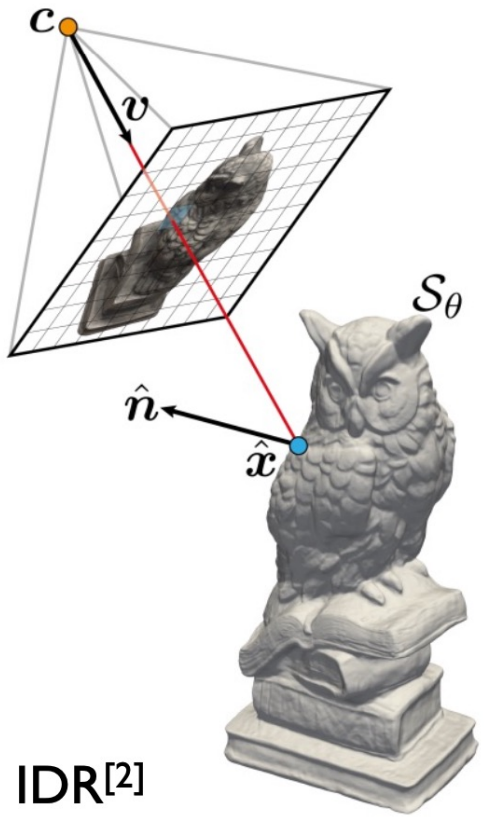


Differentiable surface rendering

- 找到表面三维点，计算颜色，并定义相对于表面三维点的梯度



DVR[1]



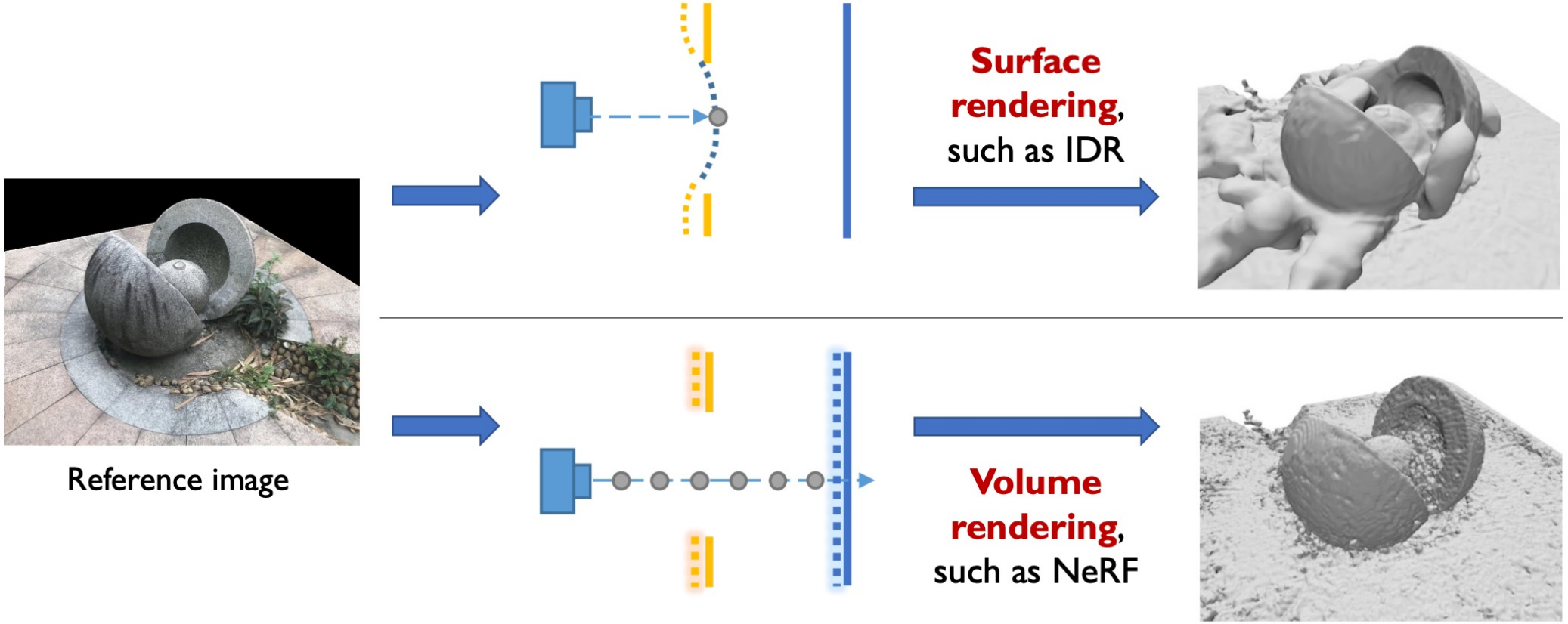
IDR[2]

Differentiable Volumetric Rendering: Learning Implicit 3D Representations without 3D Supervision, CVPR 2020.

Multiview Neural Surface Reconstruction by Disentangling Geometry and Appearance, NeurIPS 2020.

Differentiable surface rendering面临的挑战

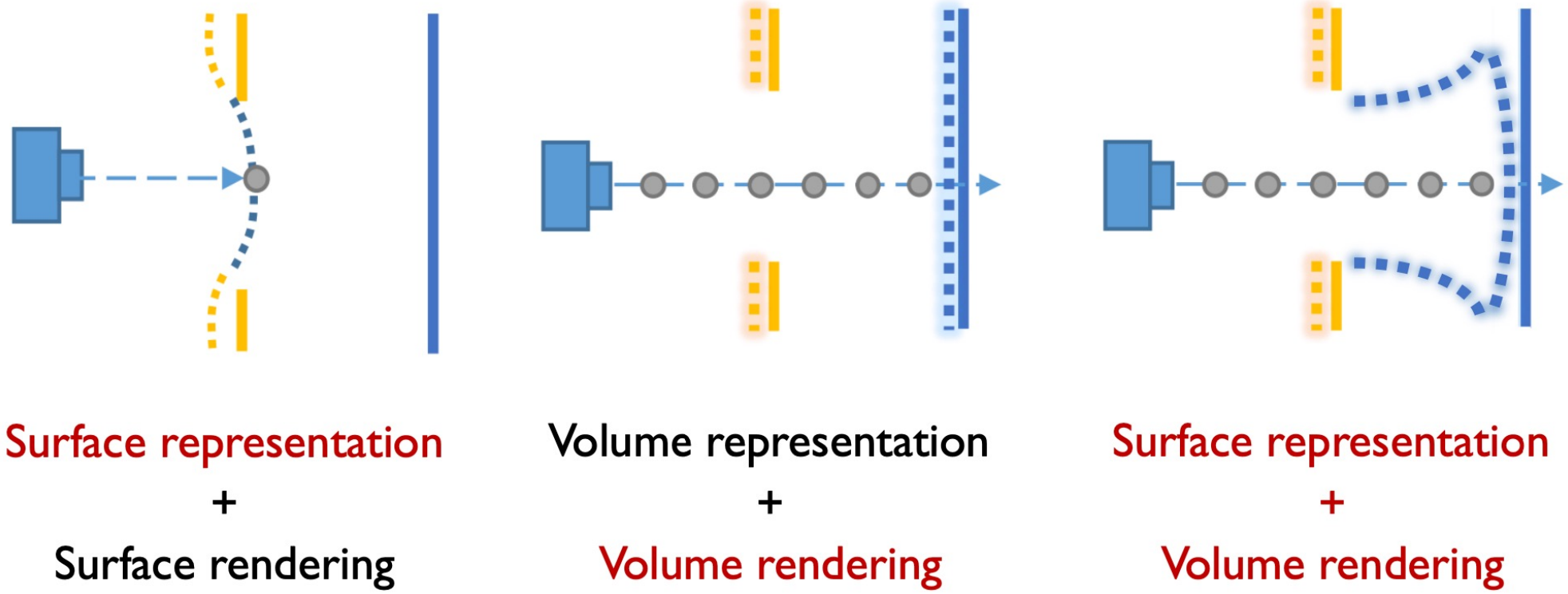
- Surface rendering只在一条射线的一个点上计算梯度



NeuS: Learning Neural Implicit Surfaces by Volume Rendering for Multi-view Reconstruction, NeurIPS 2021.

NeuS

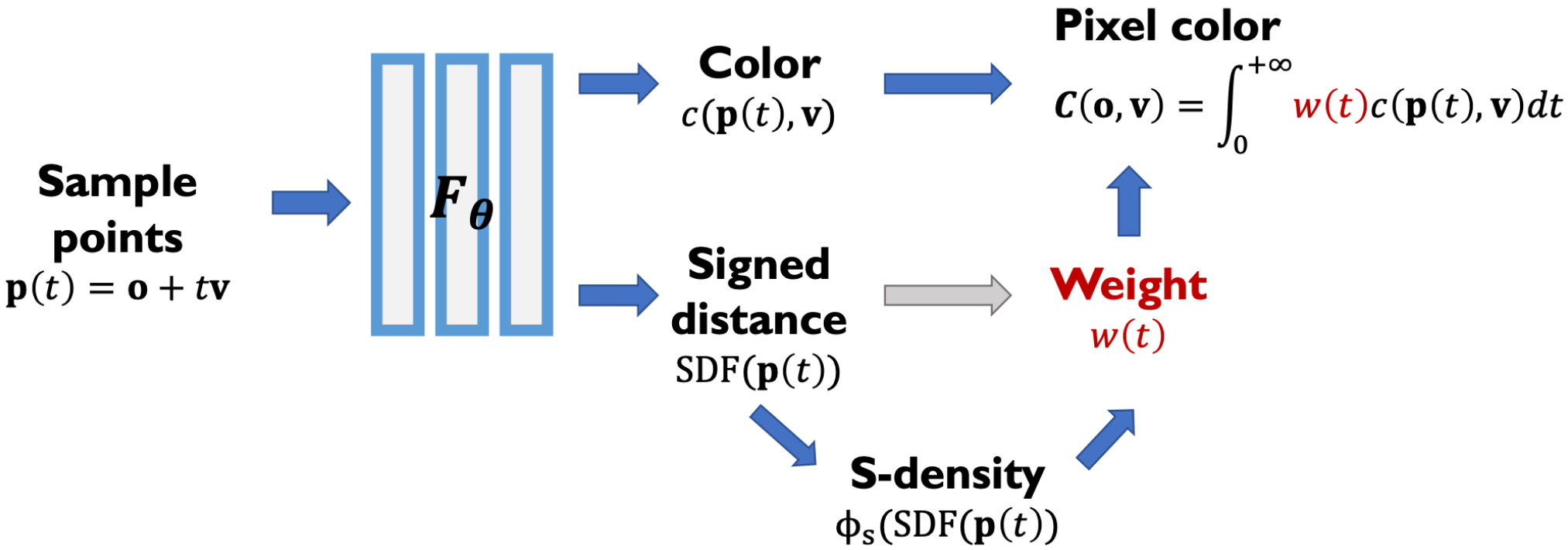
- 设计了面向Surface representation的Volume rendering technique



NeuS: Learning Neural Implicit Surfaces by Volume Rendering for Multi-view Reconstruction, NeurIPS 2021.

NeuS

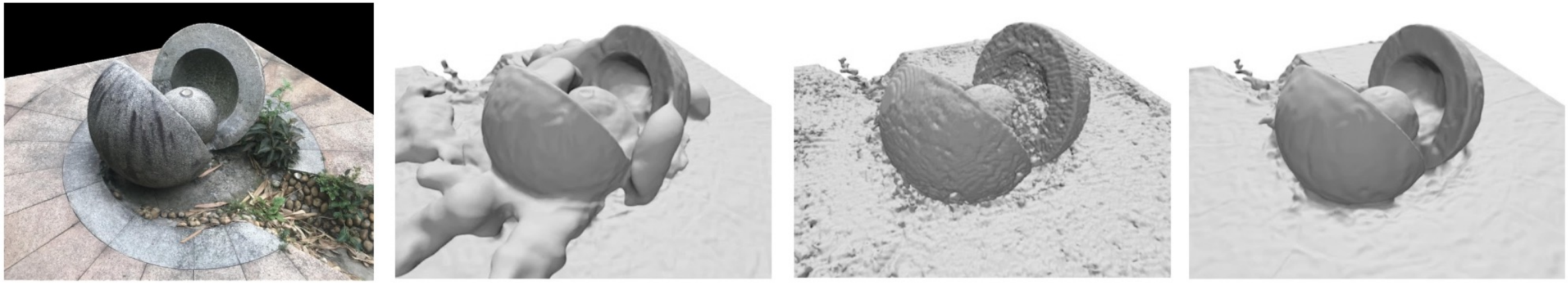
- 设计了面向Surface representation的Volume rendering technique



NeuS: Learning Neural Implicit Surfaces by Volume Rendering for Multi-view Reconstruction, NeurIPS 2021.

NeuS

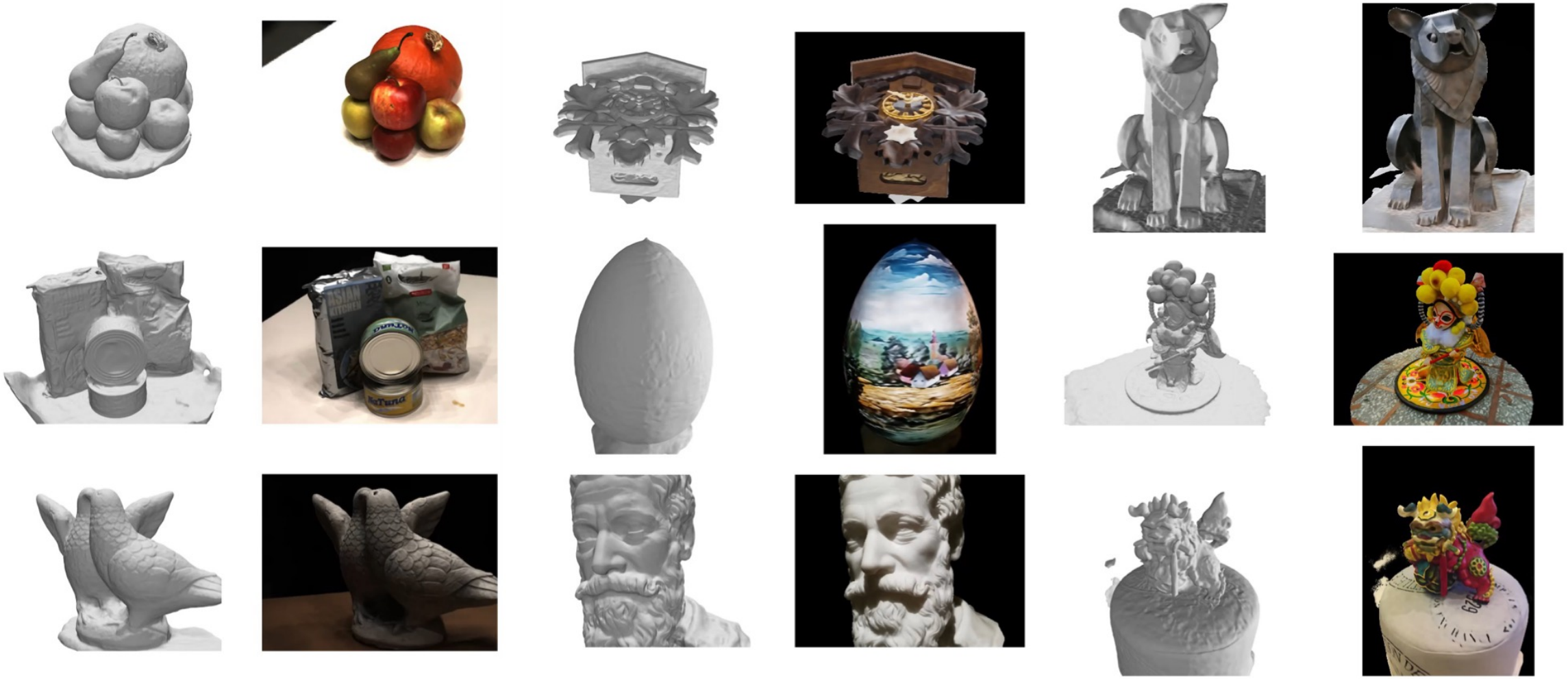
- 相比于Surface rendering的优势：



Reference image　IDR　NeRF　NeuS

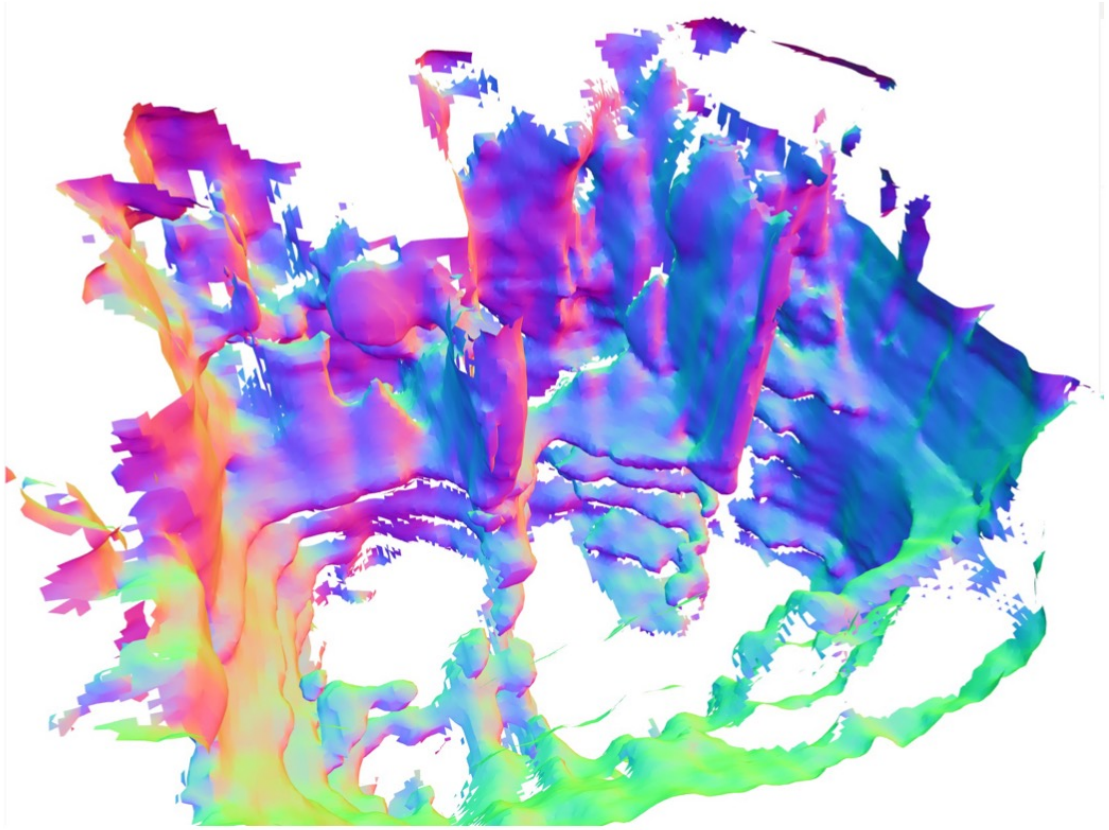
NeuS: Learning Neural Implicit Surfaces by Volume Rendering for Multi-view Reconstruction, NeurIPS 2021.

NeuS的重建结果



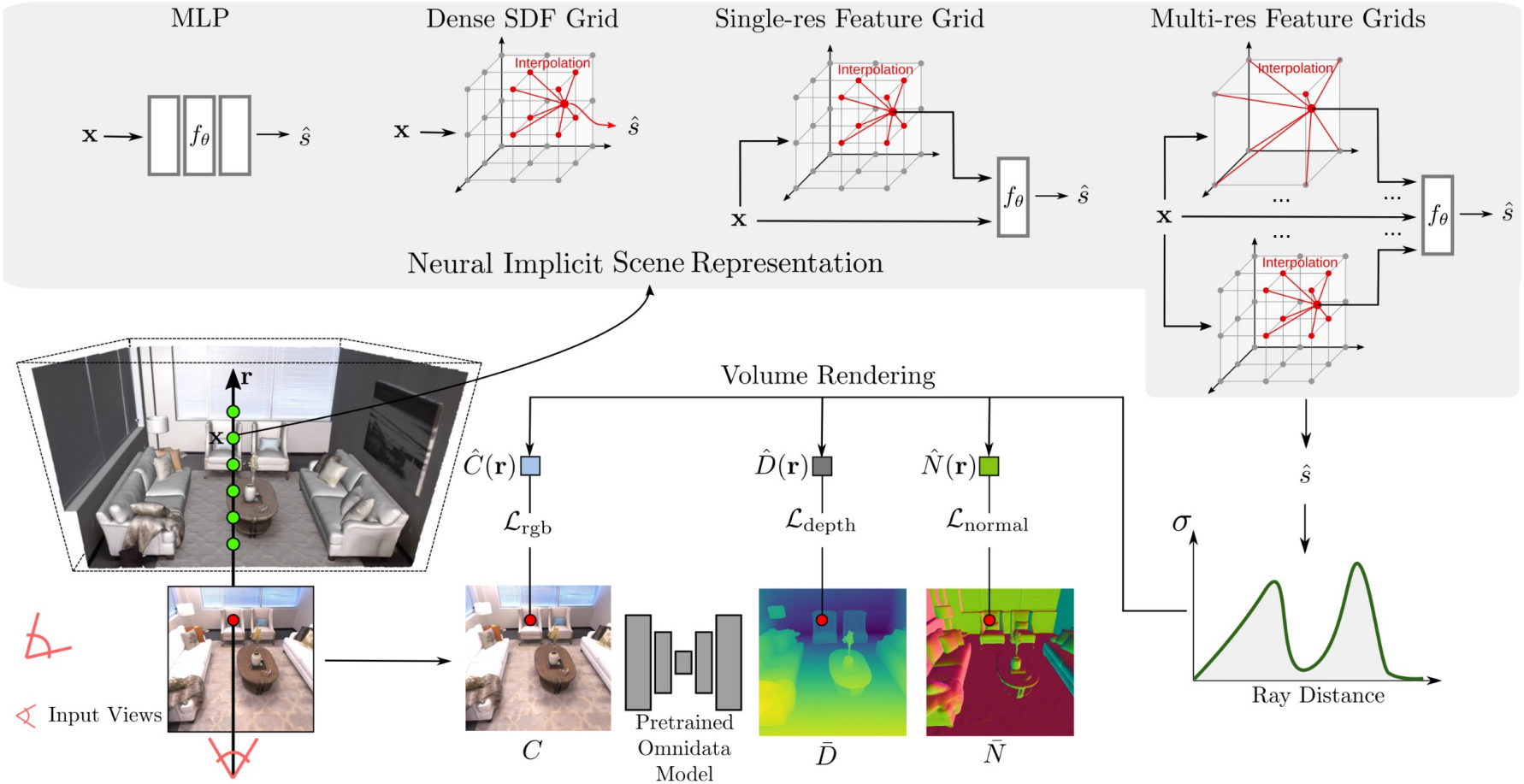
NeuS: Learning Neural Implicit Surfaces by Volume Rendering for Multi-view Reconstruction, NeurIPS 2021.

NeuS面临的挑战1：弱纹理区域的重建



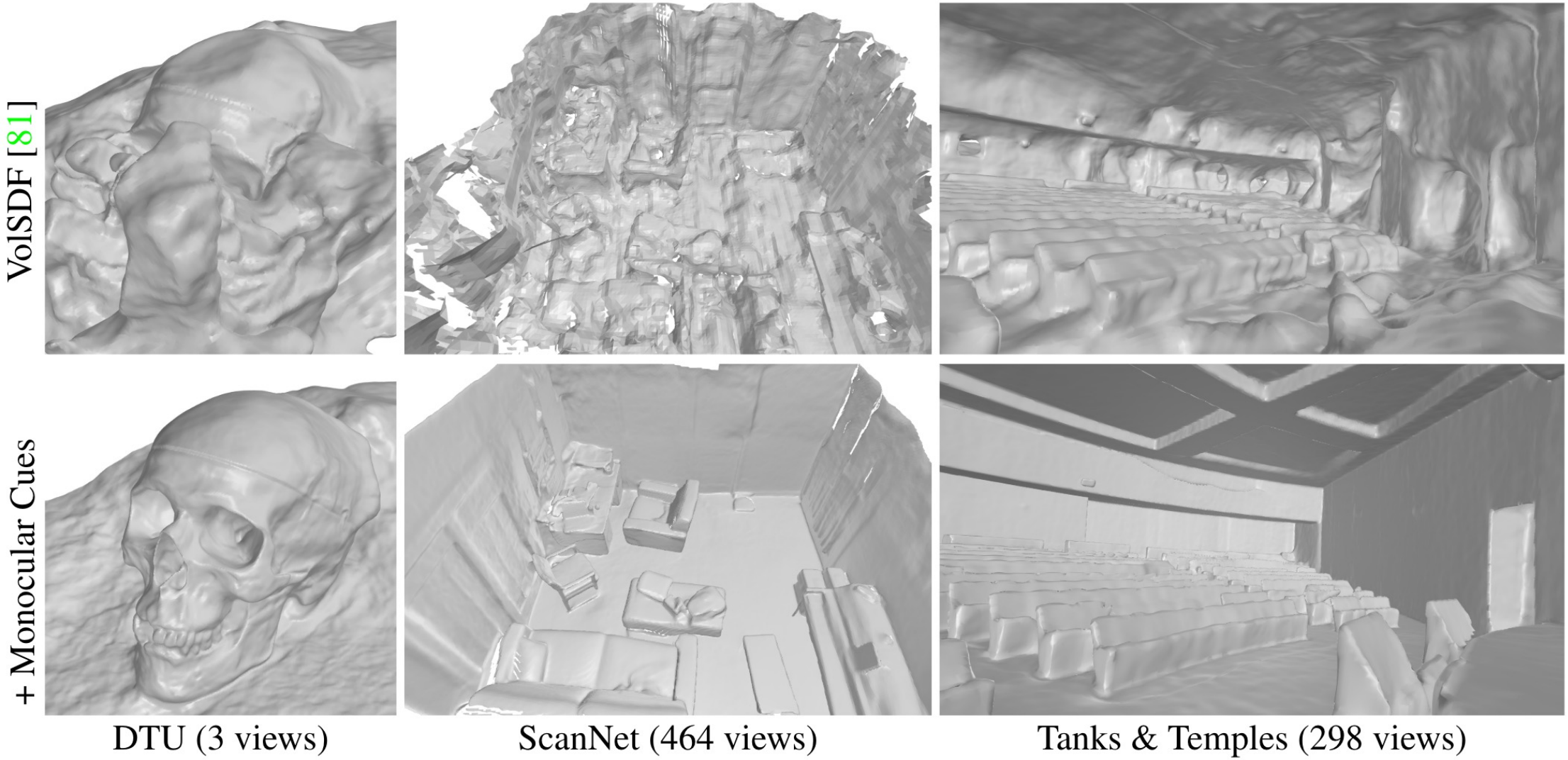
MonoSDF

- 使用Monocular depth and normal进行约束。



MonoSDF: Exploring Monocular Geometric Cues for Neural Implicit Surface Reconstruction. NeurIPS 2022.

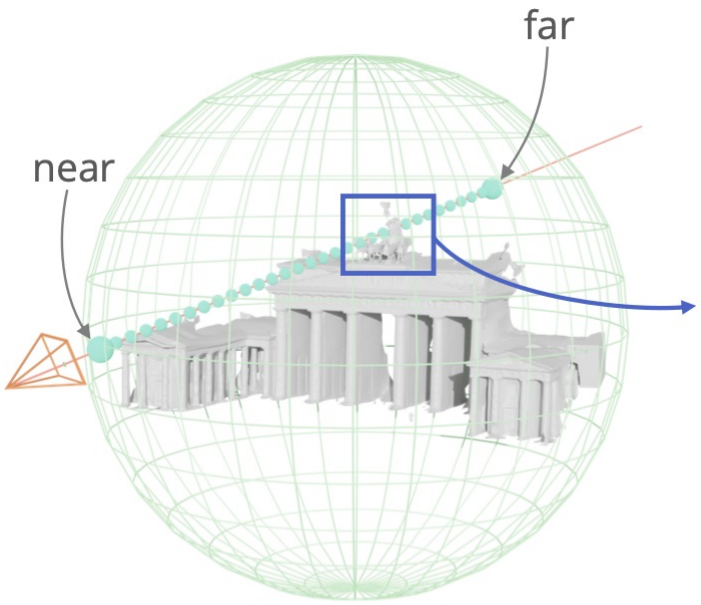
MonoSDF的效果



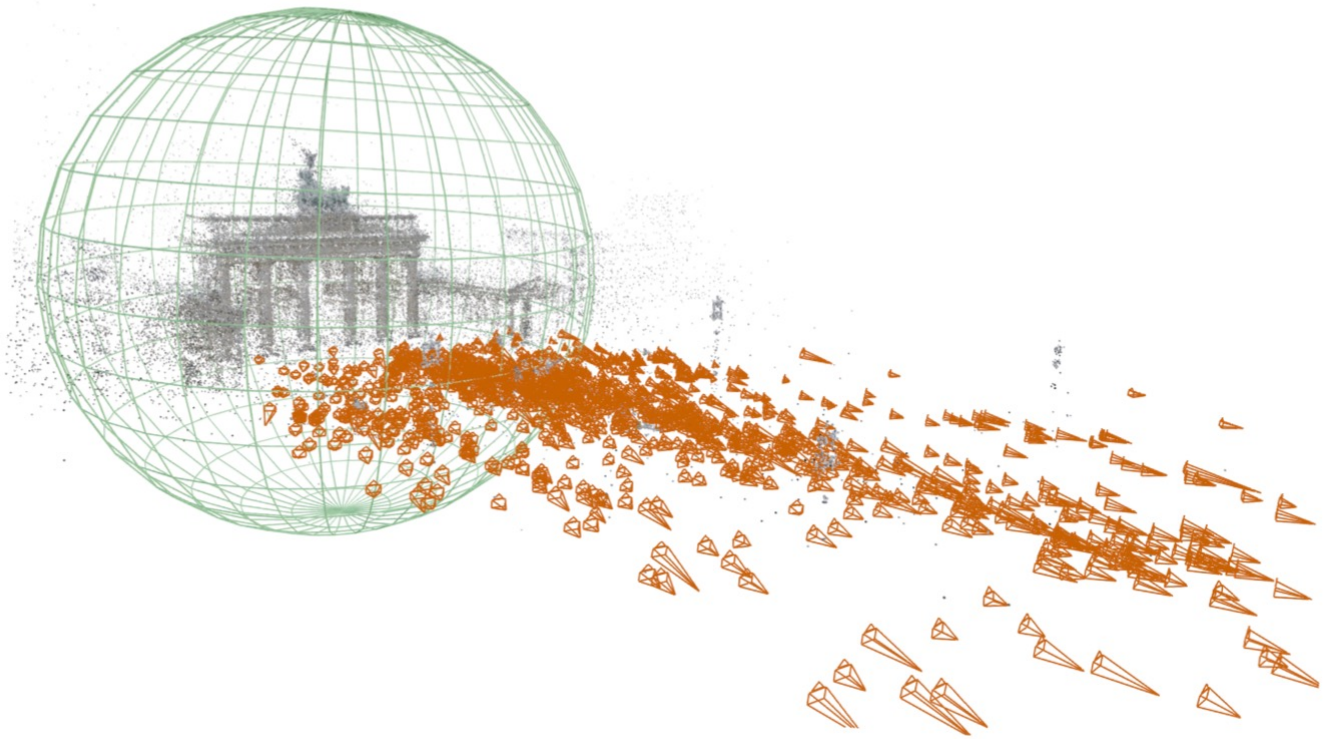
MonoSDF: Exploring Monocular Geometric Cues for Neural Implicit Surface Reconstruction. NeurIPS 2022.

NeuS面临的挑战2：大规模场景的重建

- 重建范围大，导致NeuS重建粗糙



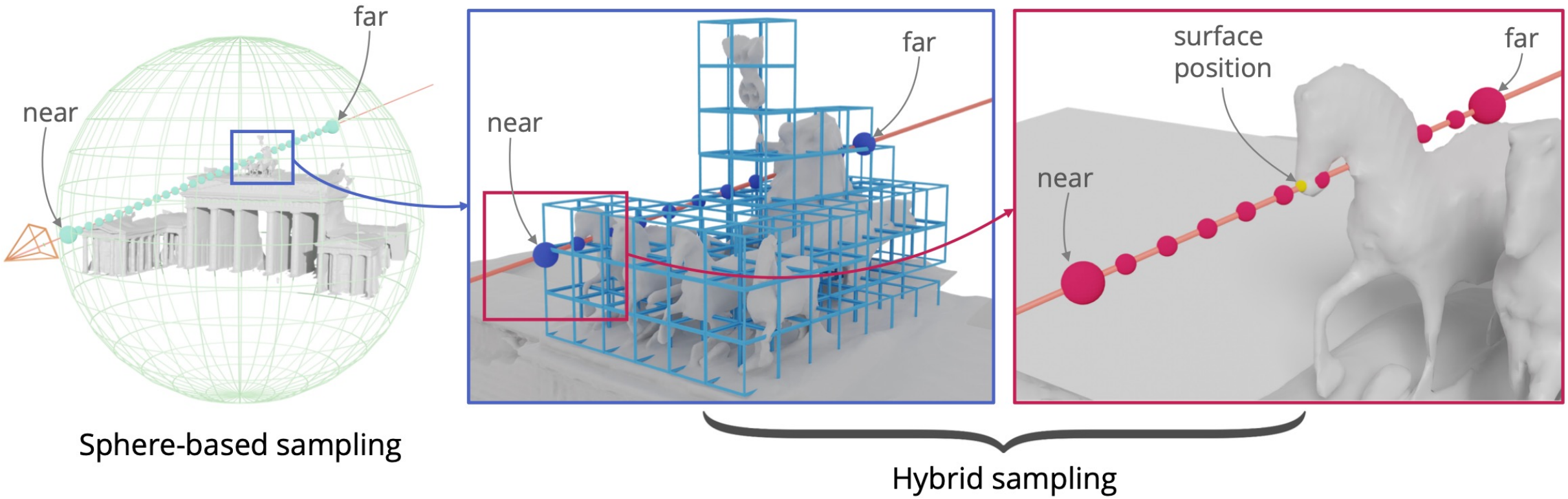
Sphere-based sampling



Neural Reconstruction in the Wild. SIGGRAPH 2022.

Neural Reconstruction in the Wild

- 利用surface–guided sampling来提升重建精度



Neural Reconstruction in the Wild. SIGGRAPH 2022.

Neural Reconstruction in the Wild的结果

Credits: Flickr

Brandenburg Gate

Neural Reconstruction in the Wild, SIGGRAPH 2022.

Neural Reconstruction in the Wild的结果

Credits: Flickr

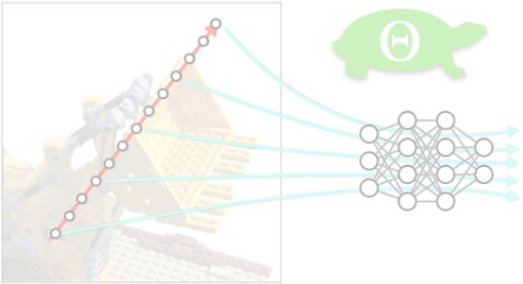
Trevi Fountain

Neural Reconstruction in the Wild, SIGGRAPH 2022.

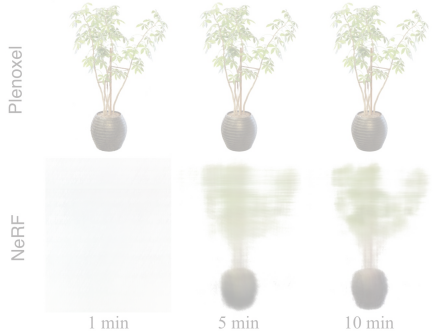
NeRF面临的问题5：几何重建质量低

NeRF在静态场景建模上遇到的问题

渲染速度慢

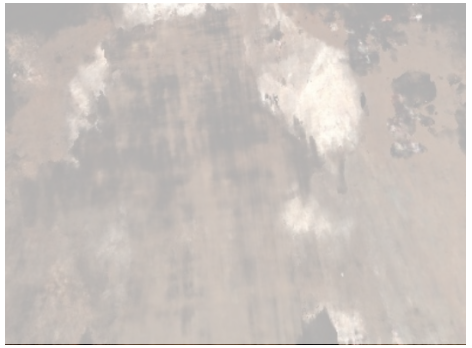


训练速度慢



场景建模能力弱

模型鲁棒性差



NeRF在其他类型场景建模上遇到的问题

几何重建质量低

无法建模多模态信号

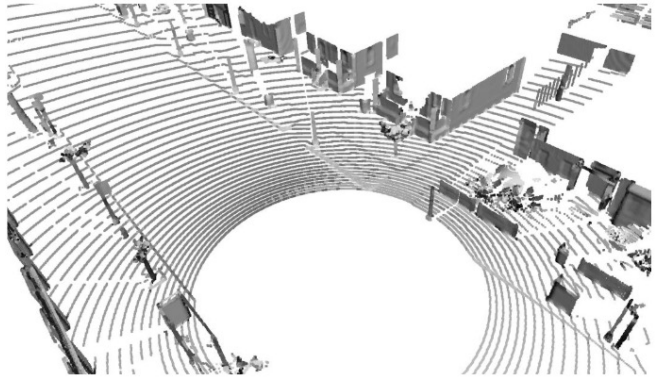
无法建模动态场景

NeRF缺少多模态信号建模

语义场

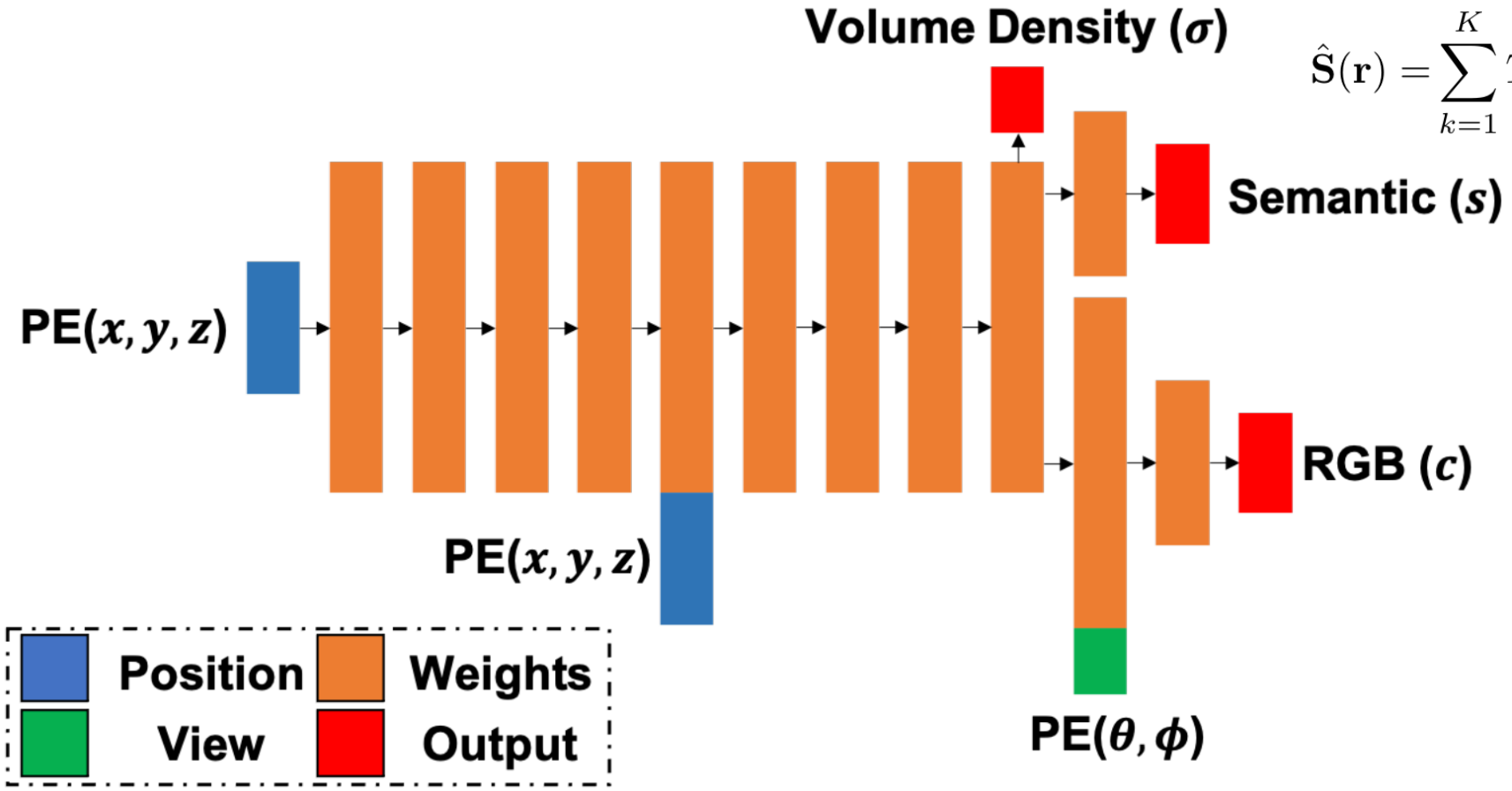
物体材质

雷达点云



建模语义场：Semantic NeRF

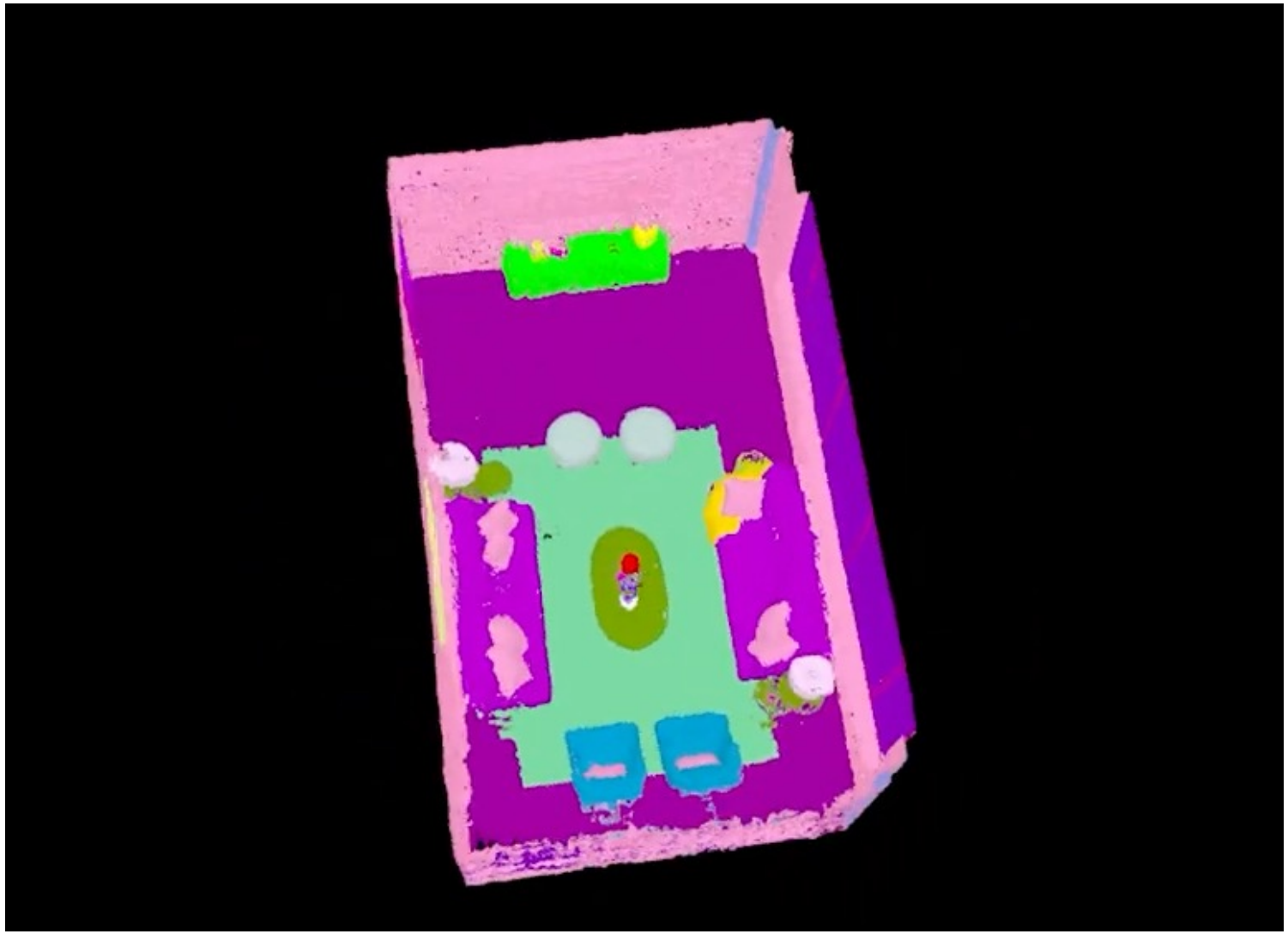
- 使用NeRF额外预测一个semantic label



$$\hat{\mathbf{S}}(\mathbf{r}) = \sum_{k=1}^{K} \hat{T}(t_k)\, \alpha\left(\sigma(t_k)\delta_k\right) \mathbf{s}(t_k)\,,$$

In-Place Scene Labelling and Understanding with Implicit Scene Representation. ICCV 2021.

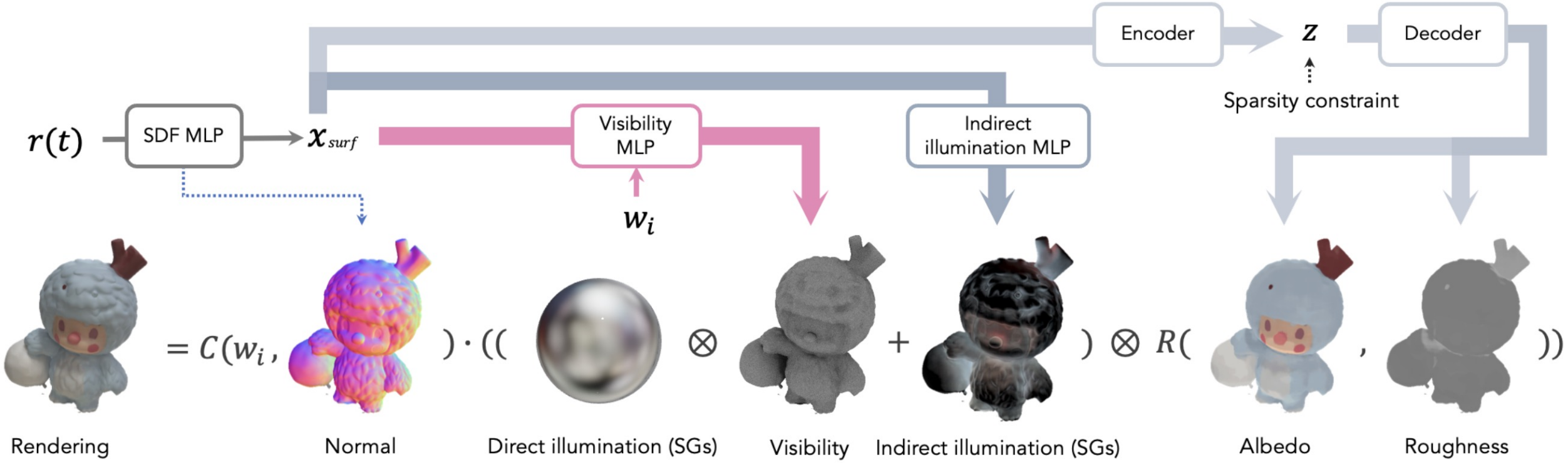
Semantic NeRF的结果



In-Place Scene Labelling and Understanding with Implicit Scene Representation. ICCV 2021.

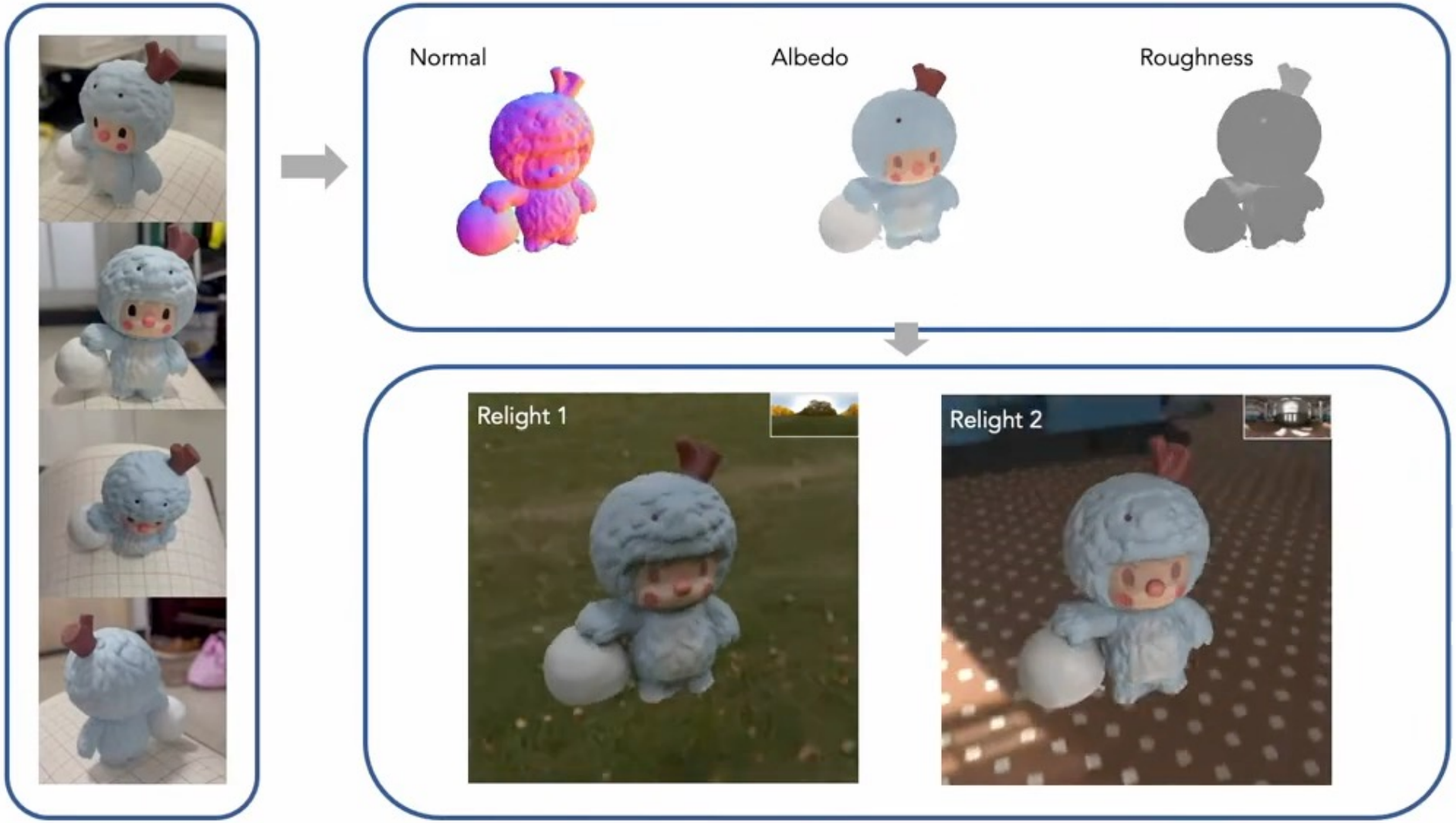
建模物体材质：InvRender

- 使用NeRF额外预测一个visibility, albedo, roughness和indirect illumination，然后使用Rendering equation渲染图片。



Modeling Indirect Illumination for Inverse Rendering. CVPR 2022.

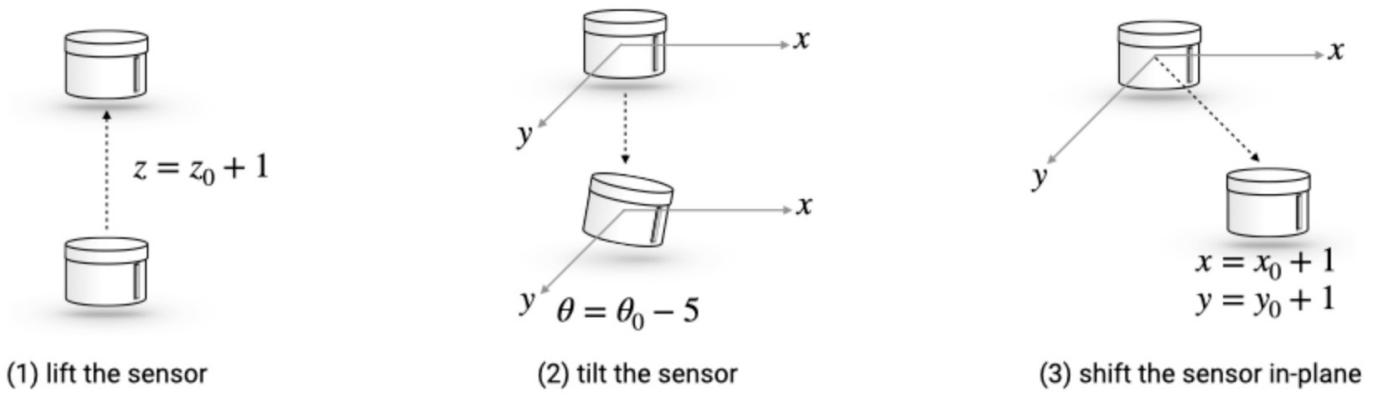
InvRender的结果



Modeling Indirect Illumination for Inverse Rendering. CVPR 2022.

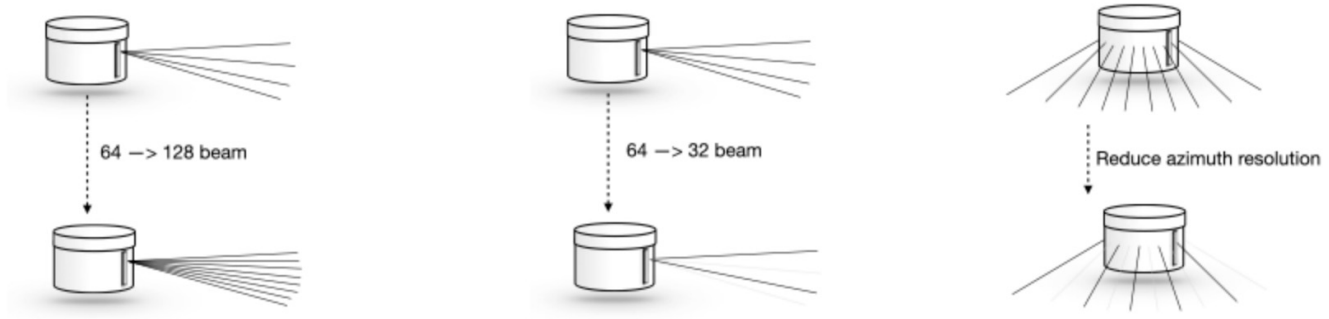
LiDAR novel view synthesis

- 获取不同camera pose下的Lidar信号



(1) lift the sensor (2) tilt the sensor (3) shift the sensor in-plane

- 获取不同属性的Lidar信号

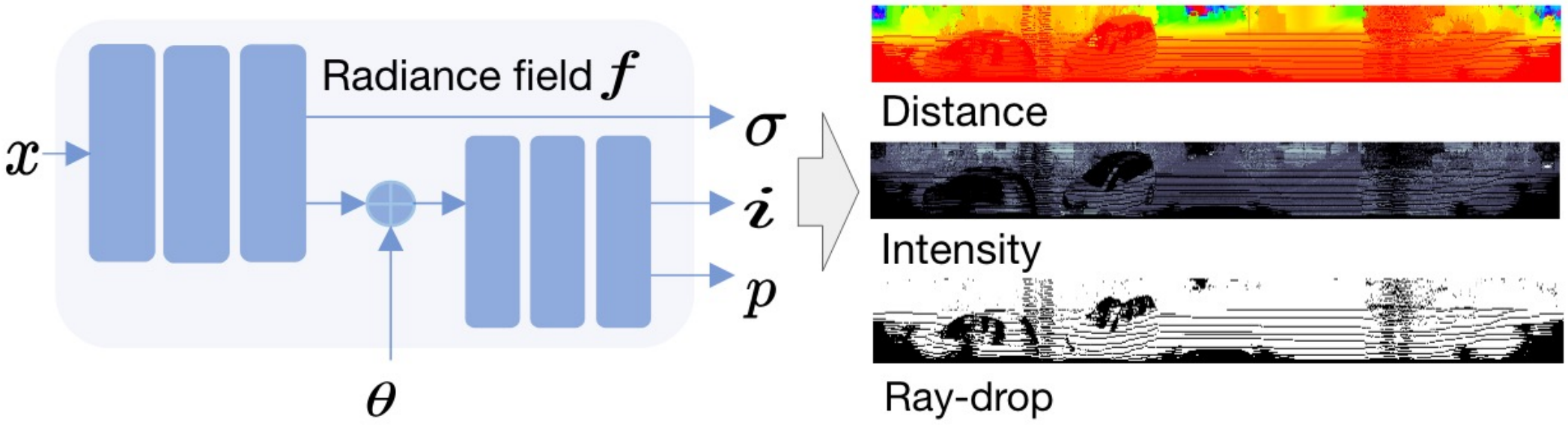


(4) upsample the beams (5) downsample the beams (6) reduce azimuth resolution

Neural LiDAR Fields for Novel View Synthesis. ICCV 2023.

Neural LiDAR Field

- 使用NeRF额外预测一个intensity和ray-drop

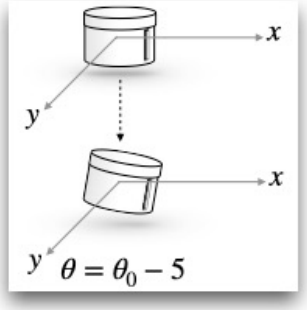


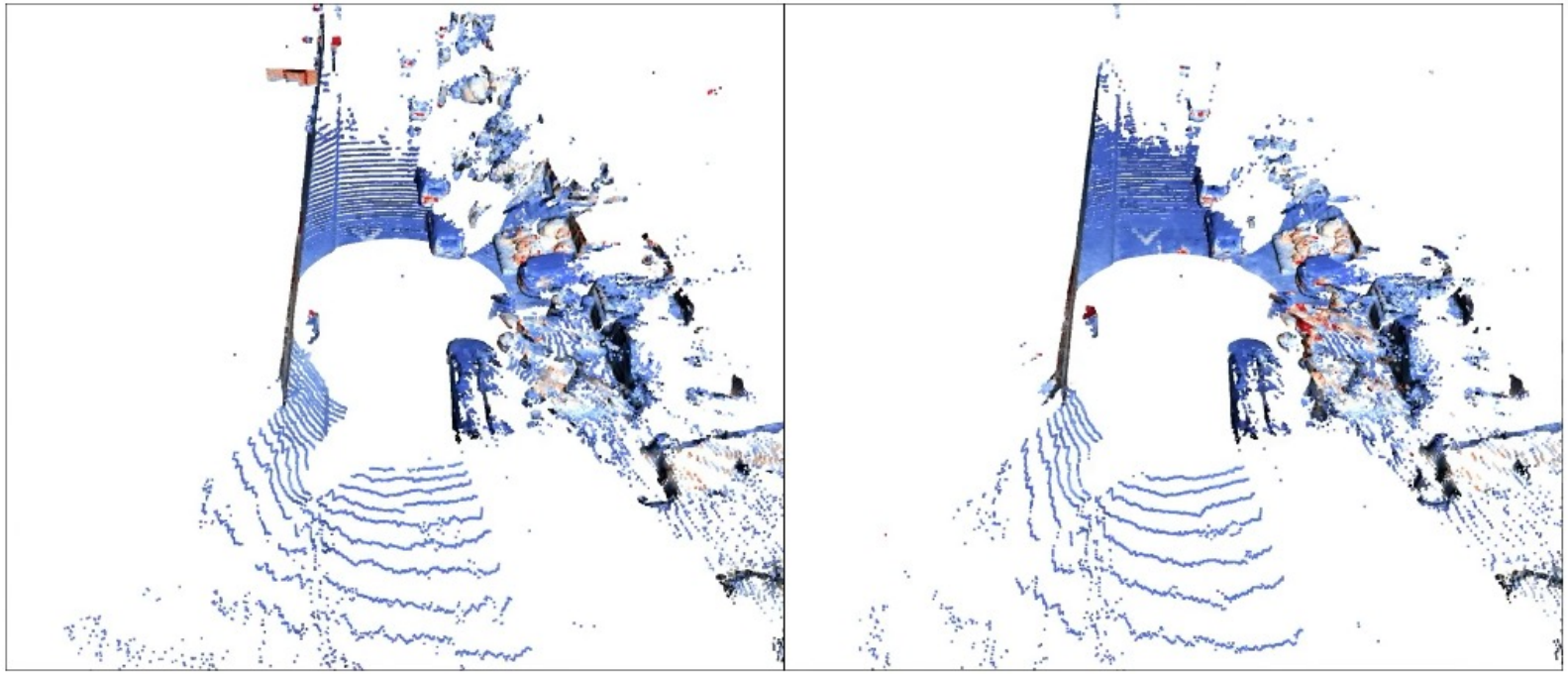
Neural LiDAR Fields for Novel View Synthesis. ICCV 2023.

LiDAR-NeRF: Novel LiDAR View Synthesis via Neural Radiance Fields. arXiv 2023.

Neural LiDAR Field的结果

LiDAR NVS by tilting the sensor





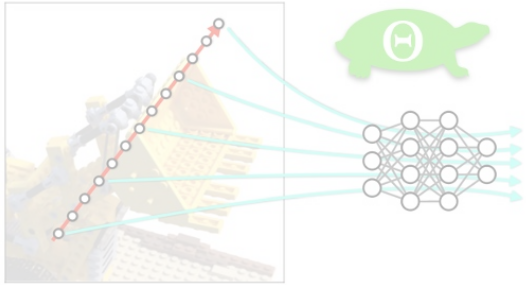
(a) raw trajectory (b) novel views

Neural LiDAR Fields for Novel View Synthesis. ICCV 2023.

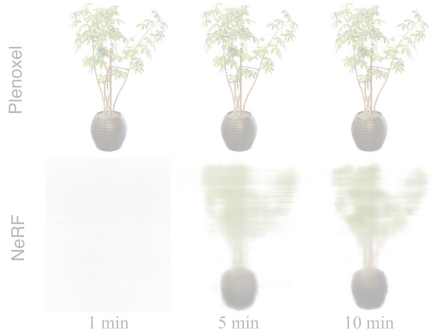
NeRF面临的问题5：几何重建质量低

NeRF在静态场景建模上遇到的问题

渲染速度慢

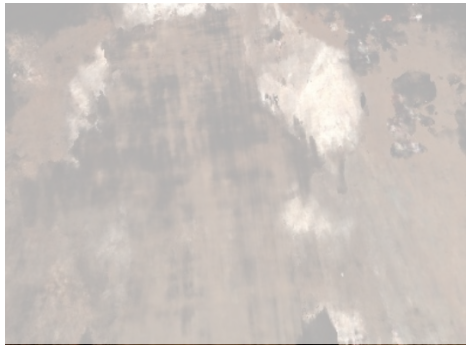


训练速度慢



场景建模能力弱

模型鲁棒性差



NeRF在其他类型场景建模上遇到的问题

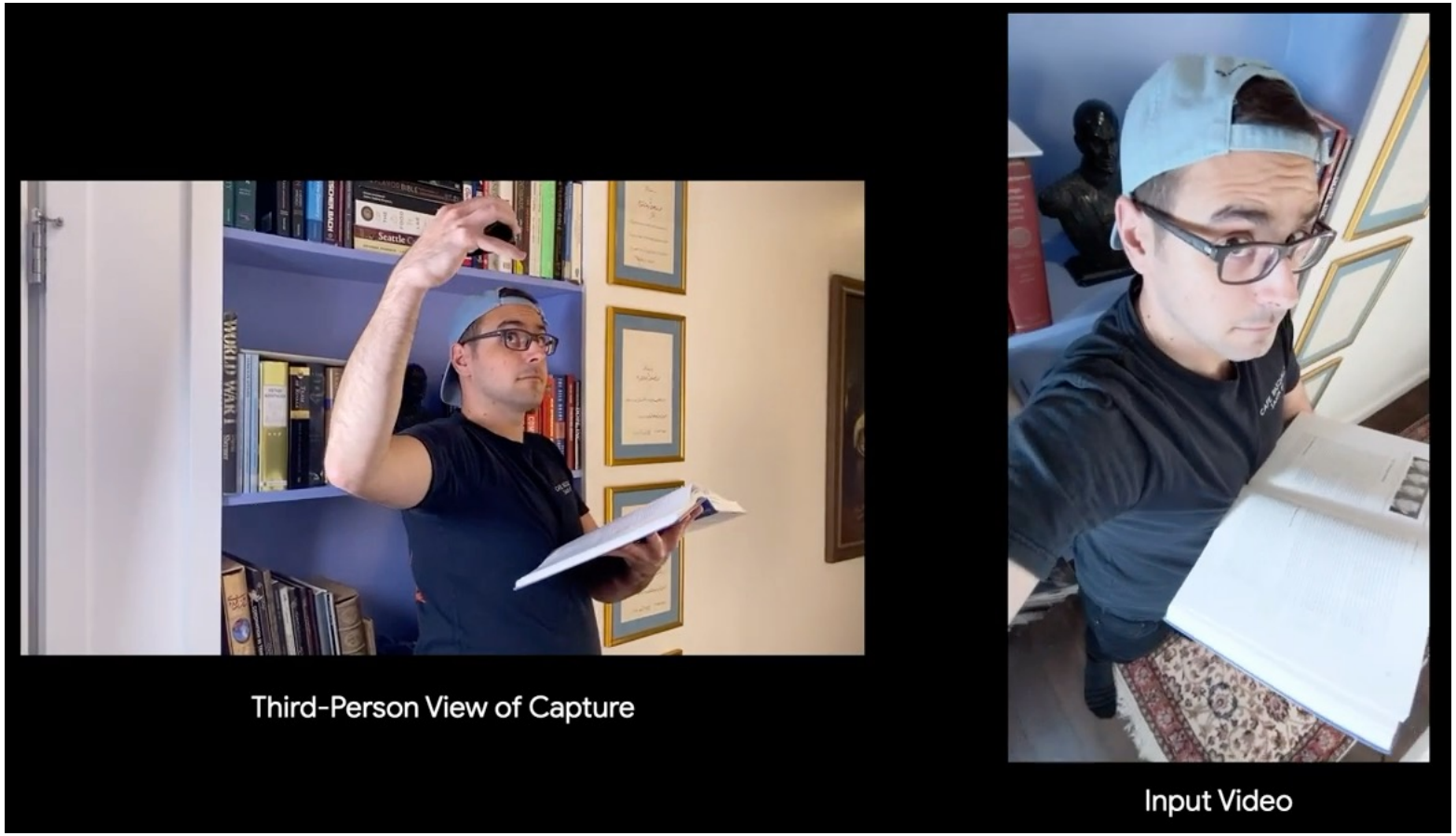
几何重建质量低

无法建模多模态信号

无法建模动态场景

NeRF无法建模动态场景

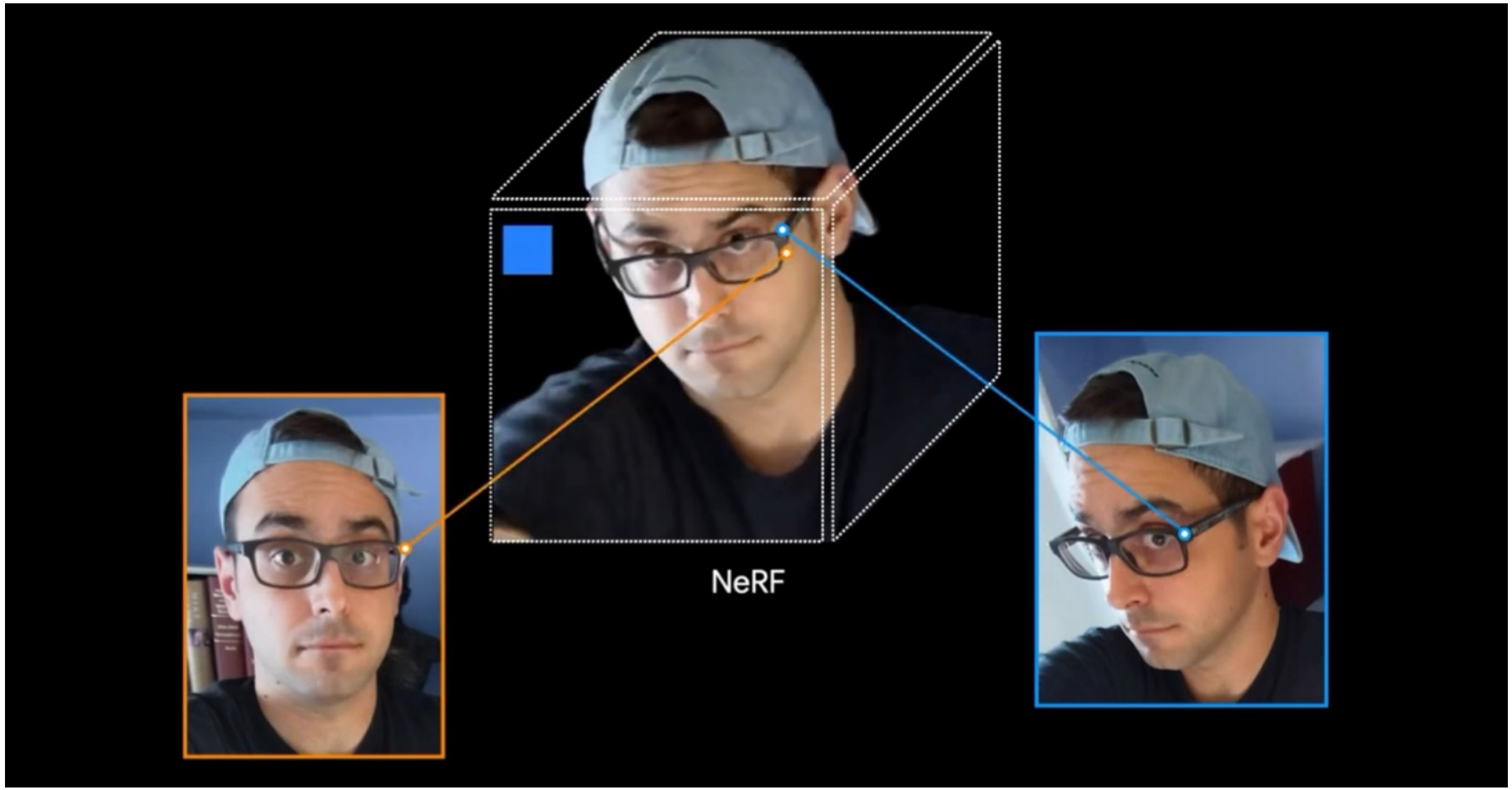
- NeRF无法从动态场景中得到合理的重建结果



Nerfies: Deformable neural radiance fields. ICCV 2021.

NeRF无法建模动态场景

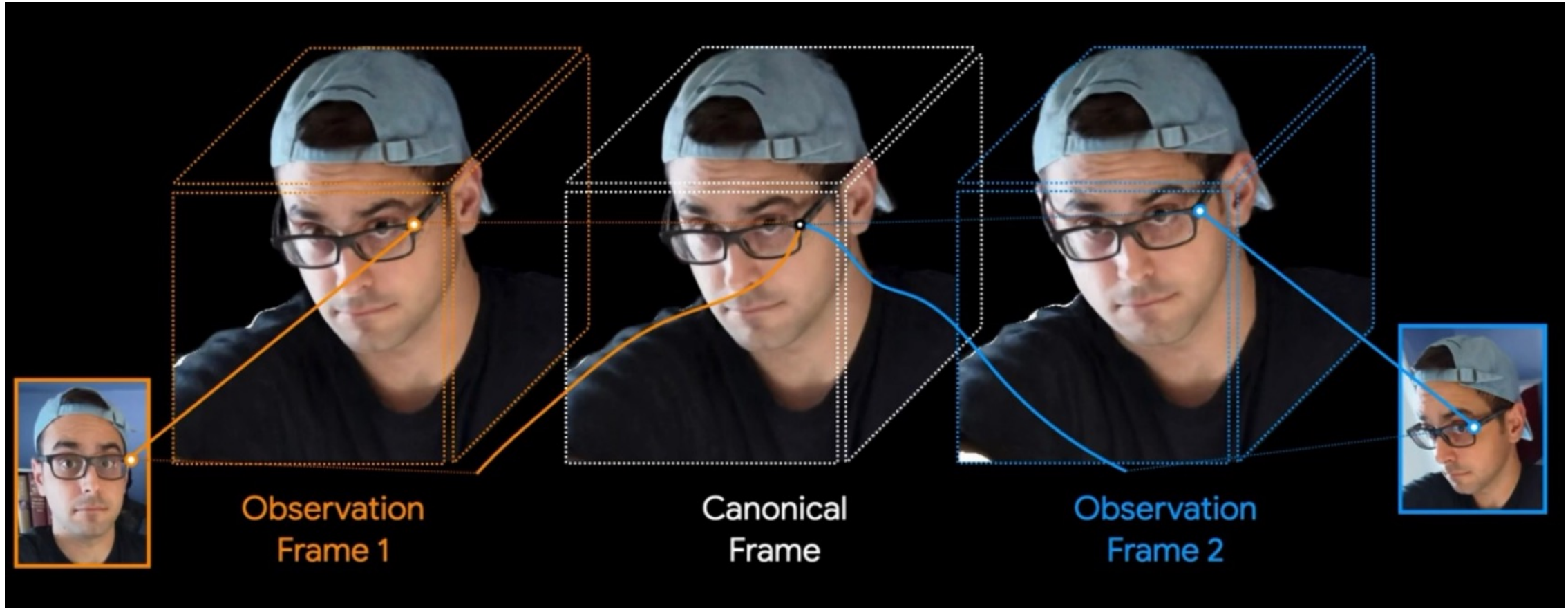
- 动态场景中的物体移动导致无法进行多视图匹配



Nerfies: Deformable neural radiance fields. ICCV 2021.

任意动态场景建模：Deformable NeRF

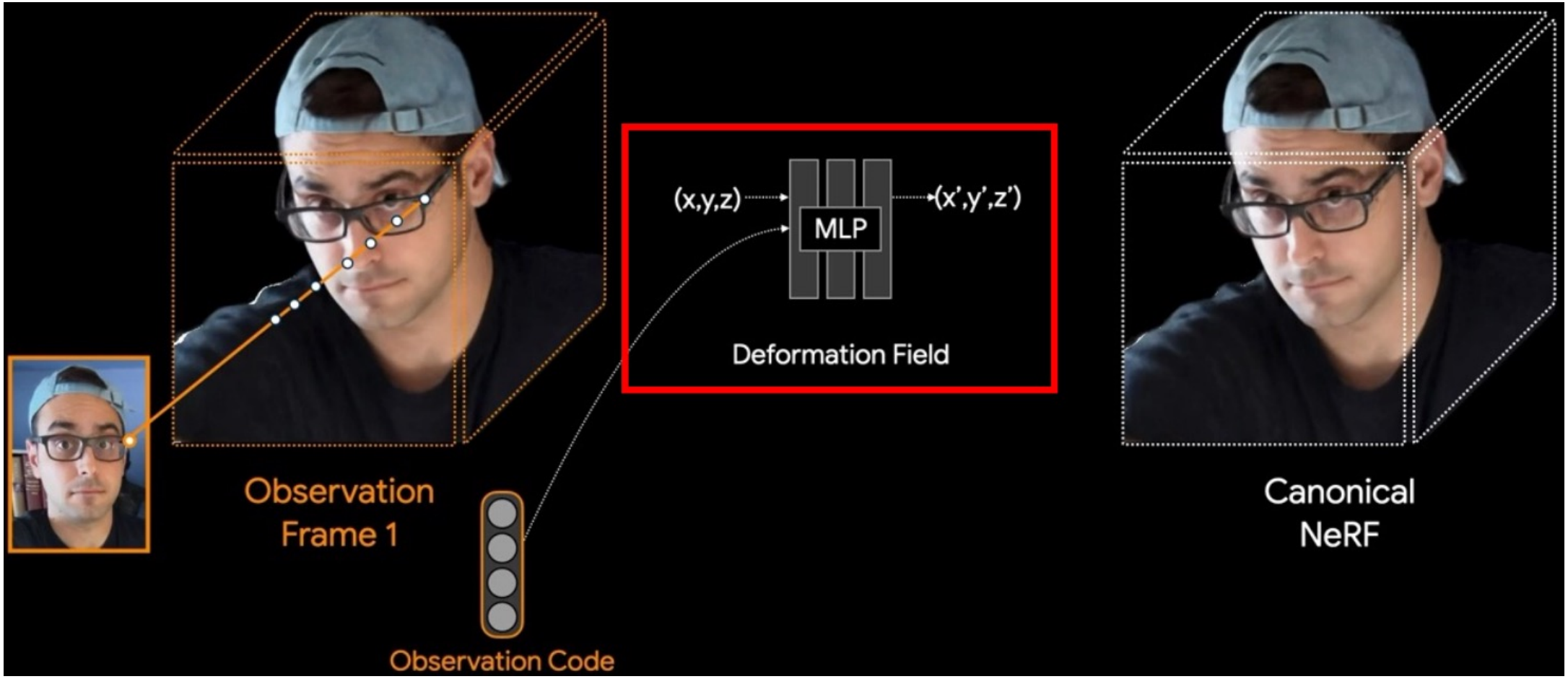
- 将动态场景建模为一个canonical nerf和一个deformation field



Nerfies: Deformable neural radiance fields. ICCV 2021.

任意动态场景建模：Deformable NeRF

- 使用MLP network表示deformation field



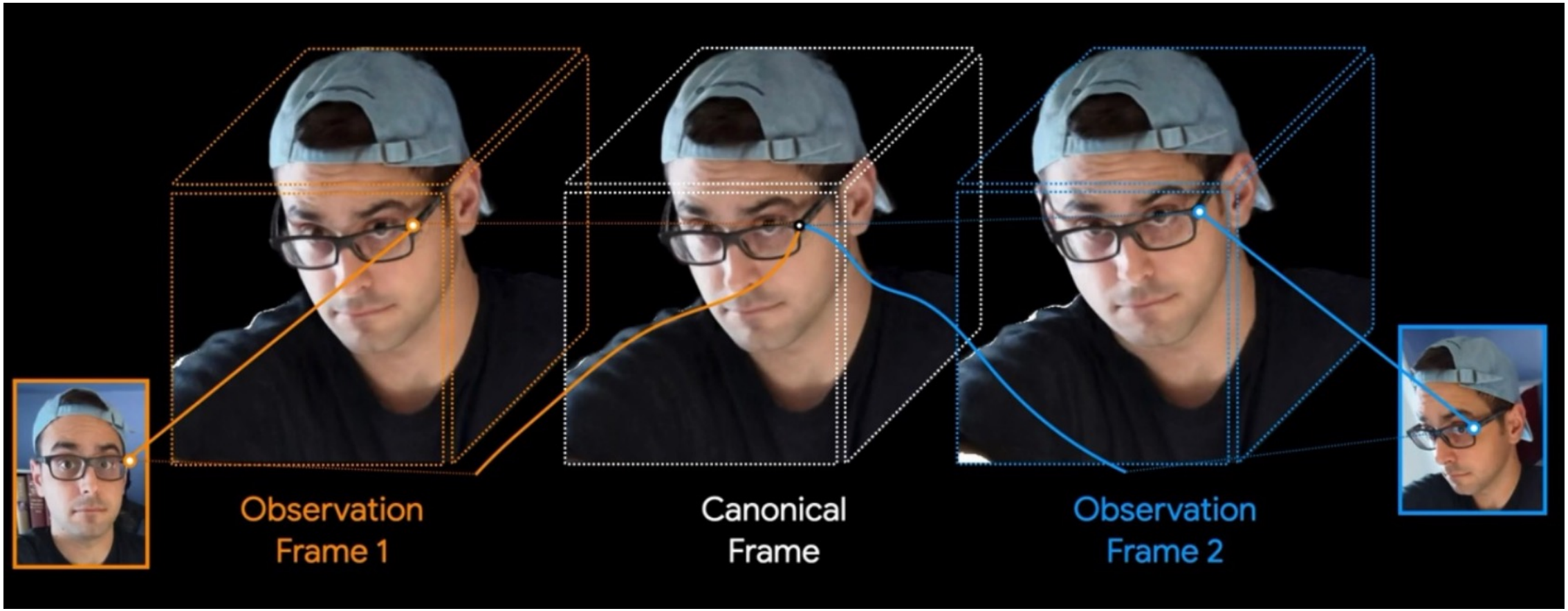
Nerfies: Deformable neural radiance fields. ICCV 2021.

Deformable NeRF的效果

Nerfies: Deformable neural radiance fields. ICCV 2021.

Deformable NeRF面临的挑战

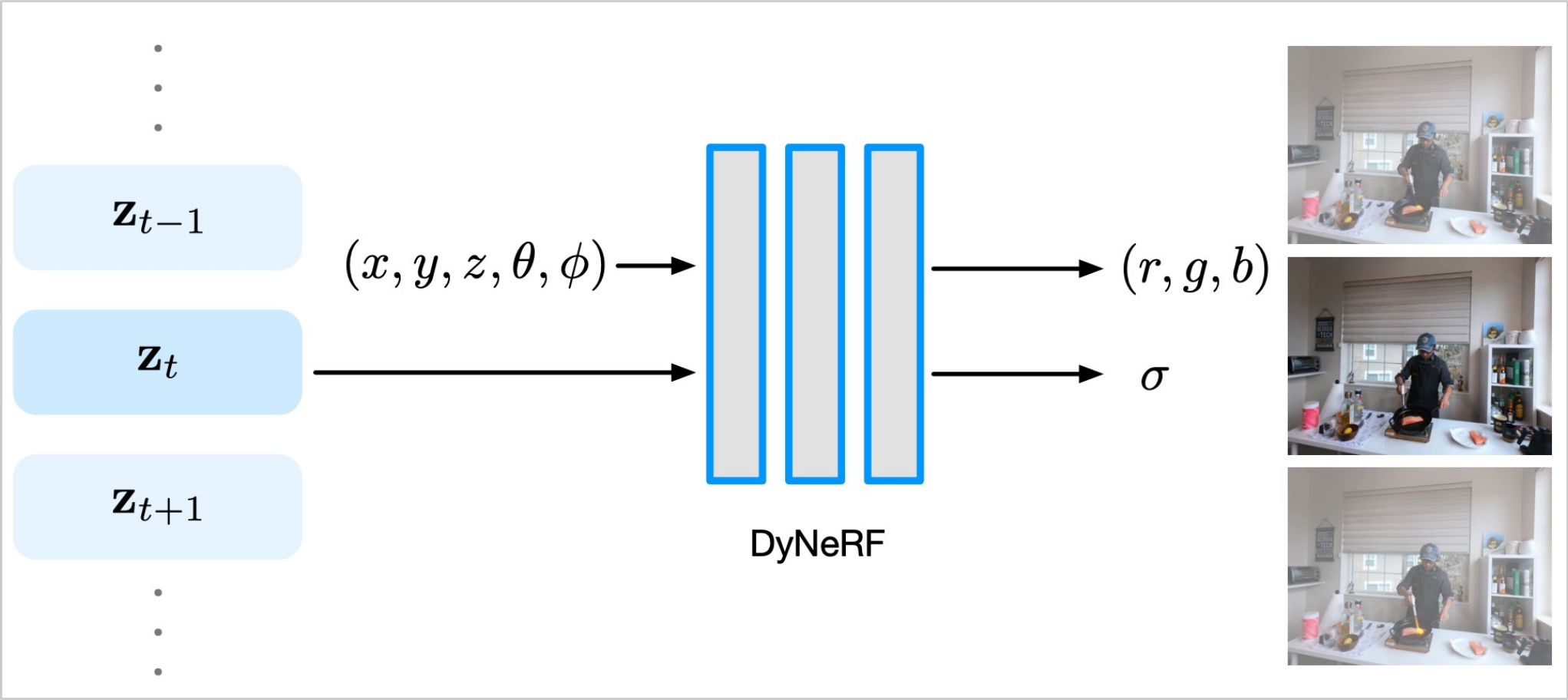
无法处理存在large motion的动态场景



Nerfies: Deformable neural radiance fields. ICCV 2021.

任意动态场景建模：DyNeRF

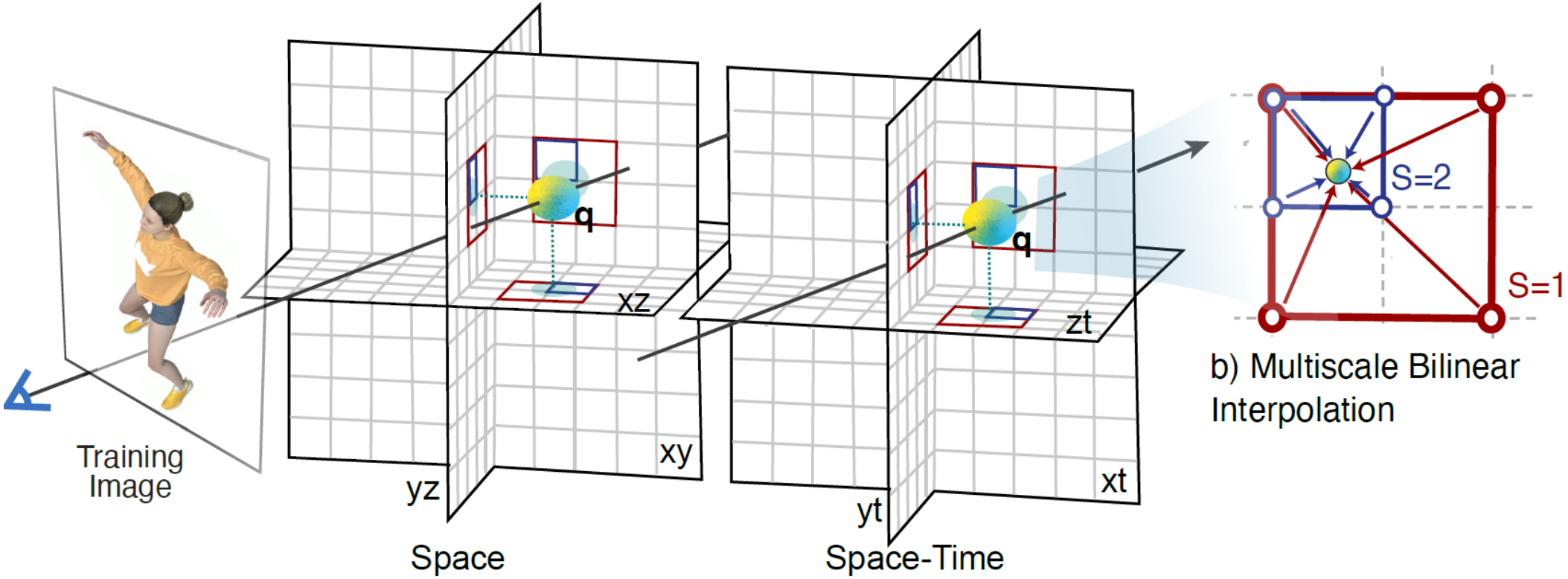
- 给NeRF增加time embedding，从而可以建模任意时间的三维场景



Neural 3D Video Synthesis from Multi-view Video. CVPR 2022.

任意动态场景建模：K-Planes

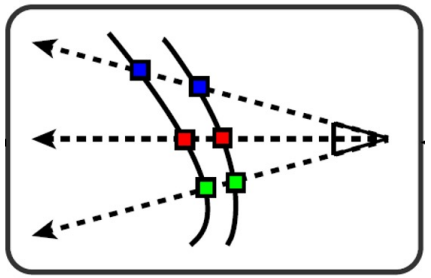
- 定义一组feature planes，给输入的3D points赋予高维feature vectors



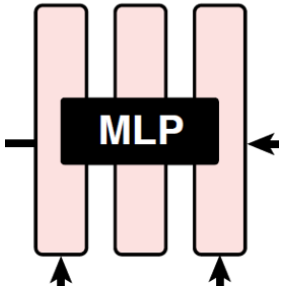
K-Planes: Explicit Radiance Fields in Space, Time, and Appearance. CVPR 2023.

K-Planes的渲染速度慢

- 网络推理成本高、推理次数多，所以渲染速度慢



*

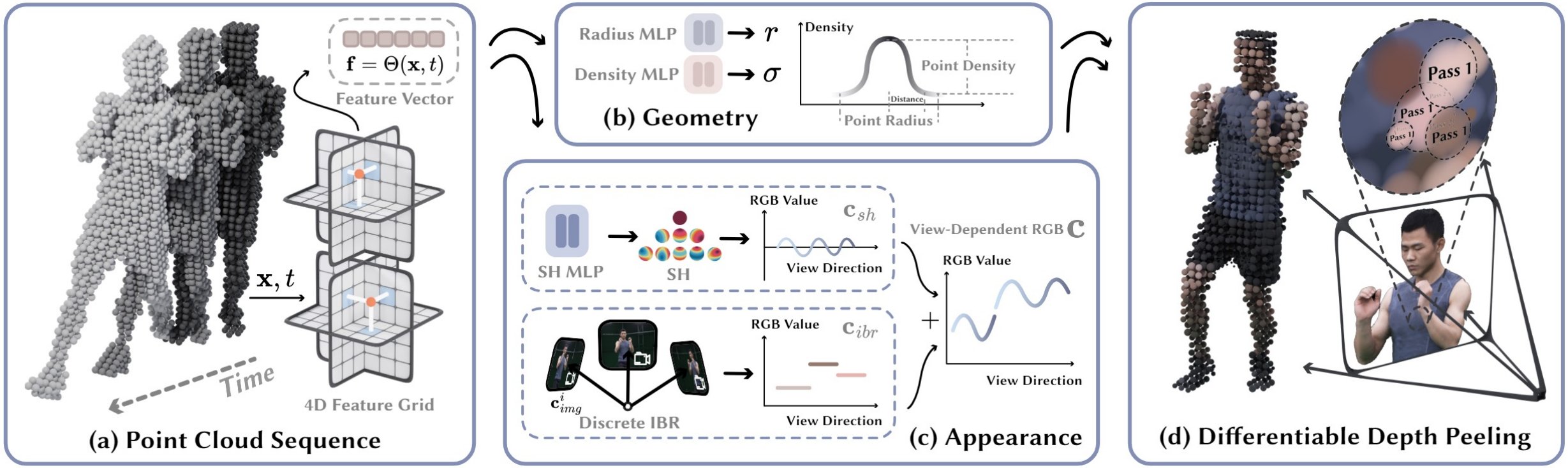


=

采样点数量 * 逐点渲染成本 = 渲染成本

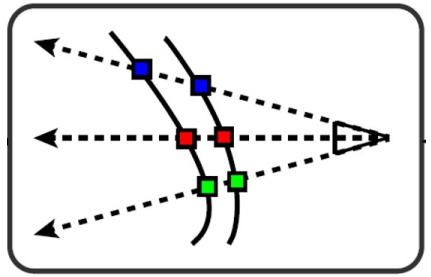
任意动态场景建模：4K4D

- 基于点云序列表示动态场景

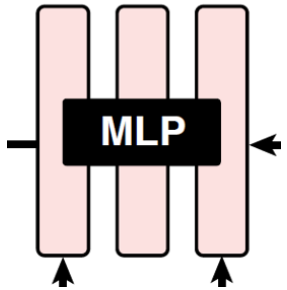


4K4D: Real-Time 4D View Synthesis at 4K Resolution. CVPR 2024.

4K4D的核心思路：降低推理次数和推理成本



*



=

采样点数量 * 逐点渲染成本 = 渲染成本

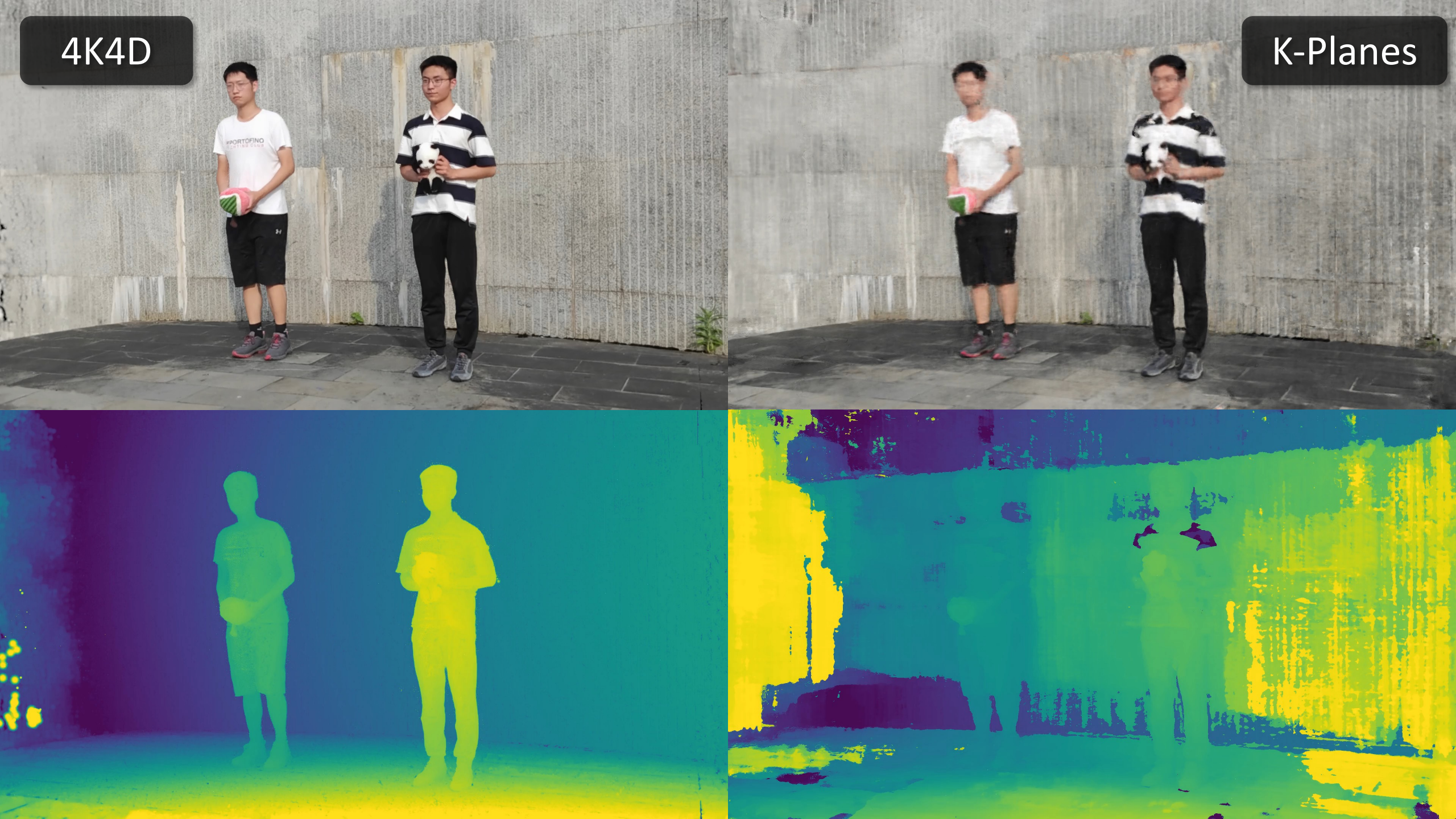
4D Point cloud representation

基于GPU的rasterization

Hybrid appearance model

Network pre-computation

Real-time on 4K


4K4D
K-Planes

动态场景渲染仍需解决的问题

采集设备

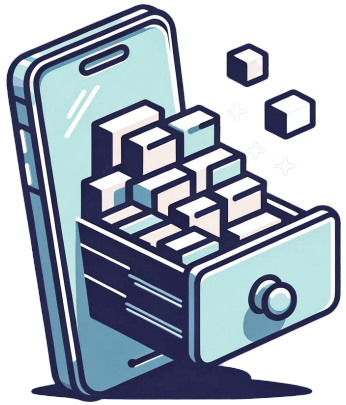
- More diverse and easier-to-use capture devices

动态场景重建

- Faster reconstruction.
- Higher quality.

模型编码

- Can be stored in mobile phones.



实时渲染

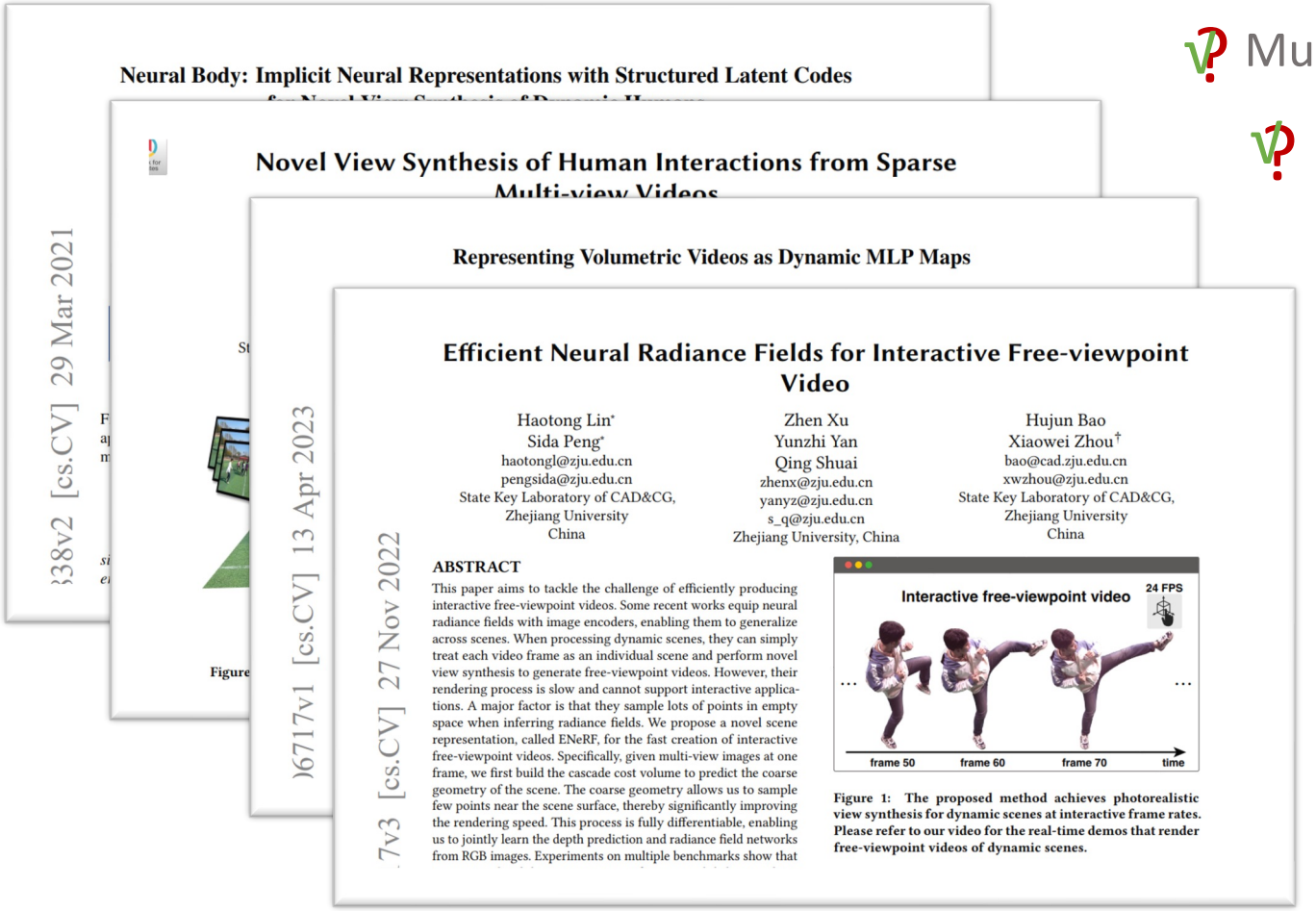
- Act like 2D video player, supporting real-time decoding and rendering.

EasyVolcap

Accelerating Volumetric Video Research

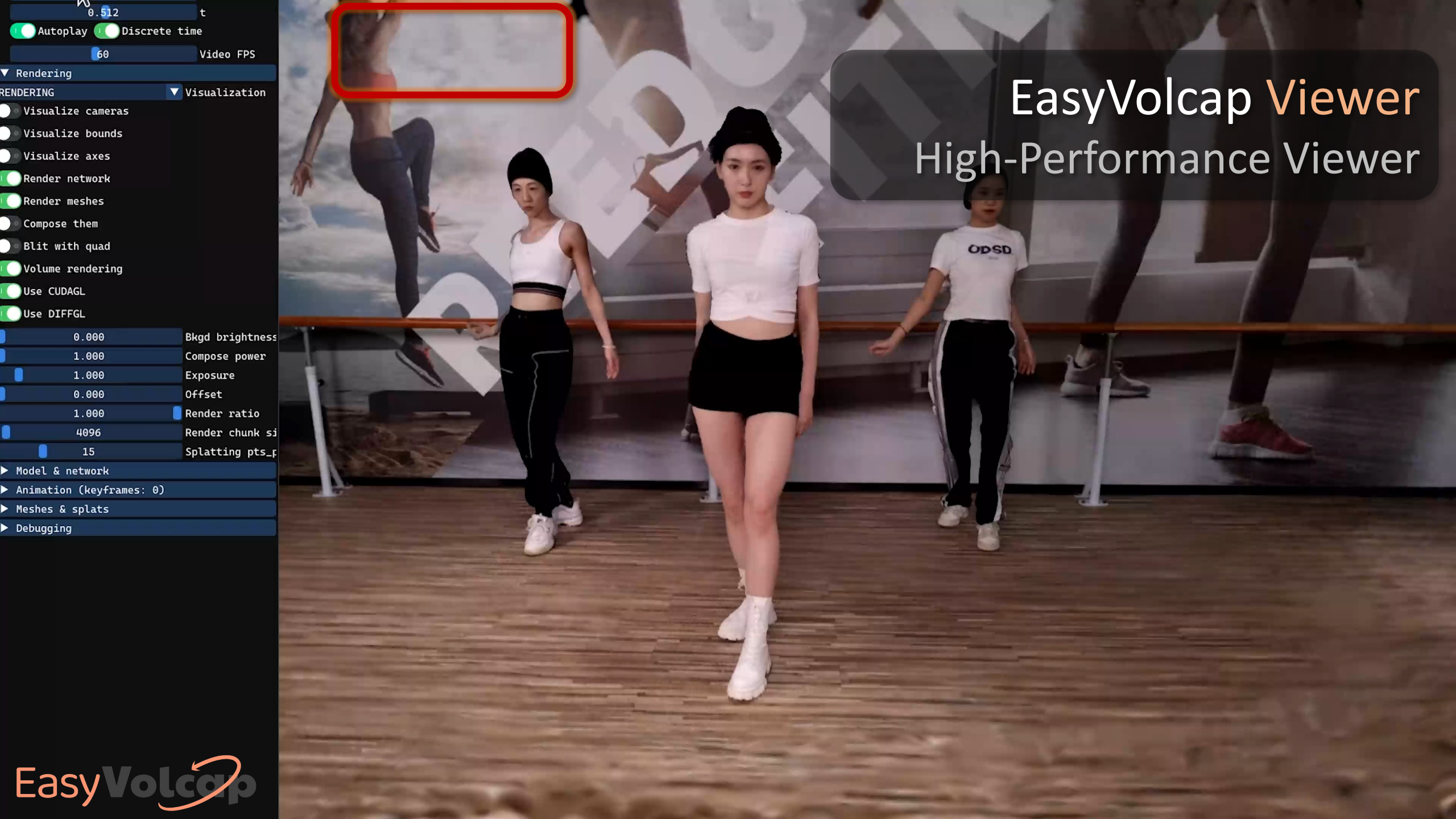
https://github.com/zju3dv/EasyVolcap

整合了动态场景数据集与动态场景渲染方法



- Multi-view dataset management
- Unified volumetric video pipeline
- Viewer for dynamic scenes

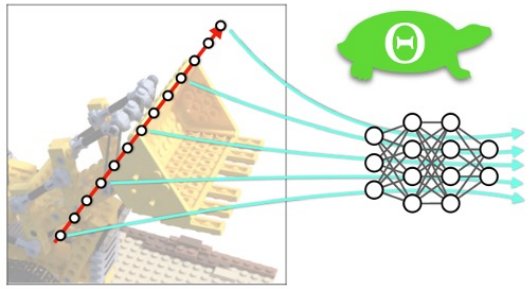
EasyVolcap


0.512
t
Autoplay
Discrete time
60
Video FPS
Rendering
RENDERING
Visualization
Visualize cameras
Visualize bounds
Visualize axes
Render network
Render meshes
Compose them
Blit with quad
Volume rendering
Use CUDAGL
Use DIFFGL
0.000
Bkgd brightness
1.000
Compose power
1.000
Exposure
0.000
Offset
1.000
Render ratio
4096
Render chunk si
15
Splatting pts_p
Model & network
Animation (keyframes: 0)
Meshes & splats
Debugging
EasyVolcap Viewer
High-Performance Viewer
EasyVolcap

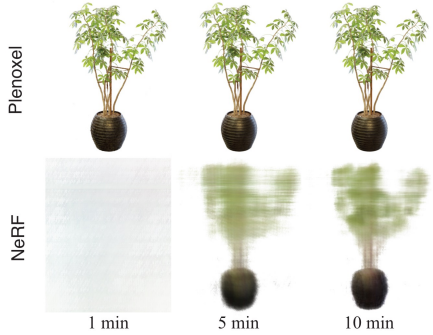
总结

在静态场景建模上对NeRF的改进

提升渲染速度

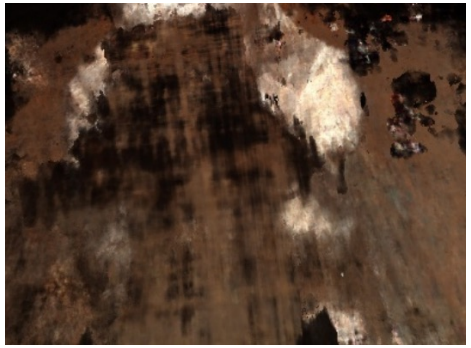


提升训练速度



大规模场景建模

提升模型鲁棒性



在其他类型场景建模上对NeRF的改进

提升几何重建质量

多模态信号建模

动态场景建模

谢谢！

彭思达
浙江大学