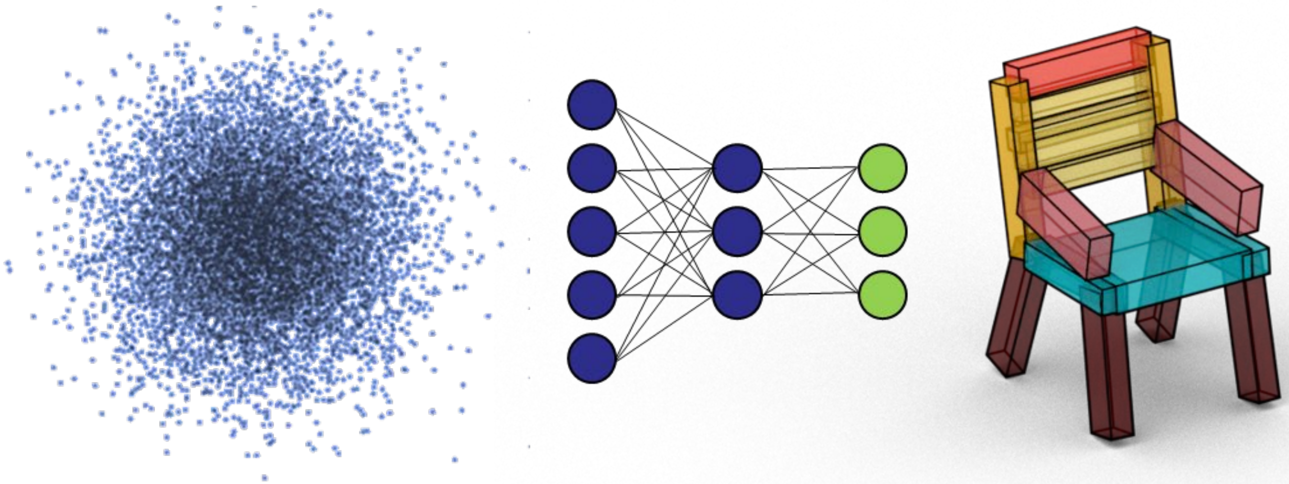
CS348n: Neural Representations and Generative Models for 3D Geometry



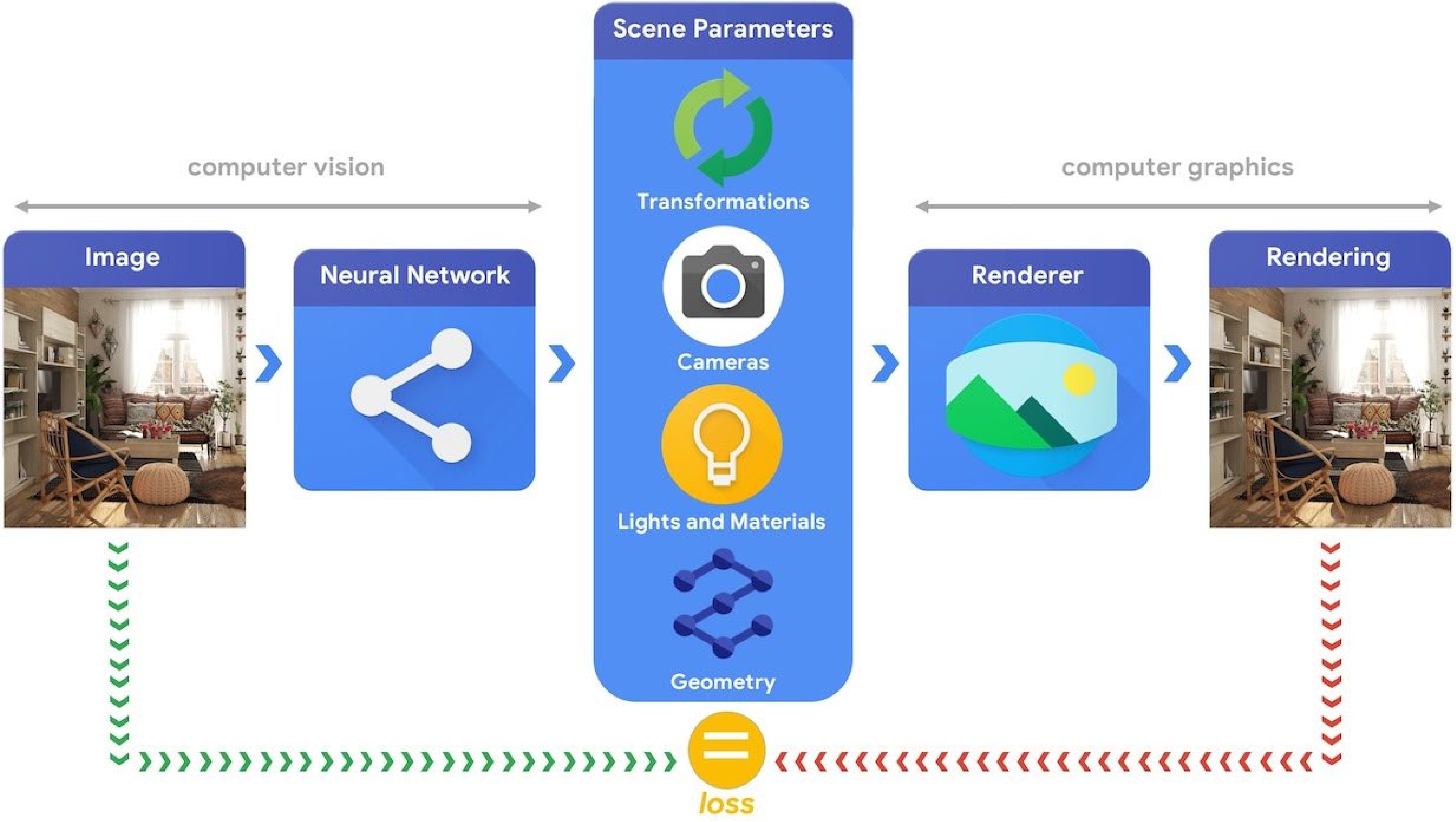
Leonidas Guibas
Computer Science Department
Stanford University

Leonidas Guibas Laboratory

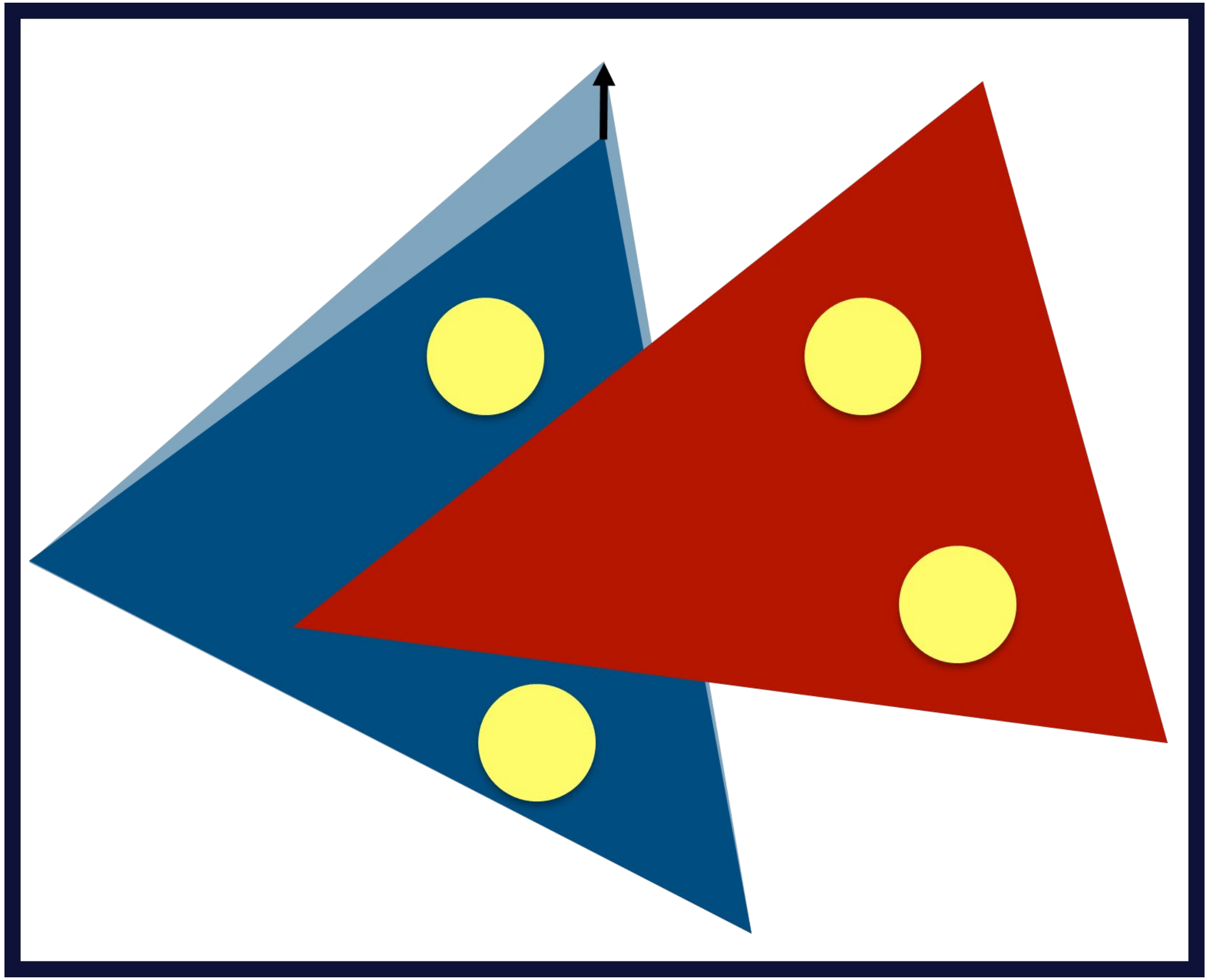
Geometric Computing

Last Time:
Neural Rendering and NeRFs

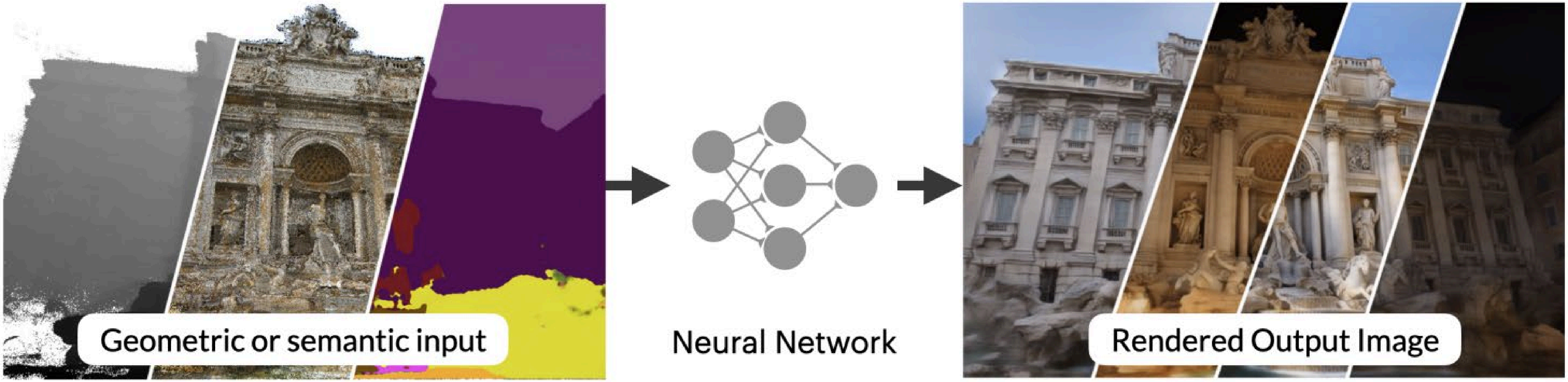
Computer Vision as Inverse Graphics



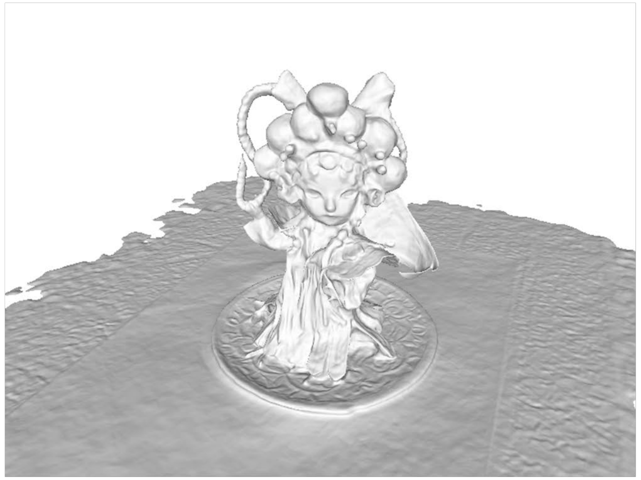
Classical Rendering not Differentiable



Neural Differentiable Rendering



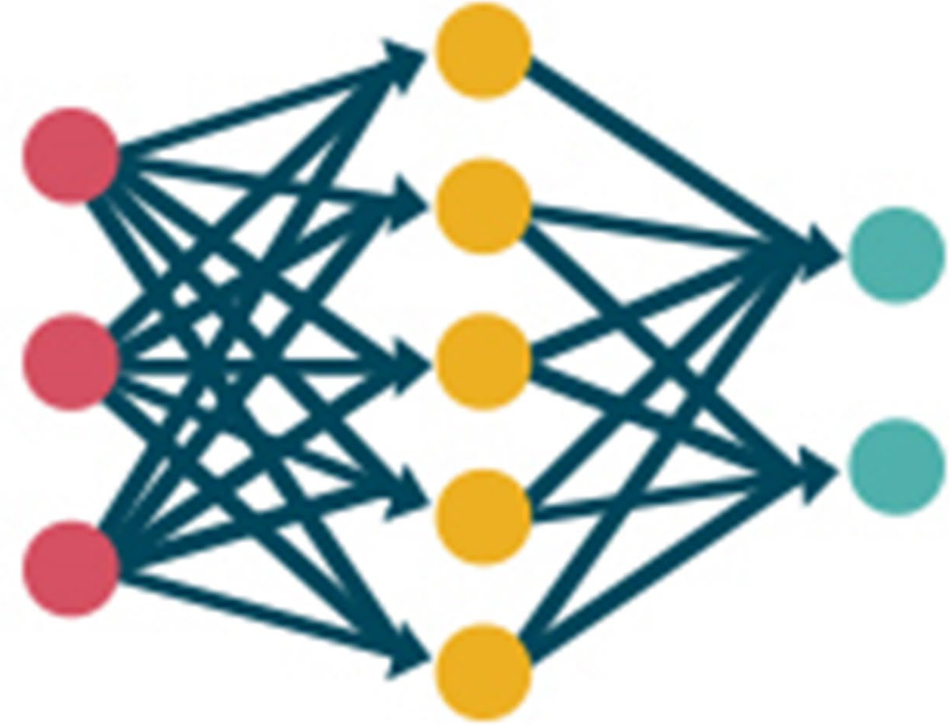
Learn 3D for 2D Supervision Alone



3D surface reconstruction

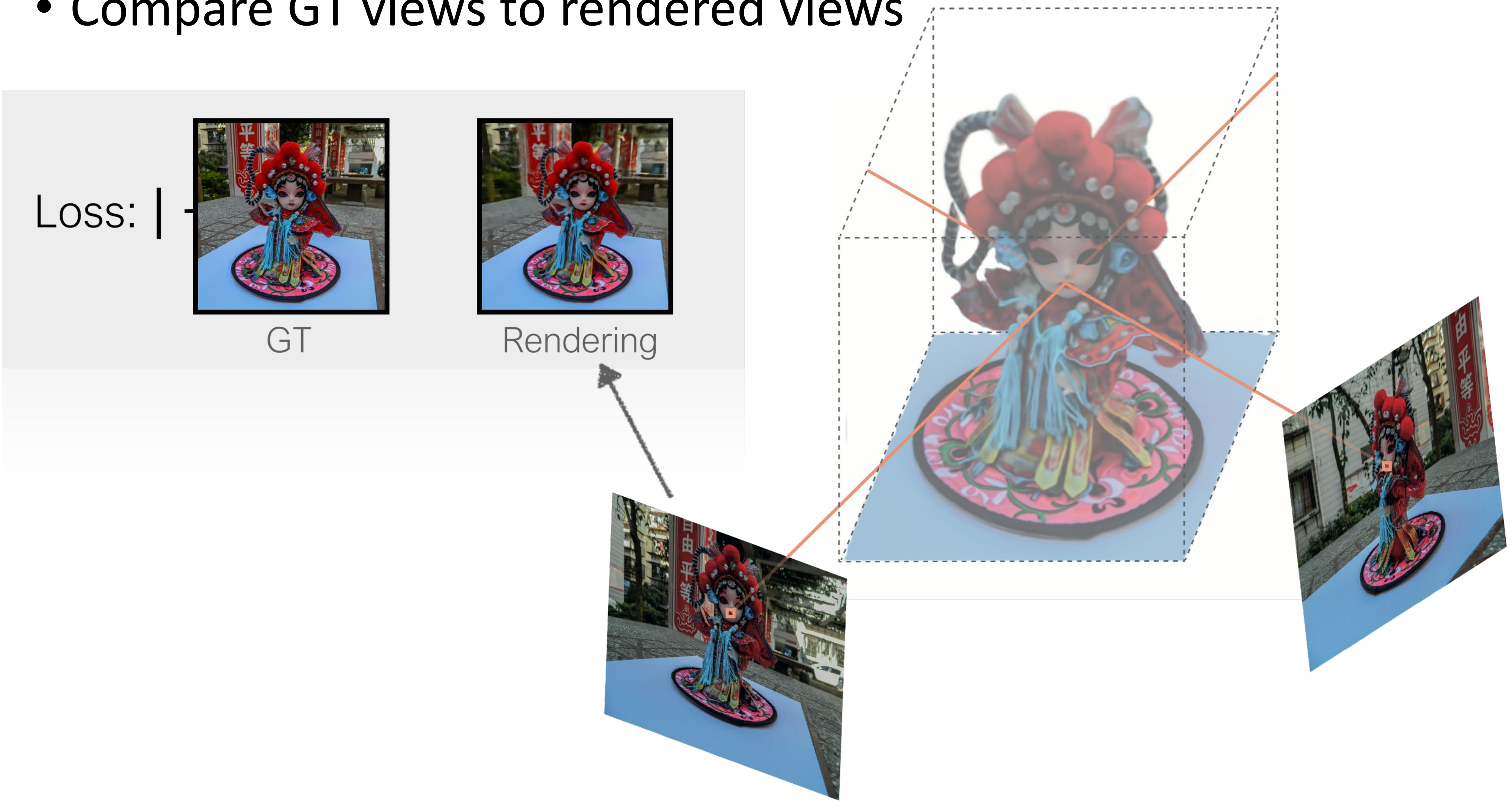
Novel views synthesis

NeRF: Encode a Scene into an MLP

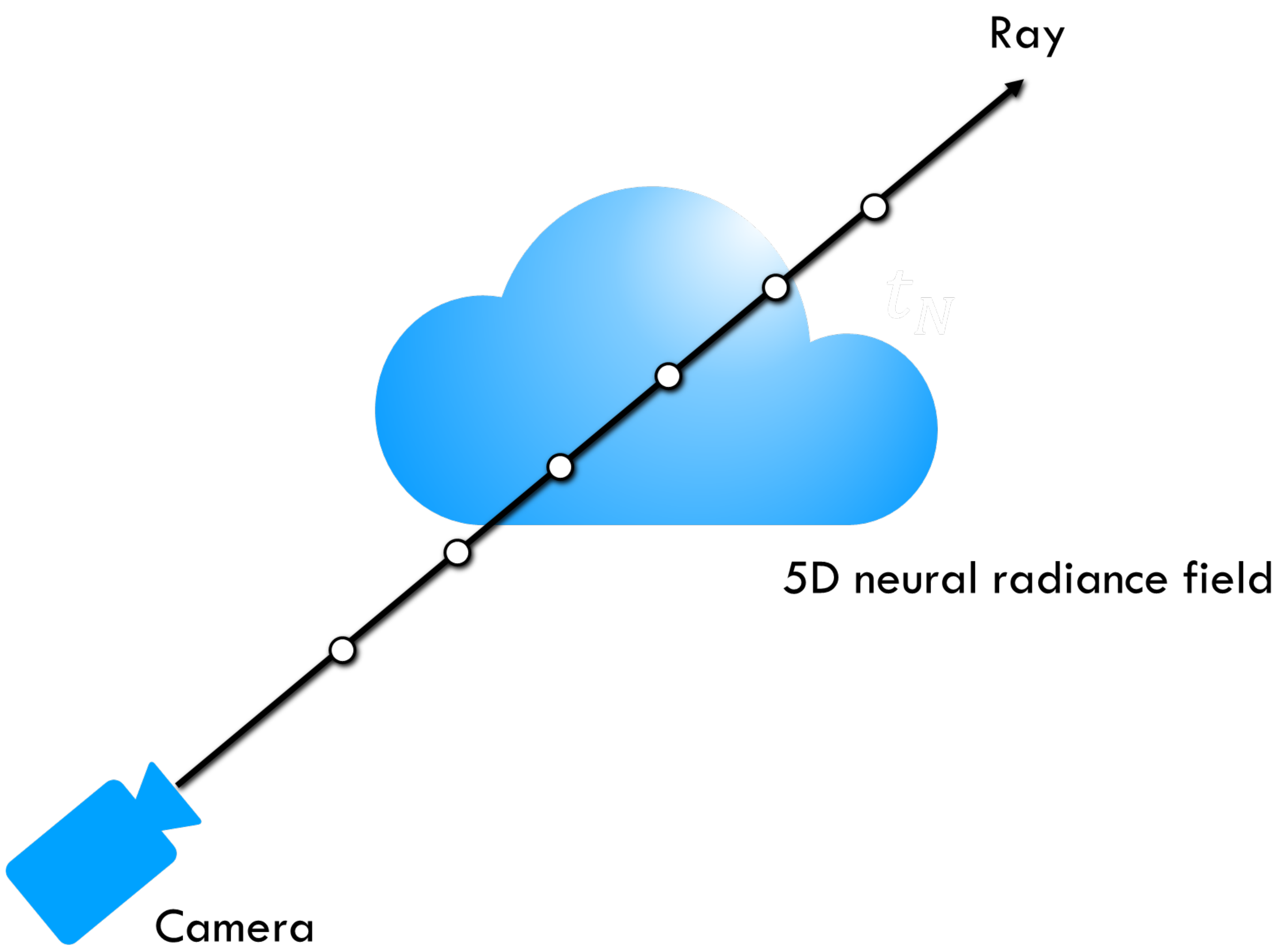


Neural Differentiable Rendering

- Compare GT views to rendered views



Volume Rendering



Volume Rendering by Integration

Rendering model for ray $\mathbf{r}(t) = \mathbf{o} + t\mathbf{d}$:

$$\mathbf{c} \approx \sum_{i=1}^{n} T_i \alpha_i \mathbf{c}_i$$

weights

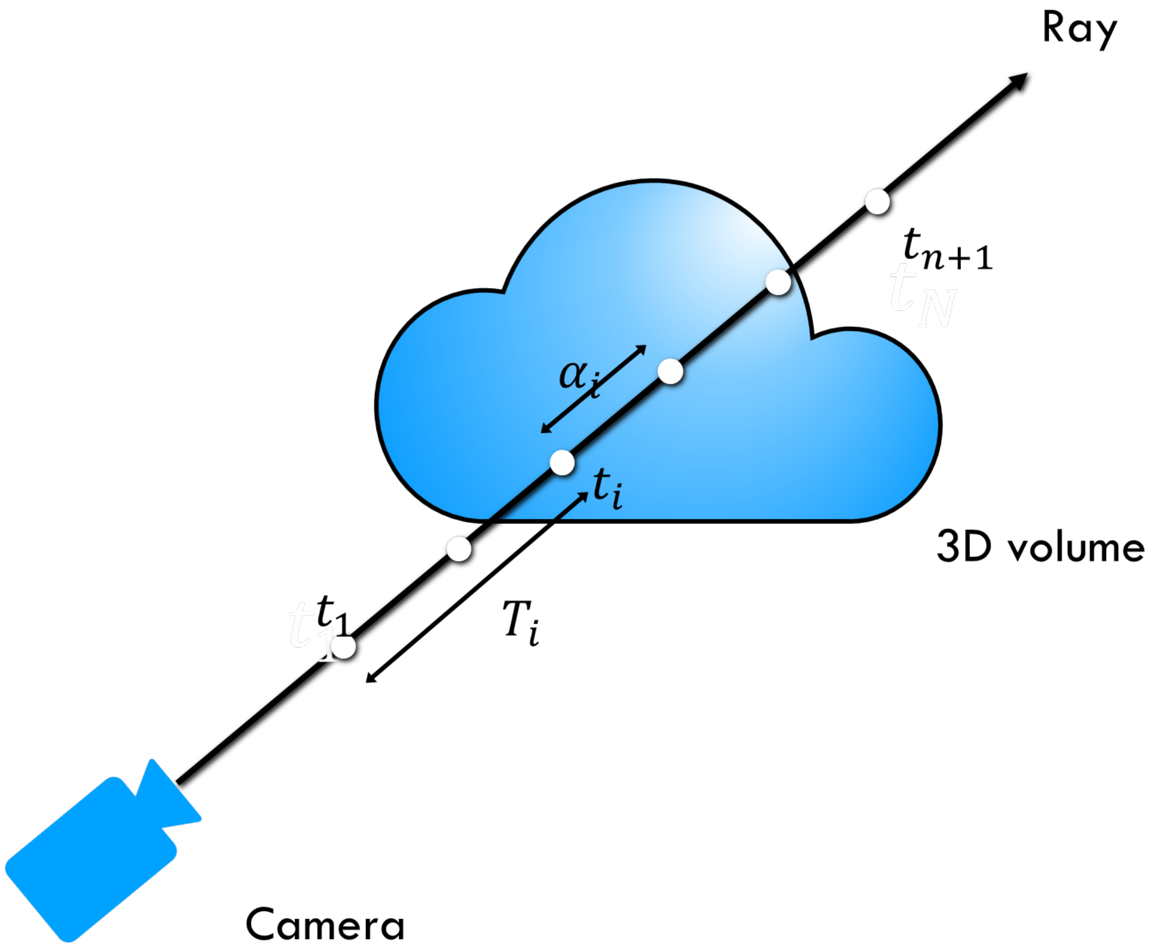
colors

How much light is blocked earlier along ray:

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

How much light is contributed by ray segment *i*:

$$\alpha_i = 1 - \exp(-\sigma_i \delta_i)$$



Scene Representation as a Continuous 5D Function

$$(\underbrace{x, y, z}_{\text{Spatial location}}, \underbrace{\theta, \phi}_{\text{Viewing direction}}) \rightarrow F_{\Omega} \rightarrow (\underbrace{r, g, b}_{\text{Output color}}, \underbrace{\sigma}_{\text{Output density}})$$

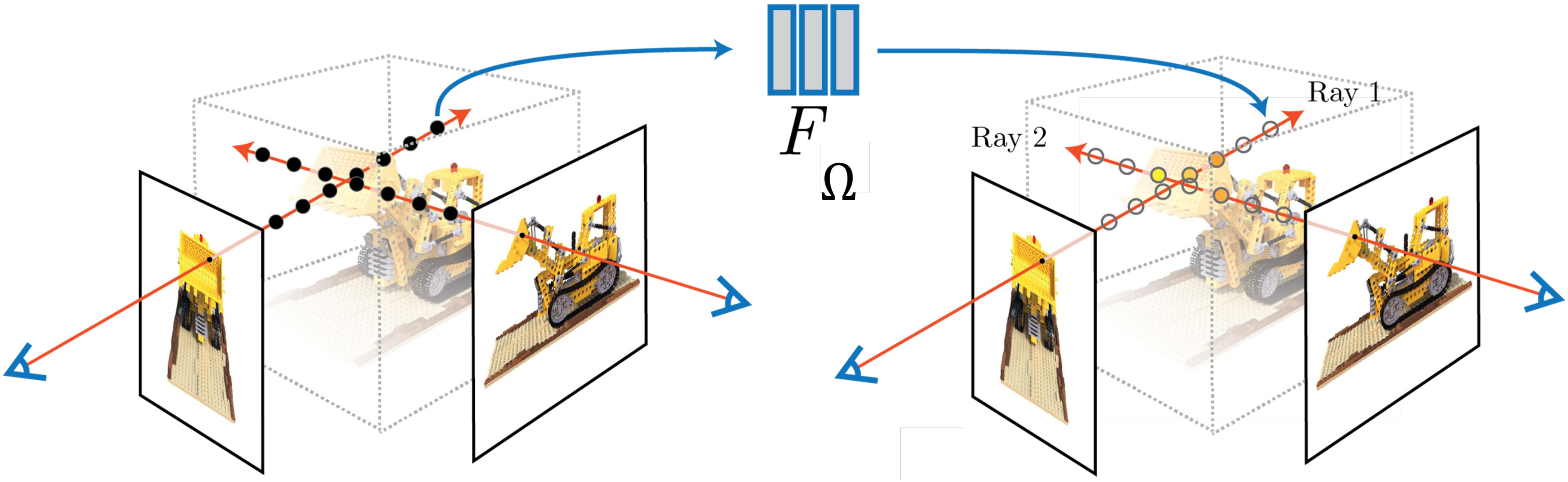
F_{Ω}

Fully-connected neural network

12 layers,

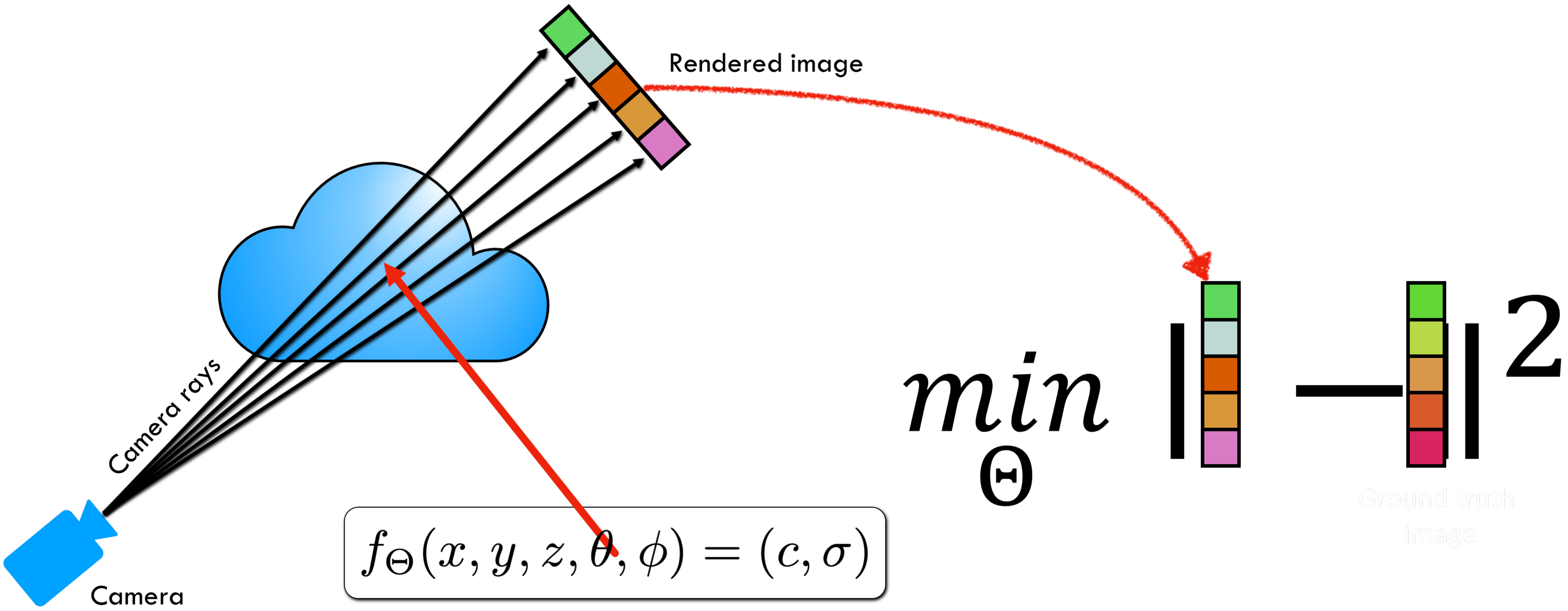
256 channels

Network Optimization



$$\min_{\Omega}\sum_{i} \| \text{render}^{(i)}(F_\Omega) - I_{\text{gt}}^{(i)} \|^2$$

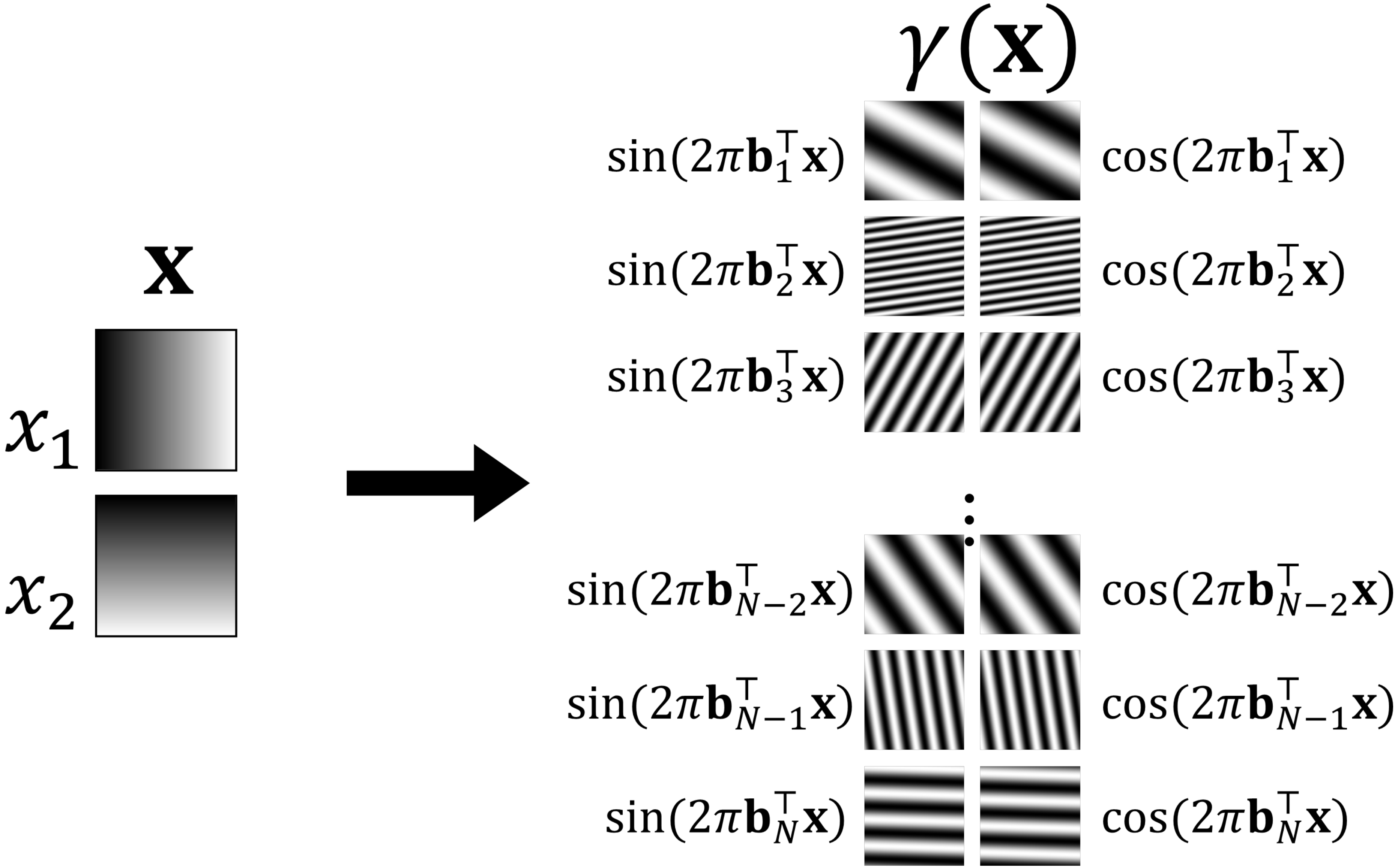
Network Optimization



1. Render volume from observed viewpoint

2. Optimize image loss to match observed image

Fourier Positional Encodings



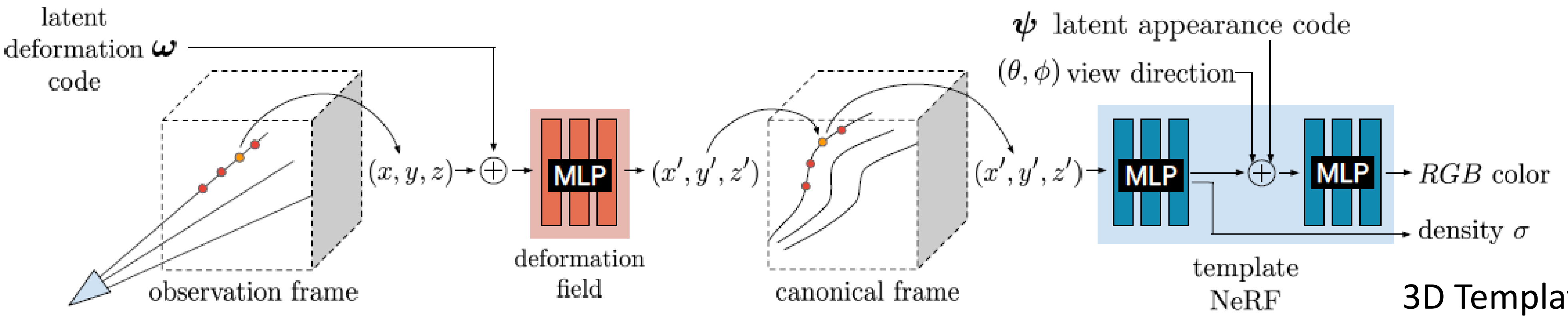
SRN [Sitzmann et al. 2019]

NeRF

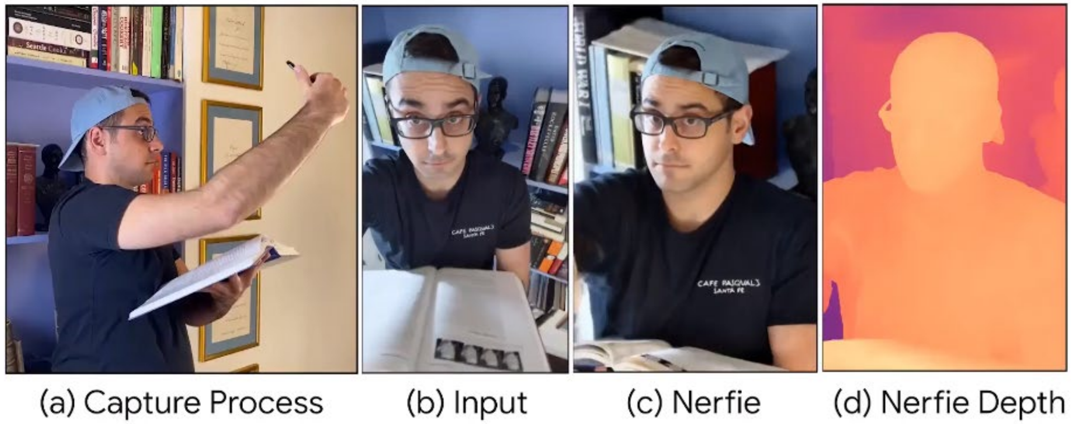
Nearest Input

NeRF Variations

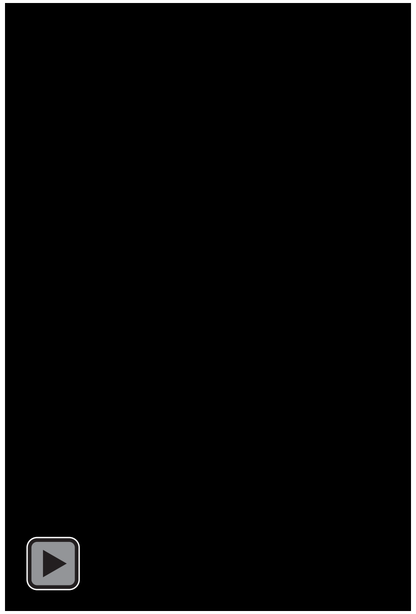
NeRFIes – NeRFs for Non-Rigid Objects



3D Template



NERFIES turns selfie videos from your phone into free-viewpoint portraits.



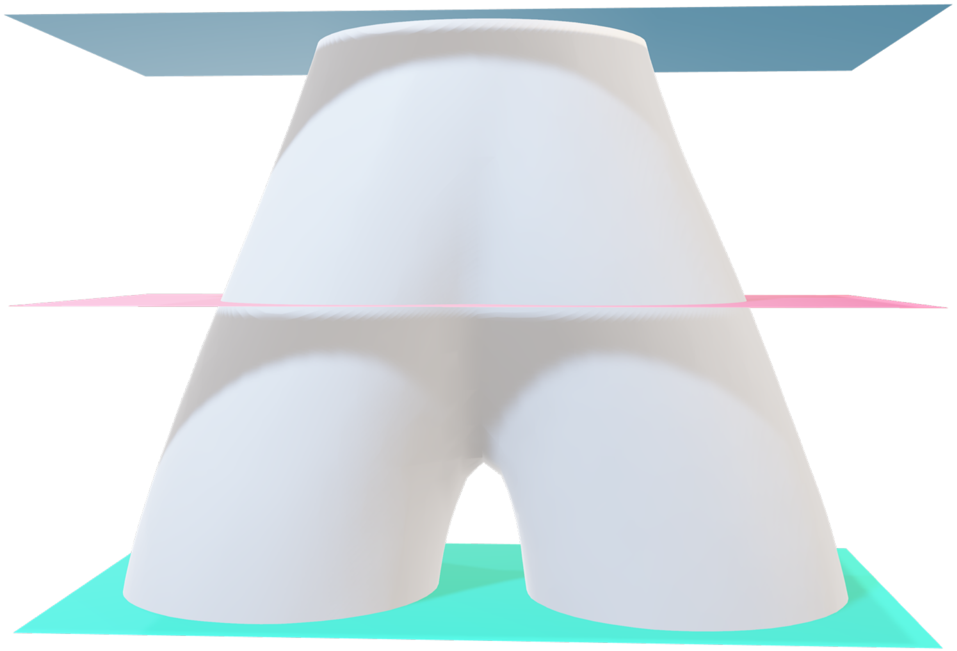
Park, K., Sinha, U., Barron, J.T., Bouaziz, S., Goldman, D.B., Seitz, S.M. and Martin-Brualla, R. **Nerfies: Deformable neural radiance fields**. ICCV 2021.

HyperNeRF – NeRFs for Non-Rigid Objects

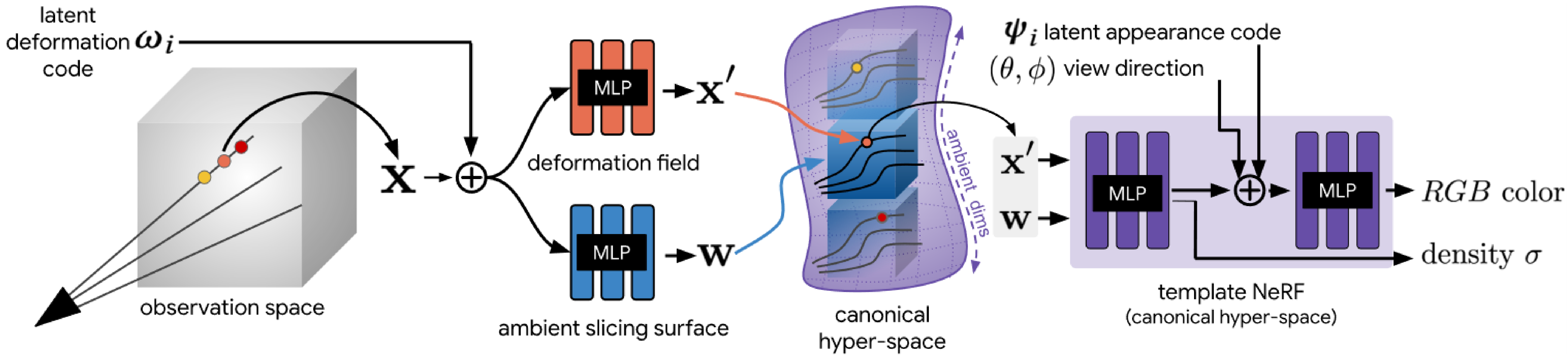
(a) Input Images (b) Nerfies $(\mathbf{c}, \sigma) = F(x, y, z)$ (c) HyperNeRF $(\mathbf{c}, \sigma) = F(x, y, z, w_1, w_2)$ (d) Hyper-Space

Park, K., Sinha, U., Hedman, P., Barron, J.T., Bouaziz, S., Goldman, D.B., Martin-Brualla, R. and Seitz, S.M., 2021. **HyperNeRF: A higher-dimensional representation for topologically varying neural radiance fields**. ACM Trans. Graph., 2021.

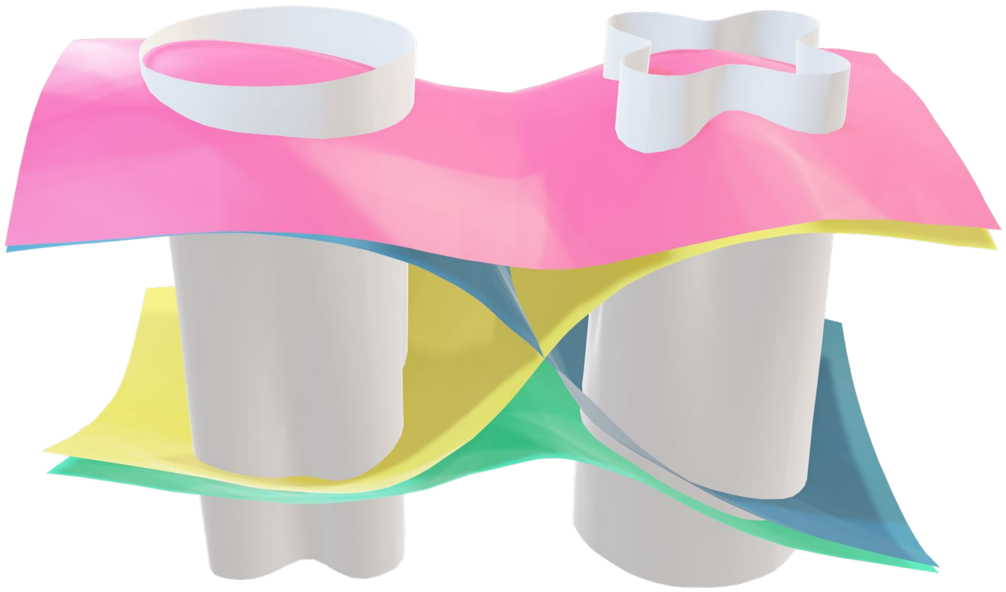
Level Sets for Topology Changes



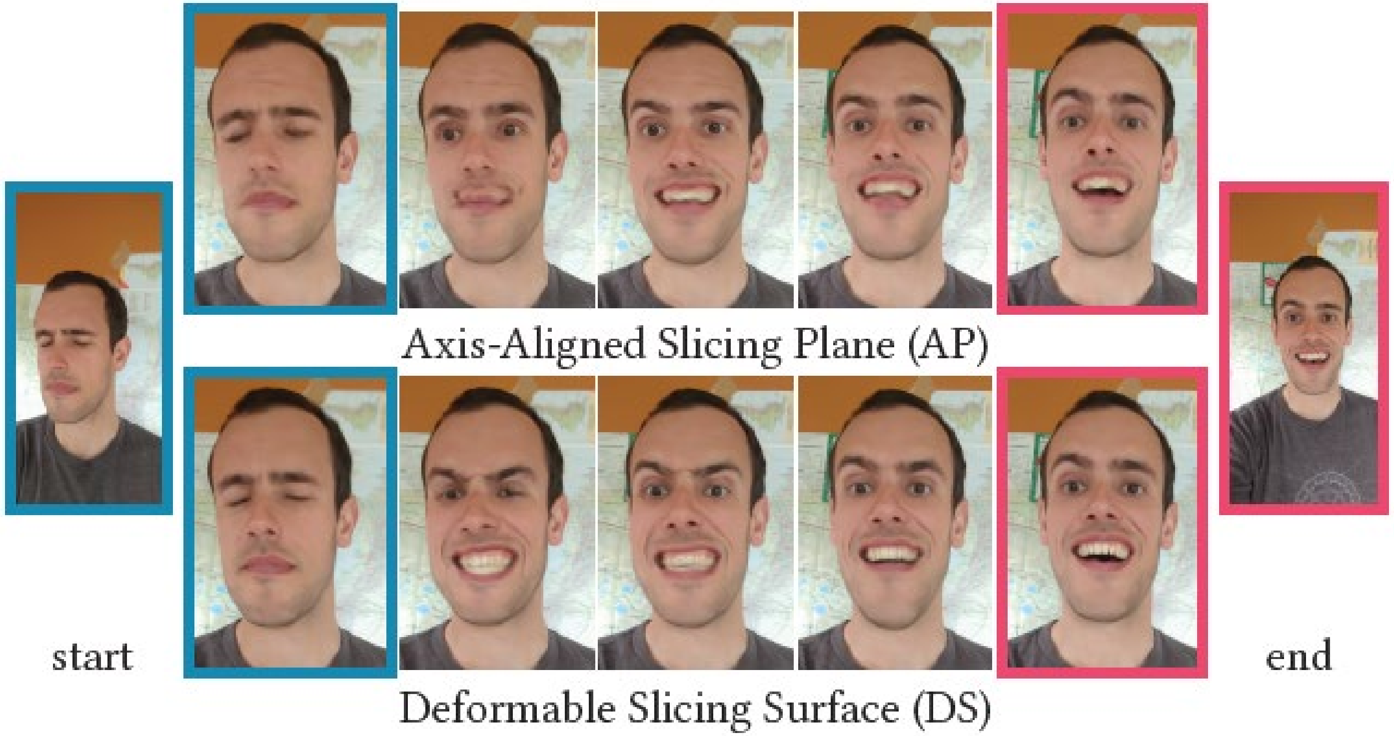
Hype-Space Template



Non-Level Slices for Topology Changes



Deformable Slicing Surface



Geometry and Appearance from Sparse Images

Surface Rendering vs Volume Rendering

Surface Rendering:

- Implicit surface
- Find intersection
- Limitation: objects mask

SRNs [Sitzmann et al. '19]

DVR [Niemeyer et al. '20]

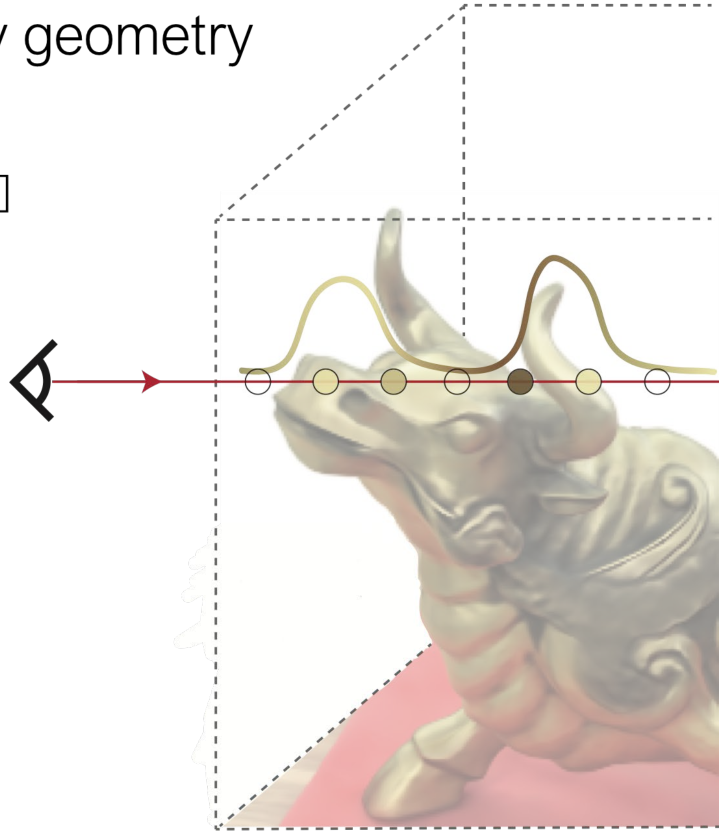
IDR [Yariv et al. '20]

Volume Rendering:

- Volume density
- Integral approximation
- Limitation: noisy geometry

NeRF [Mildenhall et al. '20]

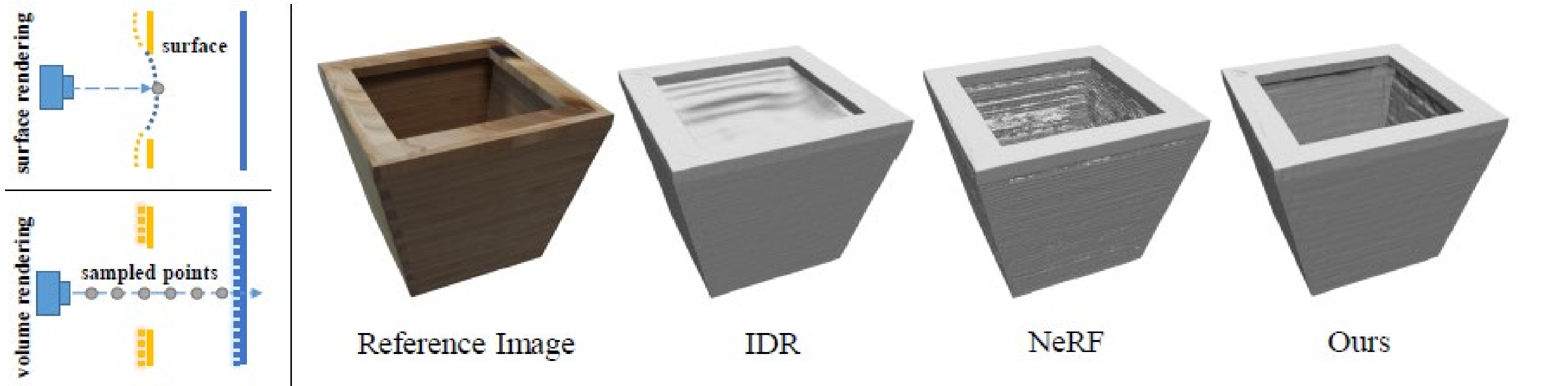
...



Combining Surface and Volume Rendering

NeuS: Learning Neural Implicit Surfaces by Volume Rendering for Multi-view Reconstruction

Peng Wang[†], Lingjie Liu[‡], Yuan Liu[†], Christian Theobalt[‡], Taku Komura[†], Wenping Wang[◇]
[†]The University of Hong Kong [‡]Max Planck Institute for Informatics
[◇]Texas A&M University
[†]{pwang3,yliu,taku}@cs.hku.hk [‡]{lliu,theobalt}@mpi-inf.mpg.de
[◇]wenping@tamu.edu



(a) Illustration (b) Example

NeurIPS 2021

NeuS: deepSDF + NeRF Combined

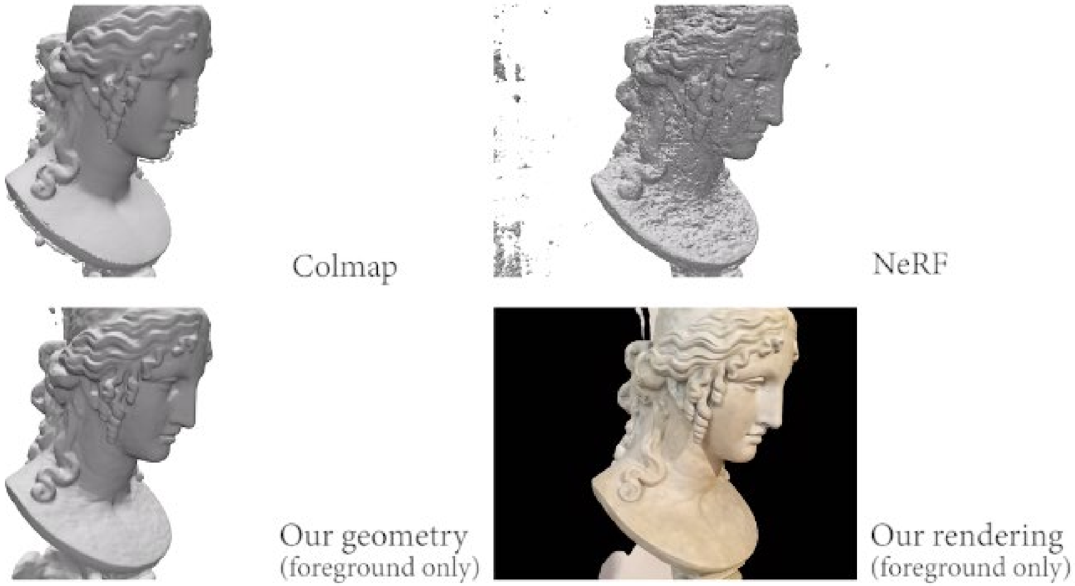
- Volume rendering using SDF data in lieu of a volume density
- Derive density σ from SDF function f: $\sigma(x) = \phi_s(f(\mathrm{x}))$

$$C(\mathbf{o}, \mathbf{v}) = \int_0^{+\infty} w(t) c(\mathbf{p}(t), \mathbf{v}) \mathrm{d}t$$

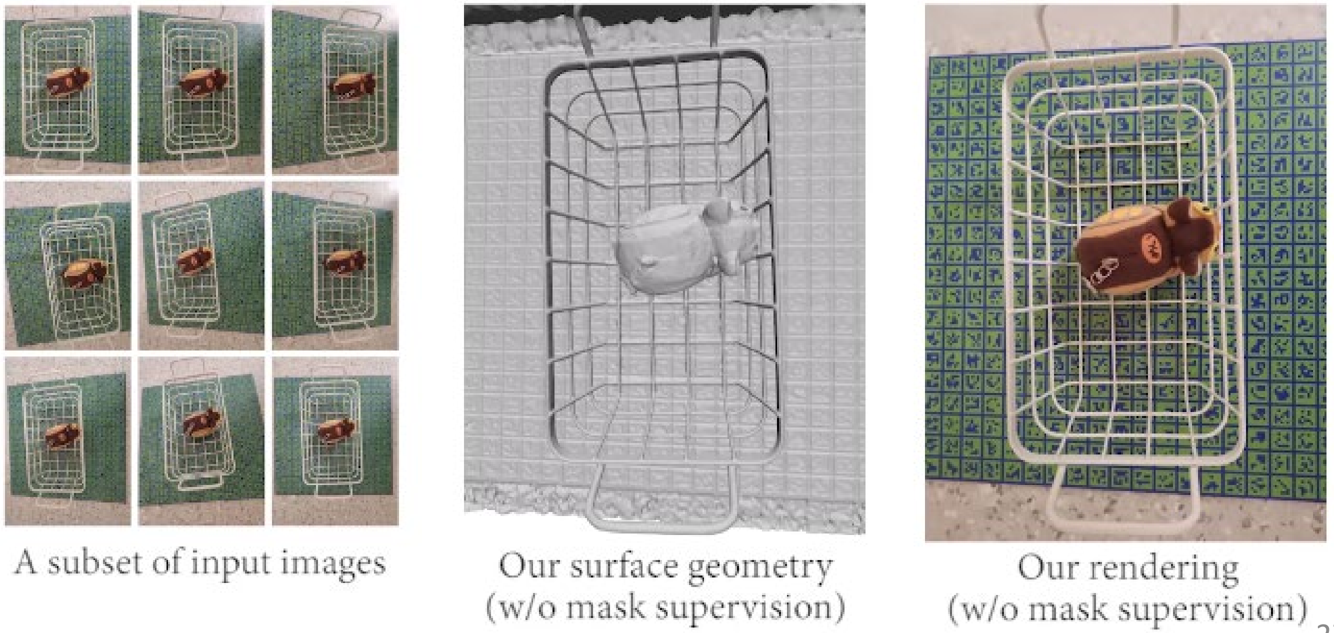
$$w(t) = \frac{\phi_s(f(\mathbf{p}(t)))}{\int_0^{+\infty} \phi_s(f(\mathbf{p}(u))) \mathrm{d}u}$$

- Allows SDF extraction from images *without object masks*

NeuS Results



Statue from the BlendedMVS dataset



VolSDF: Another Variant

Volume Rendering of Neural Implicit Surfaces

Lior Yariv[1] Jiatao Gu[2] Yoni Kasten[1] Yaron Lipman[1,2]

[1]Weizmann Institute of Science [2]Facebook AI Research

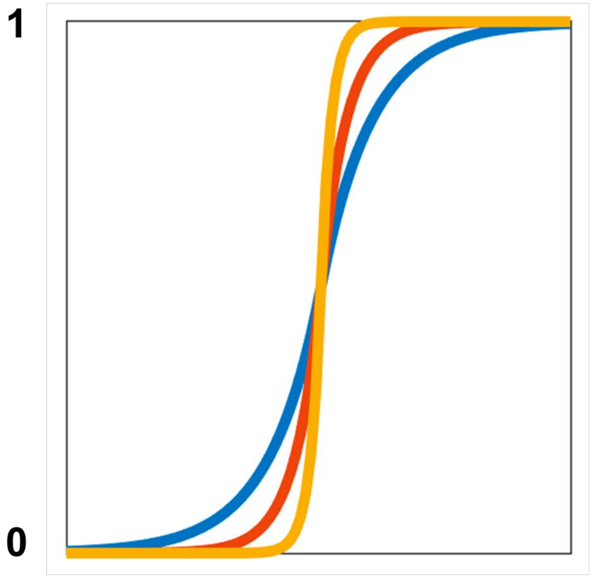
Figure 1: VolSDF: given a set of input images (left) we learn a volumetric density (center-left, sliced) defined by a signed distance function (center-right, sliced) to produce a neural rendering (right). This definition of density facilitates high quality geometry reconstruction (gray surfaces, middle).

NeurIPS 2021

VolSDF: Density σ as Transformed SDF

$$\sigma = \alpha\Phi_{\beta}(-d)$$

Density



Laplace CDF

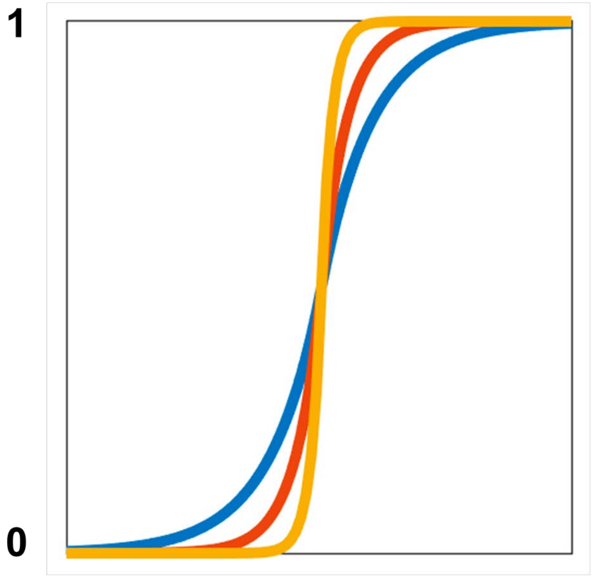
Signed Distance Function

VolSDF: Density σ as Transformed SDF

$$\Phi_\beta(s) = \begin{cases} \frac{1}{2}\exp\left(\frac{s}{\beta}\right) & \text{if } s \leq 0 \\ 1 - \frac{1}{2}\exp\left(-\frac{s}{\beta}\right) & \text{if } s > 0 \end{cases}$$

Cumulative distribution function of Laplace distribution with 0 mean and scale β

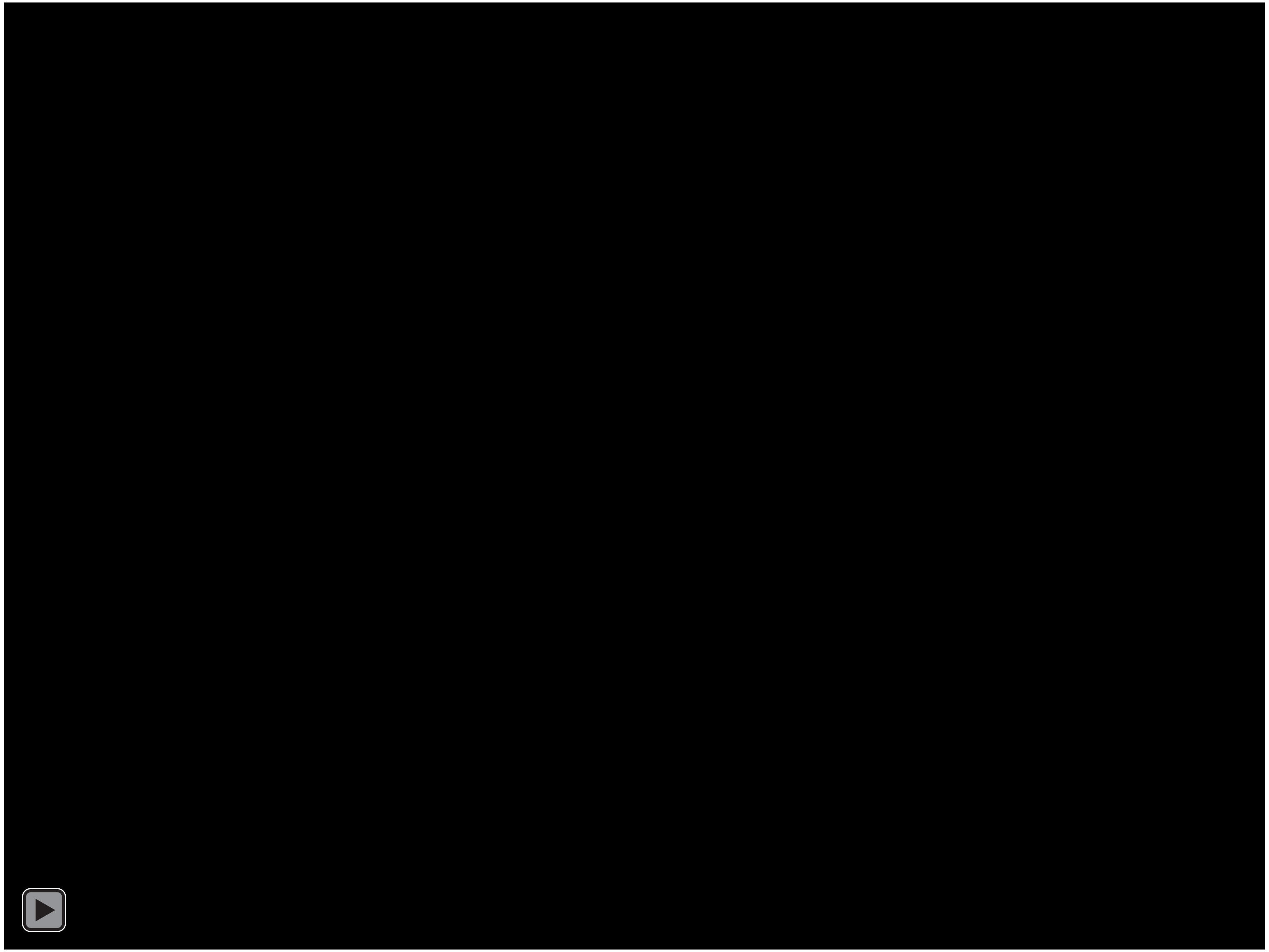
Density



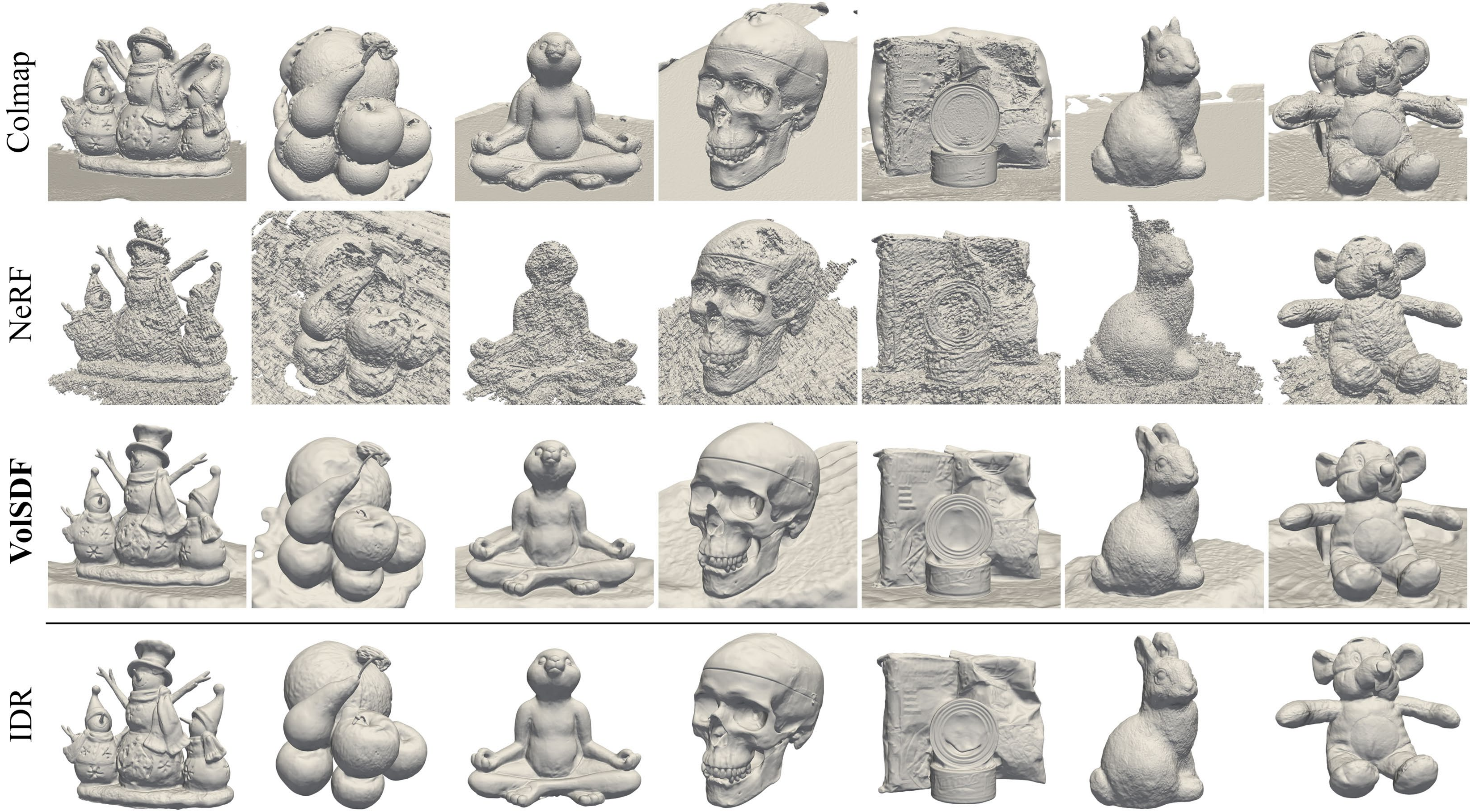
Laplace CDF

Signed Distance Function

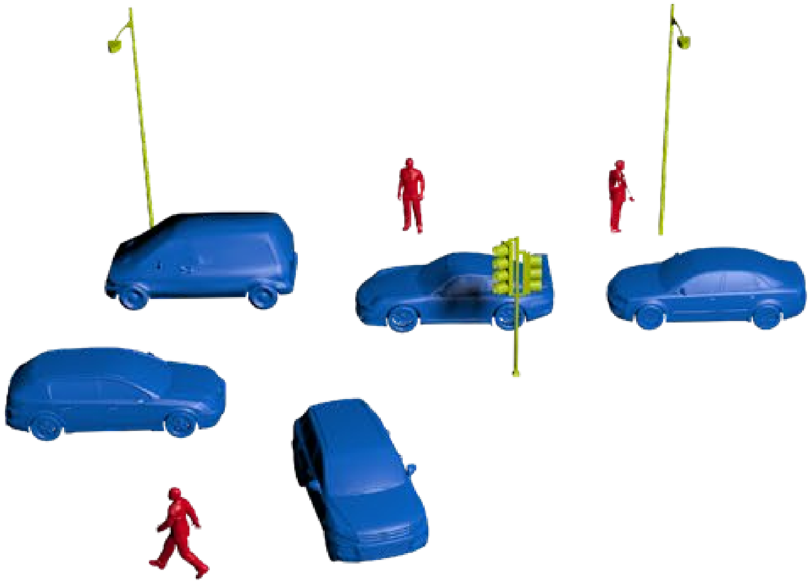
Results



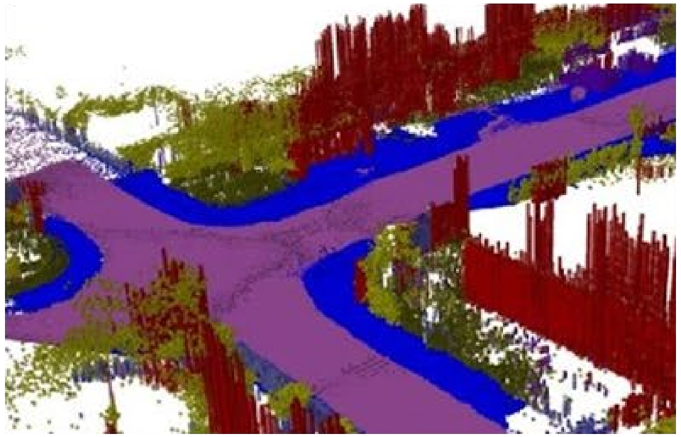
Results (DTU Data)



Compositional Structure in NeRFs



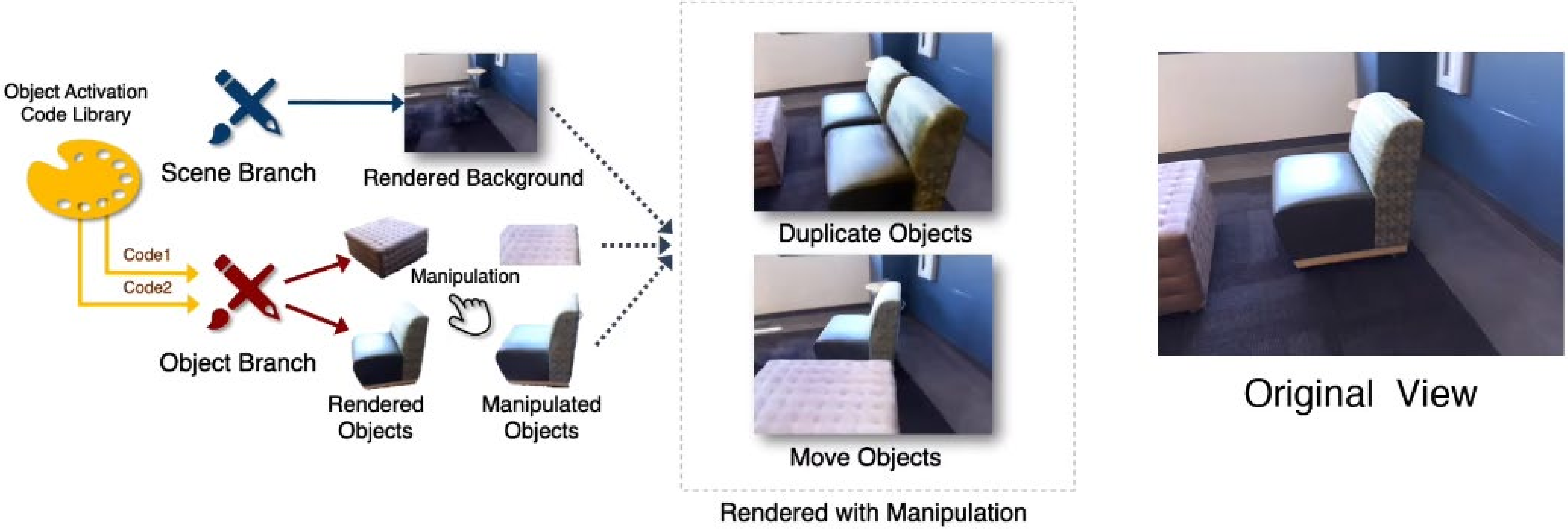
object instances (**things**)



static **stuff** background

Scene = Stuff + Thing(s)

Compositional Structure in NeRFs



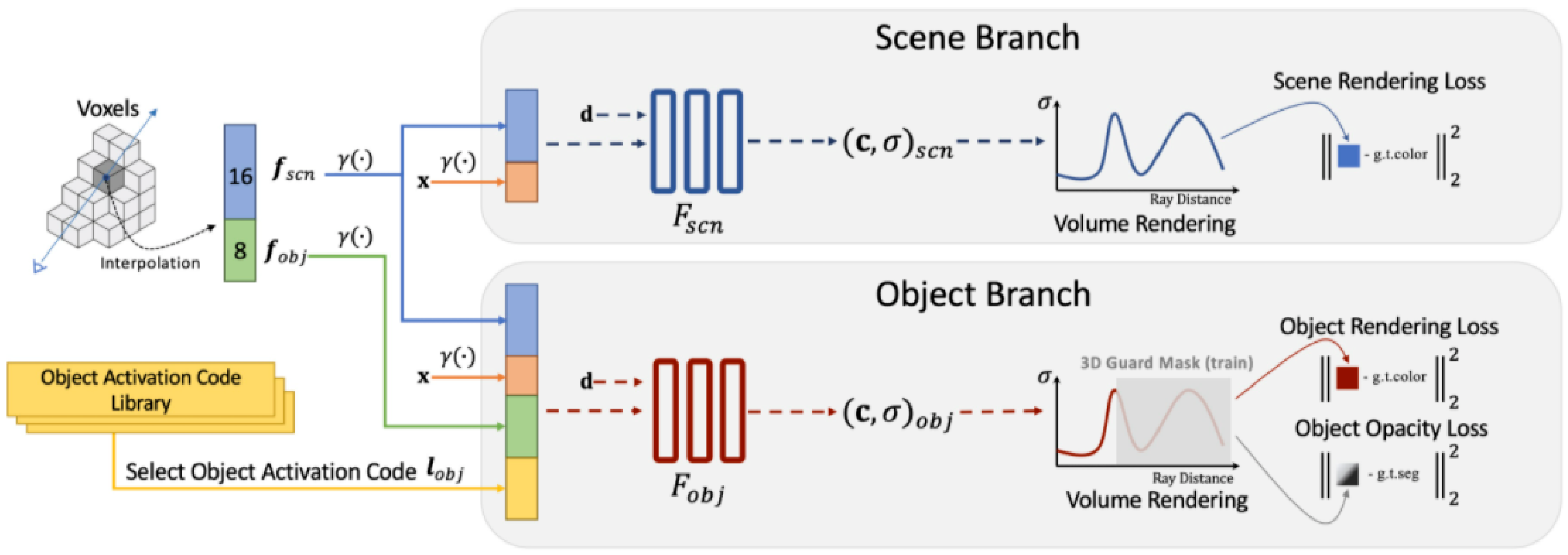
Object NeRF

Learning Object-Compositional Neural Radiance Field for Editable Scene Rendering

Bangbang Yang[1] Yinda Zhang[2] Yinghao Xu[3] Yijin Li[1]
Han Zhou[1] Hujun Bao[1] Guofeng Zhang[1] Zhaopeng Cui[1*]
[1]State Key Lab of CAD&CG, Zhejiang University [2]Google [3]The Chinese University of Hong Kong

ICCV 2021

Framework Overview



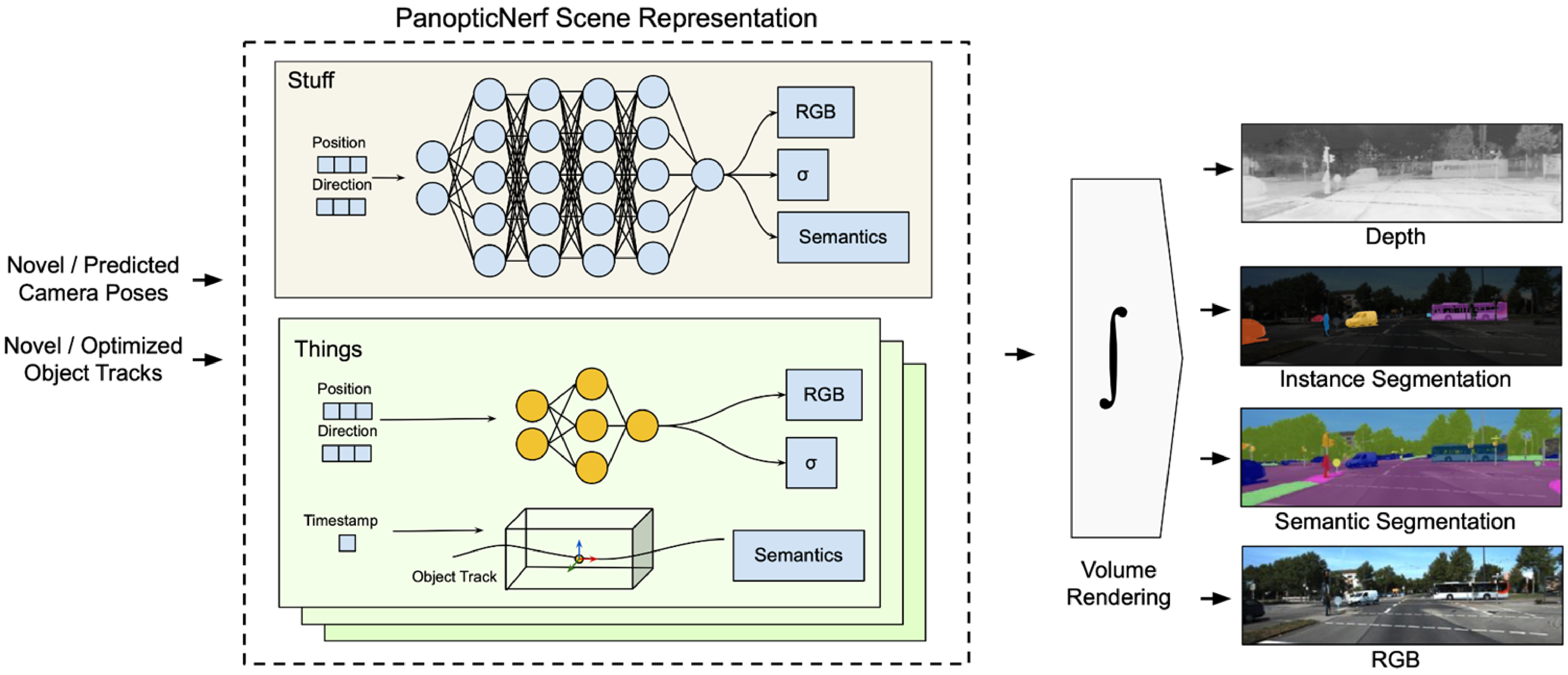
We design a two-pathway architecture for object-compositional neural radiance field. The scene branch takes the spatial coordinate $\mathbf{x}$, the interpolated scene voxel features f_{scn} at $\mathbf{x}$ and the ray direction $\mathbf{d}$ as input, and output the color $\mathbf{c}_{scn}$ and opacity σ_{scn} of the scene. The object branch takes additional object voxel features f_{obj} as well a a object activation code l_{obj} to condition the output only contains the color $\mathbf{c}_{obj}$ and opacity σ_{obj} for a specific object at its original location with everything else removed.

Object NeRF Results

Novel View Synthesis

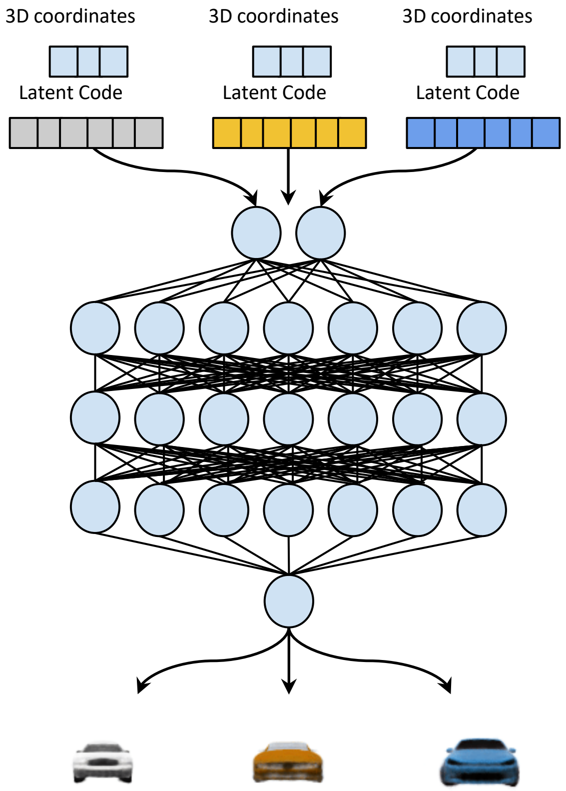
Editable Scene Rendering

Panoptic NeRF: Add Depth, Semantics, ++

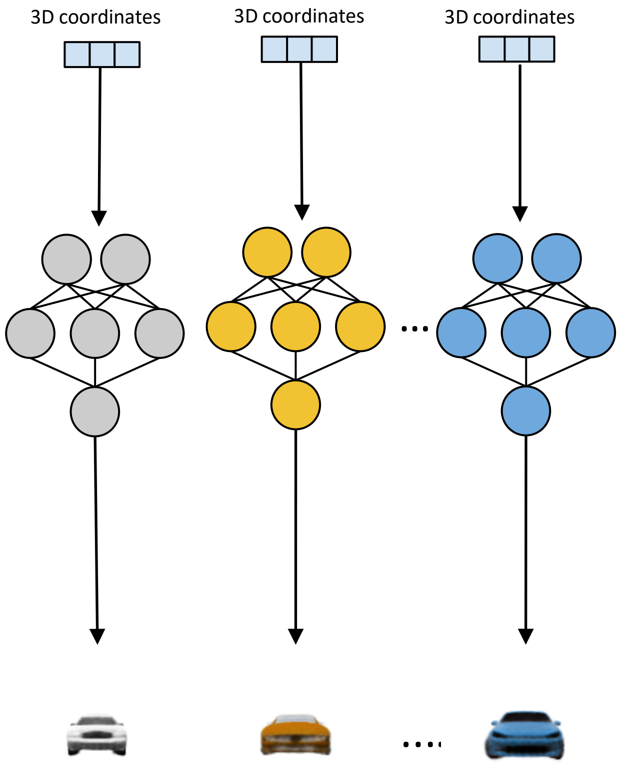


Efficient Object Representation

Shared MLP



Instance-Specific MLP

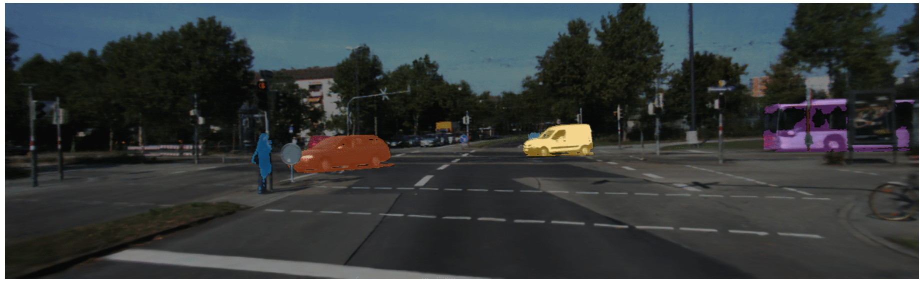


Results: Scene Understanding

Rendered RGB

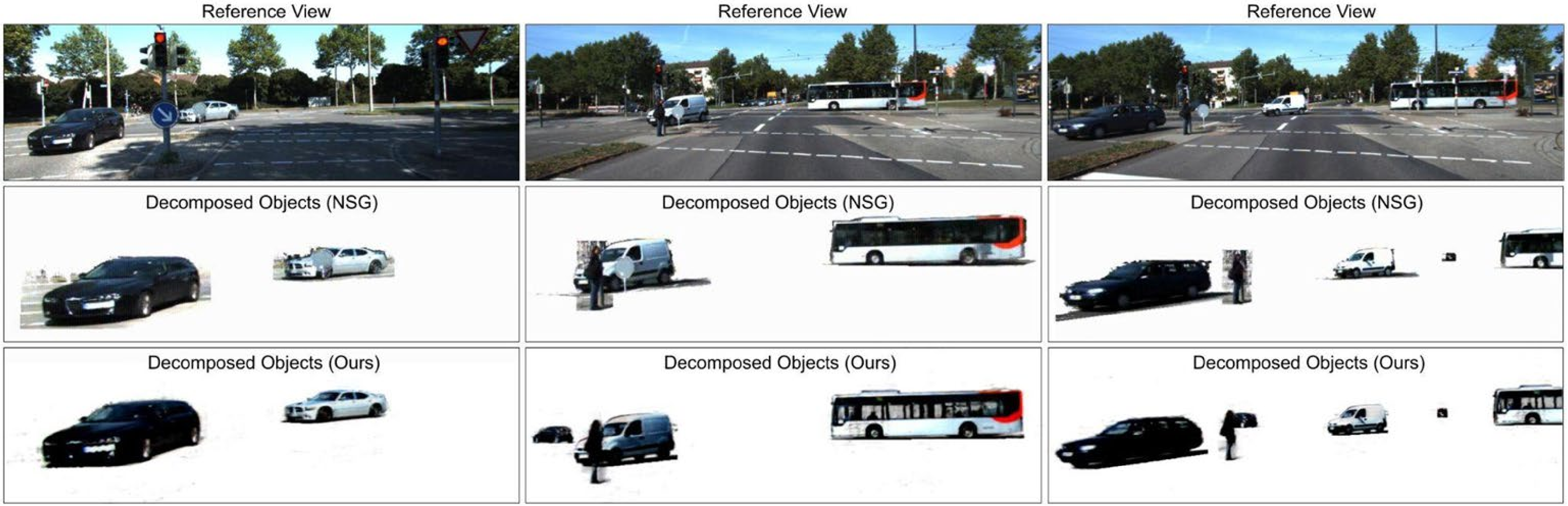
Rendered Semantic segmentation (overlaid on top of rendered color)

Rendered (expected) Depth



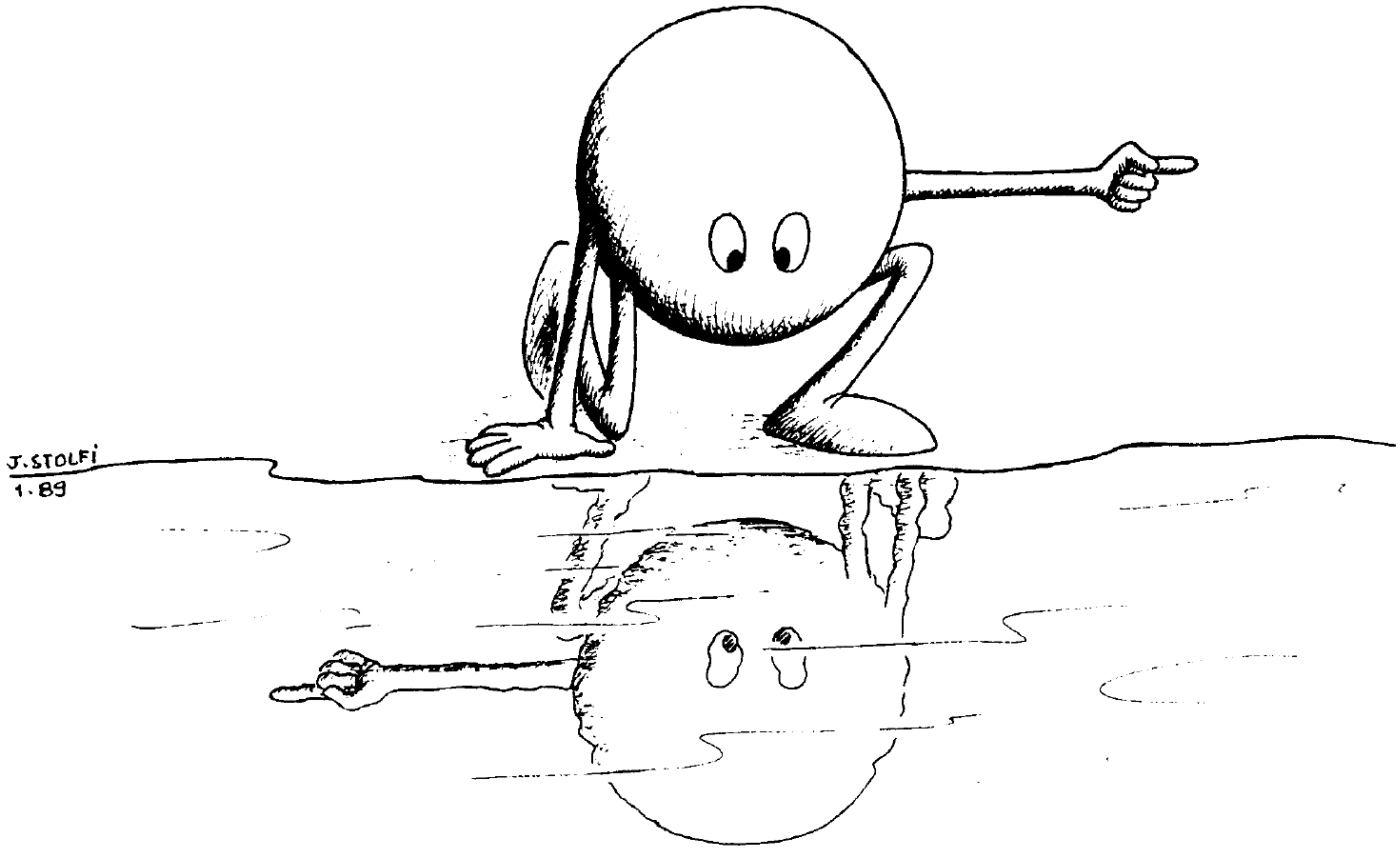
Rendered instance segmentation (overlaid on top of rendered color)

Results: Scene Decomposition



Top: reference views; middle: neural scene graphs; bottom: ours.

Results: New View Synthesis


J.STOLFI
1.89