

CIDAI

Centre of Innovation
for Data tech
and Artificial Intelligence



Microsoft

NTT DATA



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONTECH
Intelligent Data Science and
Artificial Intelligence Research Center



NeRF and Family

Ivan Huerta- i2CAT

17/10/2023

www.cidai.eu

- Introducción a la Generación de Nuevas Vistas y Representación de Superficies 3D
- NeRF: Modelo de Representación Implícita en 3D
- Extensiones y Desarrollos Avanzados de NeRF

¿Qué problema quiere solucionar?

Generación de nuevas vistas



- Virtual tourism
- Walk-throughs and photo tours
- Entertainment
- AR / VR / XR
- More engaging interaction with pictures
- Video stabilization
- Data augmentation
- Content generation

¿ Como se representa ?

Representación implícita

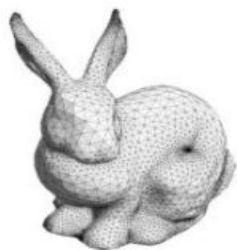
New view



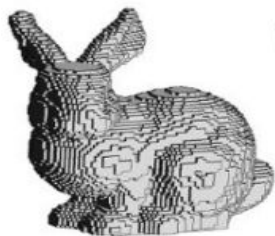
Representing 3D surfaces

Explicit and Implicit representations

Explicit:



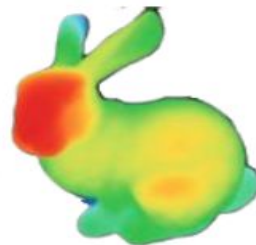
Mesh



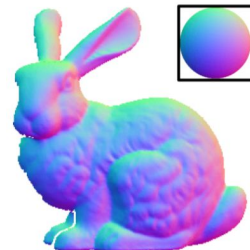
Voxels



Point clouds

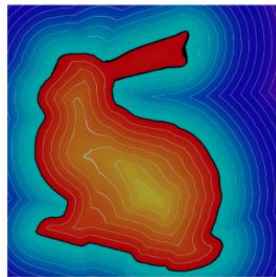


Depth

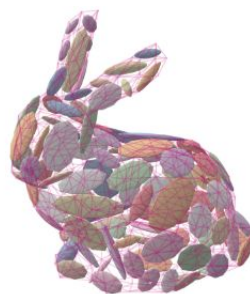


Surface Normals

Implicit:



Signed distance field

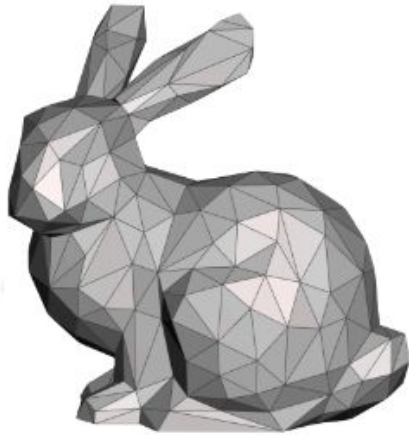


Mixture of primitives
(e.g gaussian mixtures)



Density

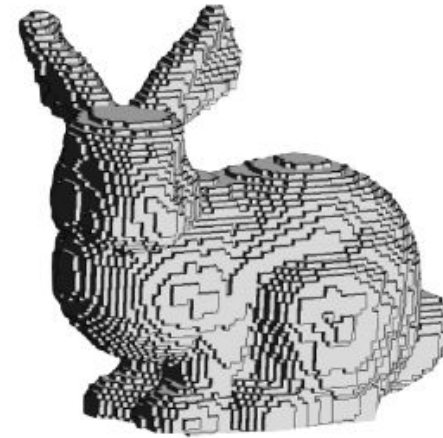
Representing 3D surfaces Geometry representations



Mesh Representation

Small memory footprint

Hard to optimize

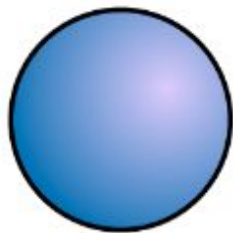


Voxel Representation

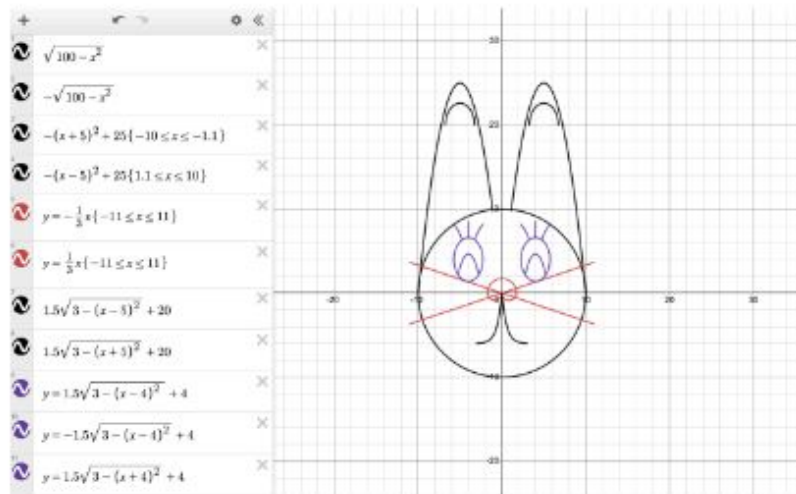
Easy to optimize

Large memory footprint

Representing 3D surfaces Implicit functions

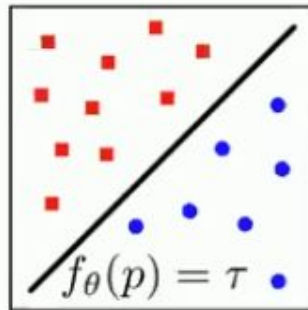


$$x^2 + y^2 + z^2 = 1$$

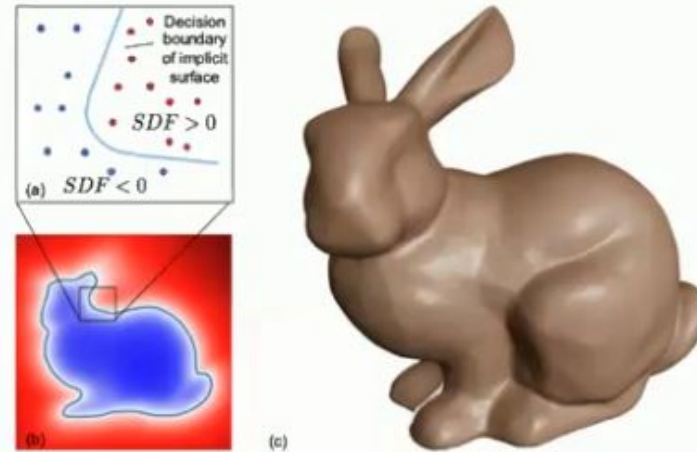


Occupancy Networks (Mescheder et al. 2019)

$(x, y, z) \rightarrow \text{occupancy}$



DeepSDF
(Park et al. 2019)
 $(x, y, z) \rightarrow \text{distance}$



Scene Representation Networks

(Sitzmann et al. 2019)

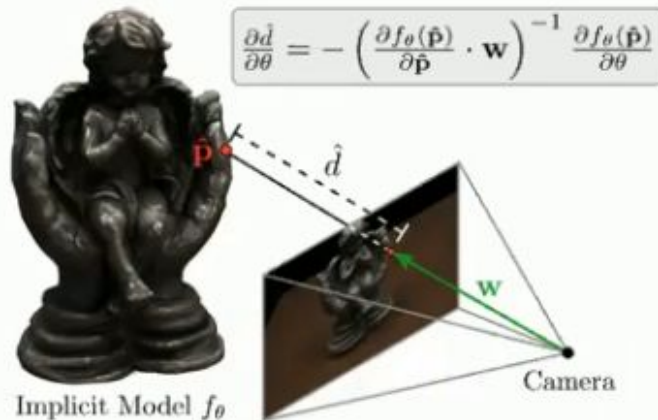
$(x, y, z) \rightarrow$ latent vec. (color, dist.)



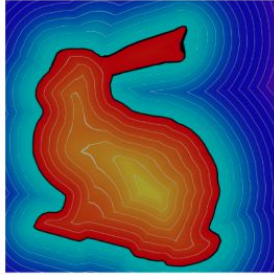
Differentiable Volumetric Rendering

(Niemeyer et al. 2020)

$(x, y, z) \rightarrow \text{color, occ.}$

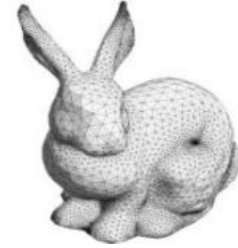


Converting Implicit Surfaces to meshes



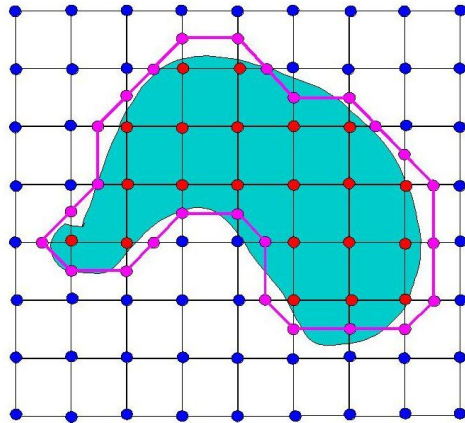
Implicit representation

Extract iso-surface

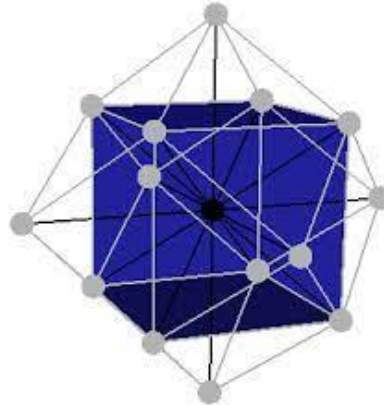


Mesh

Marching Cubes



Marching tetrahedra



Pros:

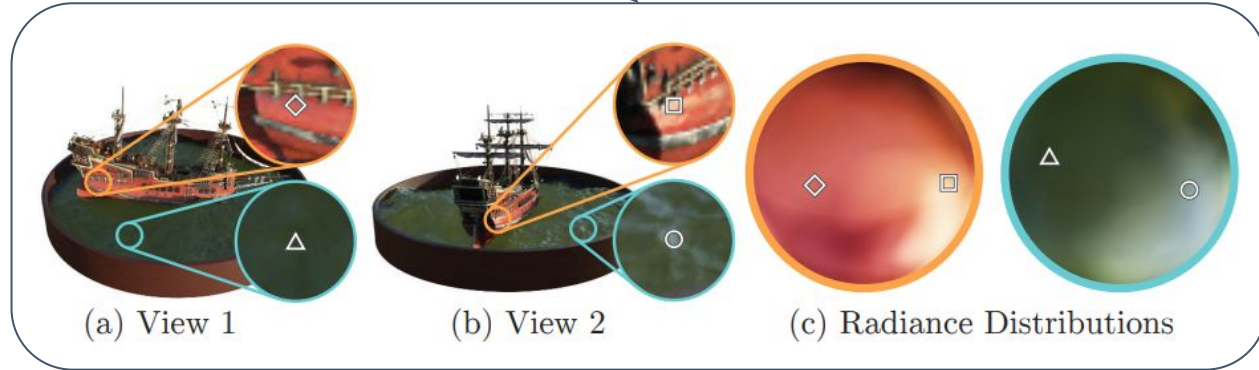
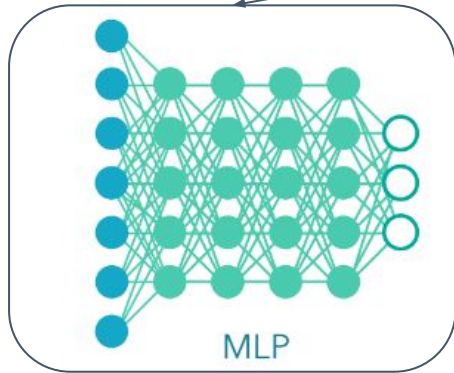
- Se puede definir la iso-surface sin agujeros (opuesto a **PC**)
- Puede adaptarse a cualquier topología (opuesto a **Meshes** e.g: de una esfera a un donut)
- Se puede representar con menos parámetros y de forma eficiente (opuesto a **Voxels**)
 - Son altamente comprimible (1-10 Mb)
- Son continuos en el espacio 3D

Cons:

- No es trivial extraer la representación extrínseca
- Necesita pasos extra para visualizar las representaciones implícitas
- ~~● No todas las formas complejas pueden representarse de manera eficiente o precisa~~
- ~~● Modelo de renderización limitado: difícil de optimizar~~

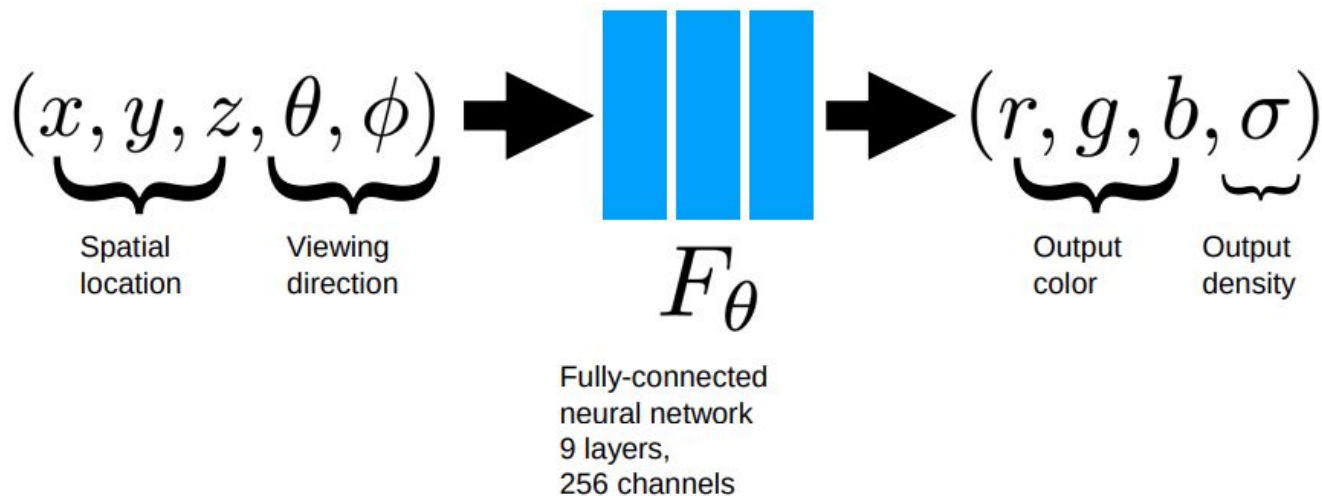
Neural Radiance fields (NeRF) Representation

NeRF: Neural Radiance fields



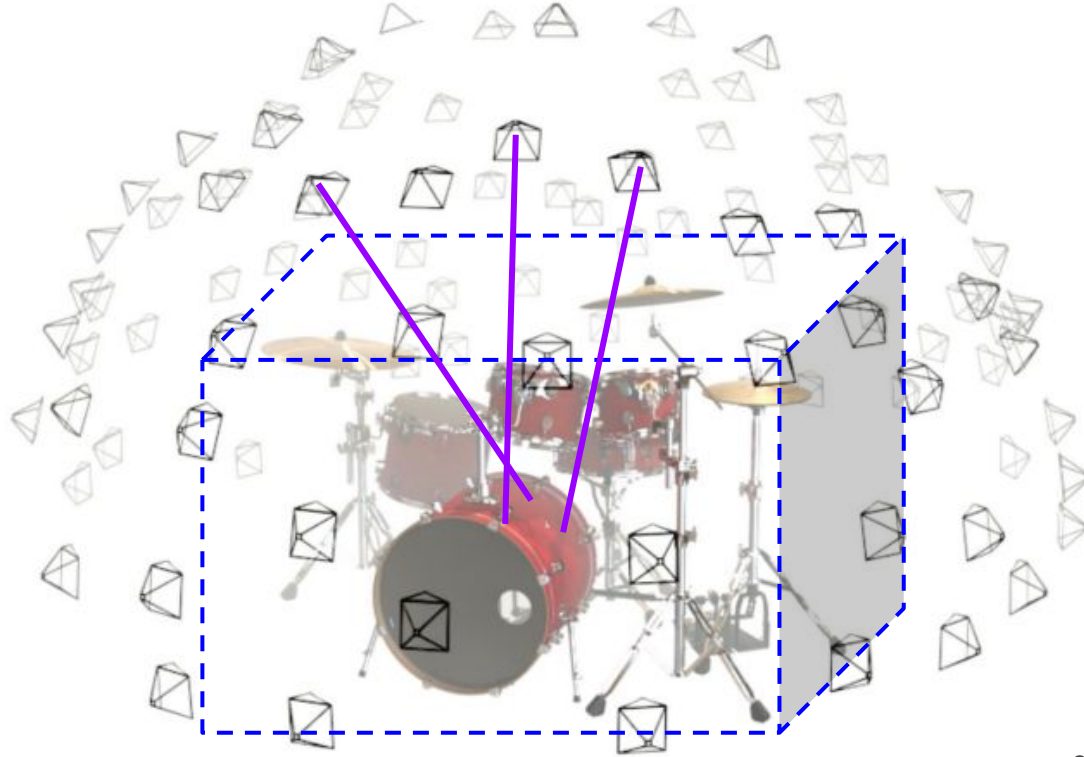
[1] Mildenhall, Srinivasan, Tancik et al., [NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis](#), ECCV 2020

Neural Radiance fields (NeRF) Representation



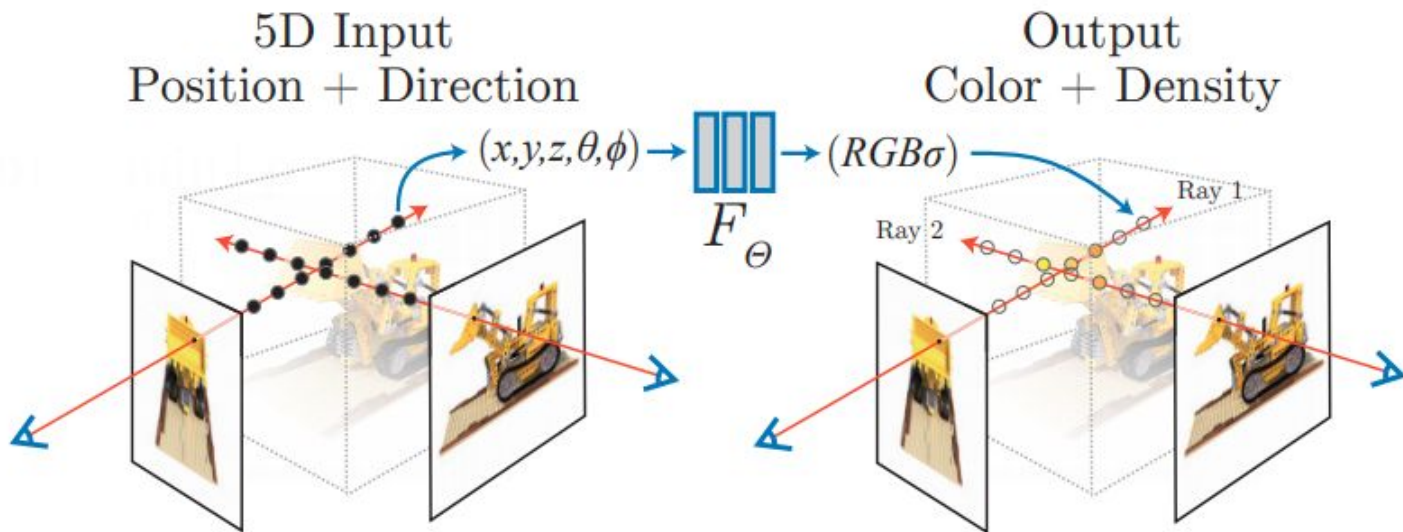
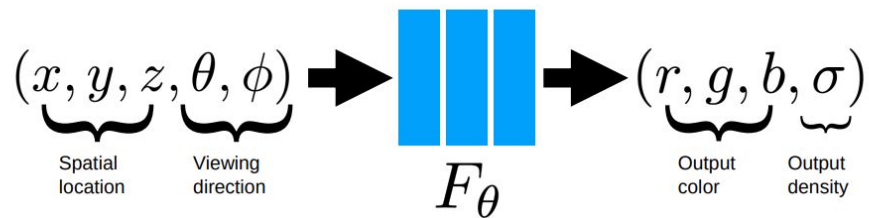
Neural Radiance fields (NeRF)

Input images



Slide credit: Jon Barron

Neural Radiance fields (NeRF) Representation



Neural Radiance field (NeRF)

Volumetric rendering with ray tracing

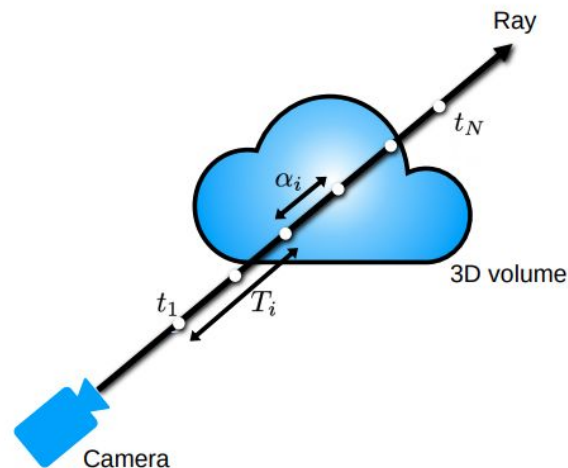
Expected color
for a ray

Opacity

Predicted colors

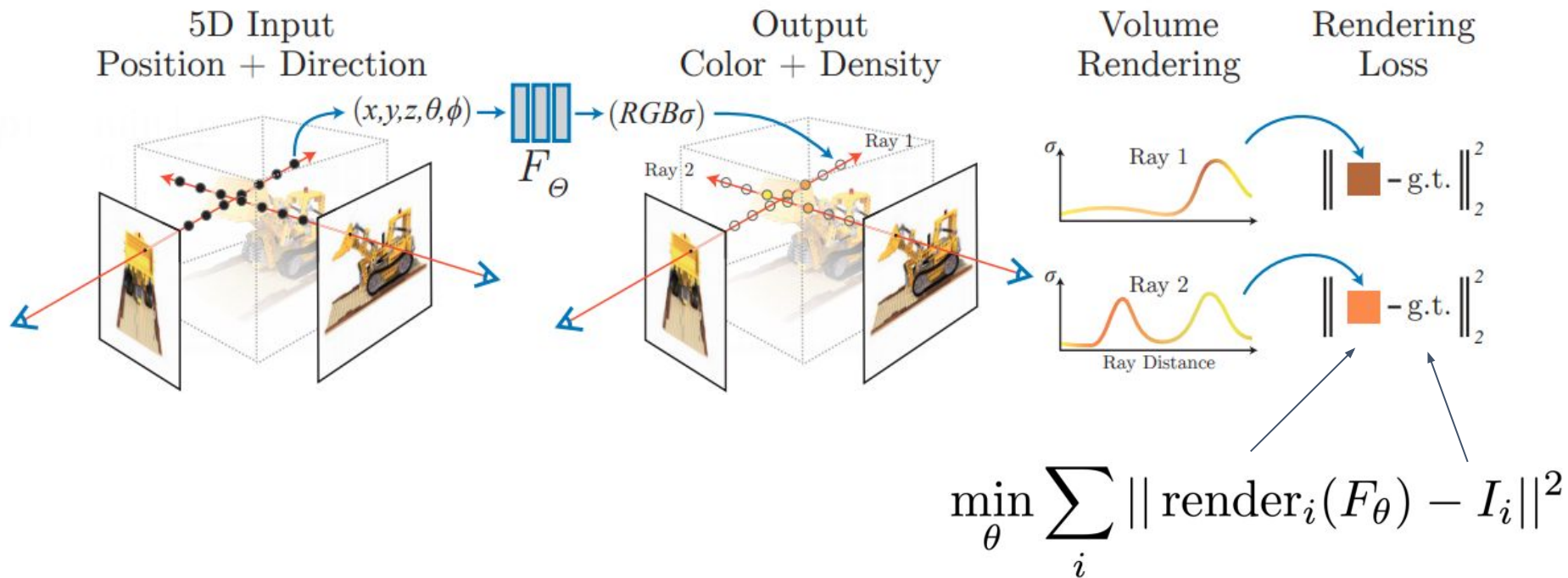
$$C(\mathbf{r}) = \int_{t_n}^{t_f} T(t) \sigma(\mathbf{r}(t)) \mathbf{c}(\mathbf{r}(t), \mathbf{d}) dt, \text{ where } T(t) = \exp\left(-\int_{t_n}^t \sigma(\mathbf{r}(s)) ds\right)$$

Volume density

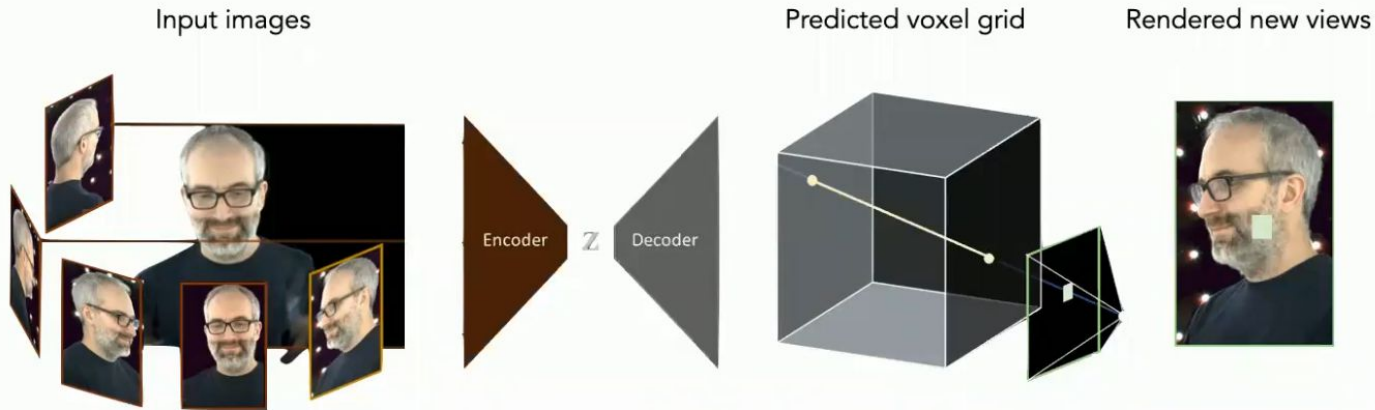


Neural Radiance field (NeRF)

Volumetric rendering with ray tracing



RGB-alpha volume rendering for view synthesis



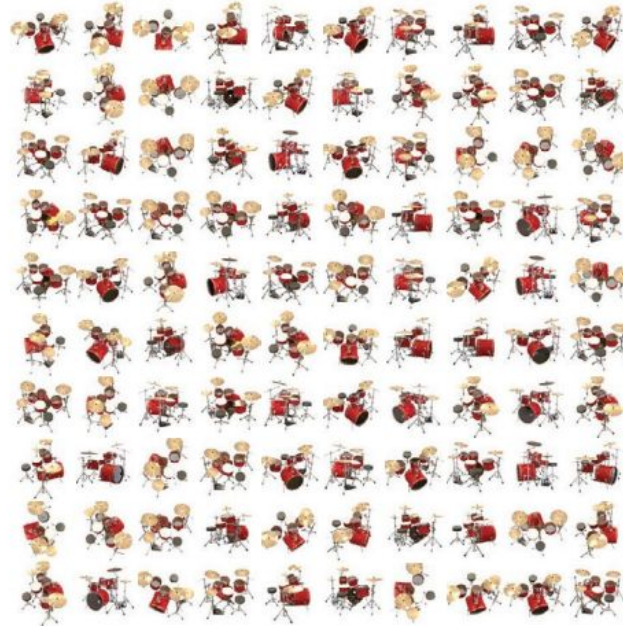
- + Great rendering model: good for optimization
- Horrible storage requirements (1-10 GB)

Neural Volumes, Lombardi et al. 2019

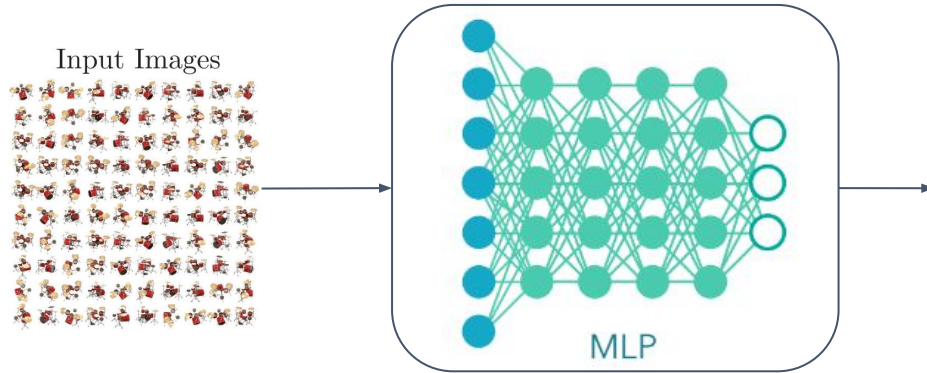
Capture Scene



Input Images



Optimize NeRF



Render New Views



Neural Radiance field (NeRF) Image Based Rendering



Rendered Camera Path



View Dependent Apperance

Neural Radiance field (NeRF) Inputs

Calibrated Camera
Intrinsics and extrinsics



Neural Radiance field (NeRF) NeRF input

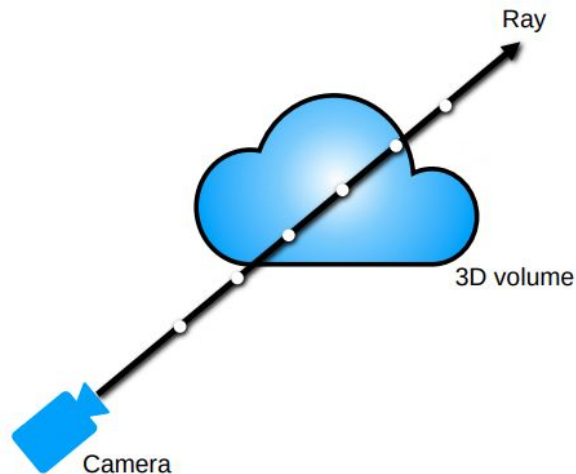


Sparse model of central Rome using 21K photos produced by COLMAP's SfM pipeline.



Dense models of several landmarks produced by COLMAP's MVS pipeline.

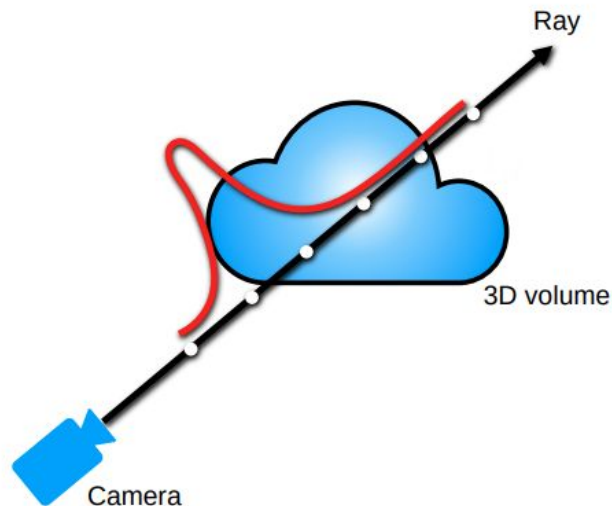
Can we allocate samples more efficiently?
Two pass rendering



Two pass rendering: coarse

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

treat weights as probability
distribution for new samples



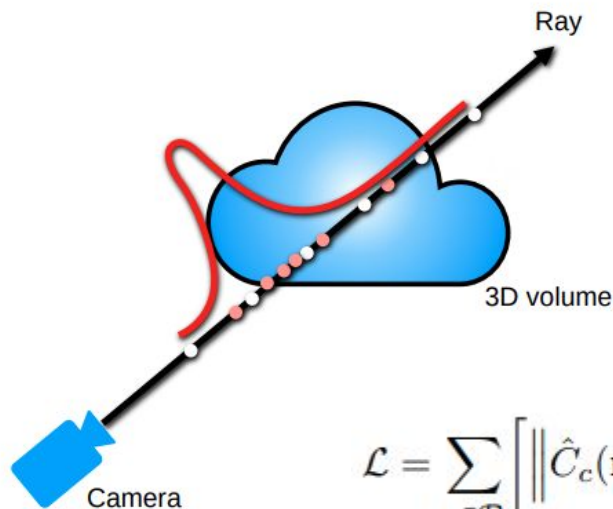
Neural Radiance field (NeRF)

Trick: Two pass filtering

Two pass rendering: fine

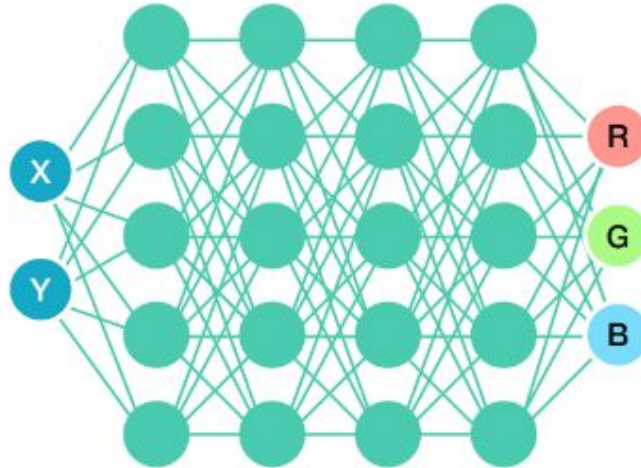
$$C \approx \sum_{i=1}^N T_i \alpha_i c_i$$

treat weights as probability
distribution for new samples



$$\mathcal{L} = \sum_{\mathbf{r} \in \mathcal{R}} \left[\left\| \hat{C}_c(\mathbf{r}) - C(\mathbf{r}) \right\|_2^2 + \left\| \hat{C}_f(\mathbf{r}) - C(\mathbf{r}) \right\|_2^2 \right]$$

Image Representation





MLP output



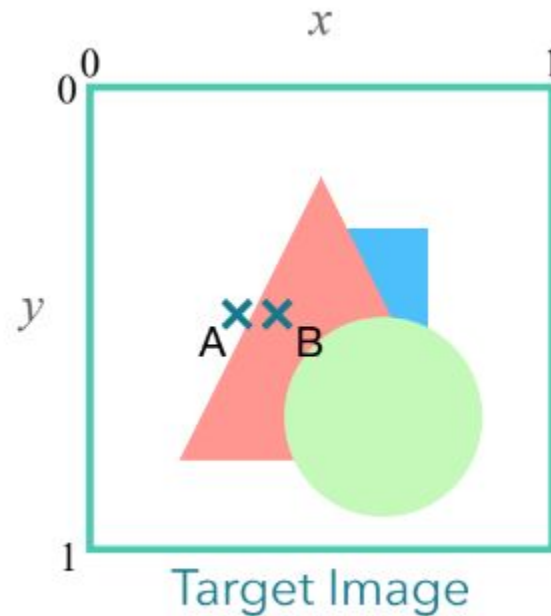
Supervision image

Neural Radiance field (NeRF) :Tricks

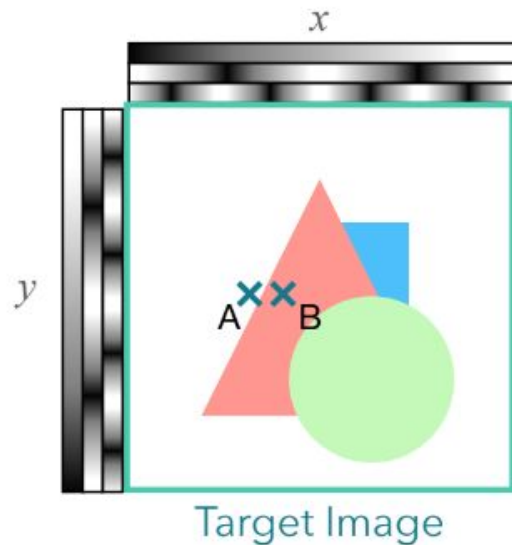
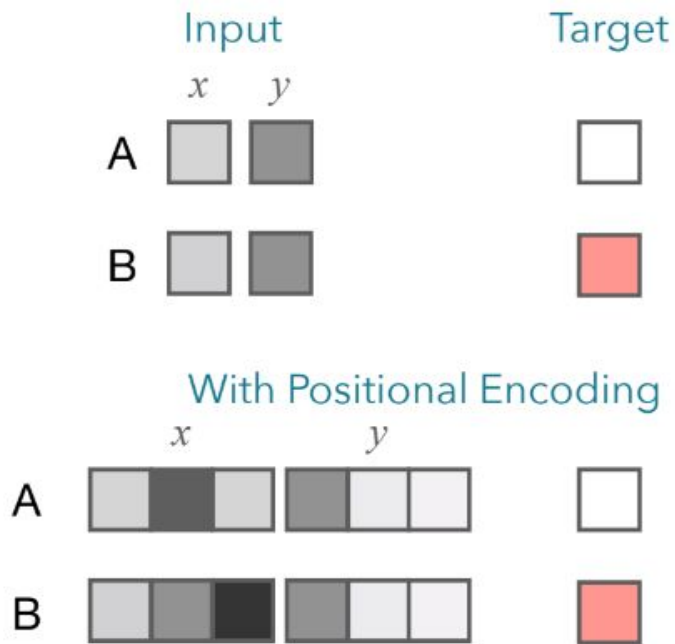
Trick: Positional encoding

	Input	
	x	y
A	.36	.5
B	.38	.5

Target



Why does positional encoding help?



NeRF with and without positional encoding



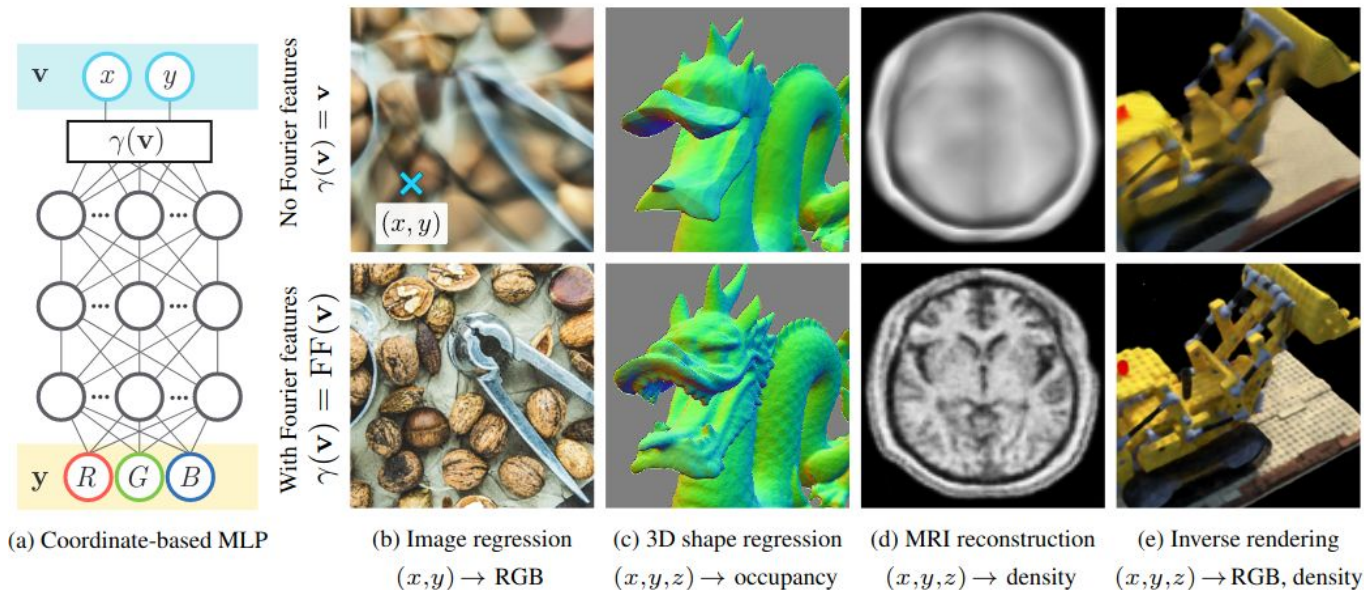
NeRF (Naive)



NeRF (with positional encoding)

Neural Radiance field (NeRF) :Tricks

Trick: Positional encoding with Fourier features



[Fourier features let networks learn high frequency functions in low dimensional domains](#), [M Tancik](#), [P Srinivasan](#), [B Mildenhall](#)

Neural Radiance field (NeRF)

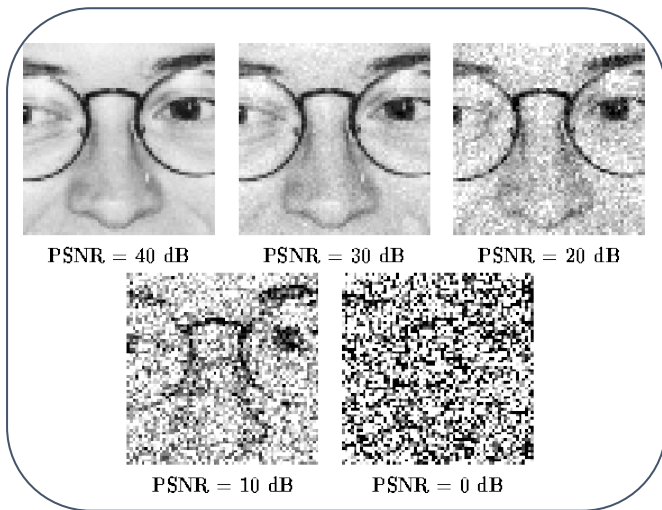
Results



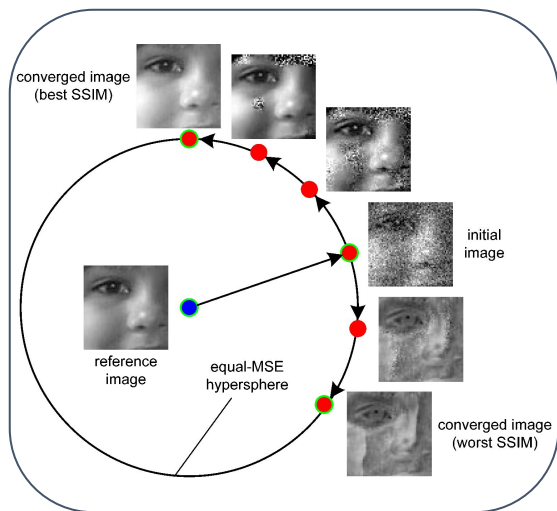
Neural Radiance field (NeRF) Results



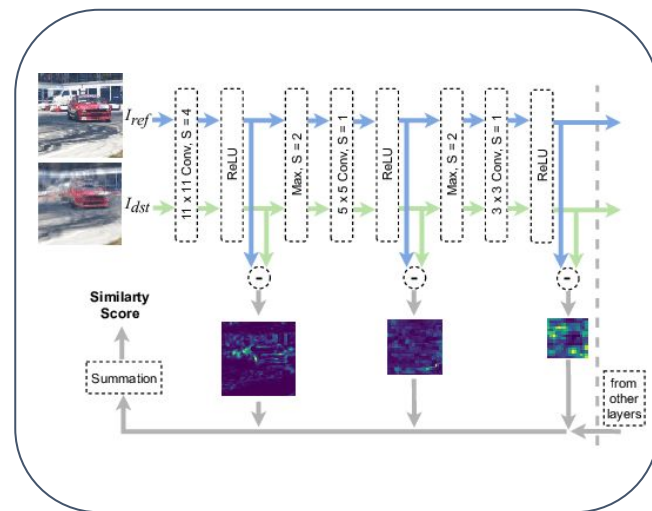
PSNR



SSIM



LPIPS



[1] Zhou Wang, Eero P Simoncelli, and Alan C Bovik. Multiscale structural similarity for image quality assessment. Asilomar Conference on Signals, Systems & Computers, 2003.

[2]. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In: CVPR (2018)

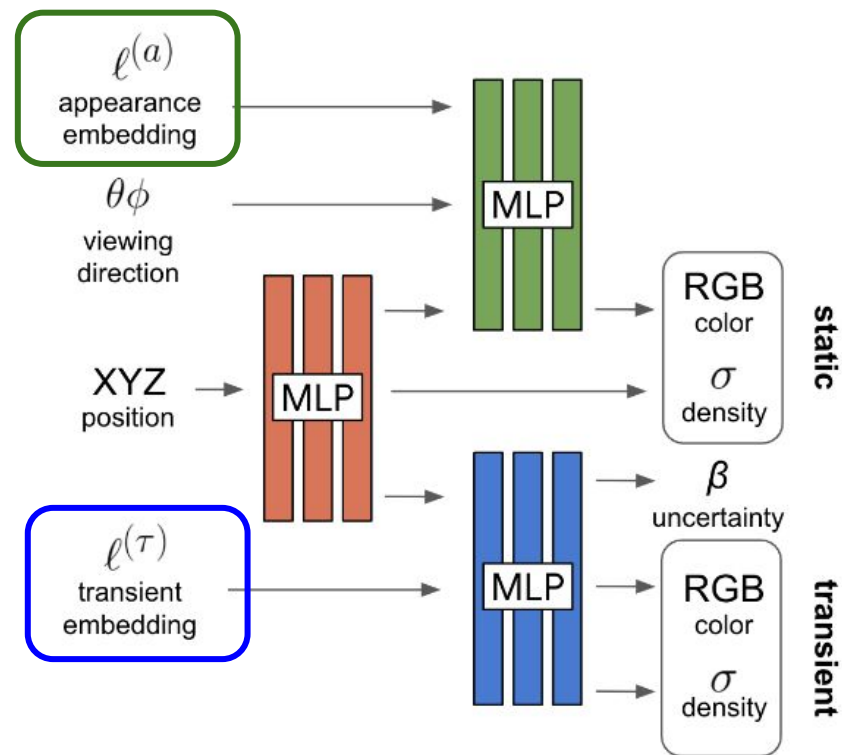
Limitations original NeRF



Family Natural Images



Family Natural Images

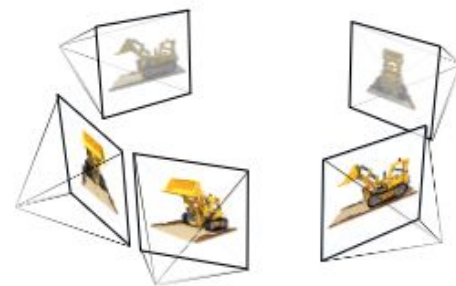


Family Natural Images



Family Pose Estimation

NeRF



Images + accurate camera poses



3D scene representation

BARF (ours)



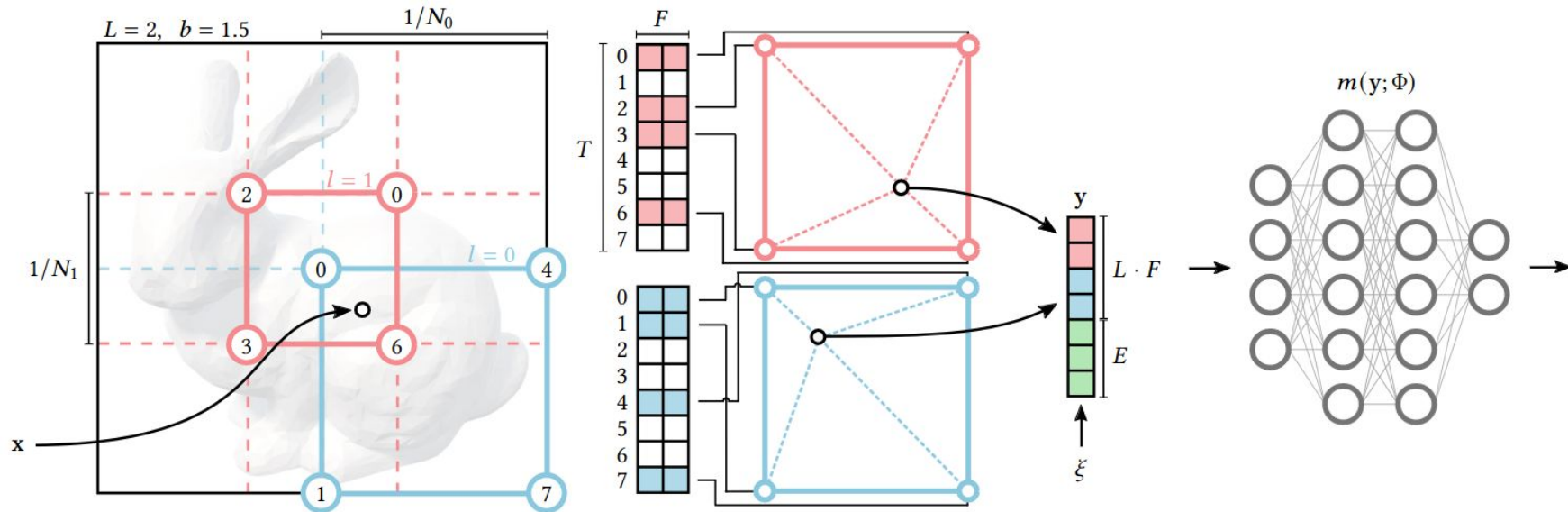
Images + **imperfect** camera poses



3D scene representation
+ **registered** camera poses

Family

Faster Training and Inference



Family

Faster Training and Inference



Family

Multi-scale: Aliasing problem

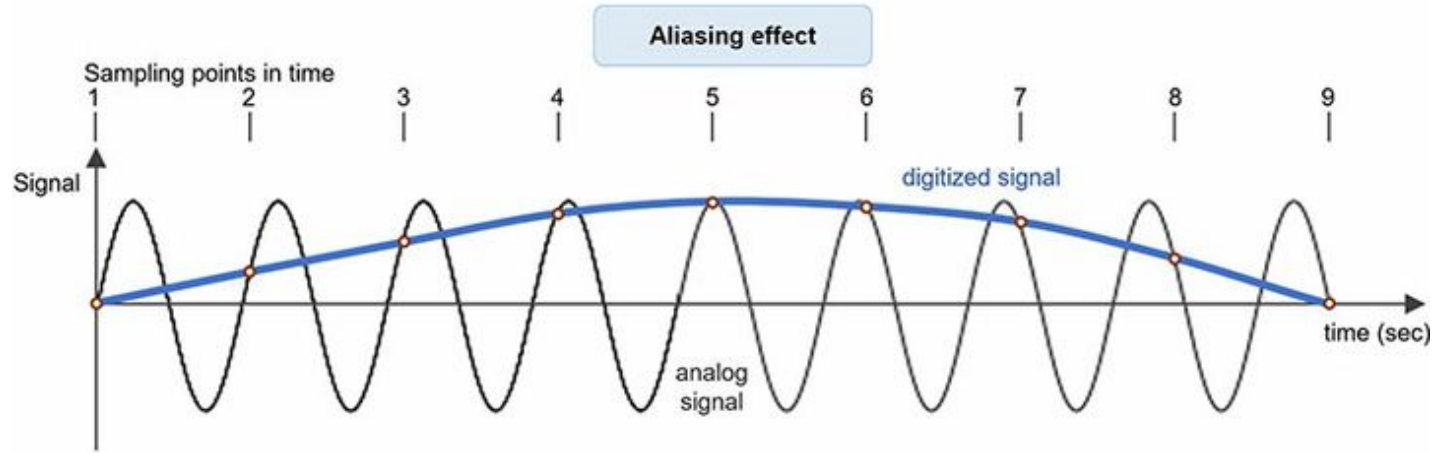


Figure credit: <https://www.kistler.com/sampling-rate-and-aliasing-effect>

Family

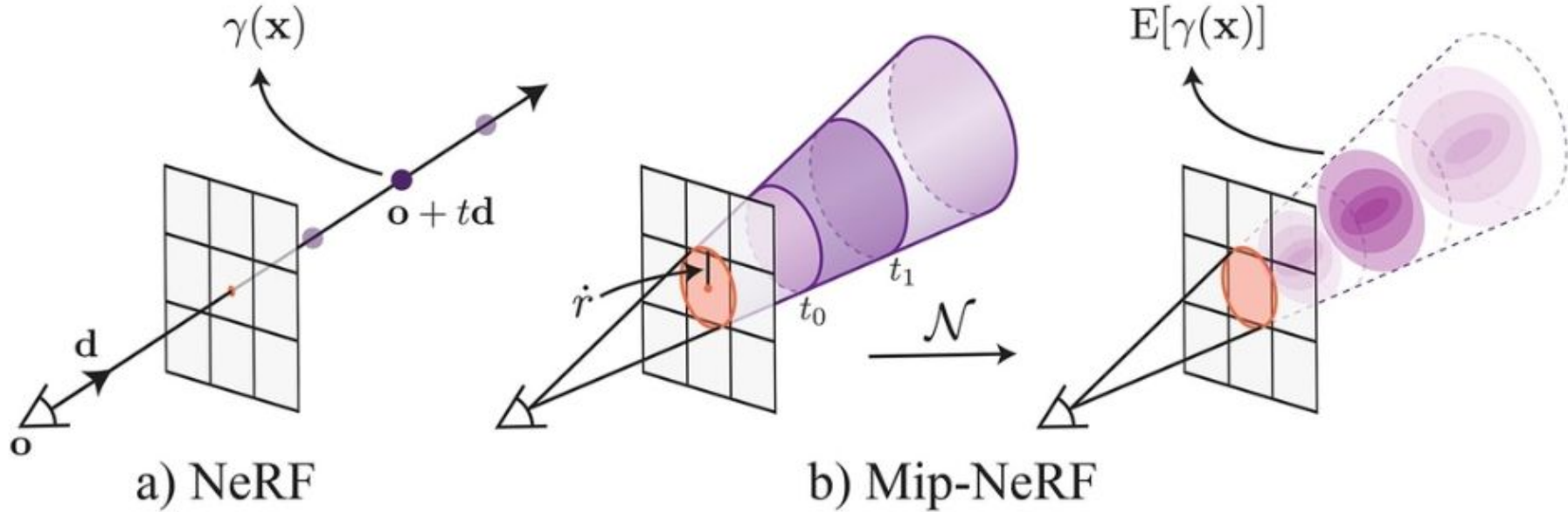
Multi-scale: Flickering problem

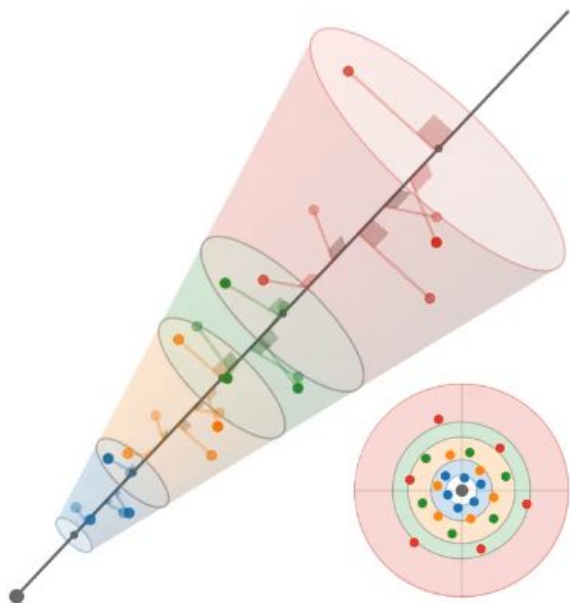


Family

Multi-scale: Aliasing or Flickering problem

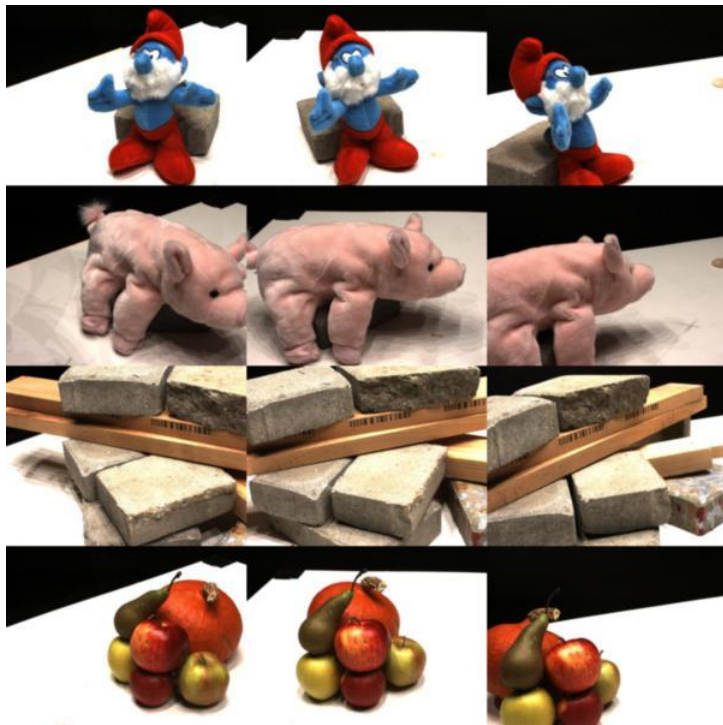






Family Generalizable

3 Input Views

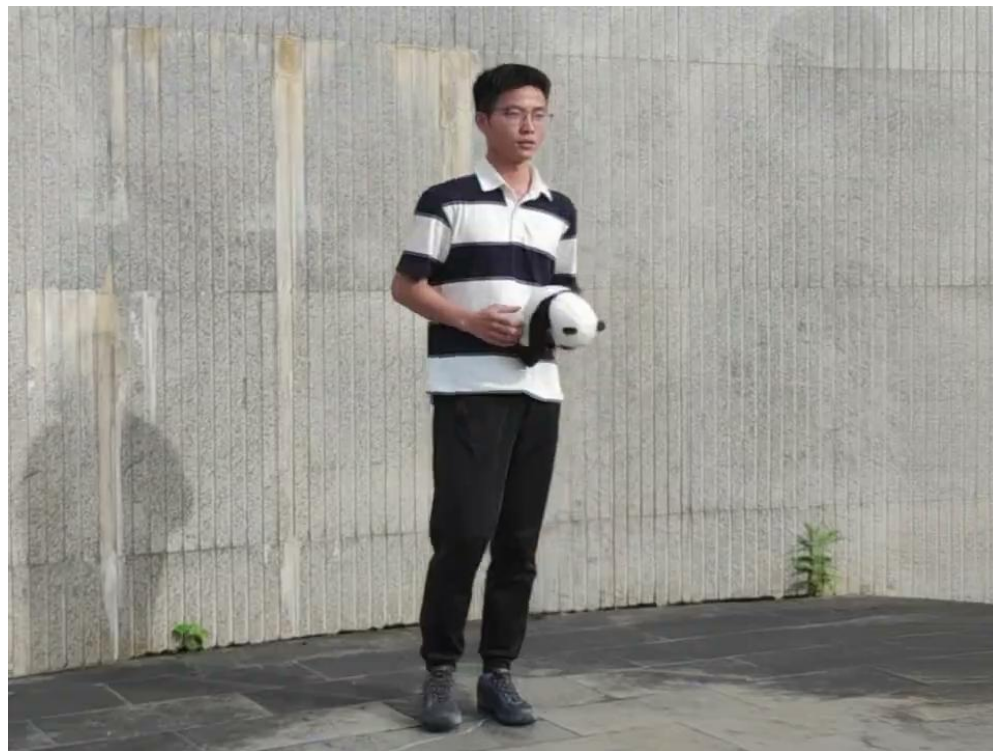
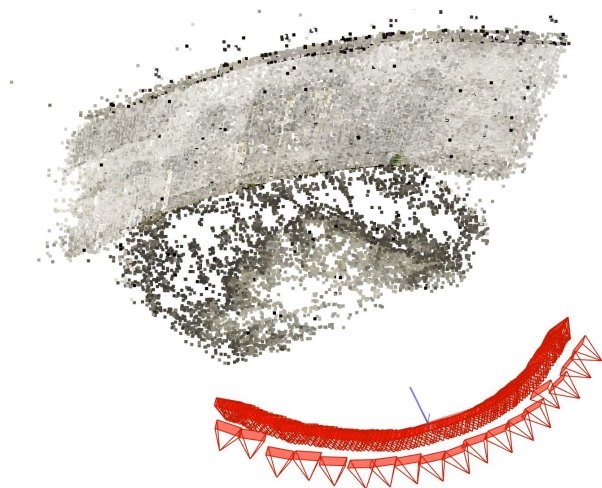


pixelNeRF



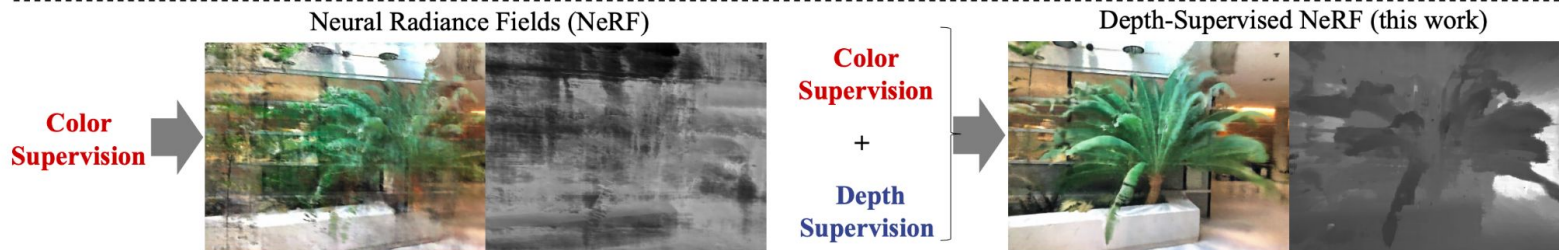
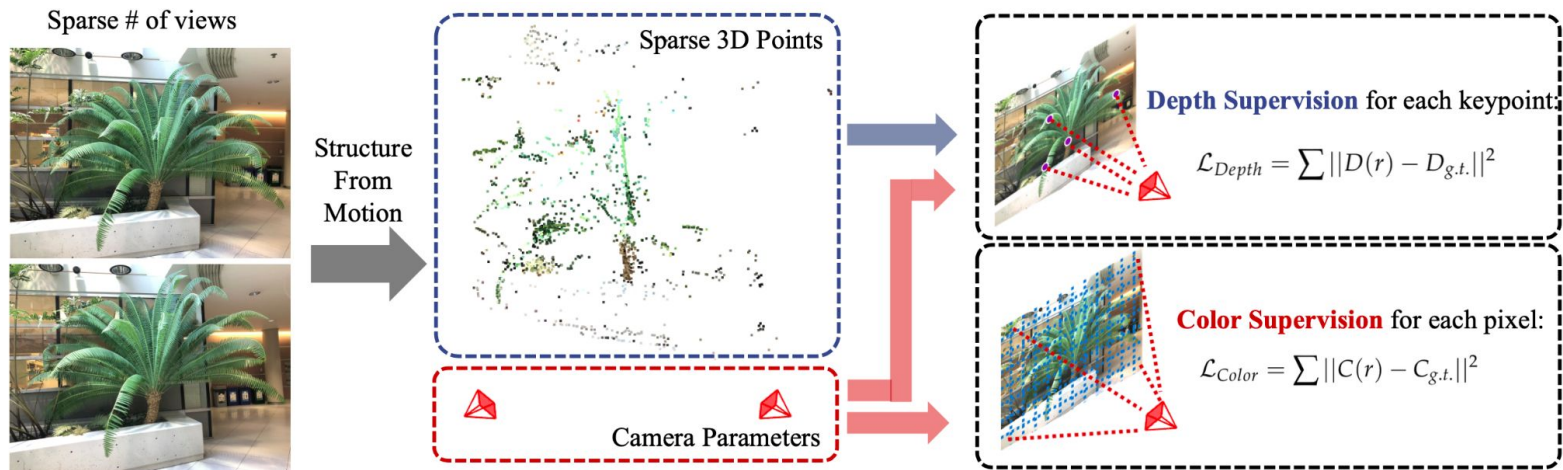
Family

Real-time Generalizable



Family

Depth input data



Family Depth input data

DSNeRF



NeRF



[6] [Depth-supervised nerf: Fewer views and faster training for free](#), K Deng, A Liu, JY Zhu... - CVPR 2022

Reference Sequences
from KITTI / Virtual KITTI 2



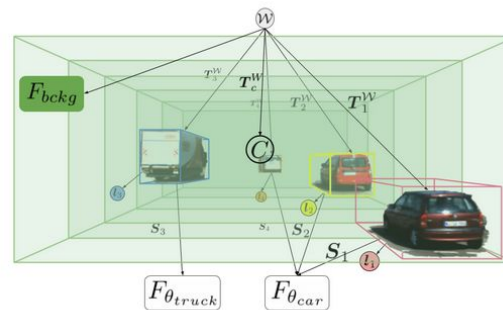
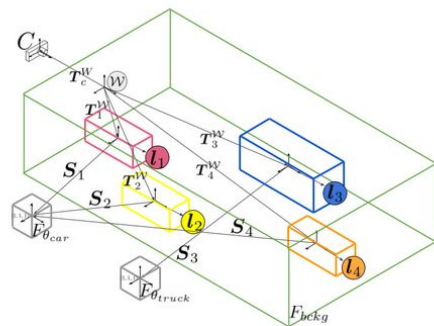
Decomposed Background



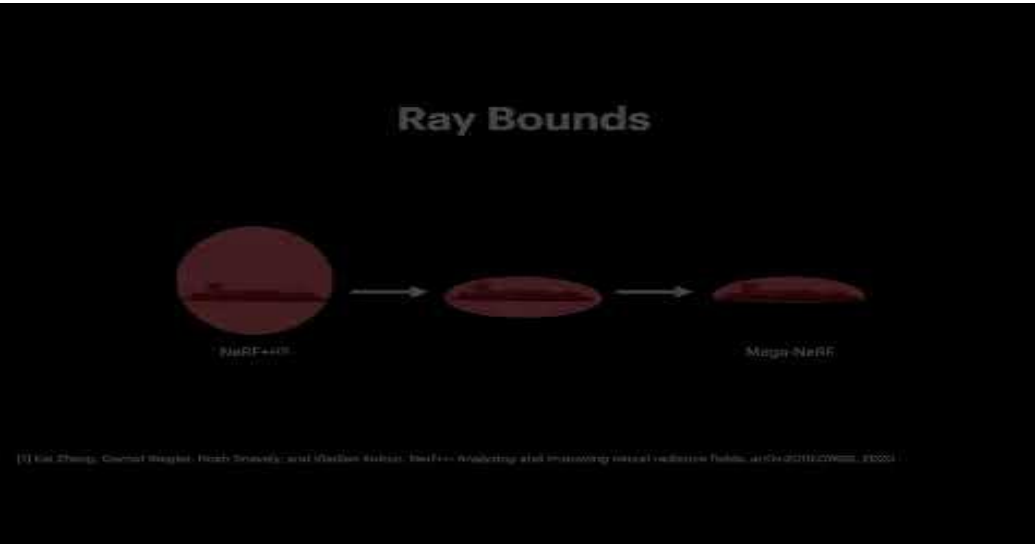
Decomposed Objects



Reconstructed Scene

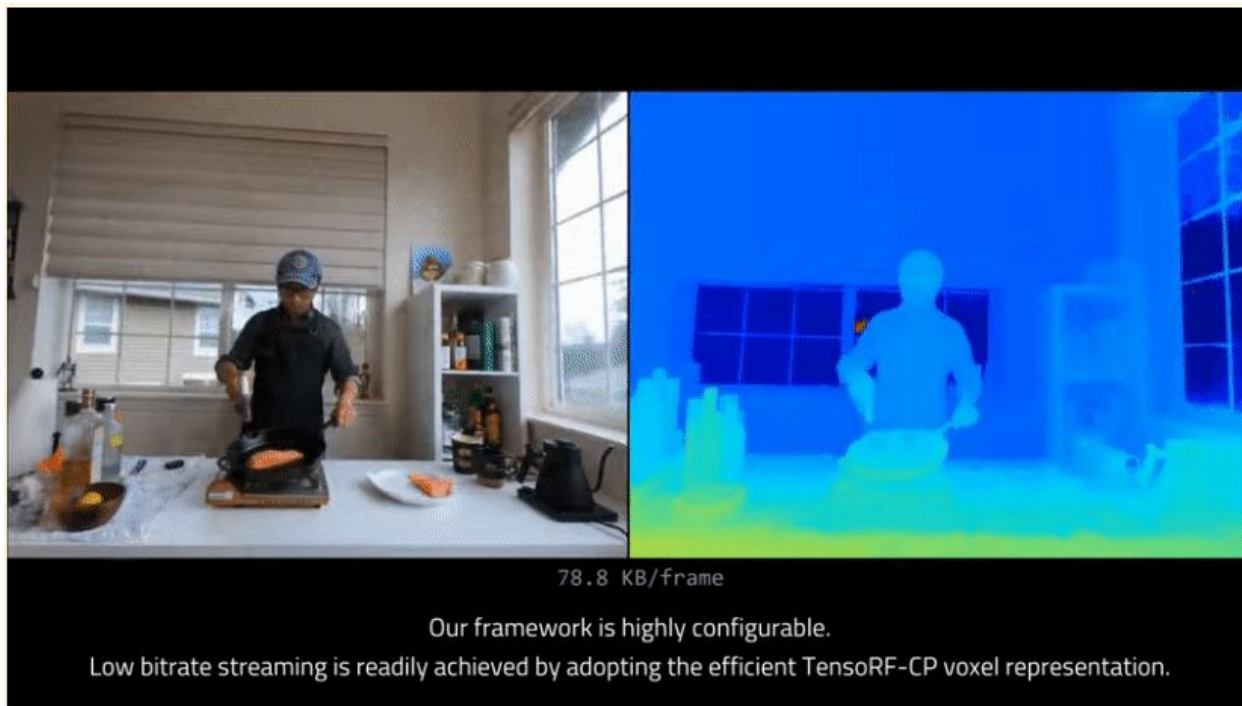






[1] [Mega-NeRF: Scalable Construction of Large-Scale NeRFs for Virtual Fly-Throughs](#), [H Turki](#), [D Ramanan](#)... - CVPR 2022

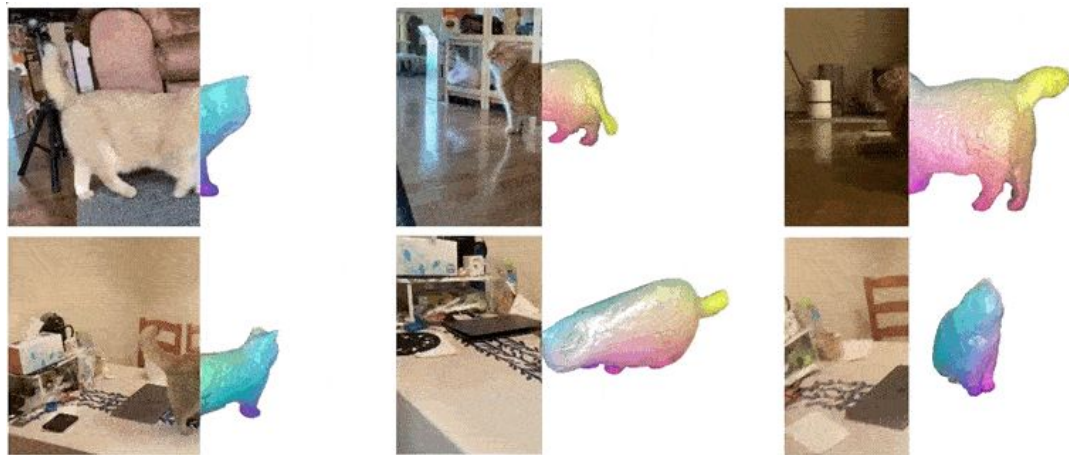
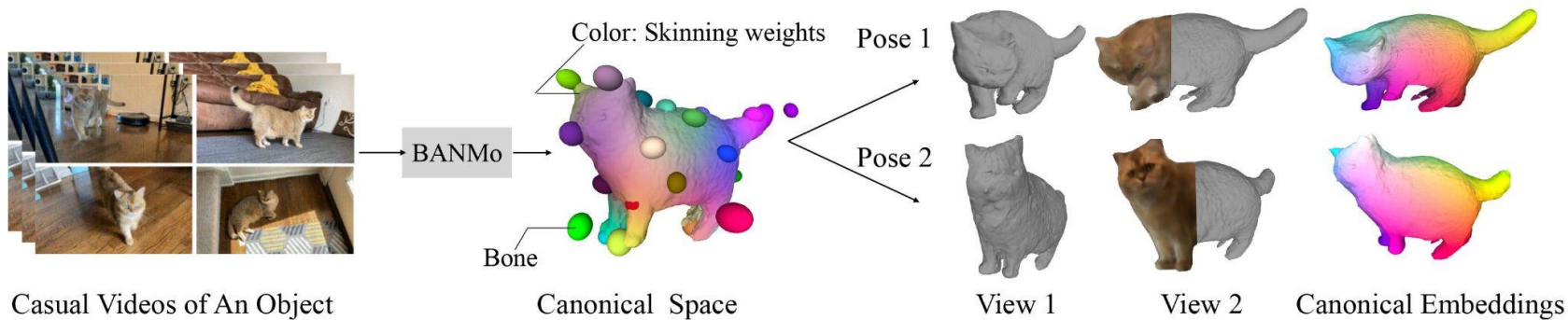
[2] [Block-nerf: Scalable large scene neural view synthesis](#), [M Tancik](#), [V Casser](#), [X Yan](#), [S Pradhan](#)... - CVPR 2022



[9] [Nerfplayer: A streamable dynamic scene representation with decomposed neural radiance fields](#), [L Song](#), [A Chen](#), [Z Li](#), [Z Chen](#), [L Chen](#)

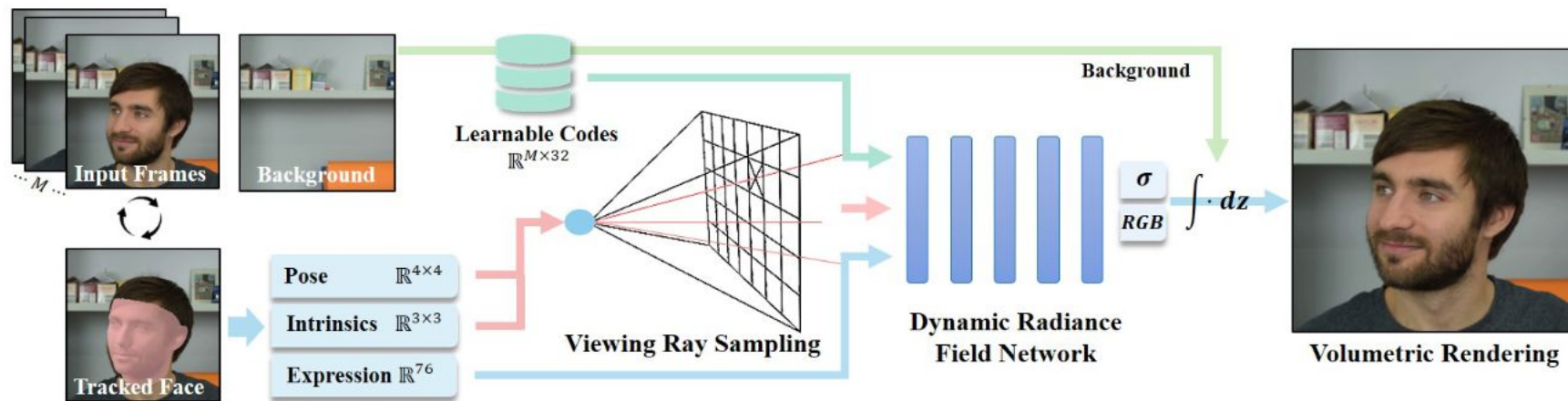
Family

Deformable (single object, complex motion)





Family Face Modelling

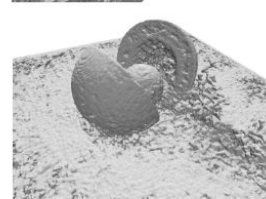
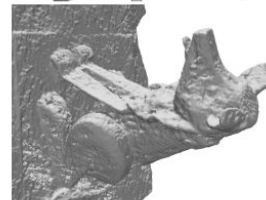
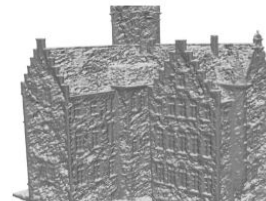


Family Surface Reconstruction



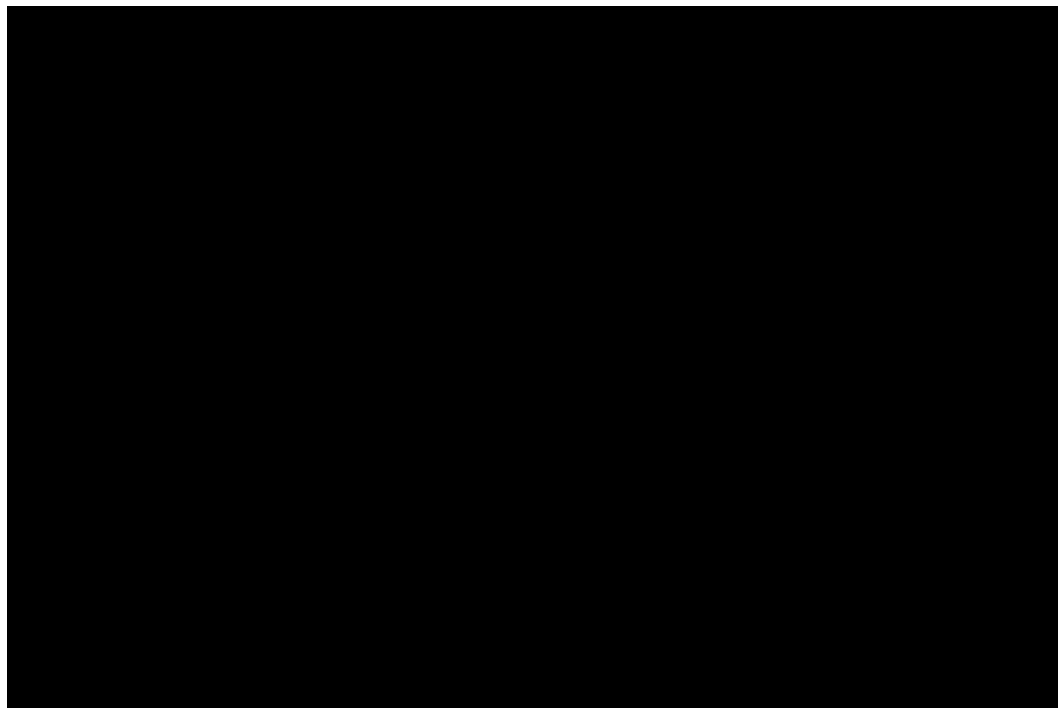
Reference Image

NeRF to Mesh



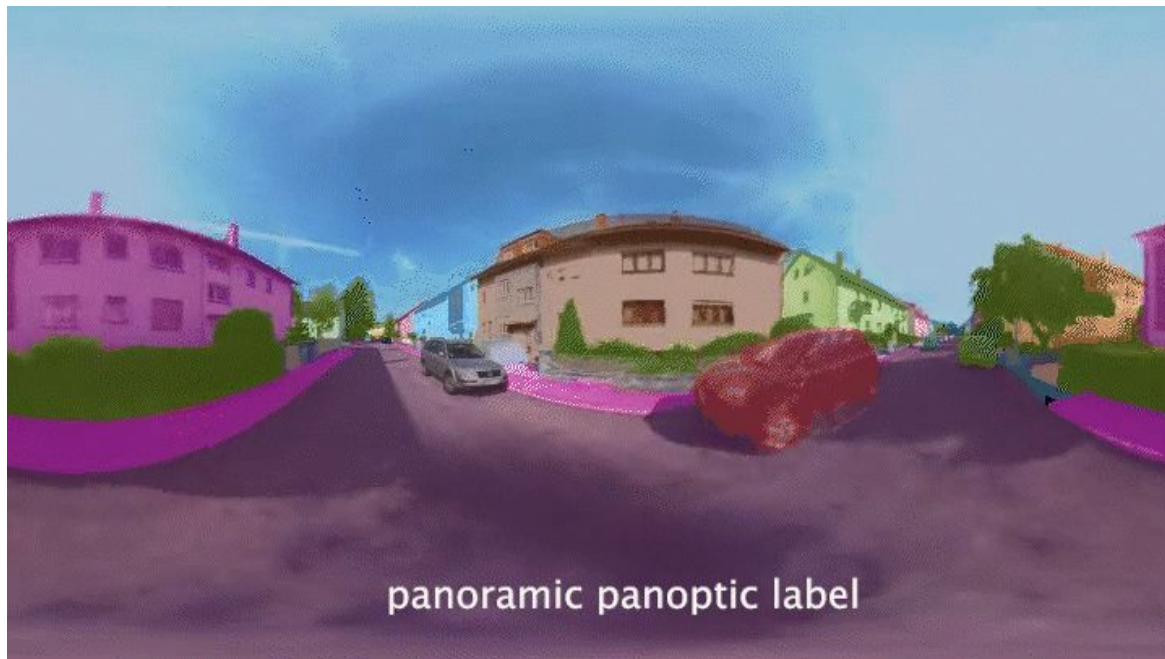
NeRF

Family Surface Reconstruction



[12] [Neuralangelo: High-Fidelity Neural Surface Reconstruction](#), [Z Li](#), [T Müller](#), [A Evans](#), [RH Taylor](#) - CVPR2023

Family Panoptic classification

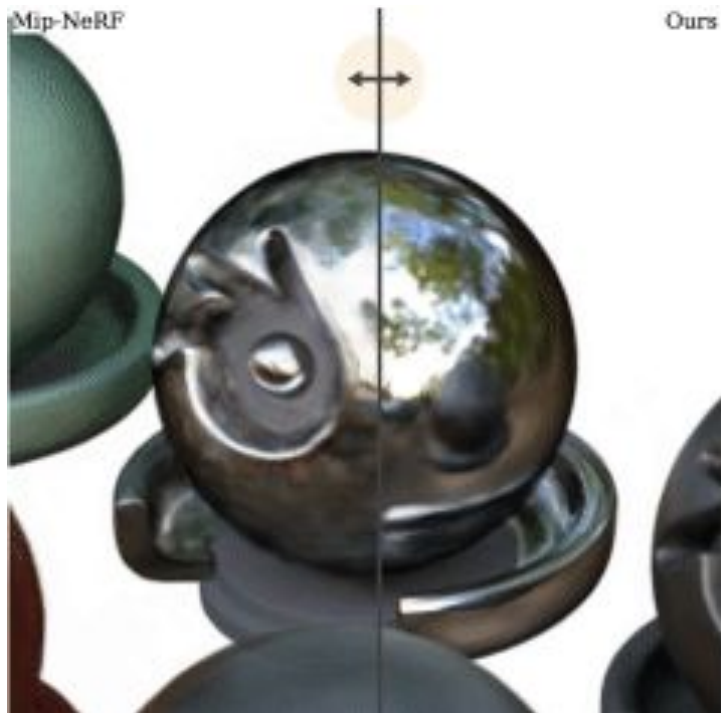


[14] [PanopticNeRF-360: Panoramic 3D-to-2D Label Transfer in Urban Scenes](#), X Fu, S Zhang, T Chen, Y Lu, X Zhou, A Geiger, Y Liao

Family

Reflection modeling and surface consistency

Reflections -> Foggy volumes



Texture and normals



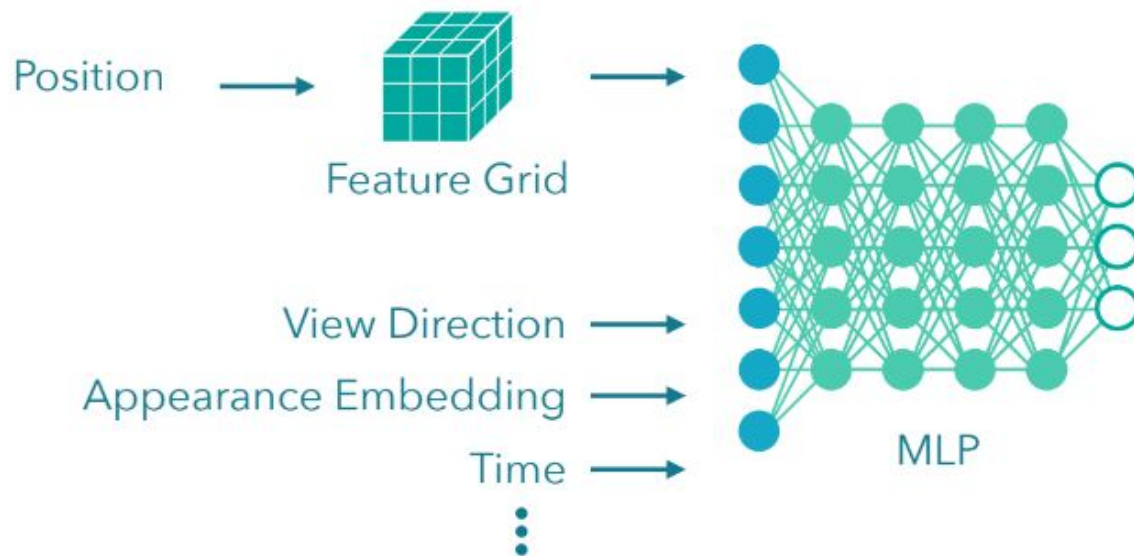
Family

Inverse rendering



[1] Neural Microfacet Fields for Inverse Rendering, [A Mai](#), [D Verbin](#), [F Kuester](#), [S Fridovich-Keil](#), ICCV 2023

Conclusions



- Talks
 - Jon Barron - Understanding and Extending Neural Radiance Field, 2020 [[Talk at MIT](#)][[Talk at TUM](#)]
 - Matthew Tancik: Neural Radiance Fields for View Synthesis, 2020 [[Talk at Tübingen](#)]
- Tutorials
 - [CVPR2022: Neural Fields in Computer Vision](#)
 - [ECCV2022: Neural Volumetric Rendering for Computer Vision](#)
 - [SIGGRAPH 2021 Course: Advances in Neural Rendering](#)
 - [CVPR 2020: Tutorial on Neural Rendering](#)
- Compilations of papers
 - <https://github.com/awesome-NeRF/awesome-NeRF>
 - <https://paperswithcode.com/task/novel-view-synthesis>
- Libraries
 - [NeRF Studio](#)

CIDAI

Centre of Innovation
for Data tech
and Artificial Intelligence



www.cidai.eu