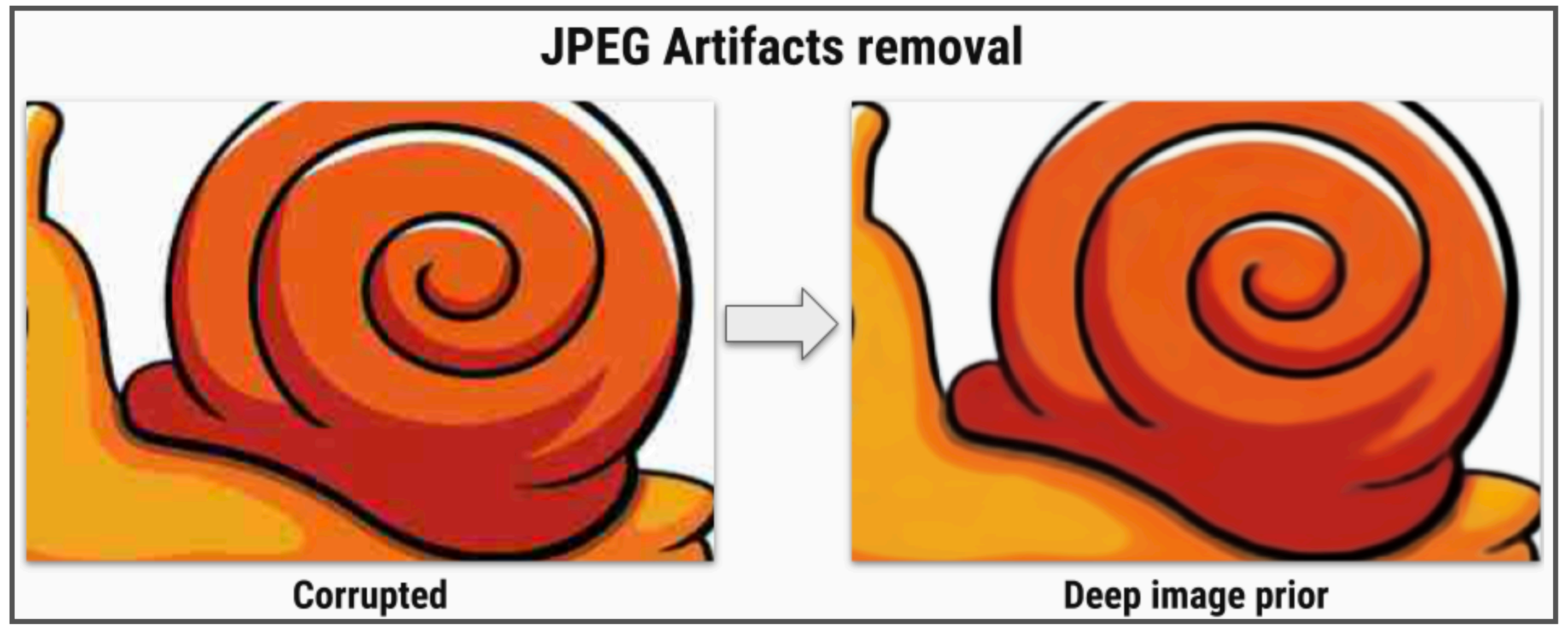
Neural Surface Reconstruction

February 14th, 2024

Previous lectures …

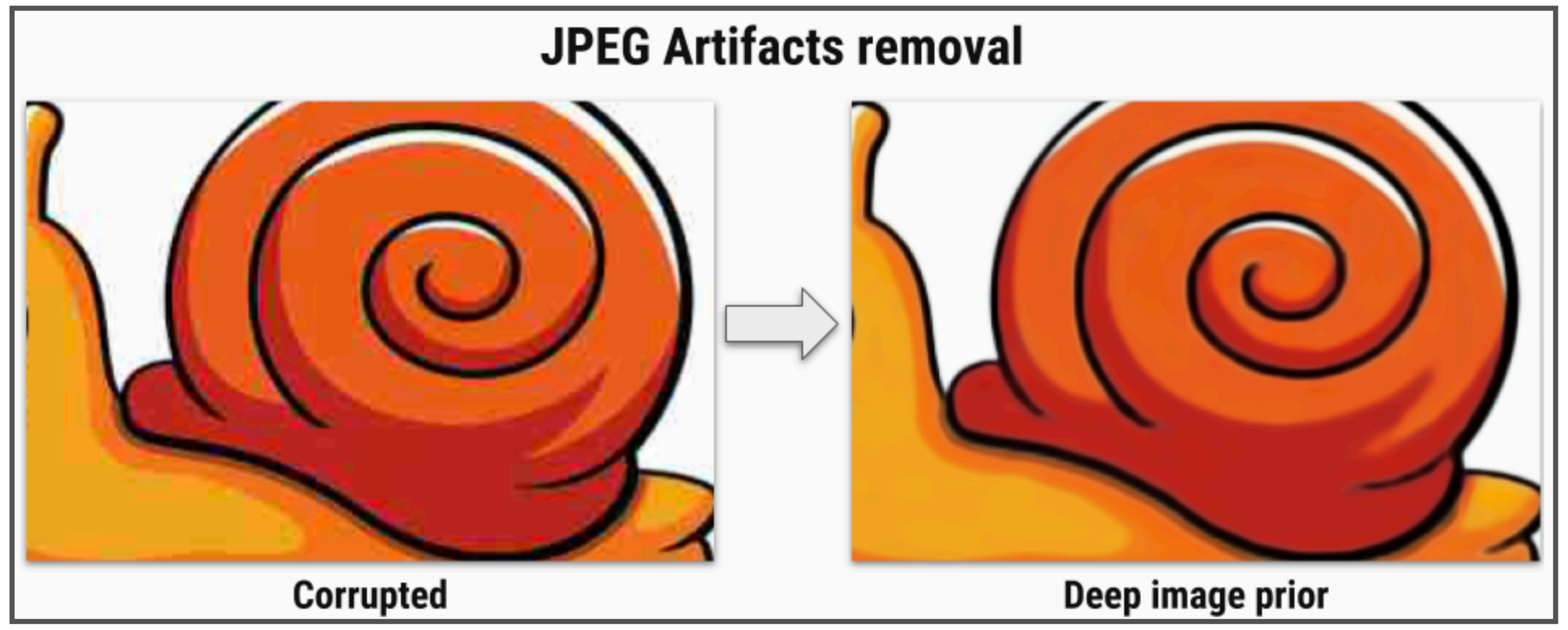
- Inverse problems



Deep Image Prior [Ulyanov et al. 2018]

Previous lectures …

- Inverse problems



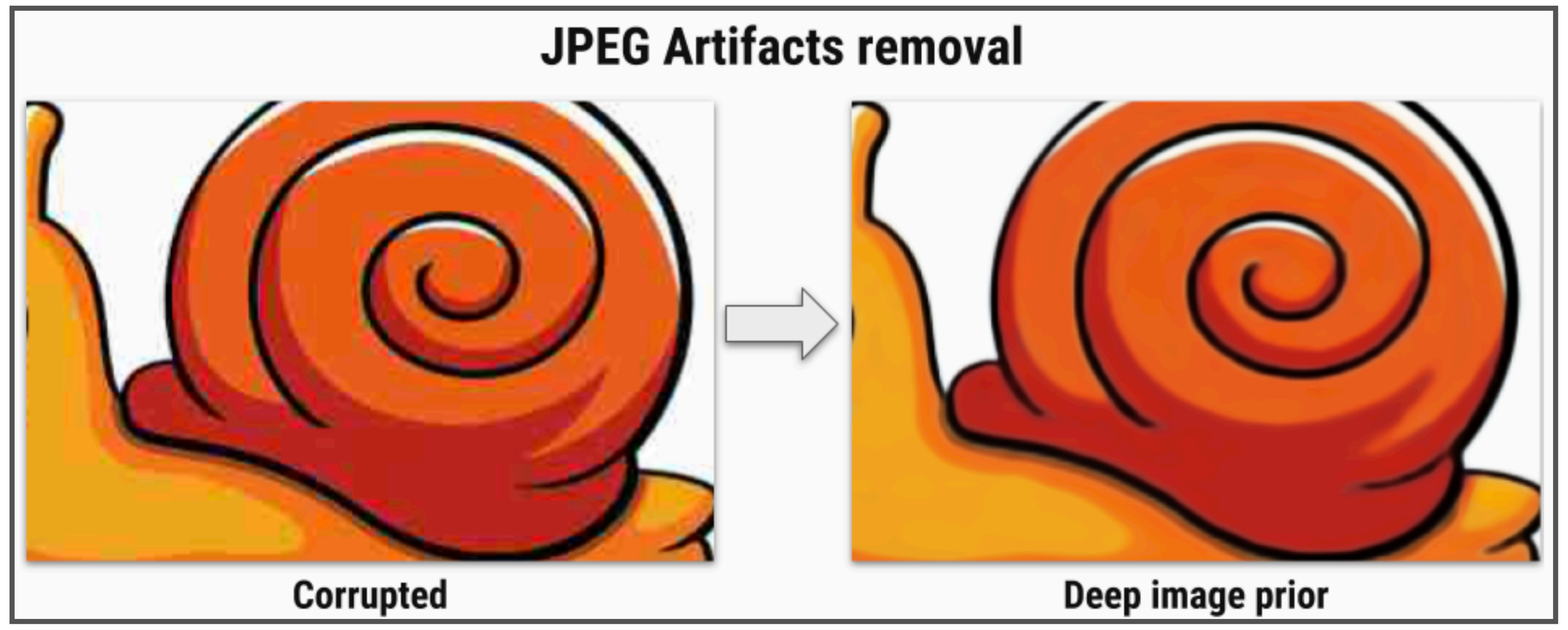
Deep Image Prior [Ulyanov et al. 2018]

- Structure from motion (2D to 3D)

Photo Tourism: [Snavely et al. 2006]

Previous lectures ...

- Inverse problems

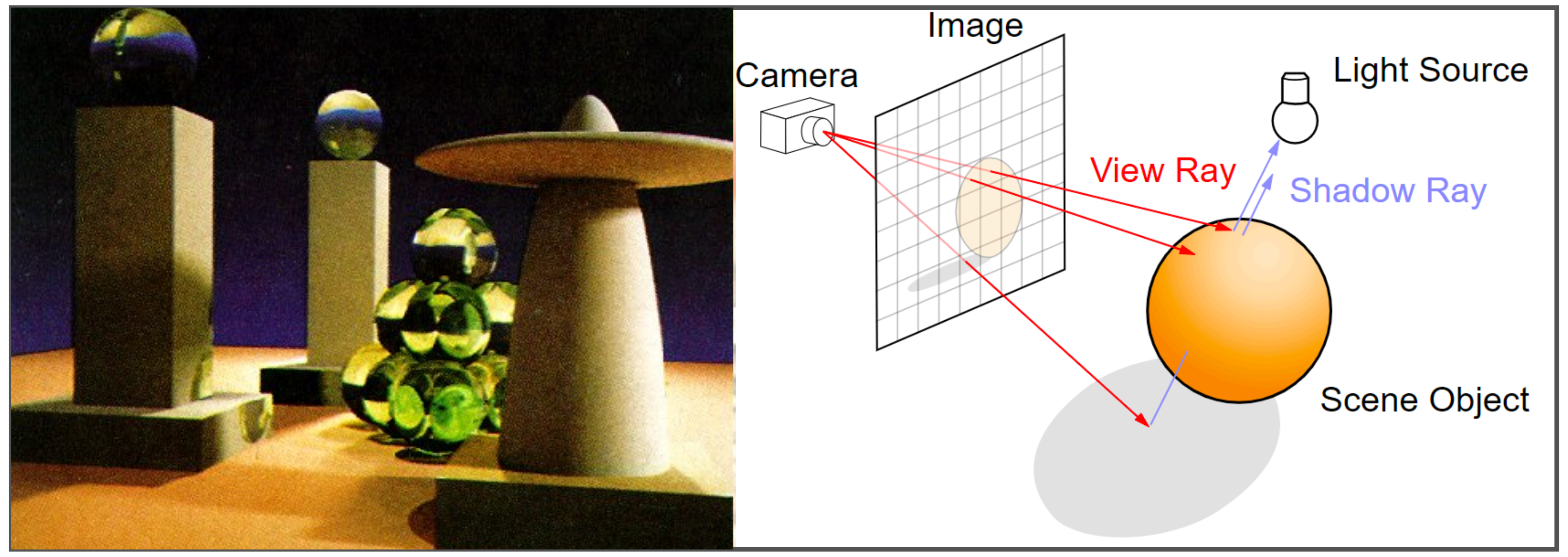


Deep Image Prior [Ulyanov et al. 2018]

- Structure from motion (2D to 3D)

Photo Tourism: [Snavely et al. 2006]

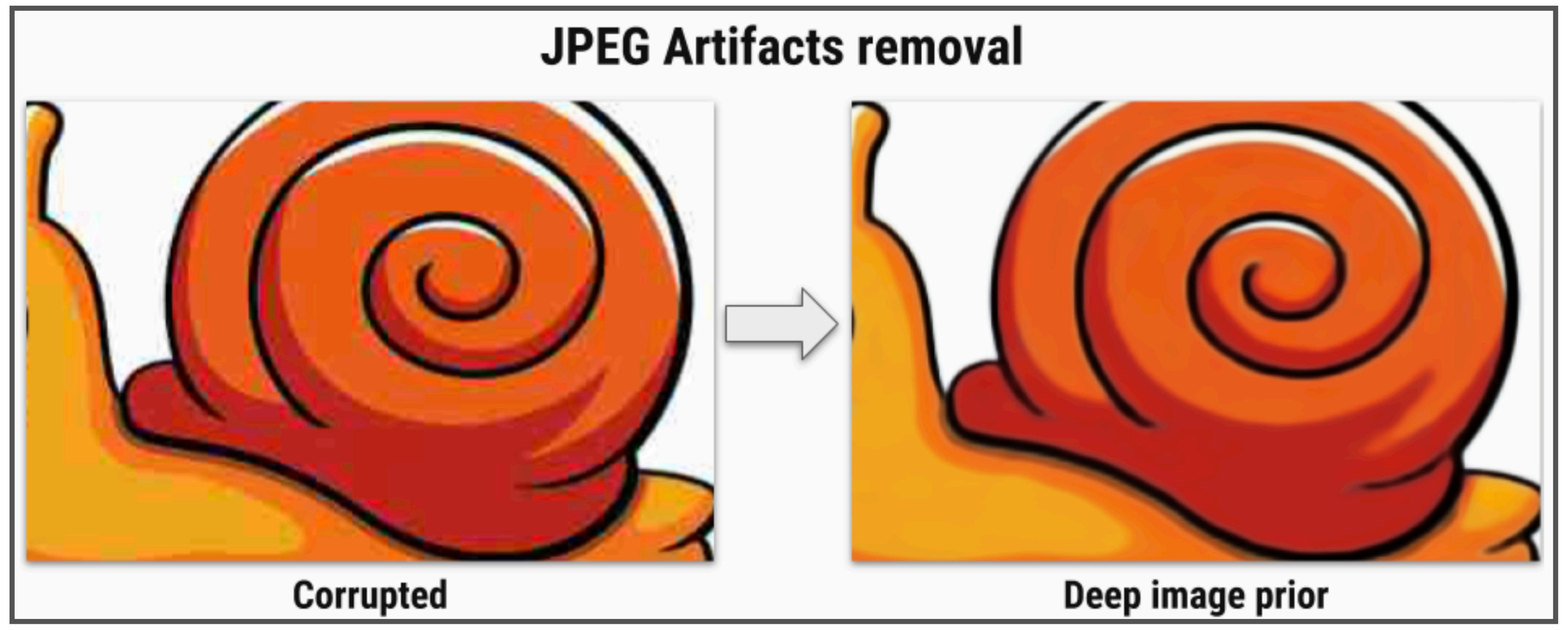
- Rendering (3D to 2D)



The Rendering Equation [Kajiya 1986]

Previous lectures …

- Inverse problems

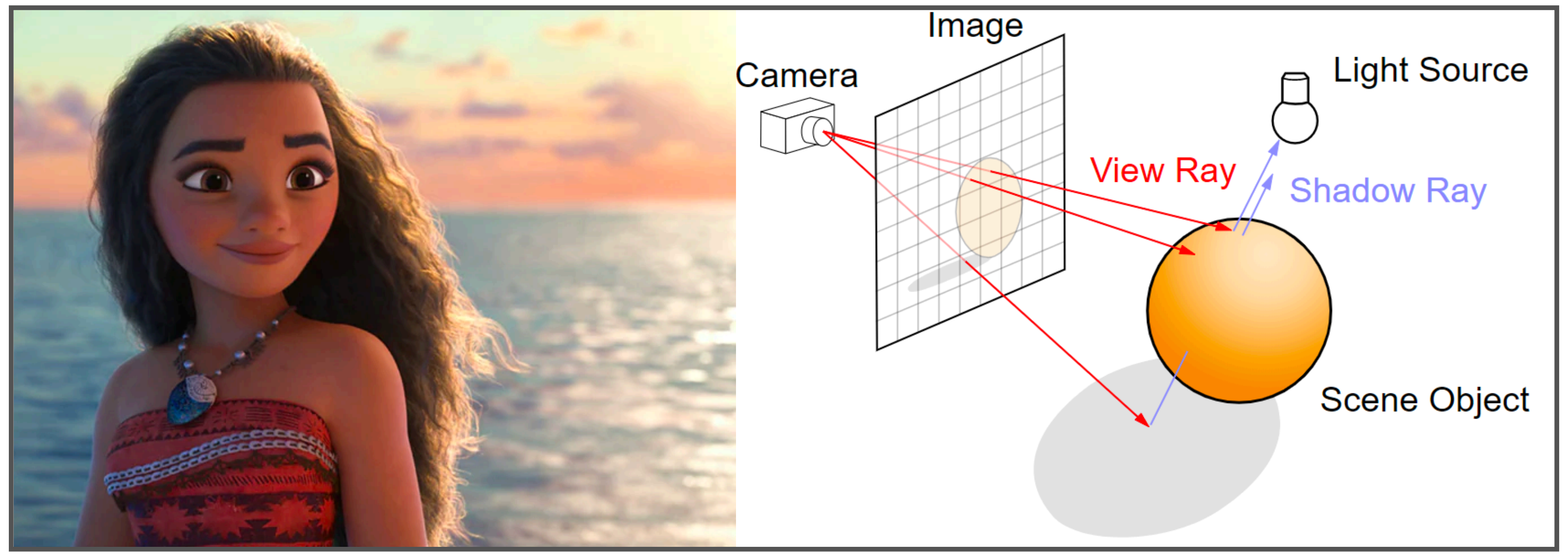


Deep Image Prior [Ulyanov et al. 2018]

- Structure from motion (2D to 3D)

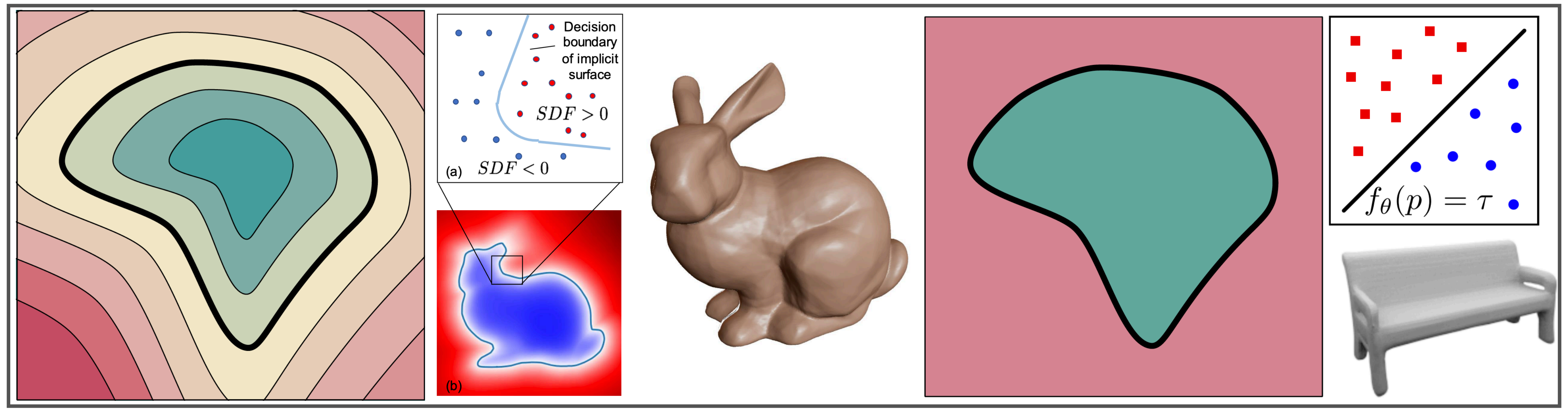
Photo Tourism: [Snavely et al. 2006]

- Rendering (3D to 2D)



The Rendering Equation [Kajiya 1986]

- Implicit Neual Representations

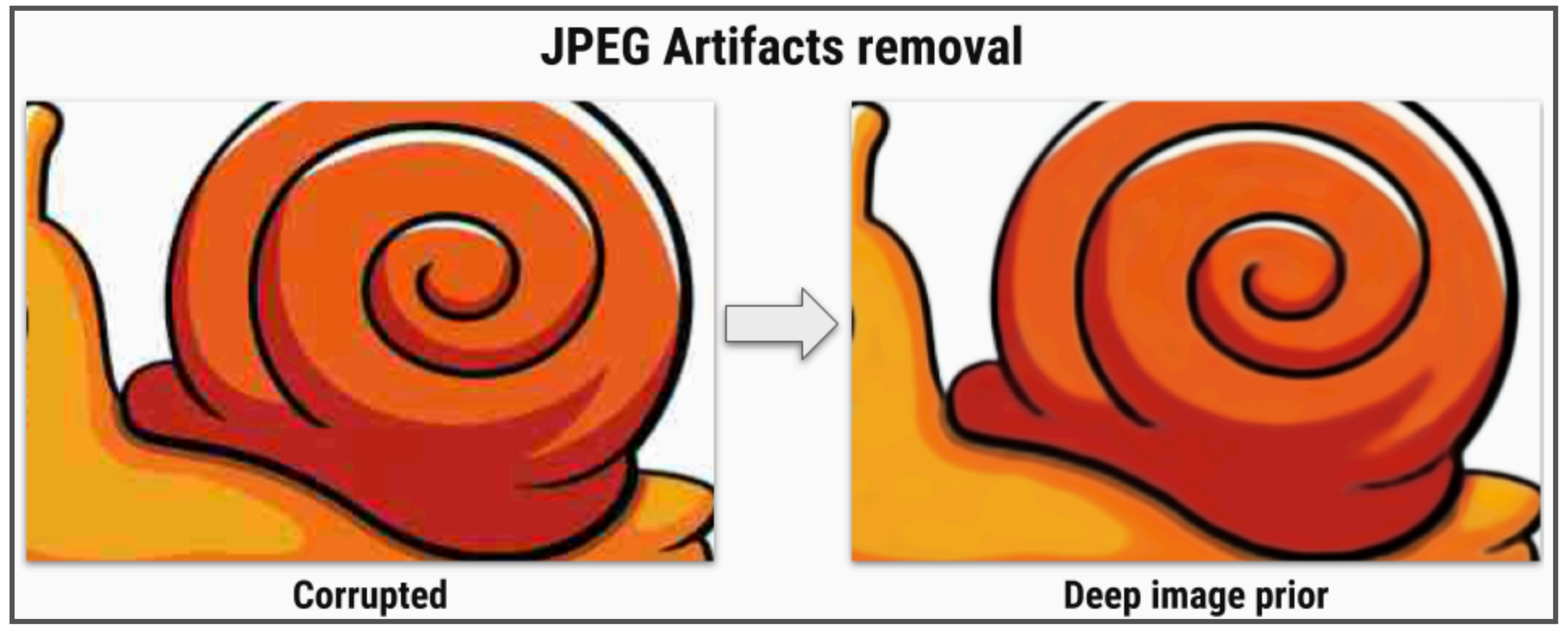


DeepSDF [Park et al. 2019]

Occupancy Networks Mescheder et al. 2019]

Previous lectures …

- Inverse problems

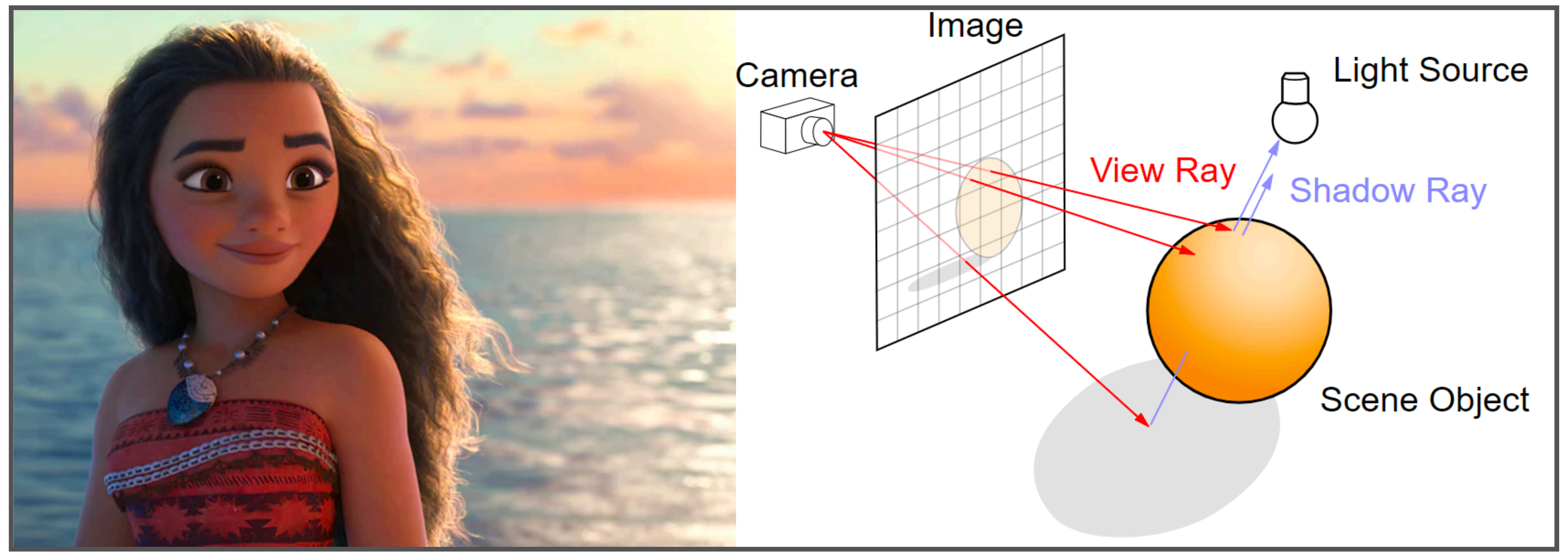


Deep Image Prior [Ulyanov et al. 2018]

- Structure from motion (2D to 3D)

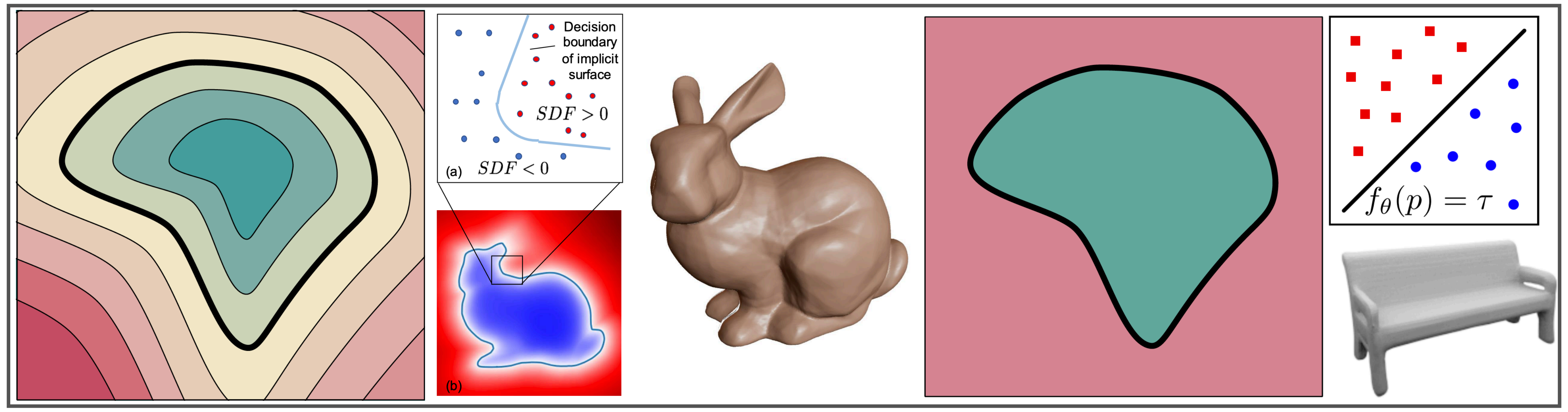
Photo Tourism: [Snavely et al. 2006]

- Rendering (3D to 2D)



The Rendering Equation [Kajiya 1986]

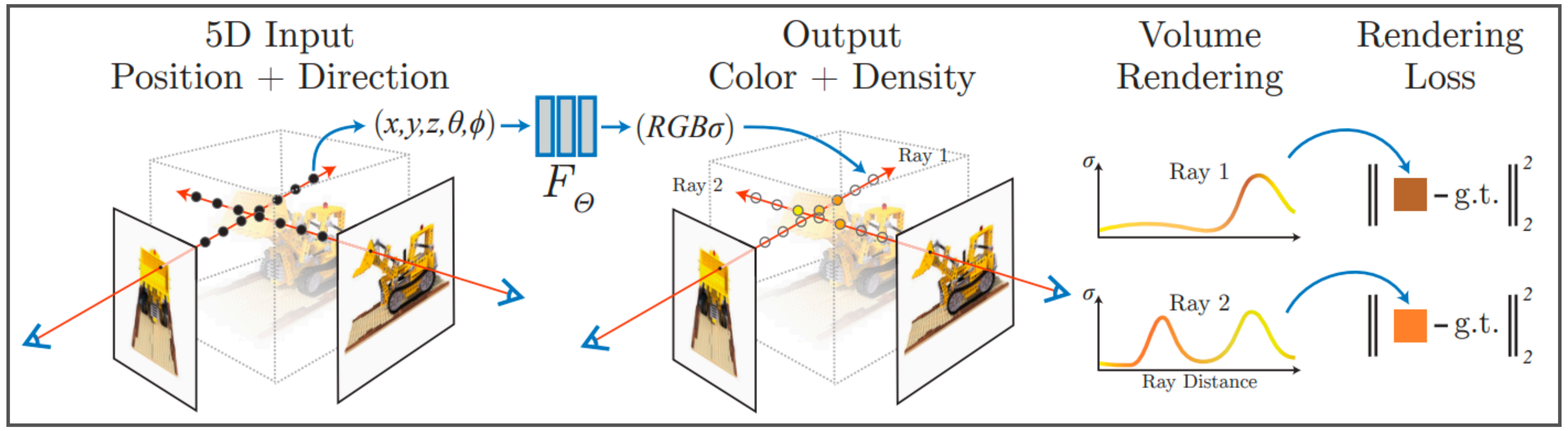
- Implicit Neual Representations



DeepSDF [Park et al. 2019]

Occupancy Networks Mescheder et al. 2019]

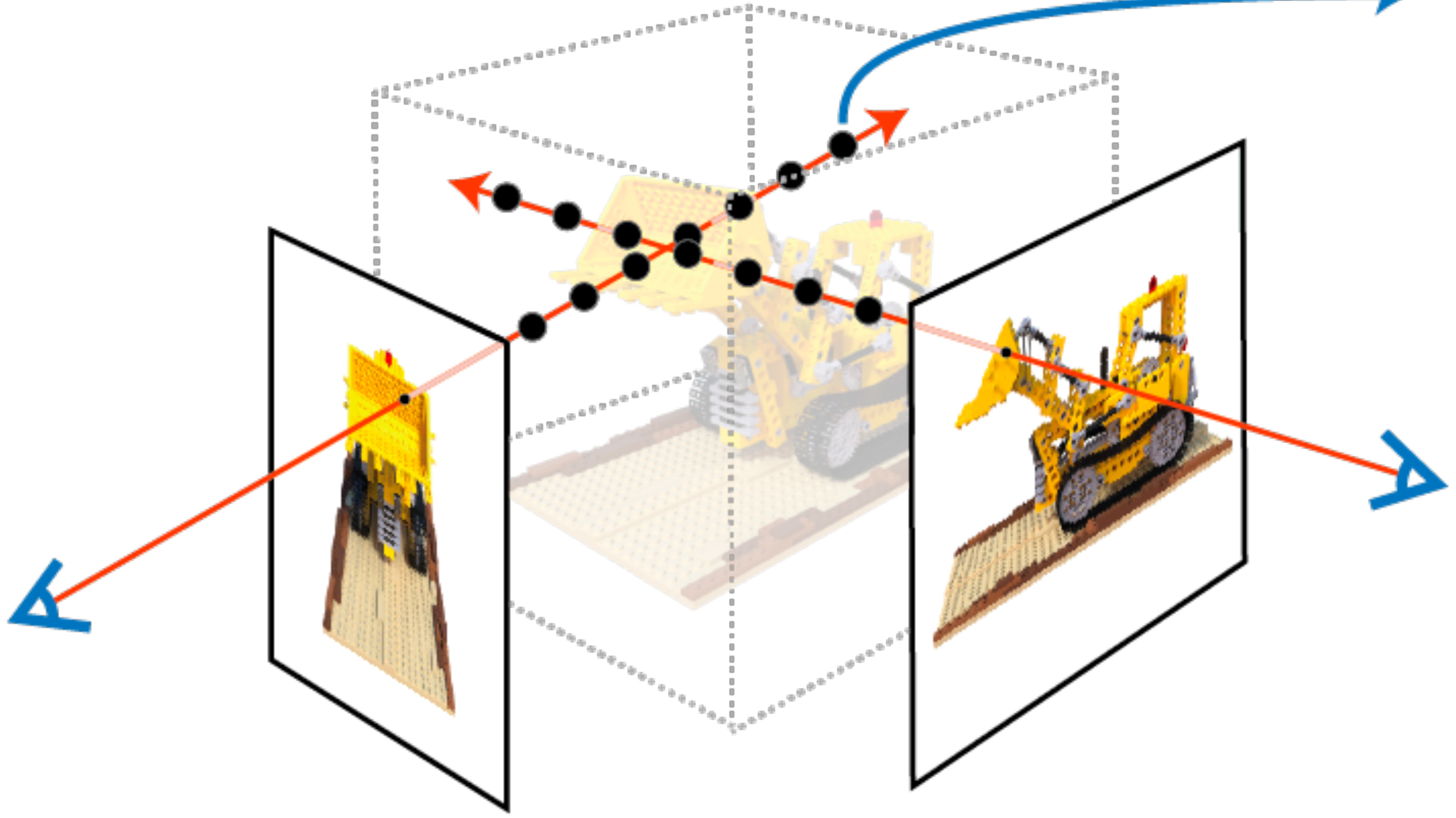
- Neural Rendering



Neural Radiance Fields [Mildenhall et al. '20]

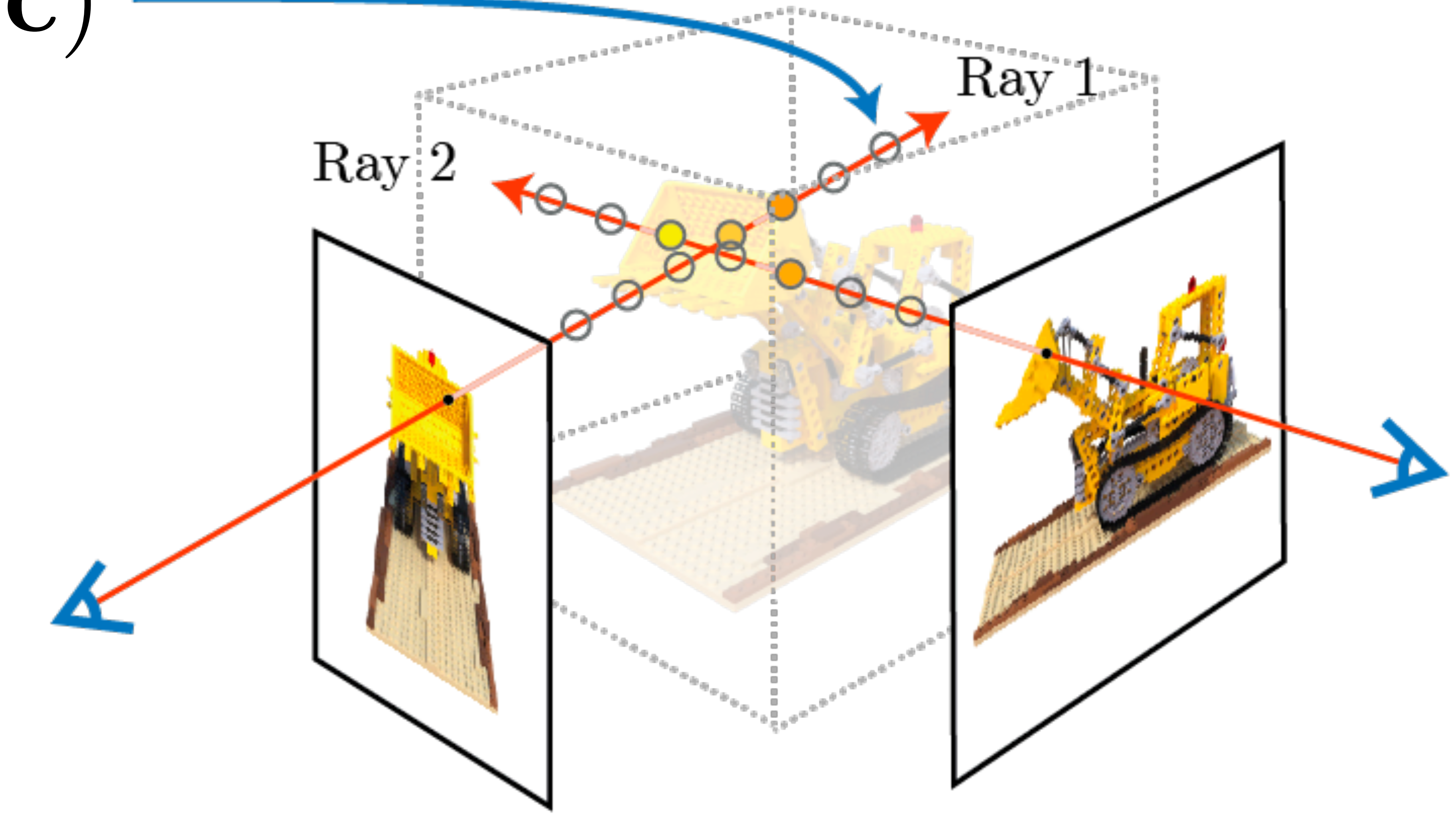
Neural Radiance Fields (NeRF)

Input:
Position + Direction
$(\mathbf{x}, \hat{\mathbf{d}})$



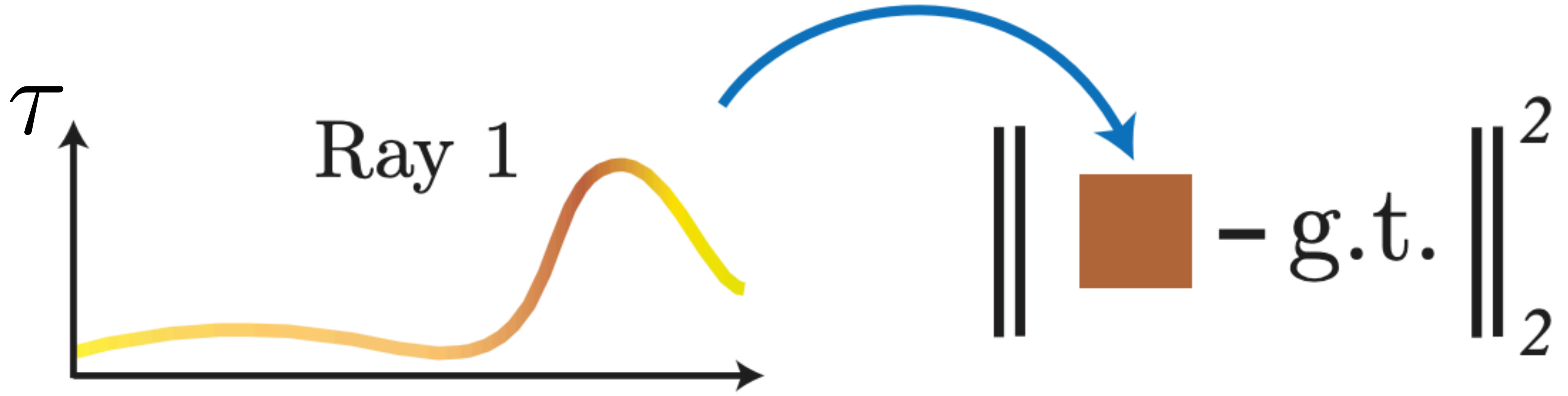
F_Θ

Output:
Density + Color
$(\tau, \mathbf{c})$



Volume rendering

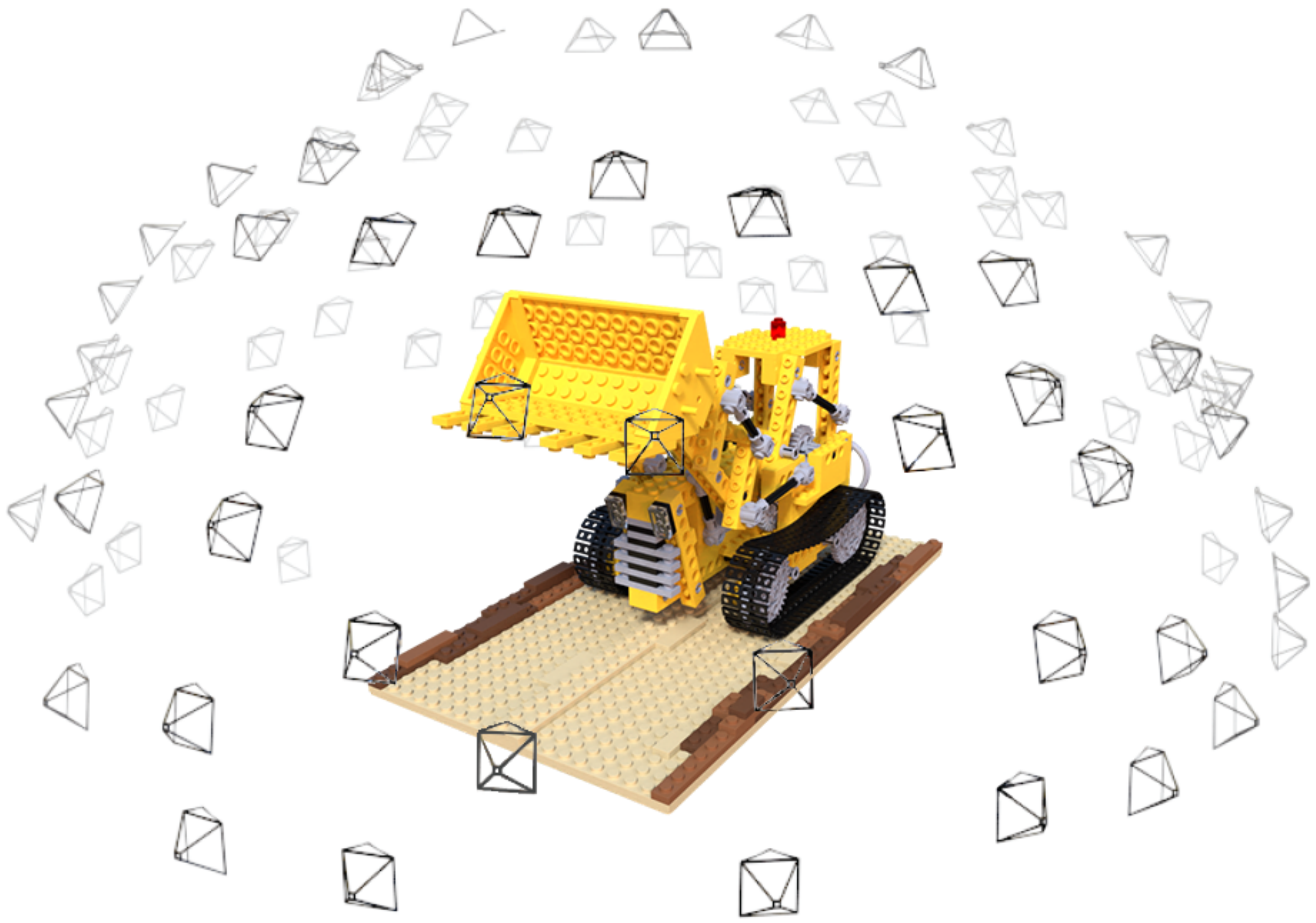
Rendering loss



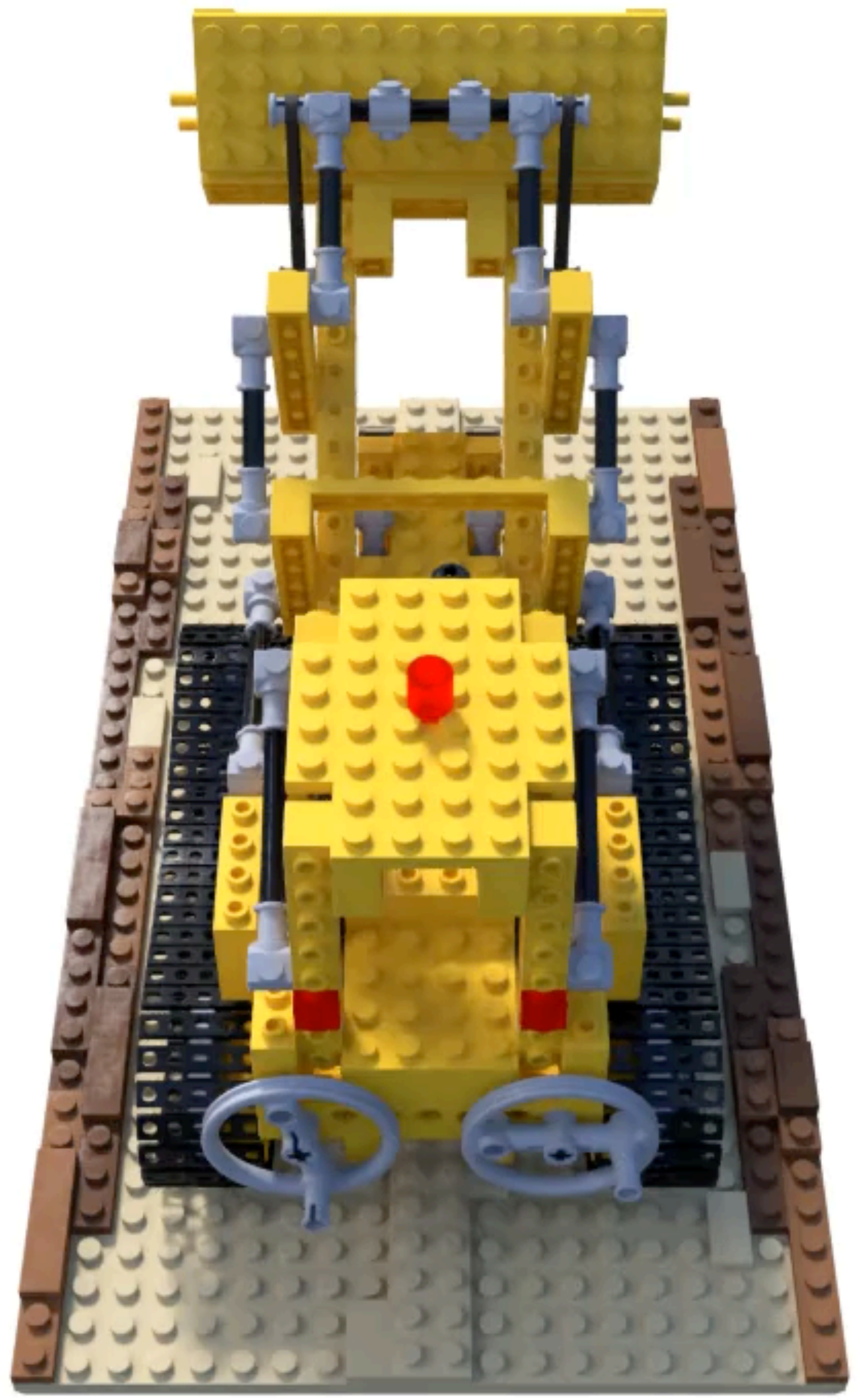
NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, Mildenhall, Srinivasan, Tancik, et al., 2020

Neural Radiance Fields (NeRF)

Input images + cameras



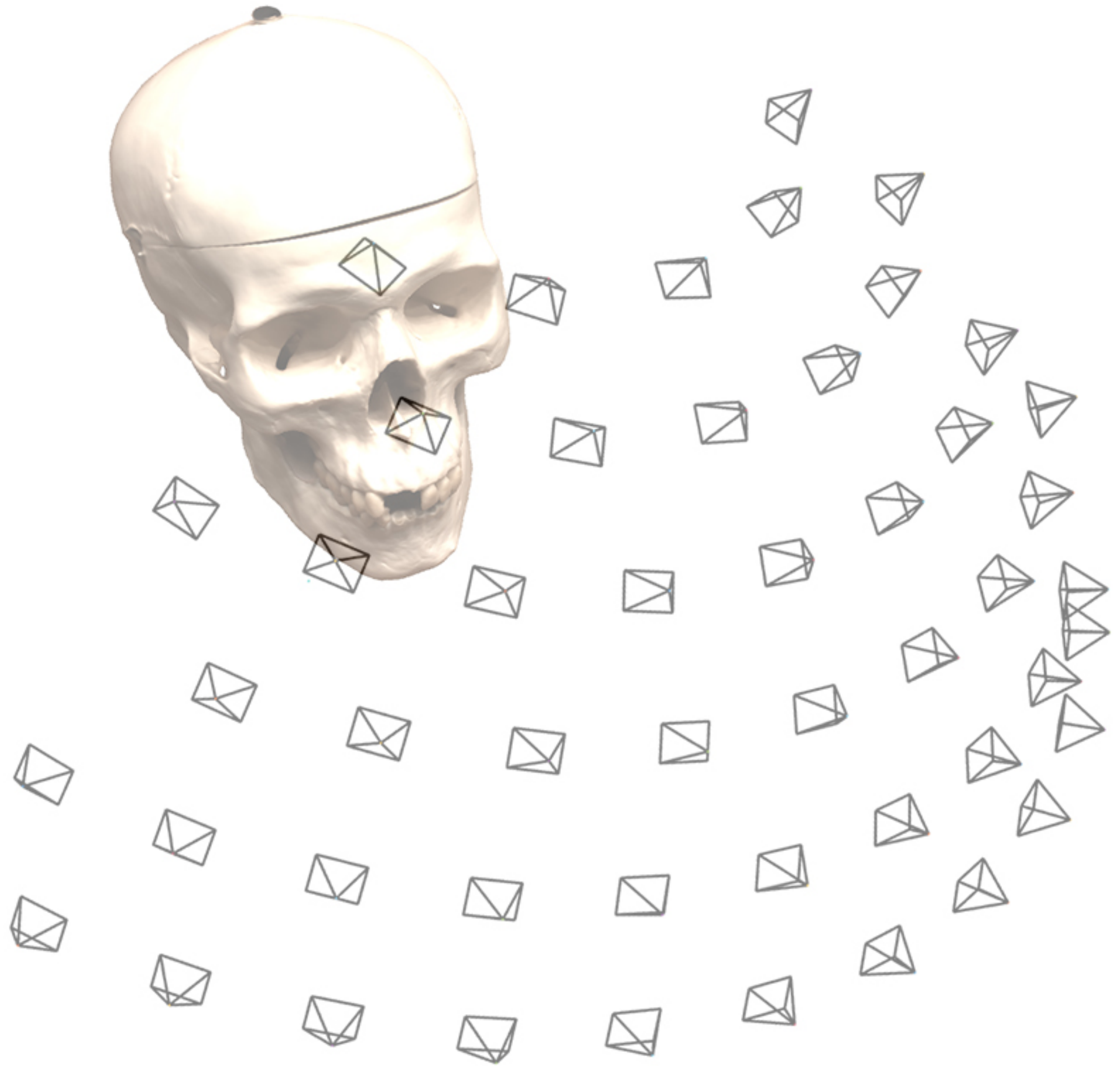
Novel views synthesis



NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, Mildenhall, Srinivasan, Tancik, et al., 2020

Today's lecture

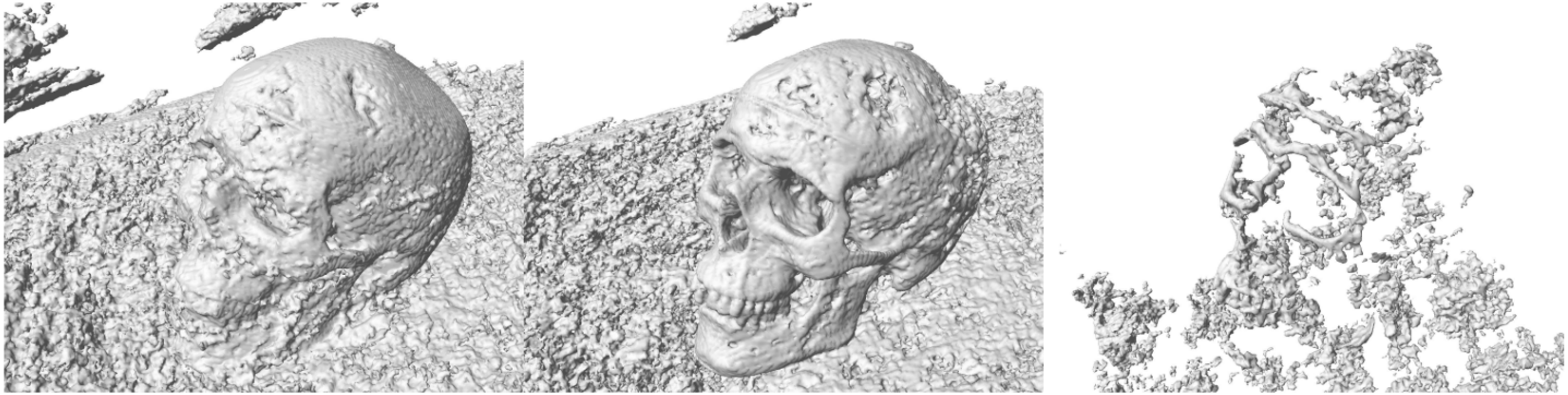
Multiview images



Camera poses

NeRF - Differential **Volume** Rendering

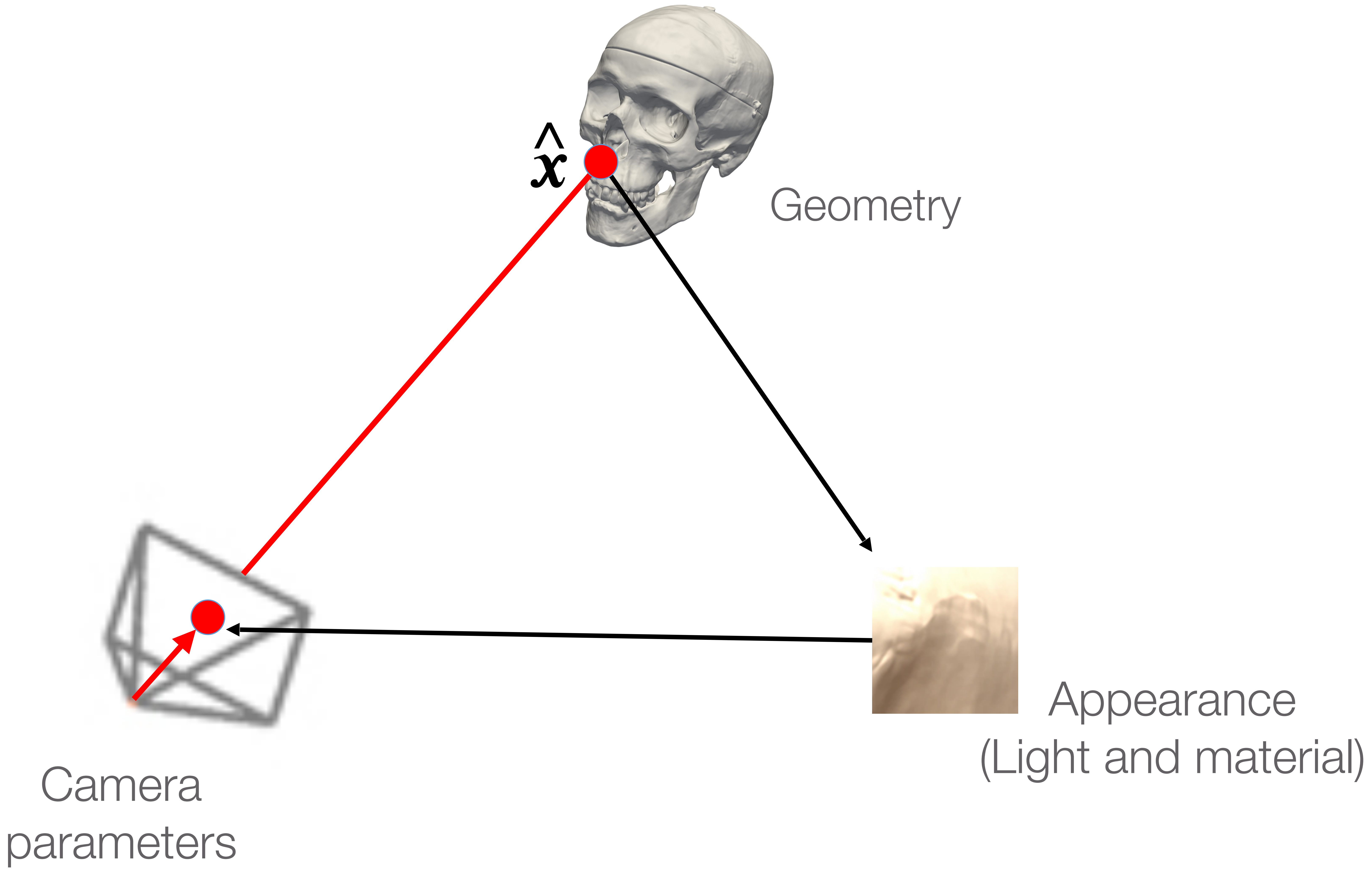
Volume density thresholds of NeRF



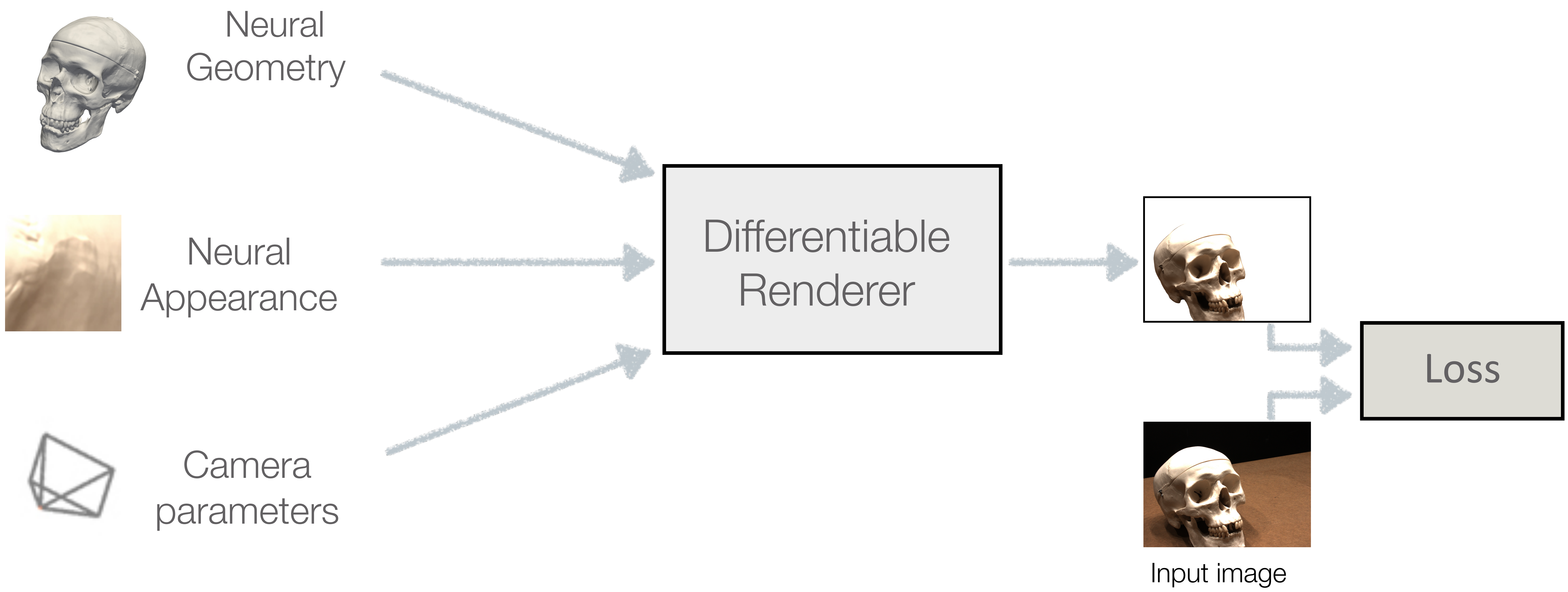
$\sigma = 1$ $\sigma = 50$ $\sigma = 500$

UNISURF: Unifying Neural Implicit Surfaces and Radiance Fields for Multi-View Reconstruction, Oechsle et al., 2022

Differential **Surface** Rendering

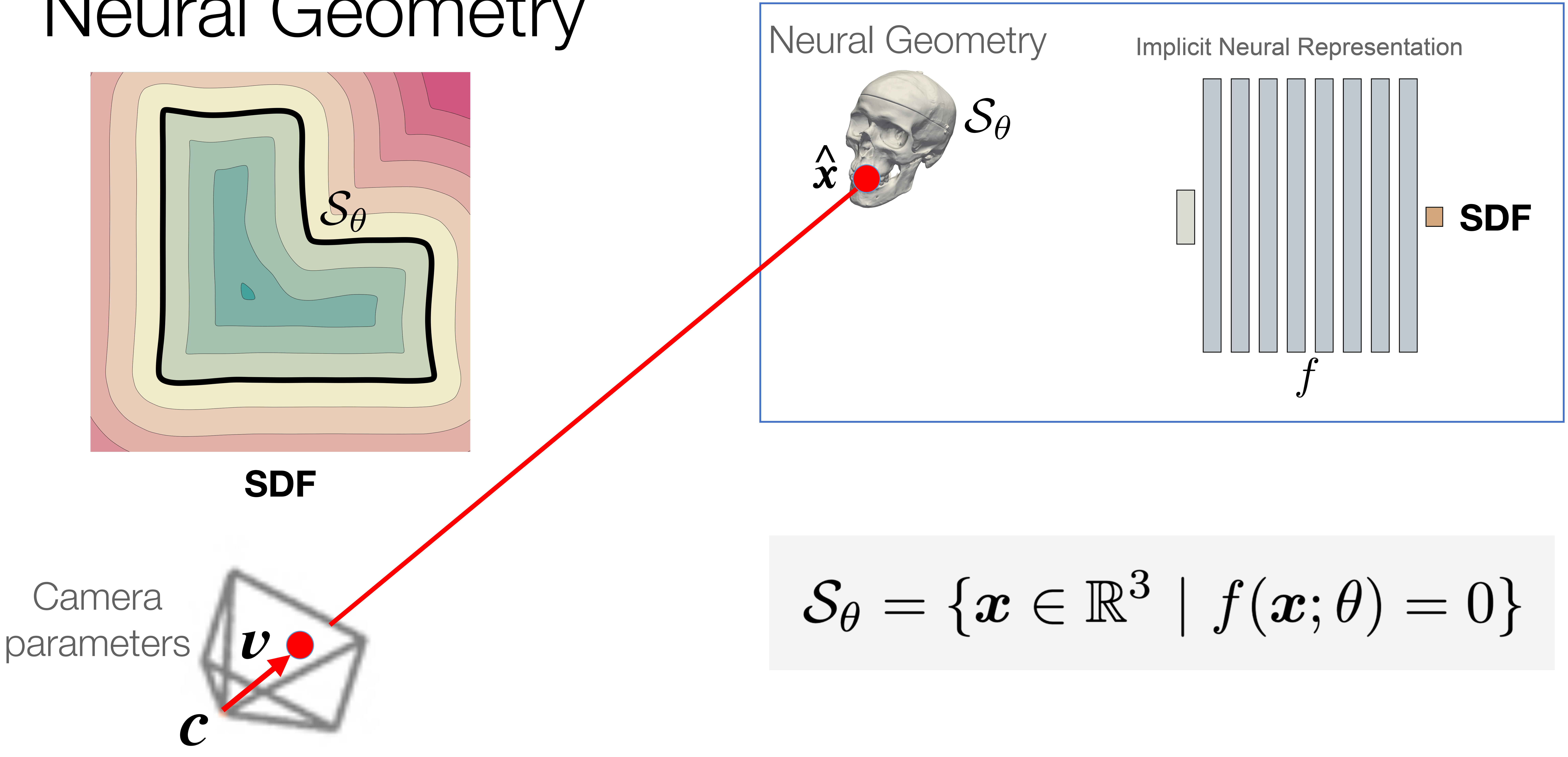


Implicit Differentiable Renderer (IDR)



Multiview neural surface reconstruction by disentangling geometry and appearance (IDR), Yariv et. al., NeurIPS 2020

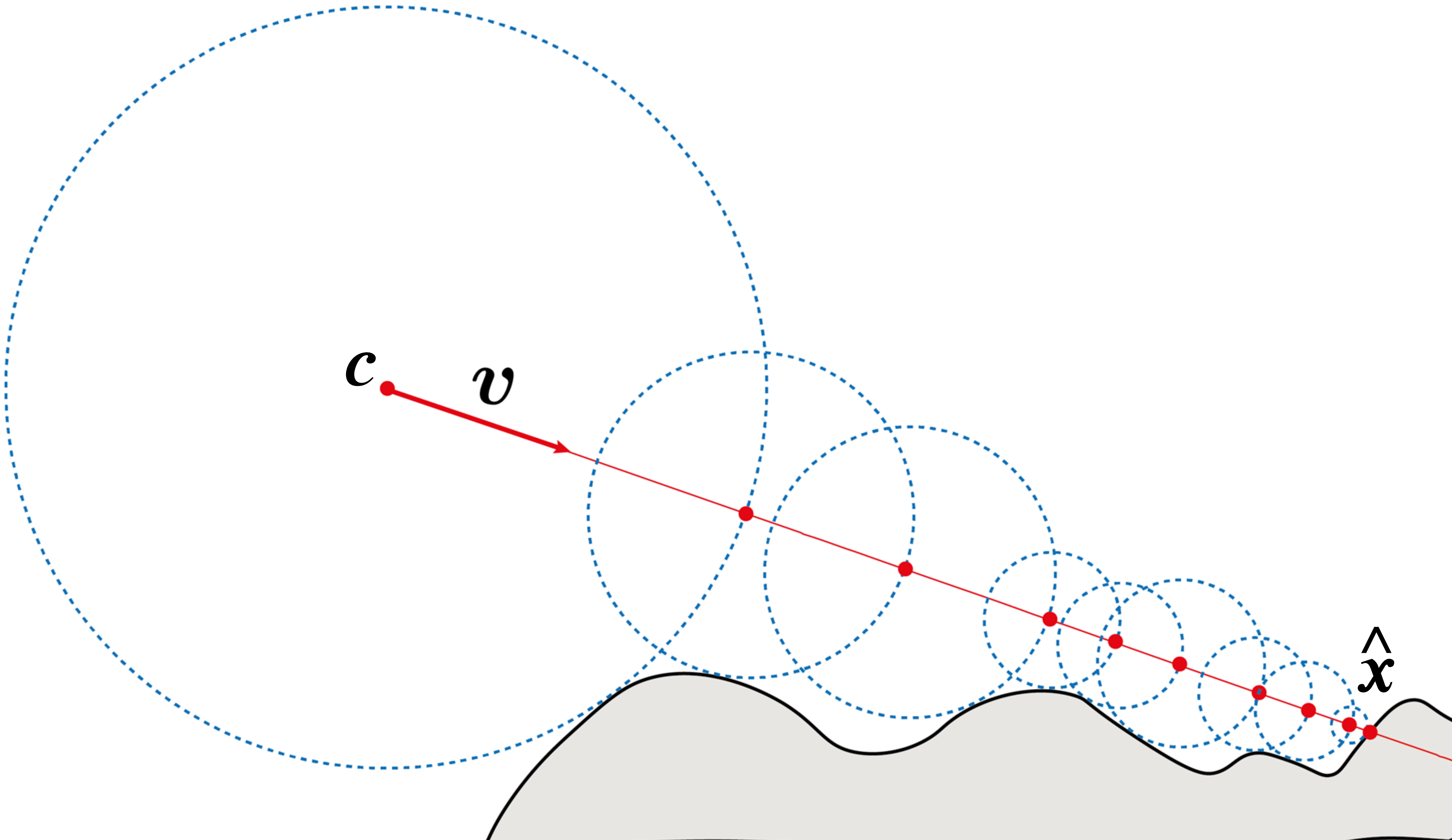
Neural Geometry



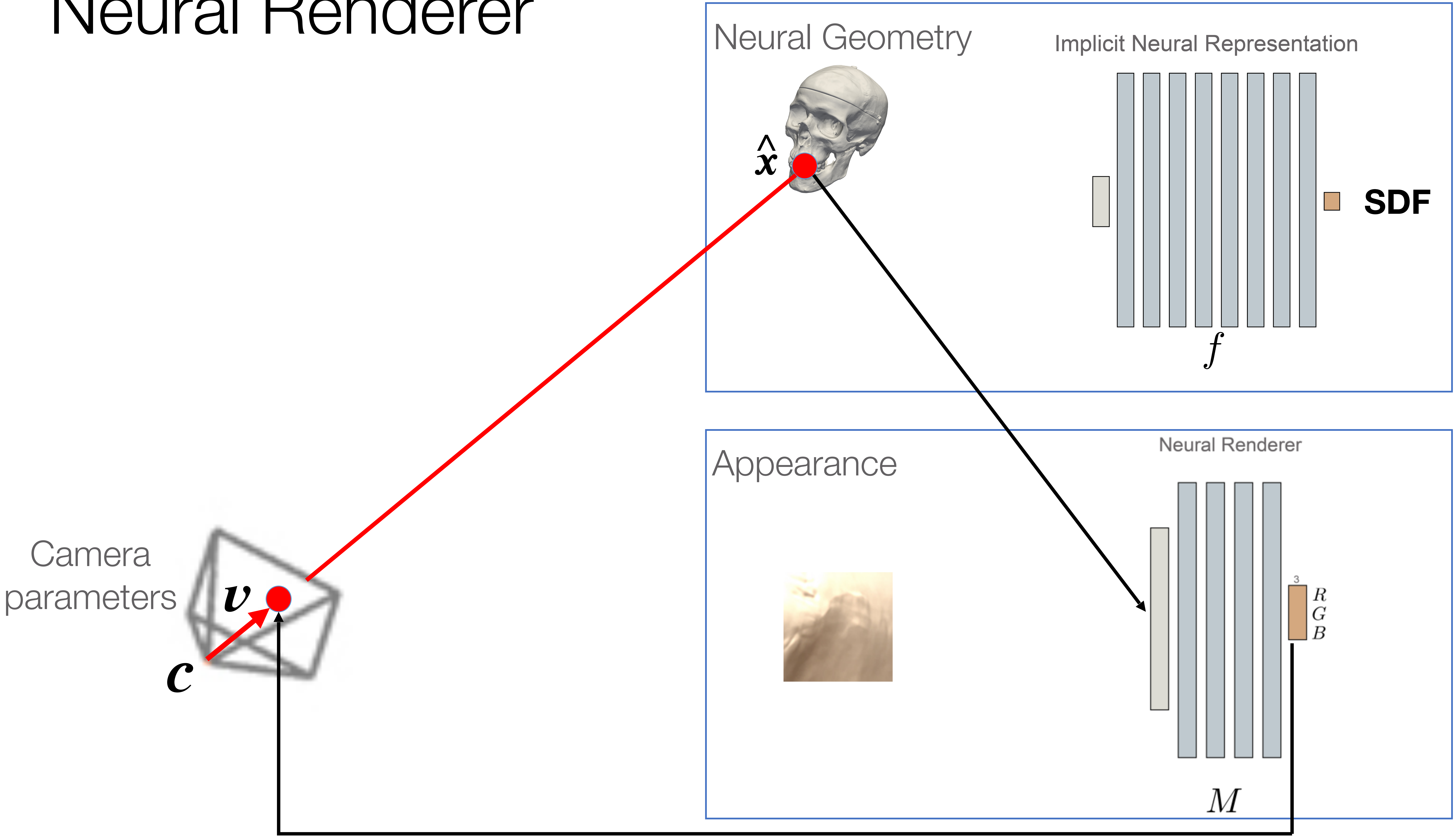
$$\mathcal{S}_\theta = \{\boldsymbol{x} \in \mathbb{R}^3 \mid f(\boldsymbol{x}; \theta) = 0\}$$

Sphere Tracing

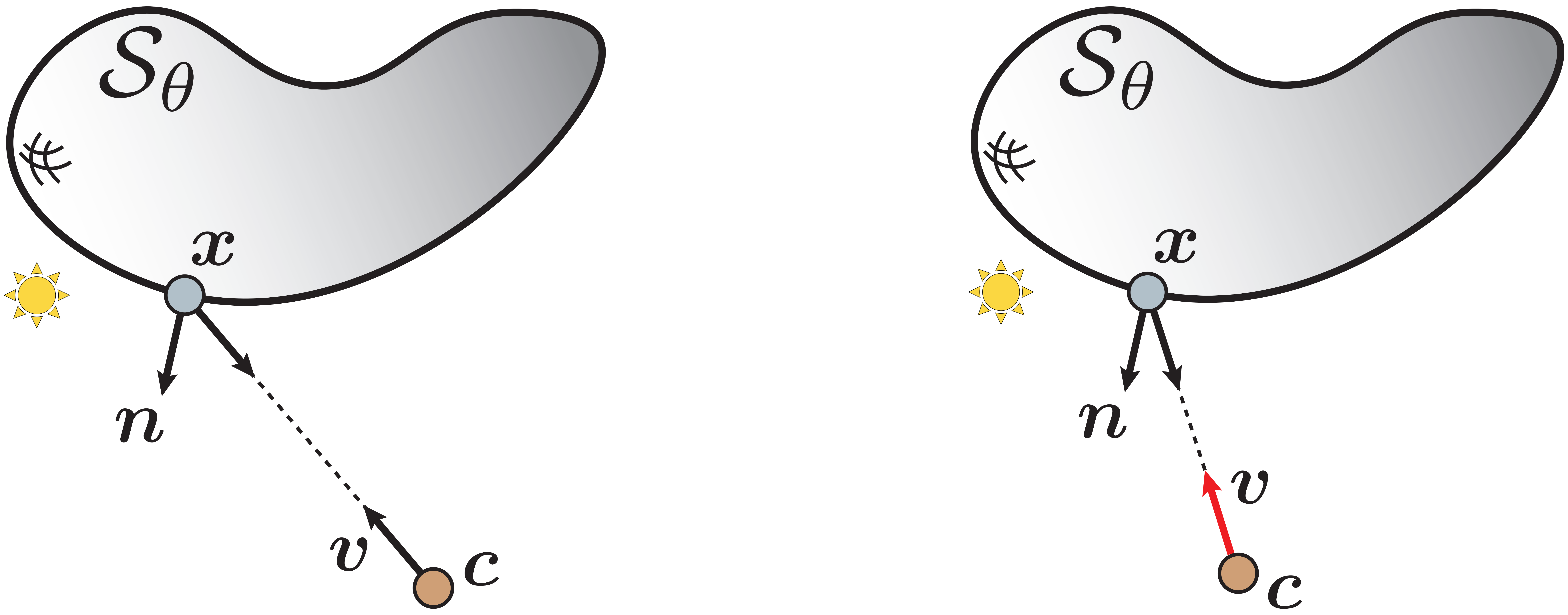
- Finding intersection point



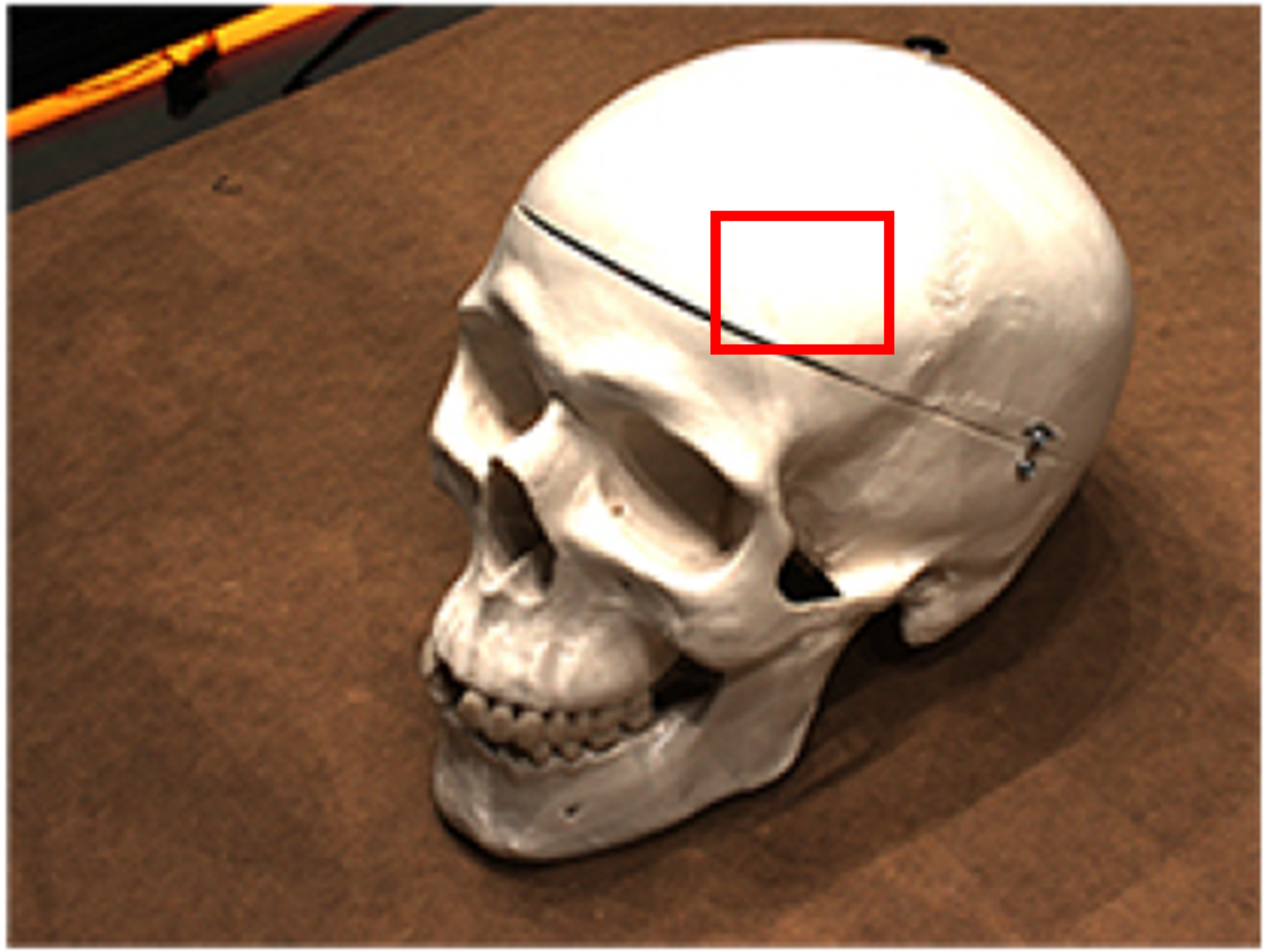
Neural Renderer

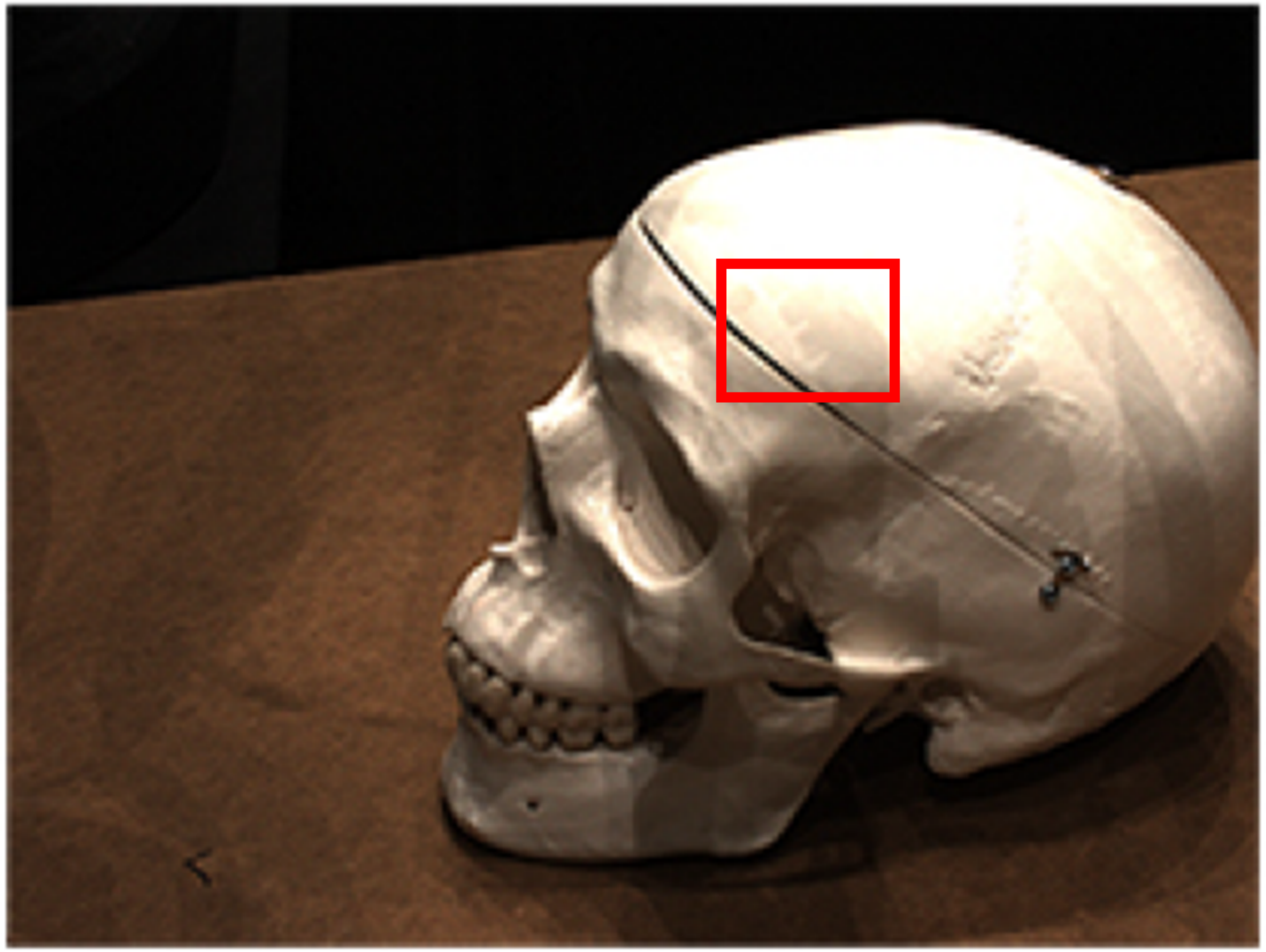


View dependent color

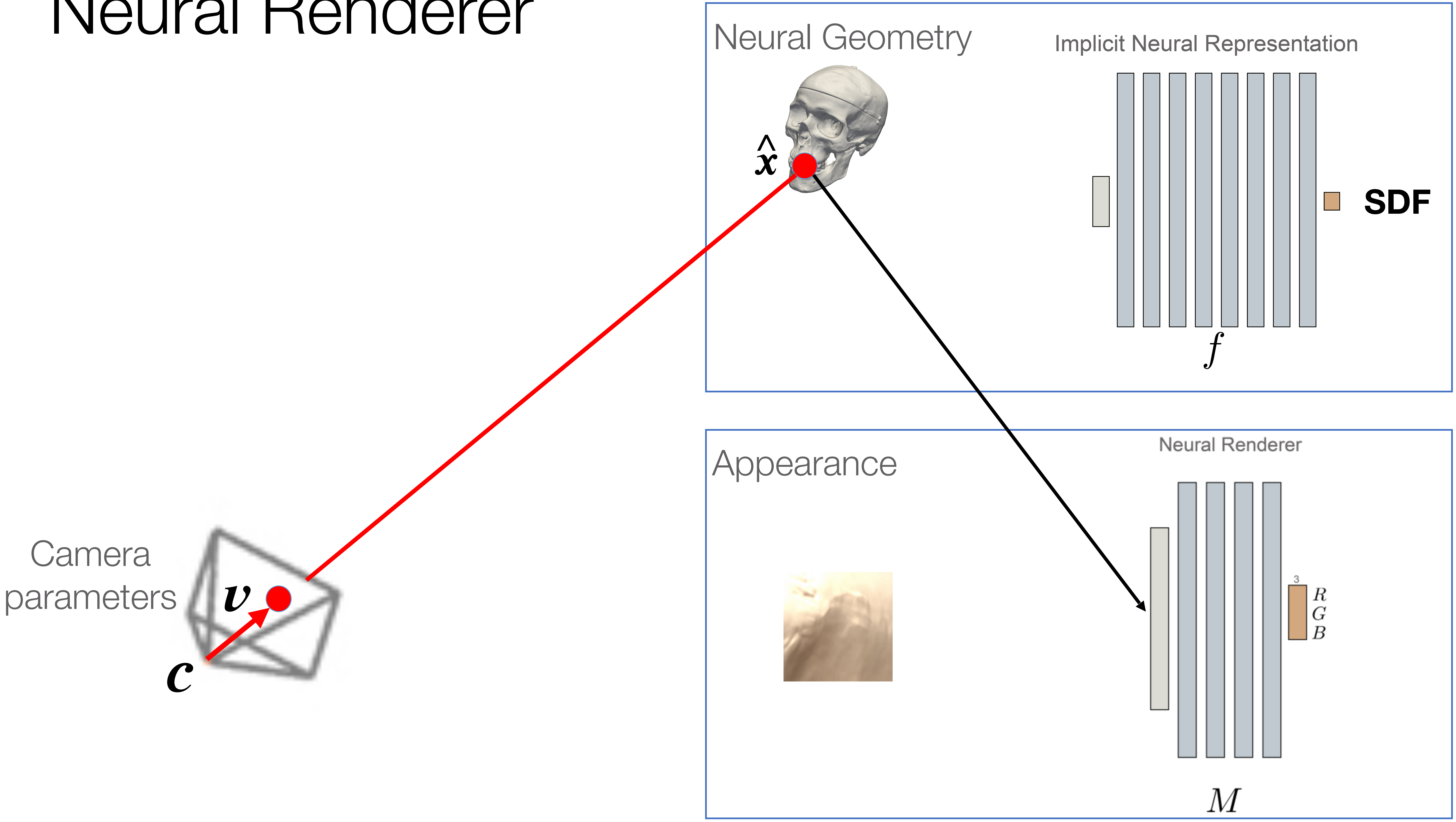


View dependent color

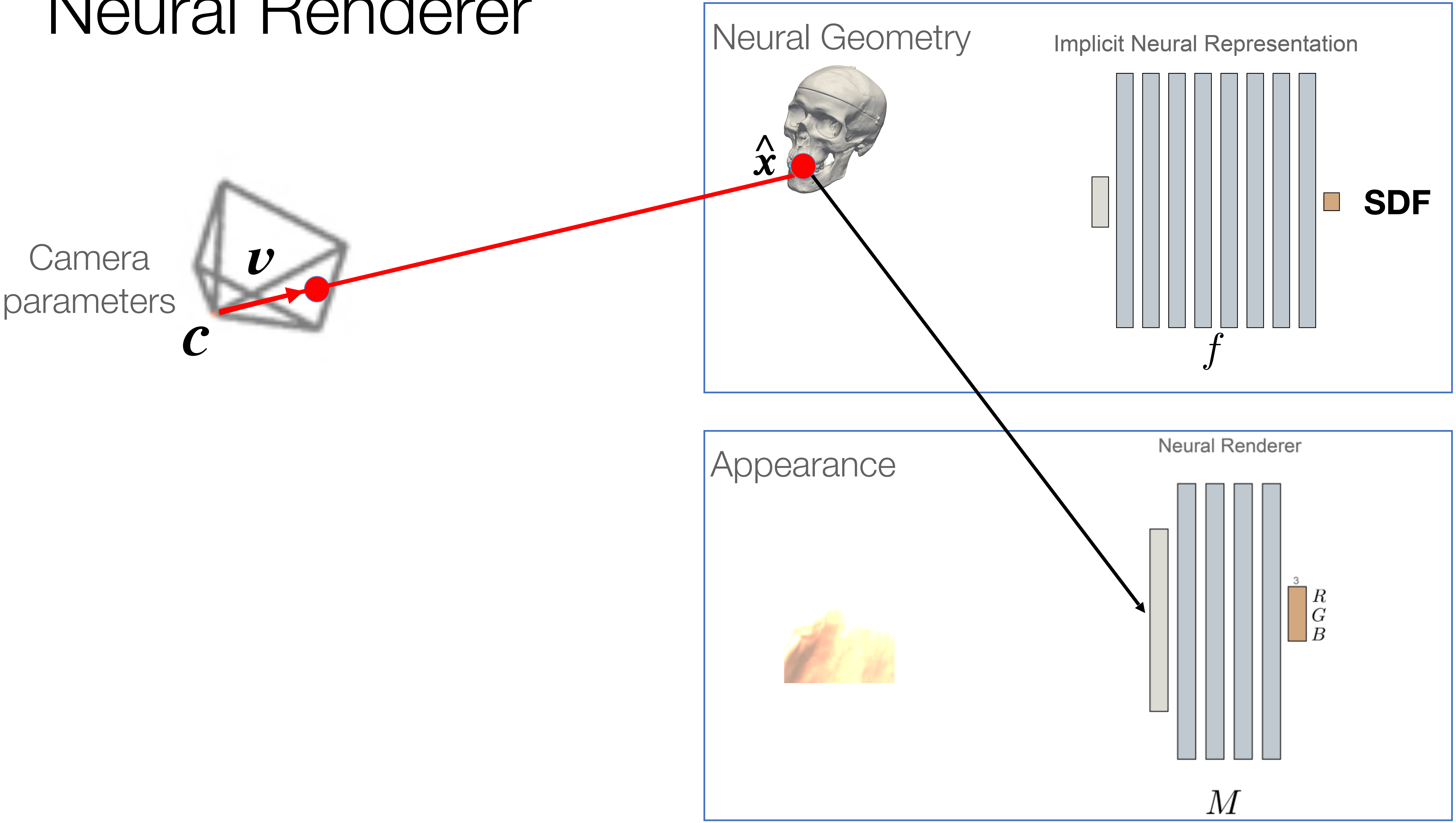




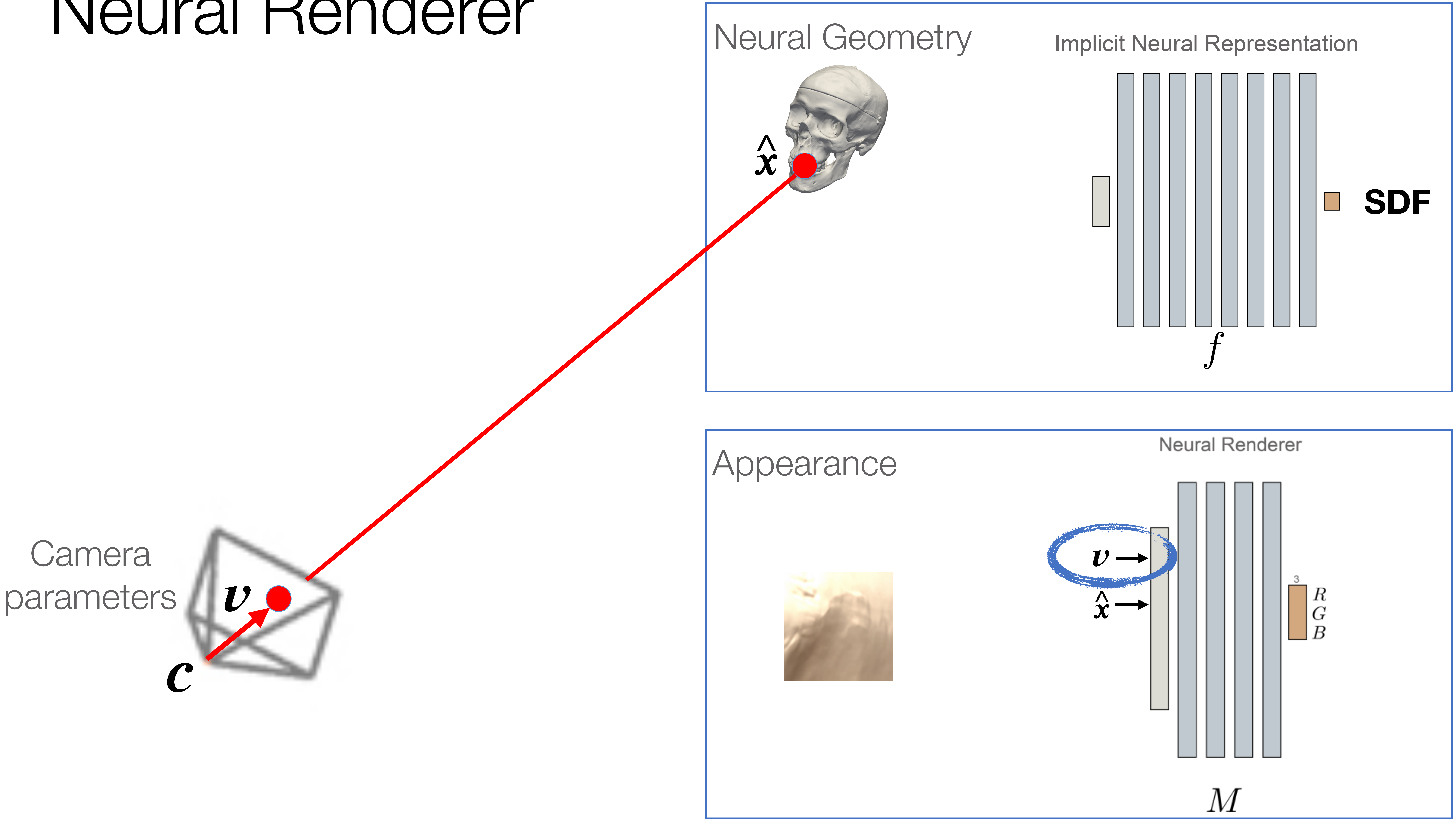
Neural Renderer



Neural Renderer

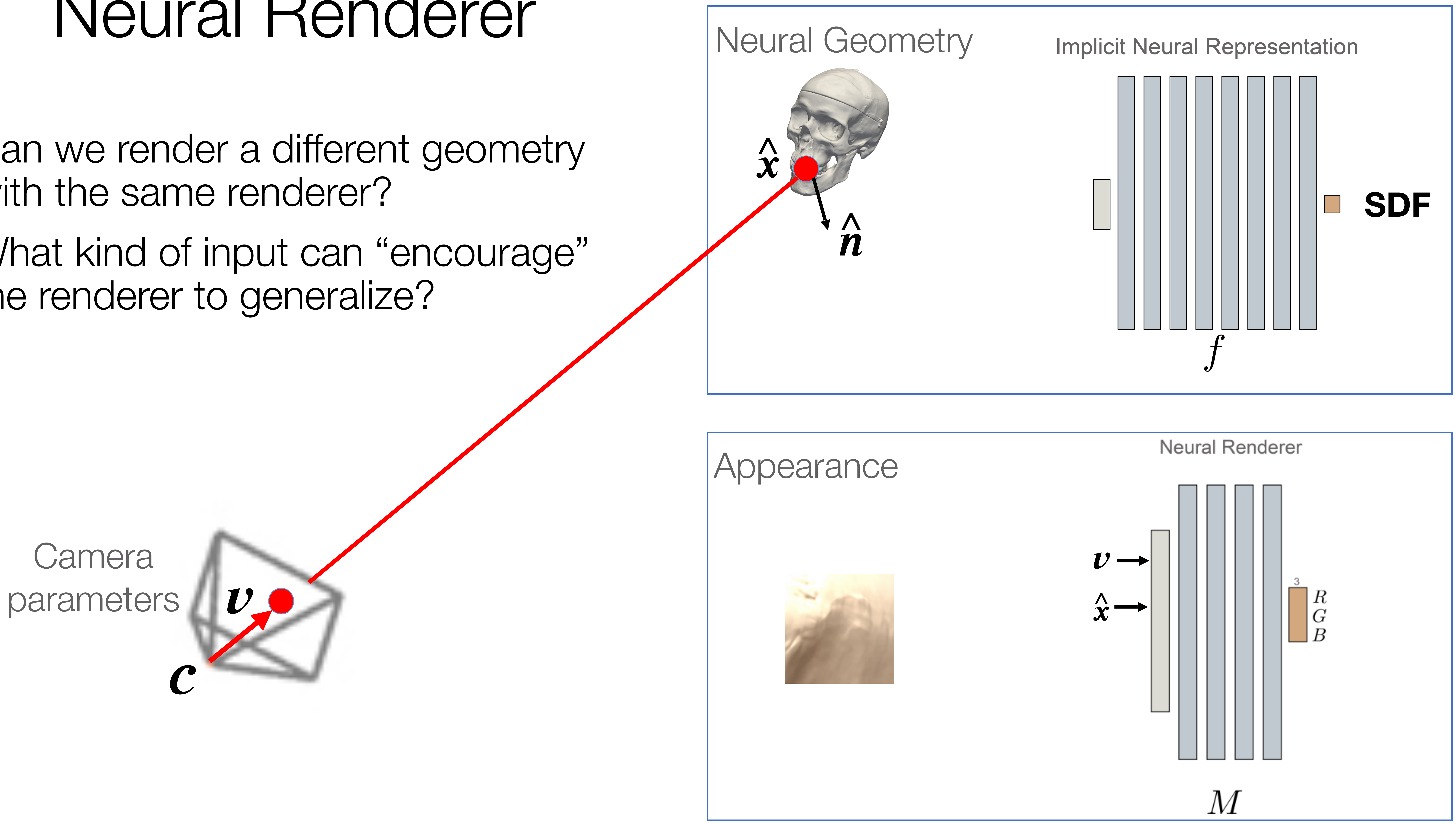


Neural Renderer

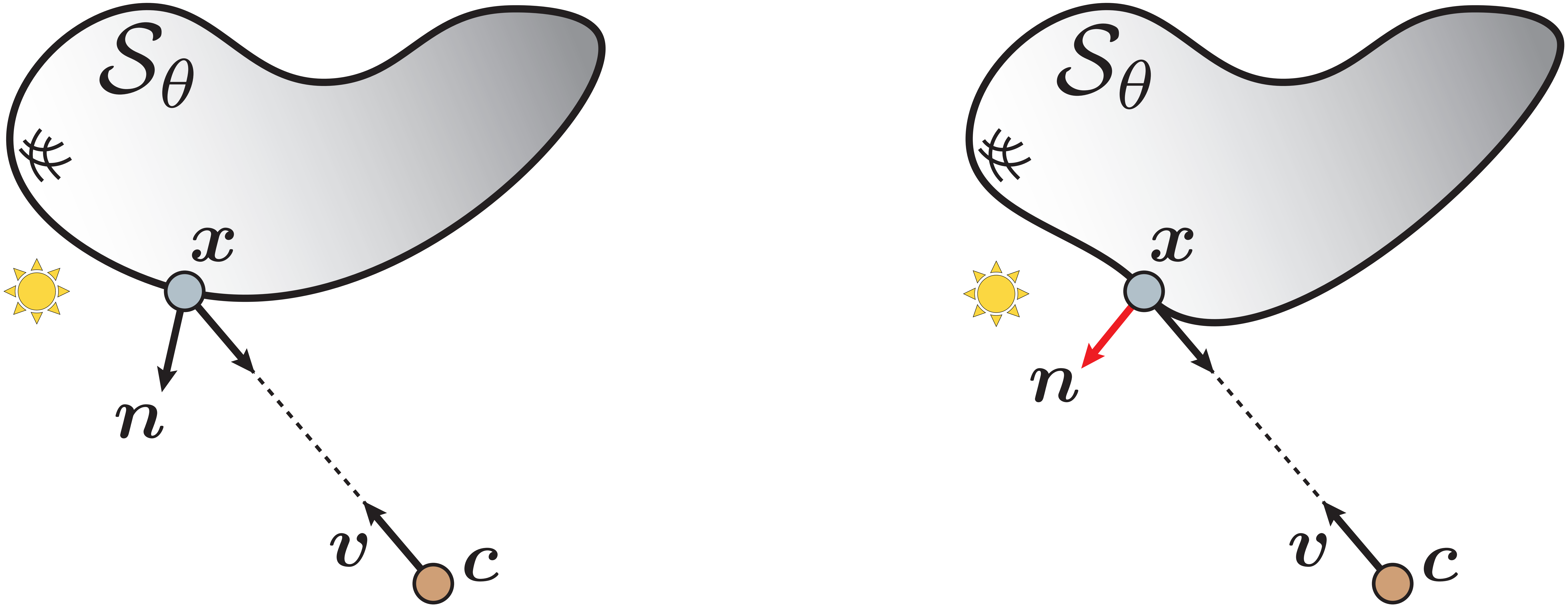


Neural Renderer

- Can we render a different geometry with the same renderer?
- What kind of input can “encourage” the renderer to generalize?

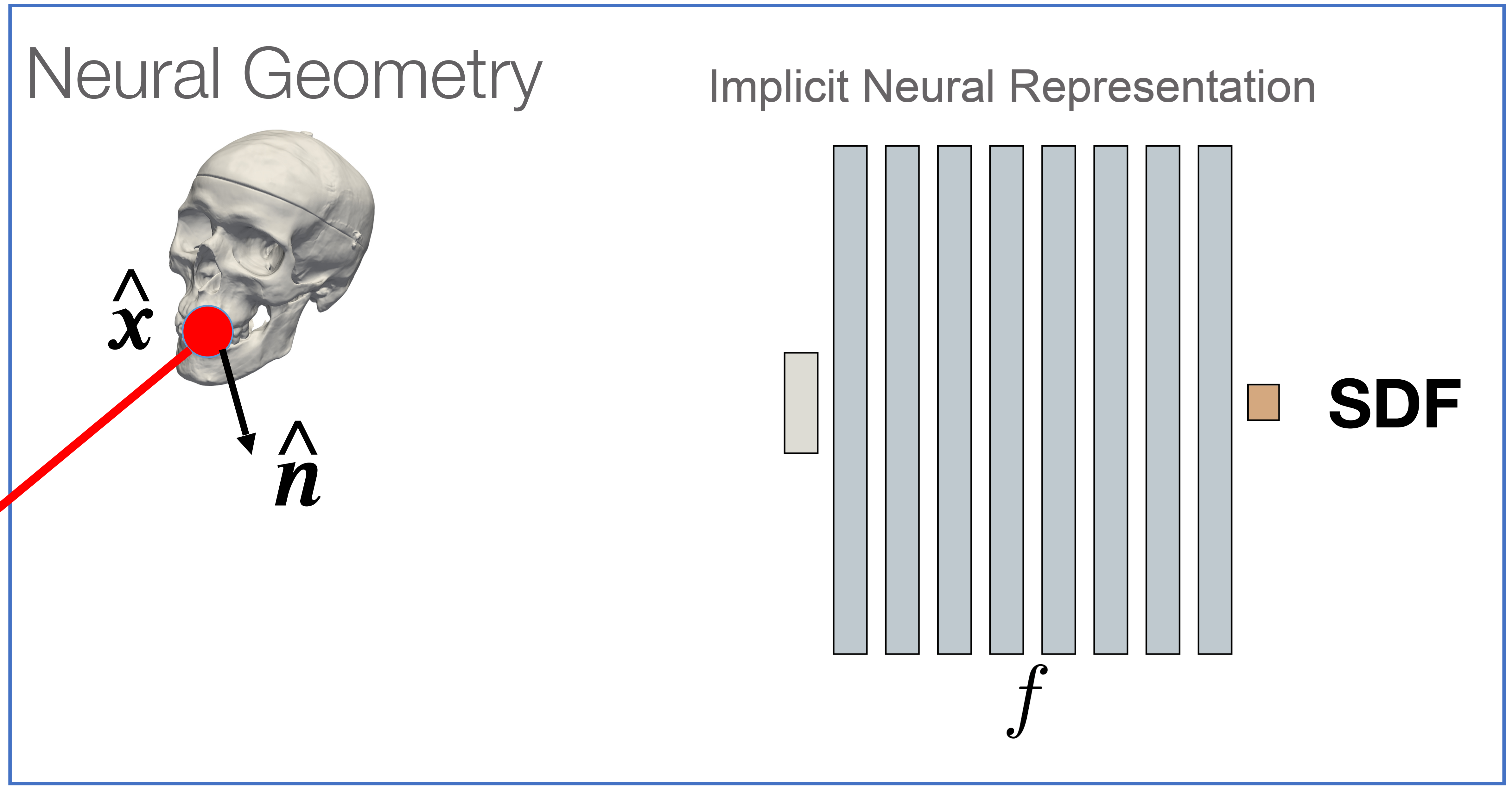


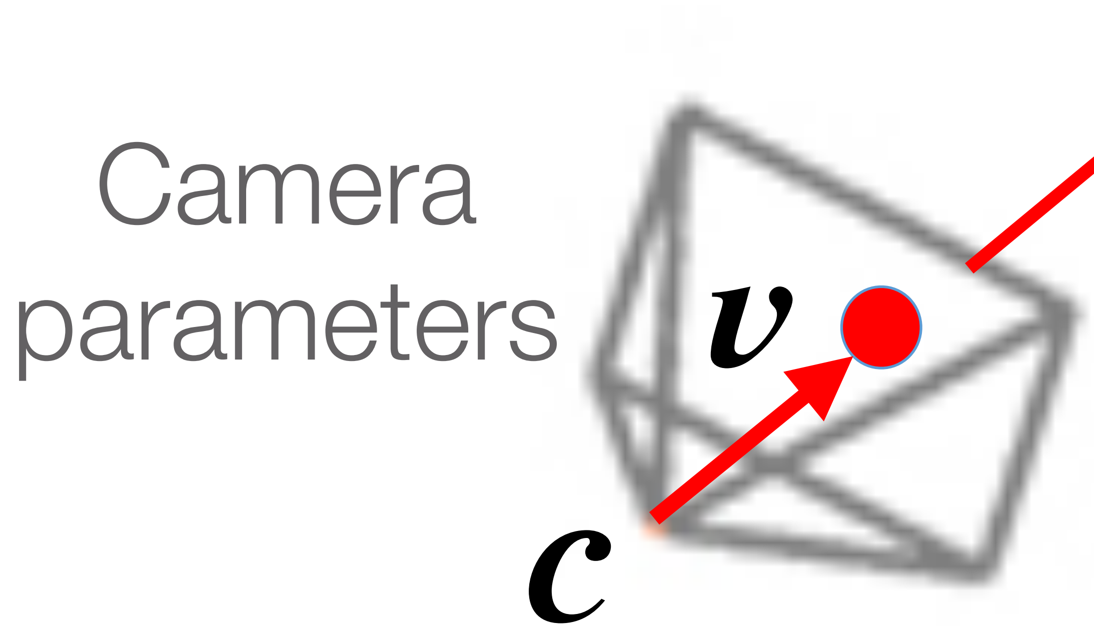
“Geometry” dependent color

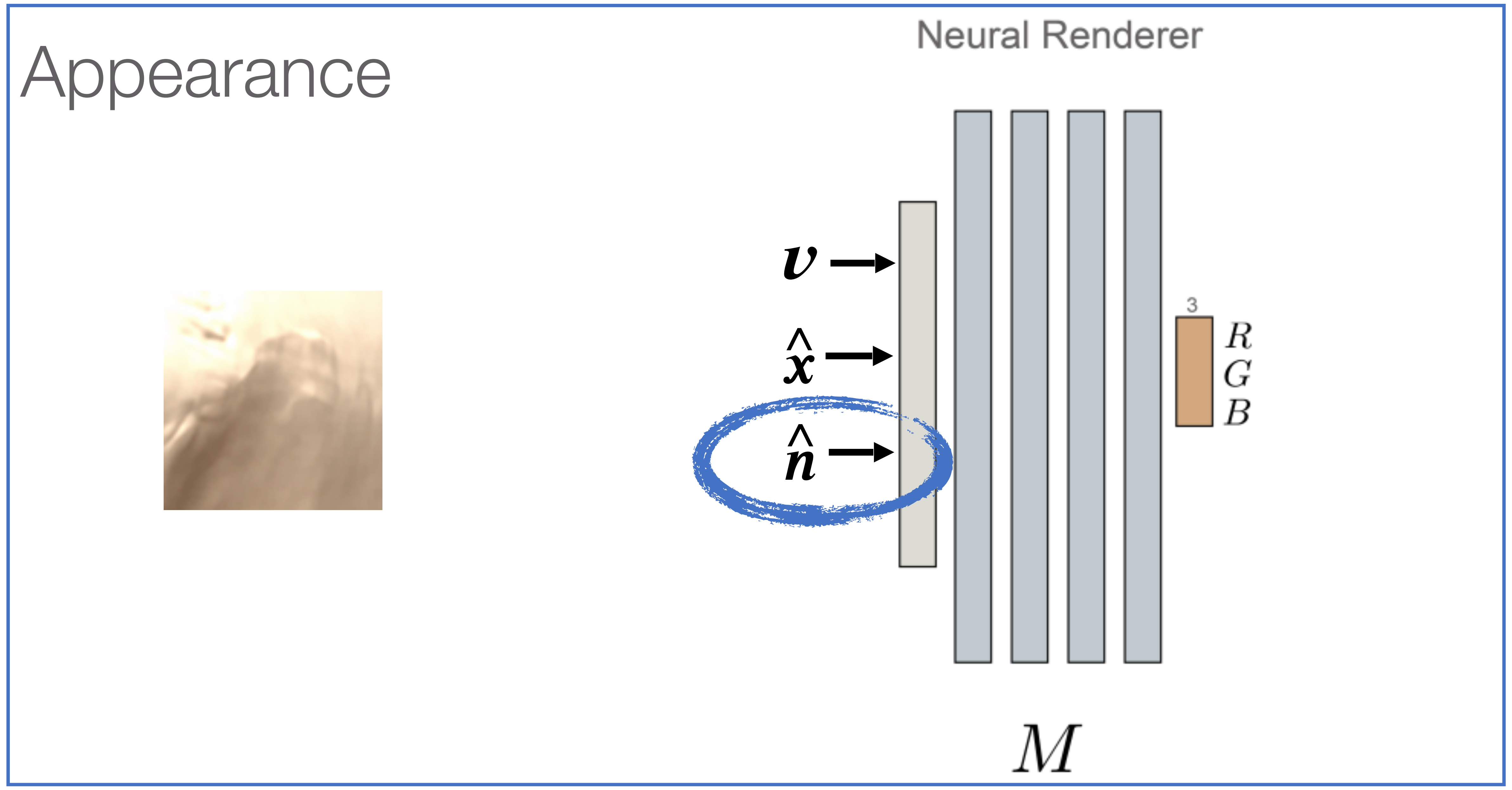


Neural Renderer

- Can we render a different geometry with the same renderer?
- What kind of input can “encourage” the renderer to generalize?

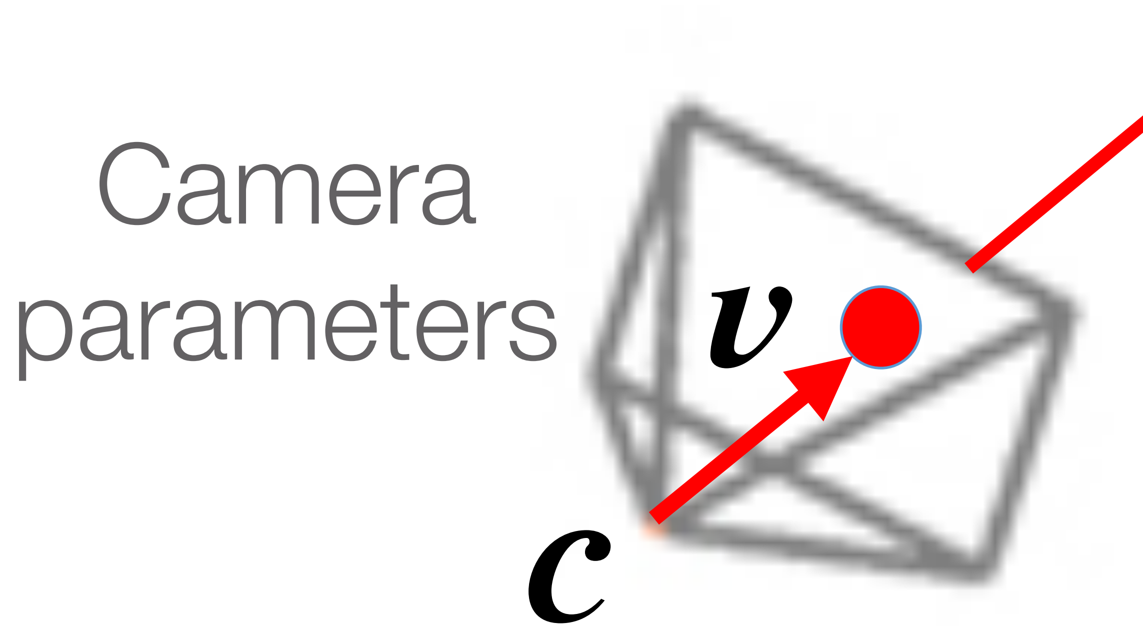


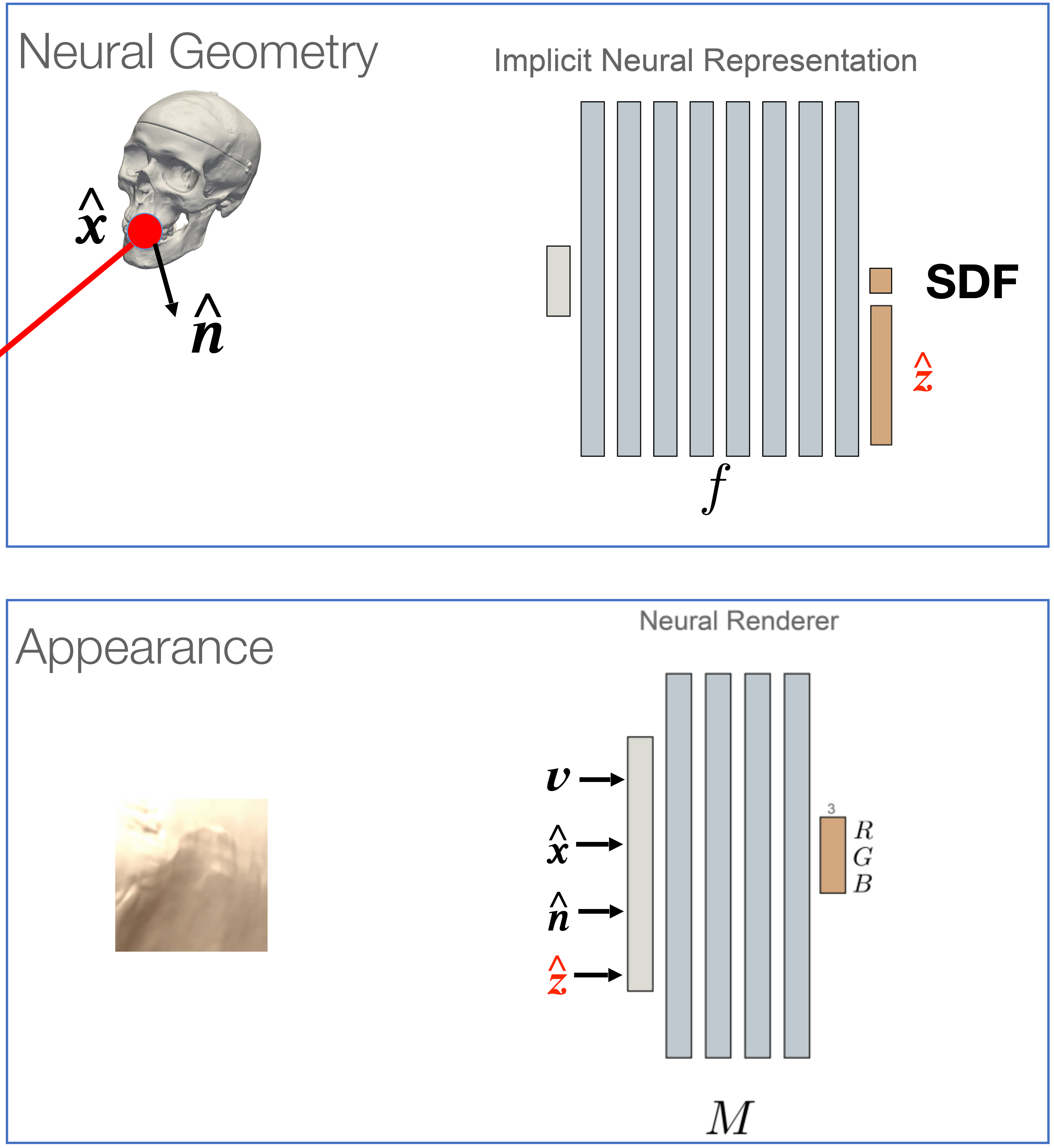




Neural Renderer

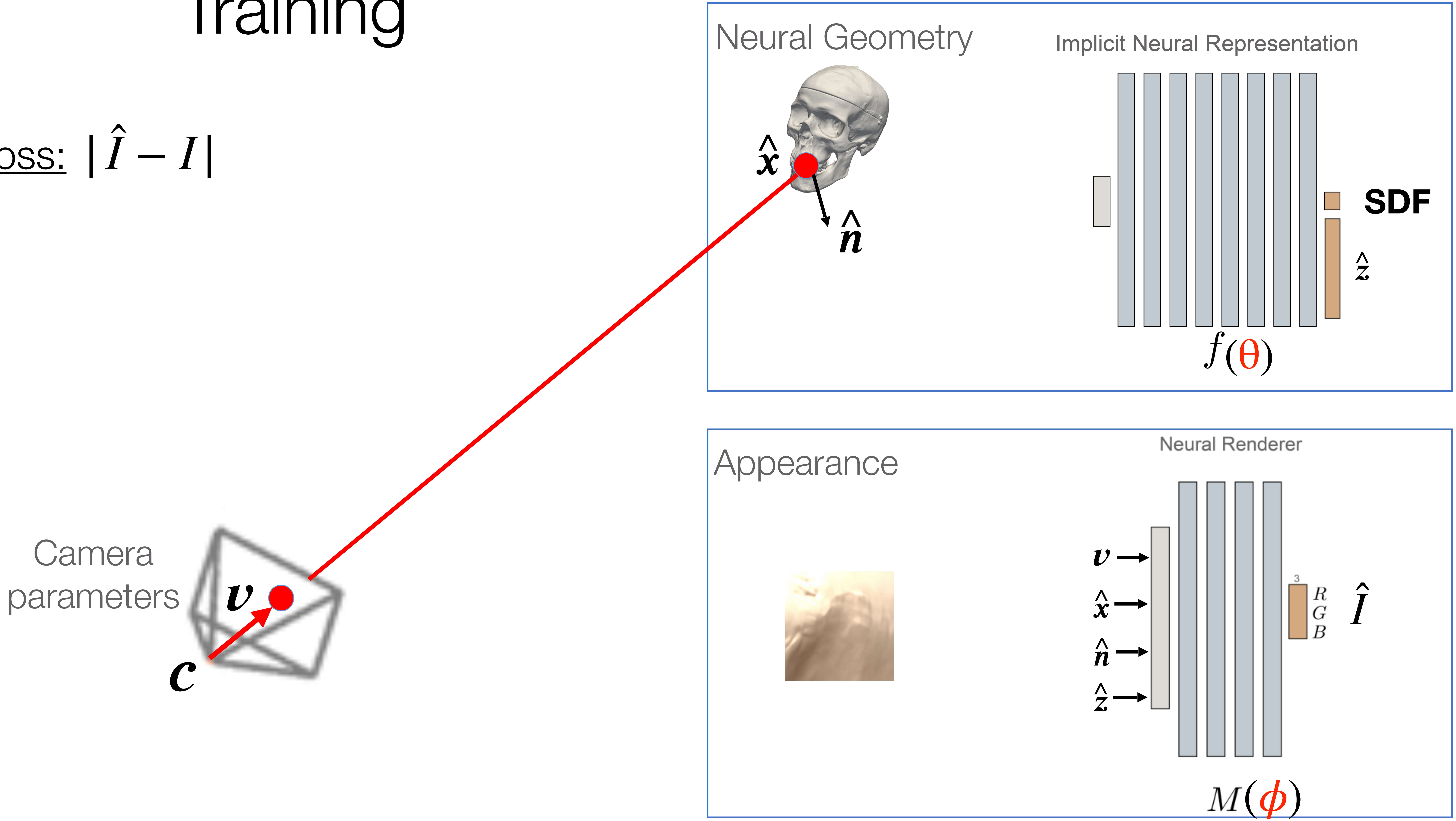
- Adding a global feature to allow secondary lighting effects and self shadows





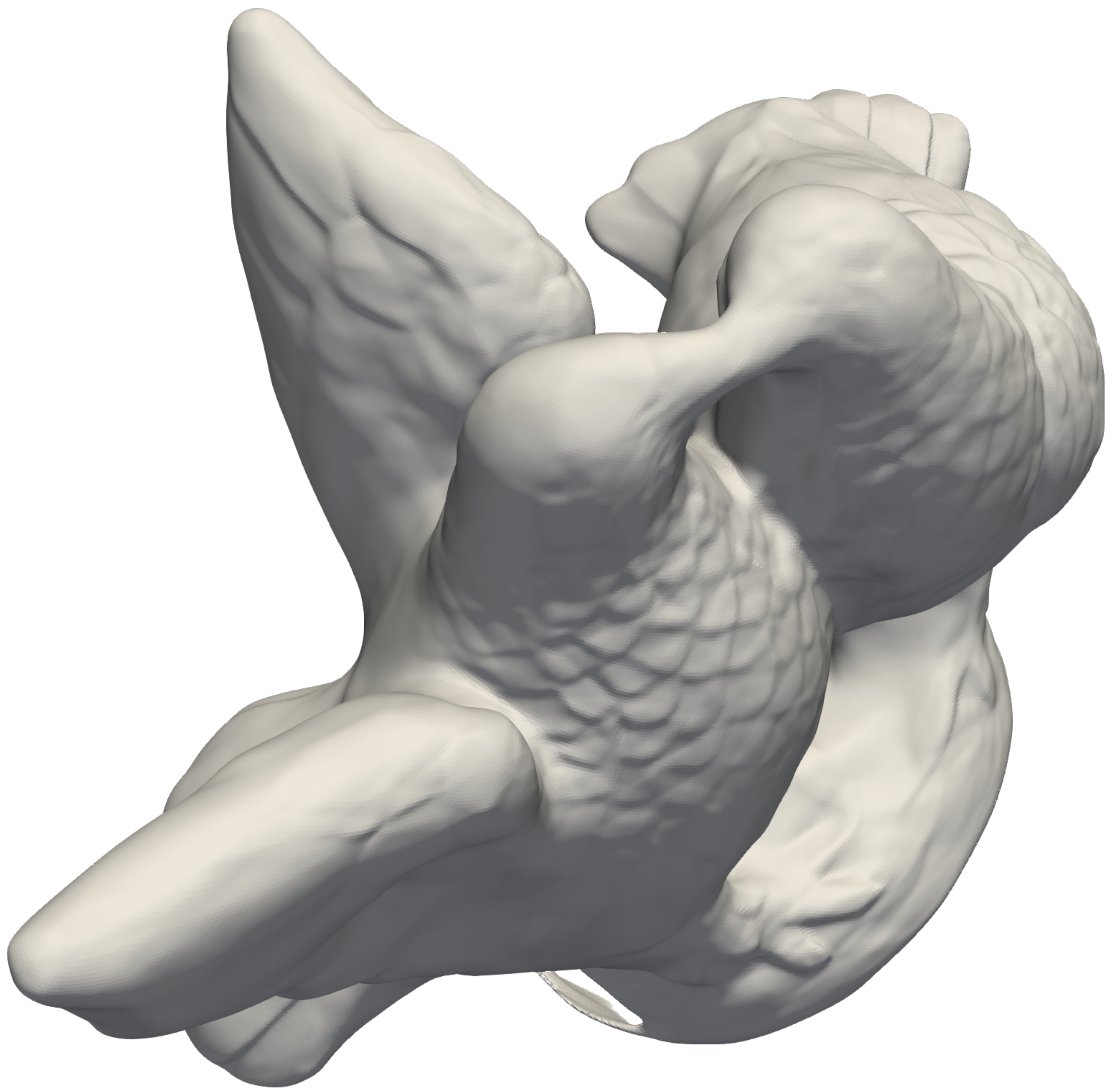
Training

- <u>Loss:</u> $|\hat{I} - I|$

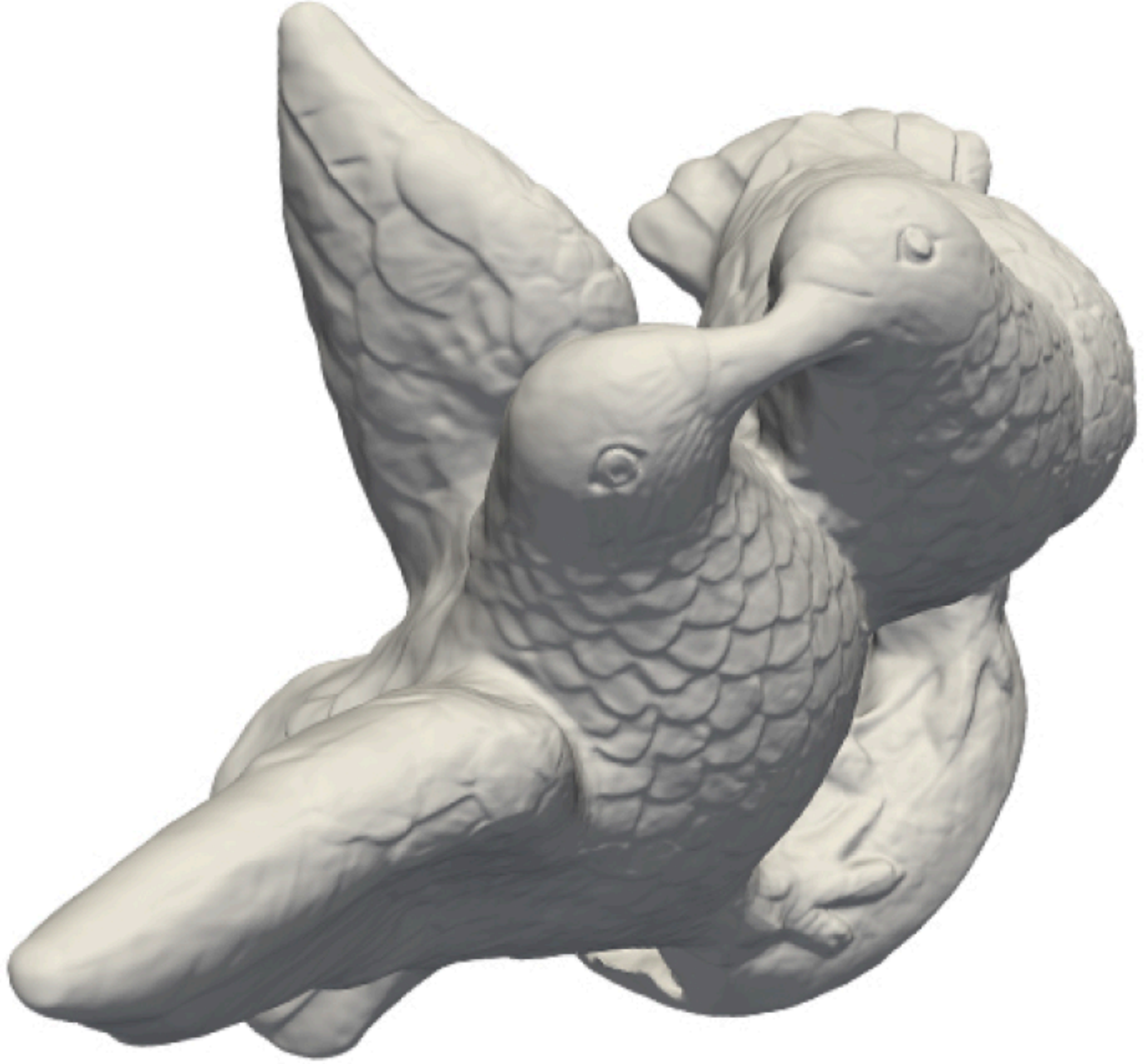


Positional encoding

$$\mathrm{PE}(\boldsymbol{y}) = \left(\sin(2^0\pi\boldsymbol{y}), \cos(2^0\pi\boldsymbol{y}), \cdots, \sin(2^{L-1}\pi\boldsymbol{y}), \cos(2^{L-1}\pi\boldsymbol{y})\right)$$



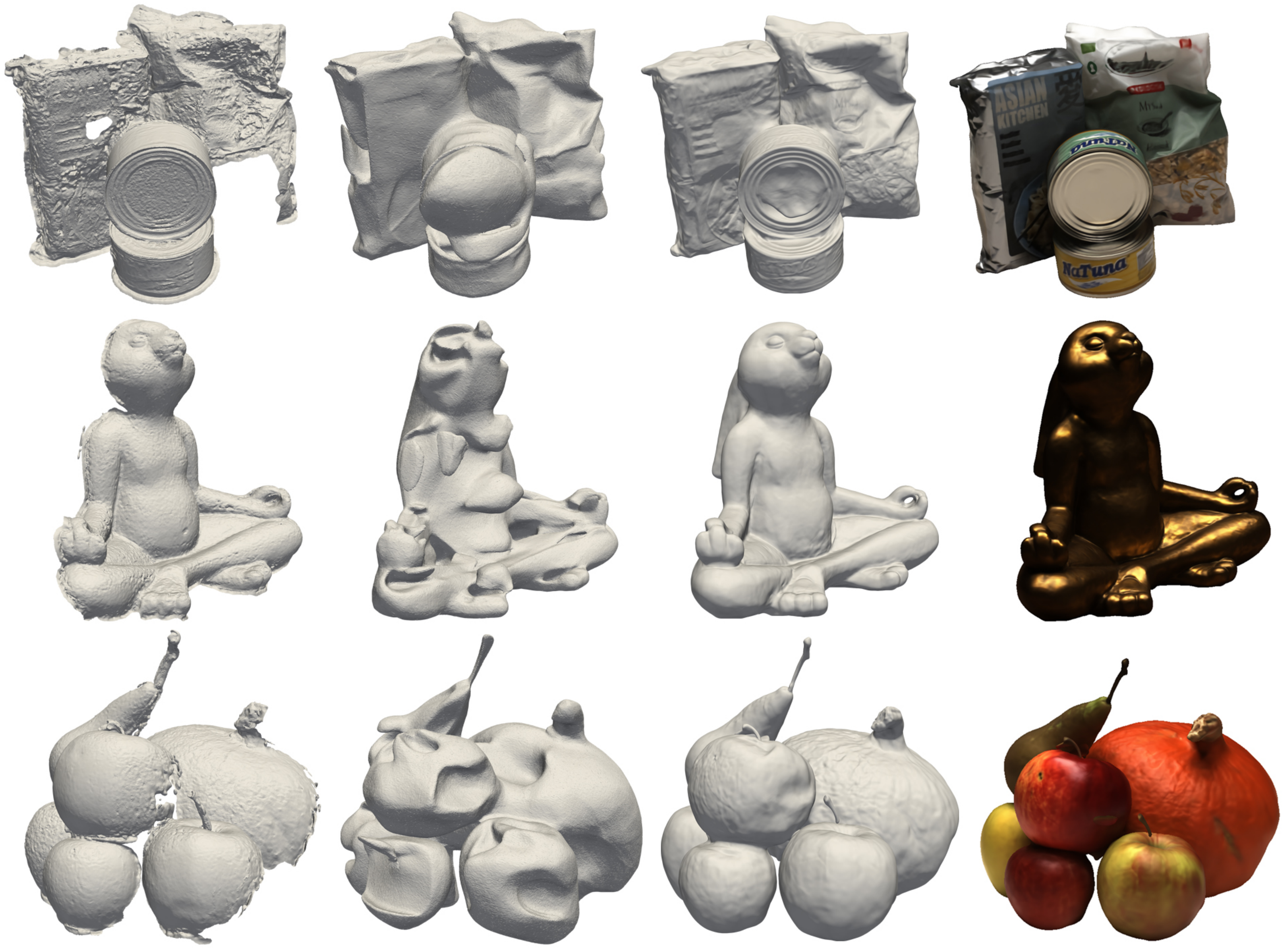
No PE, 5000 epochs



2000 epochs

NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, Mildenhall, Srinivasan, Tancik, et al., 2020

Results: comparisons



Colmap + sPSR | DVR $(\hat{x}, \cancel{\hat{n}}, \cancel{p})$ | IDR | IDR - rendering

Results: ablation study

Input images

$M(\hat{\boldsymbol{x}}, \cancel{\hat{\boldsymbol{n}}}, \boldsymbol{v}, \boldsymbol{z})$

$M(\hat{\boldsymbol{x}}, \hat{\boldsymbol{n}}, \cancel{\boldsymbol{v}}, \boldsymbol{z})$

$M(\hat{\boldsymbol{x}}, \hat{\boldsymbol{n}}, \boldsymbol{v}, \cancel{\boldsymbol{z}})$

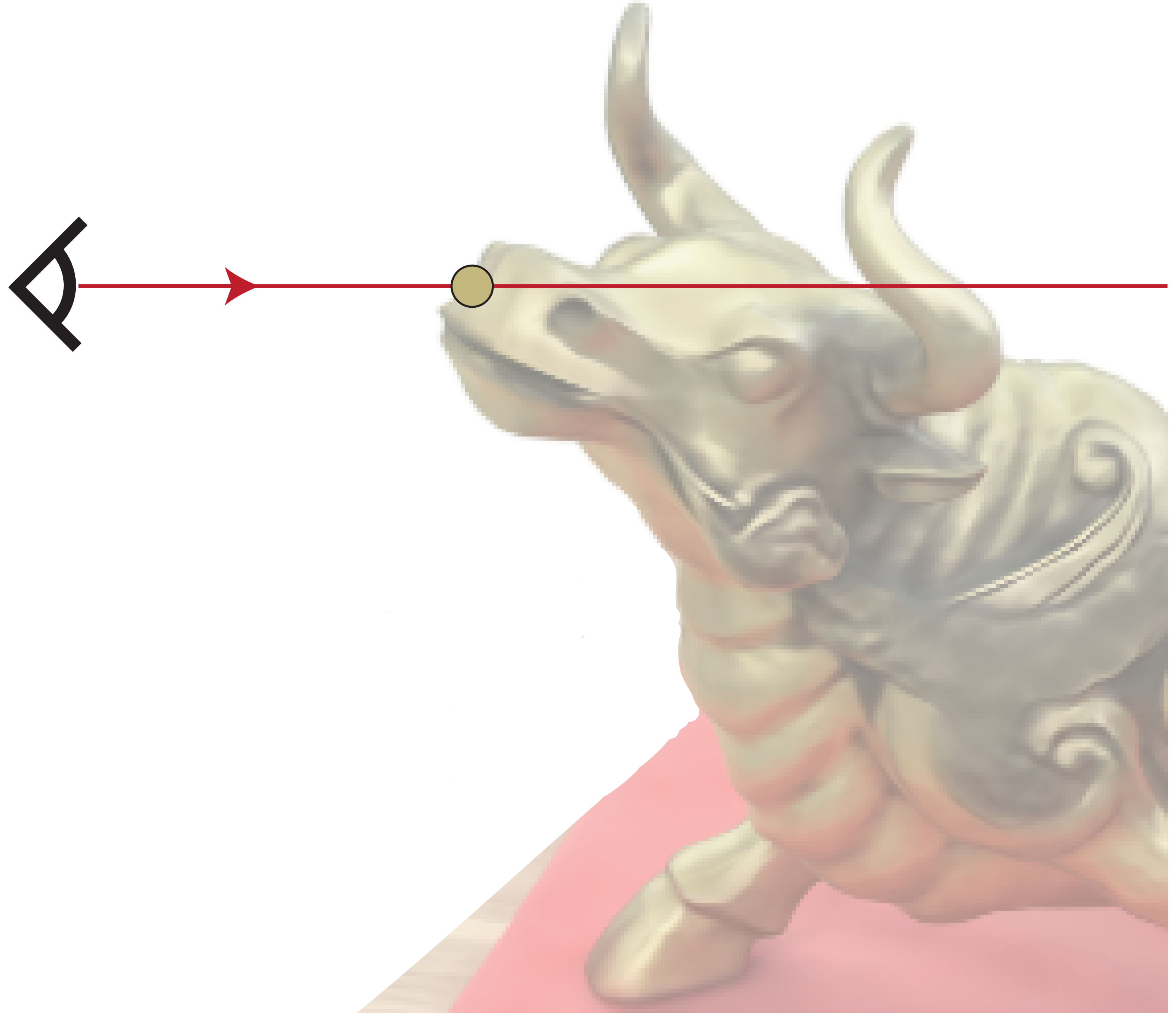
$M(\hat{\boldsymbol{x}}, \hat{\boldsymbol{n}}, \boldsymbol{v}, \boldsymbol{z})$

Surface Rendering v.s. Volume Rendering

Surface Rendering:

DVR [Niemeyer et al. '20]

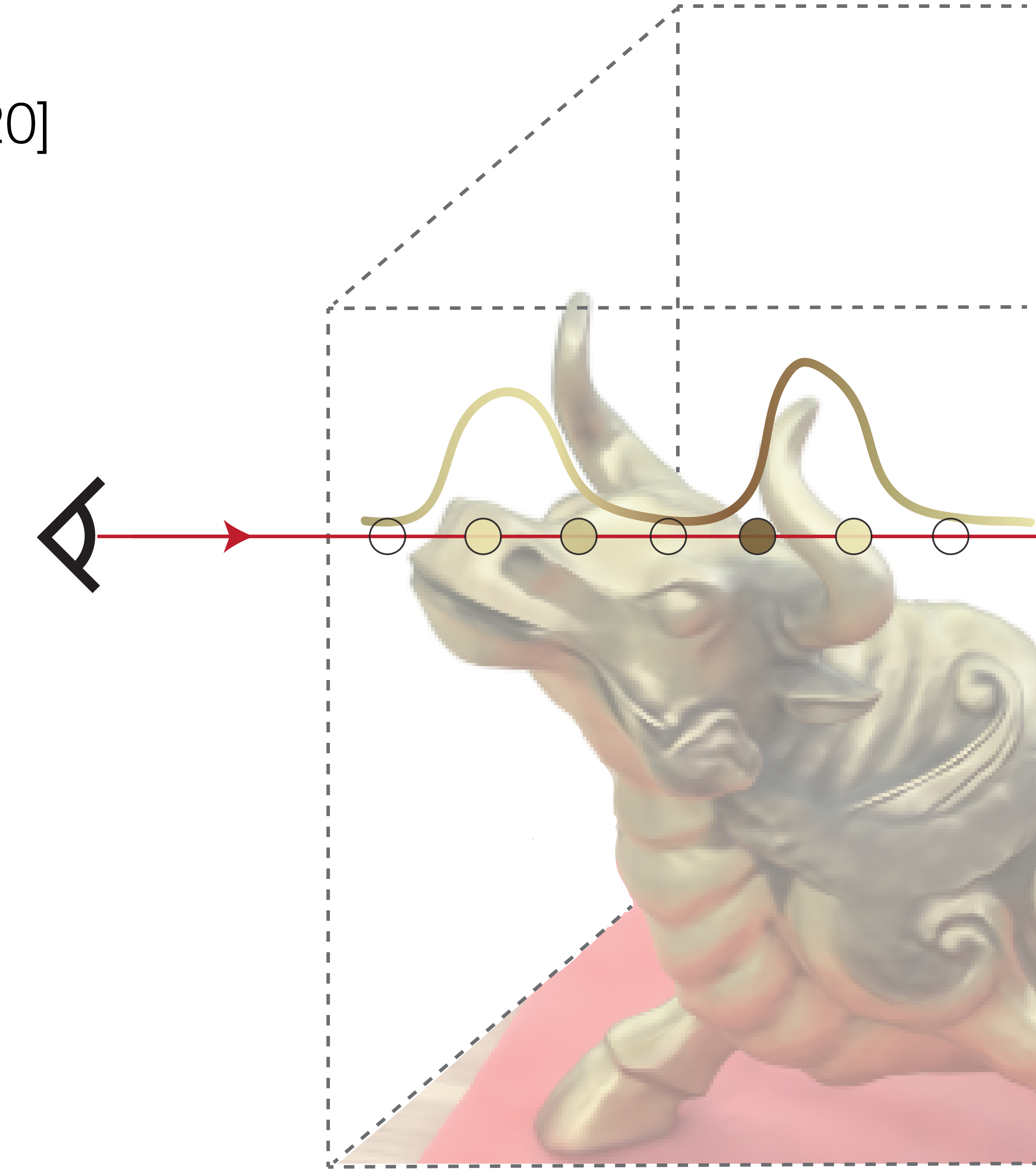
IDR [Yariv et al. '20]



Volume Rendering:

NeRF [Mildenhall et al. '20]

…



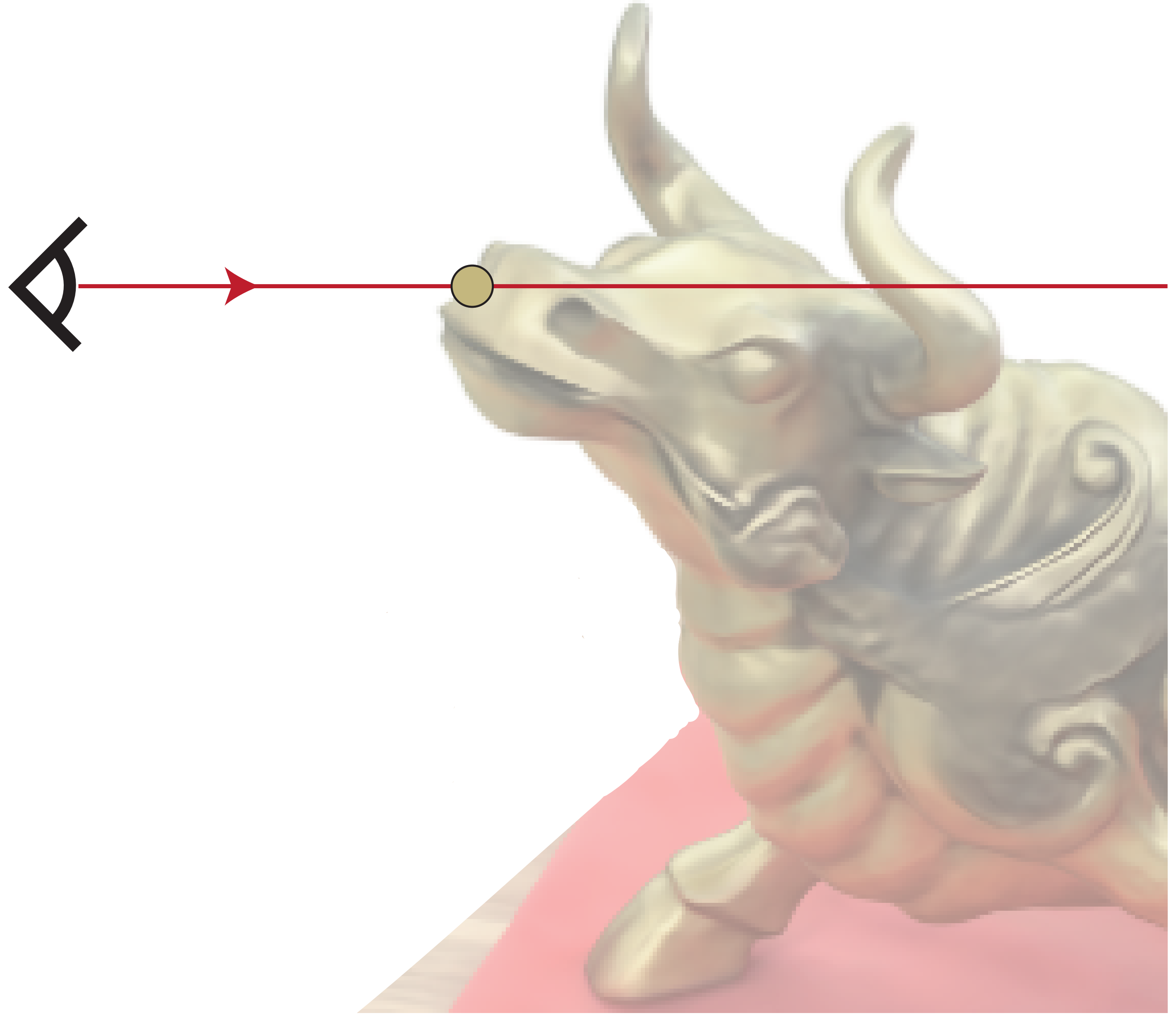
Surface Rendering v.s. Volume Rendering

Surface Rendering:

- Representation: Implicit surface

DVR [Niemeyer et al. ’20]

IDR [Yariv et al. ’20]

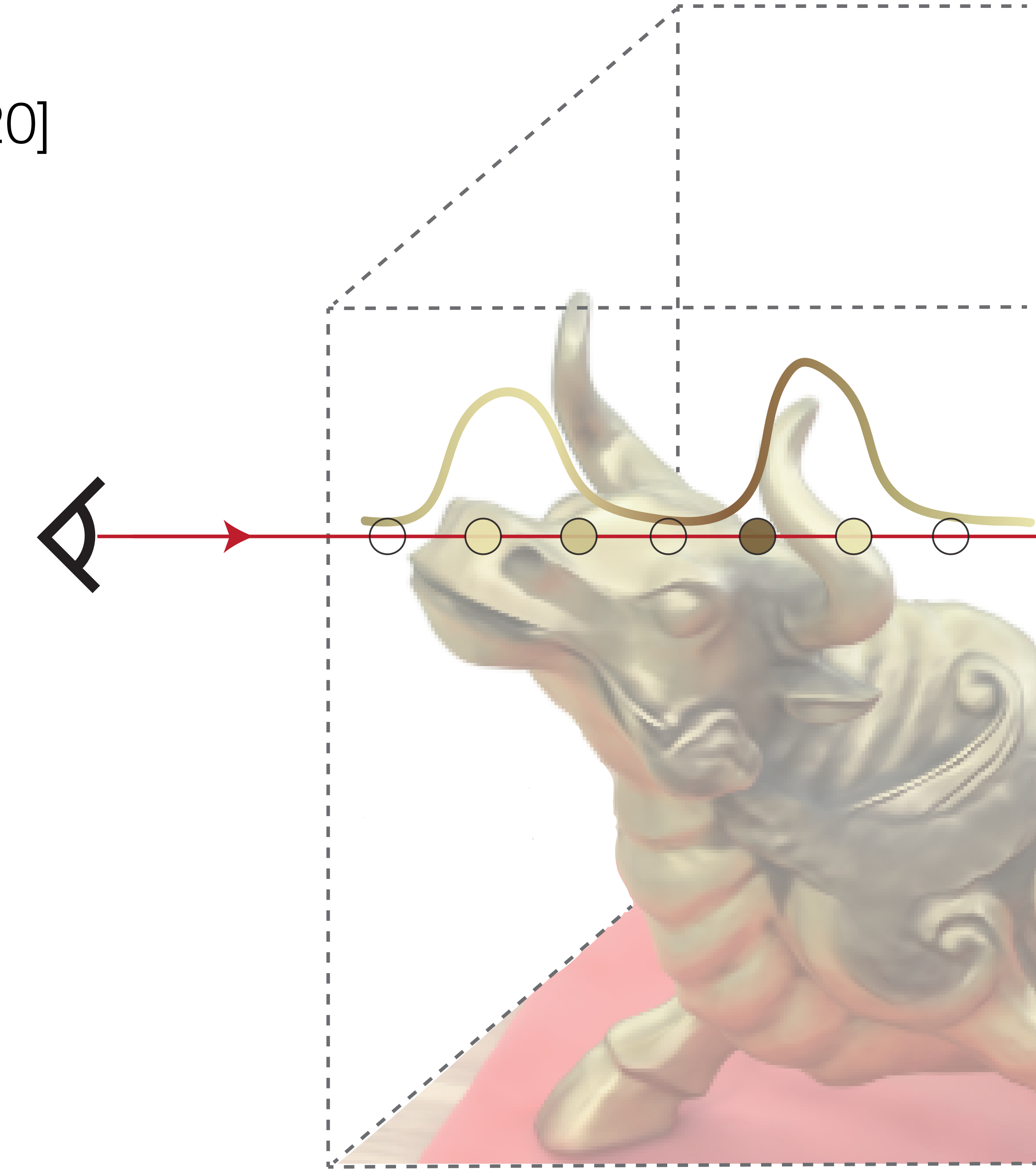


Volume Rendering:

- Representation: Volume density

NeRF [Mildenhall et al. ’20]

…



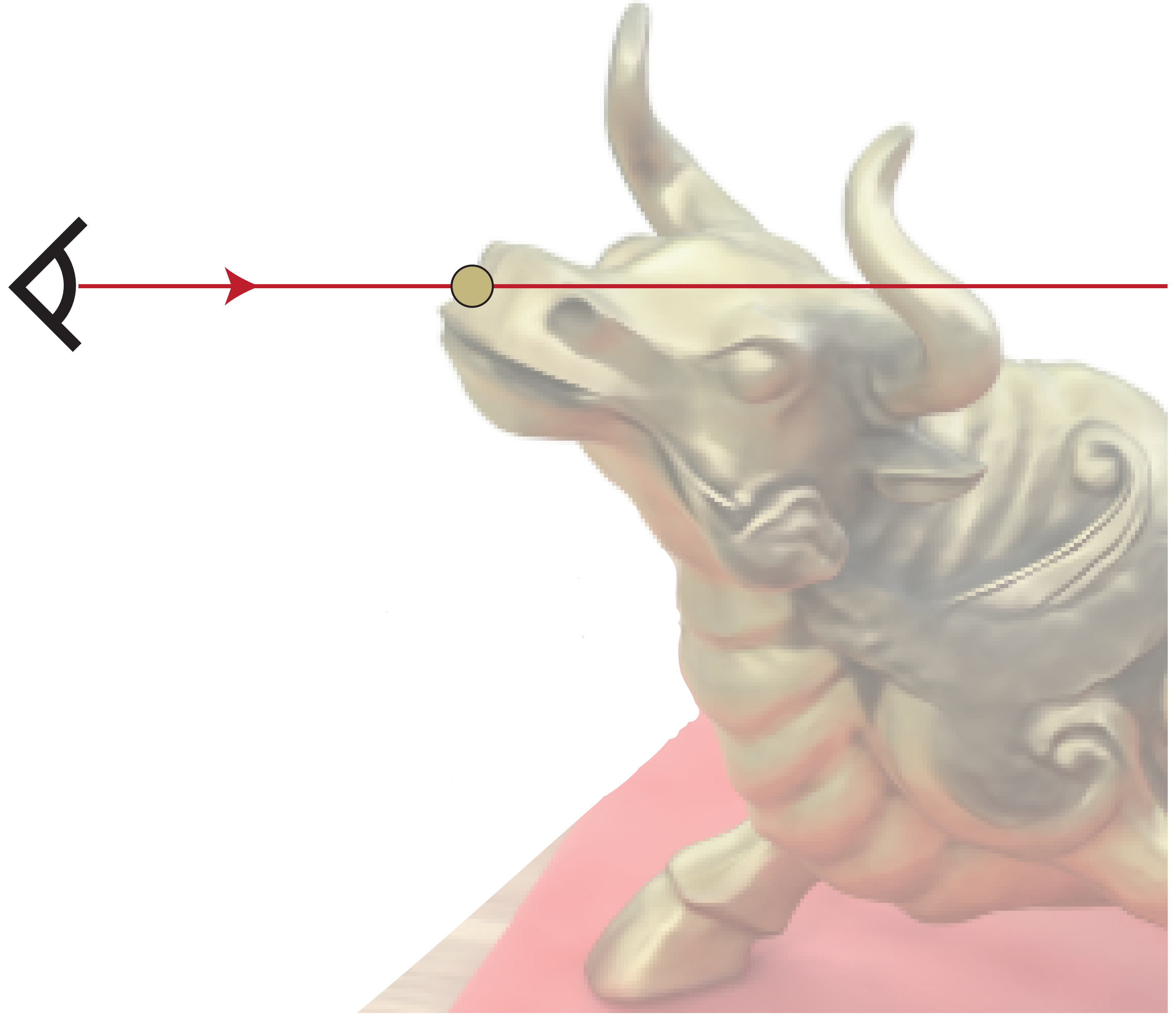
Surface Rendering v.s. Volume Rendering

Surface Rendering:

- Representation: Implicit surface
- Rendering: Find intersection

DVR [Niemeyer et al. '20]

IDR [Yariv et al. '20]

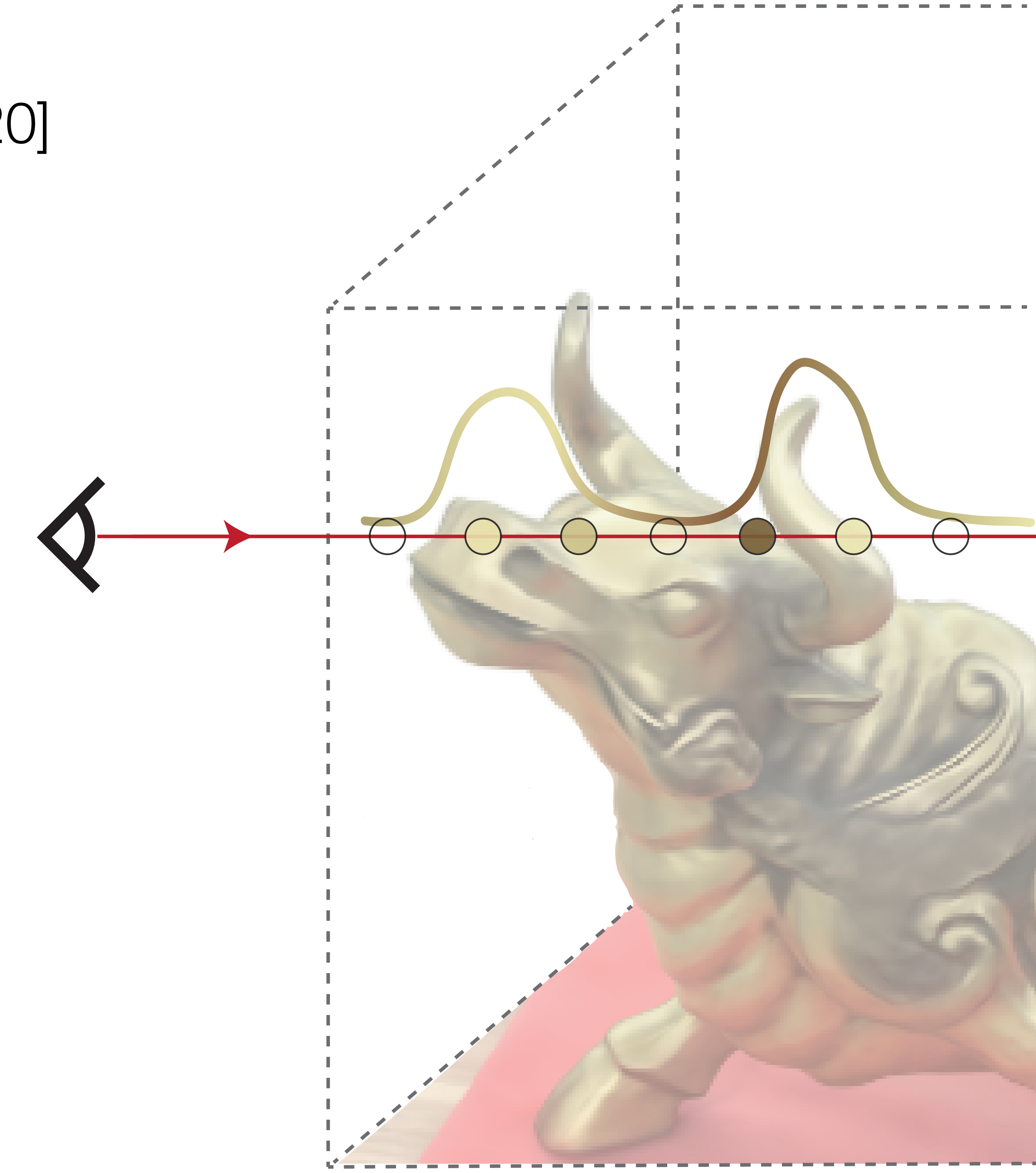


Volume Rendering:

- Representation: Volume density
- Rendering: Integral approximation

NeRF [Mildenhall et al. '20]

…



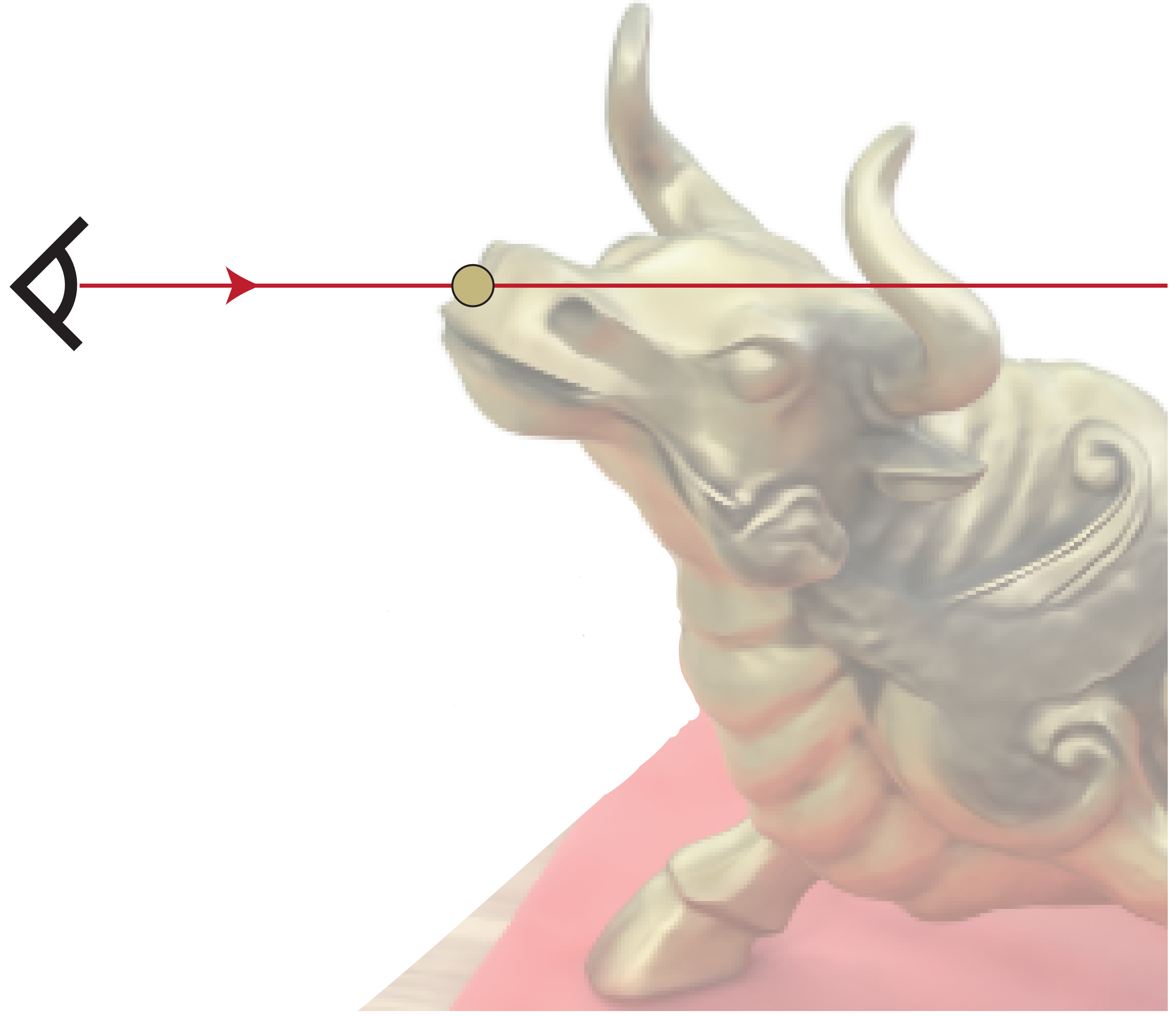
Surface Rendering v.s. Volume Rendering

<u>Surface Rendering:</u>

- Representation: Implicit surface
- Rendering: Find intersection
- Back-propagate: 1 sample

DVR [Niemeyer et al. '20]

IDR [Yariv et al. '20]

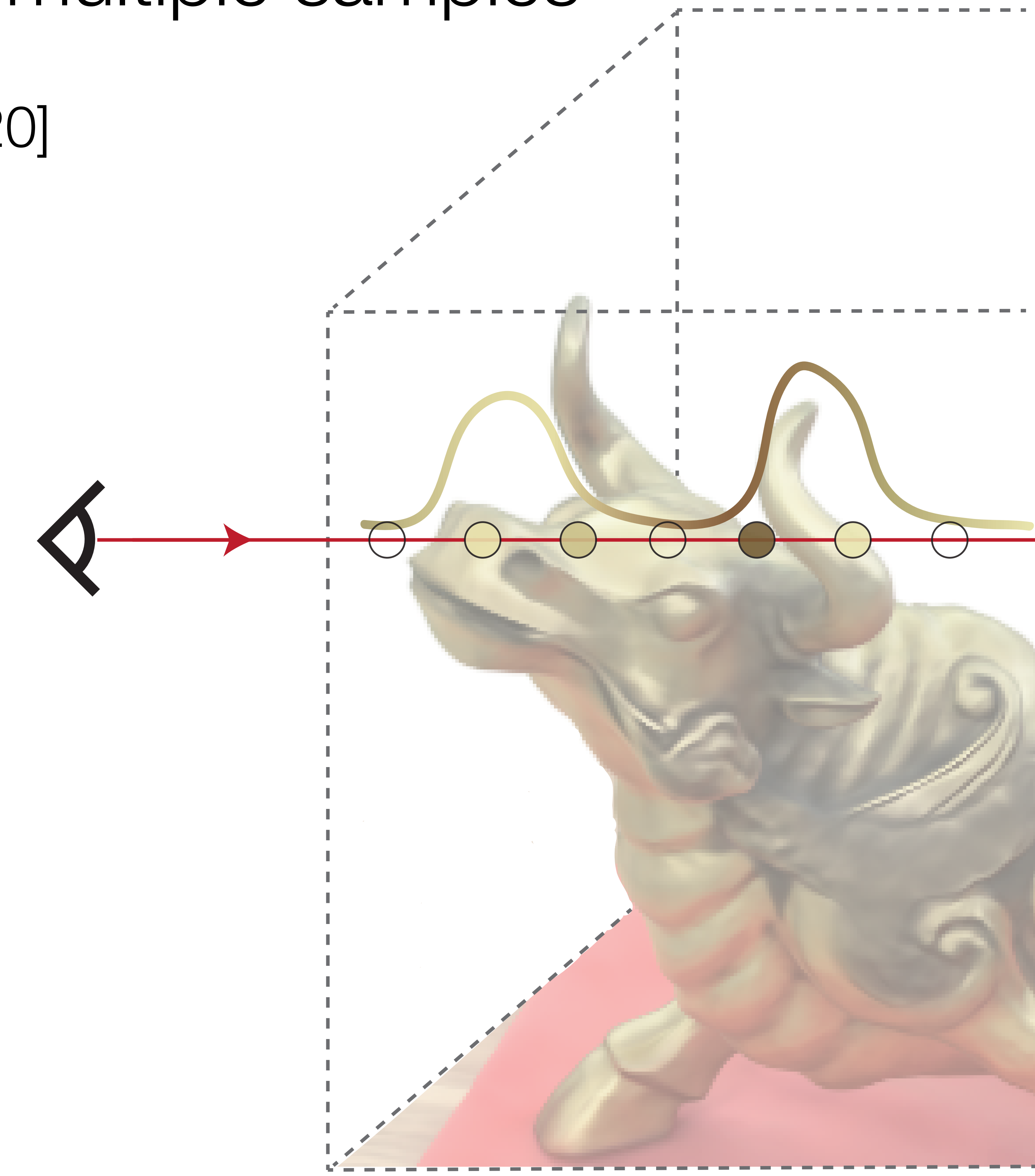


<u>Volume Rendering:</u>

- Representation: Volume density
- Rendering: Integral approximation
- Back-propagate: multiple samples

NeRF [Mildenhall et al. '20]

…



Surface Rendering v.s. Volume Rendering

<u>Surface Rendering:</u>

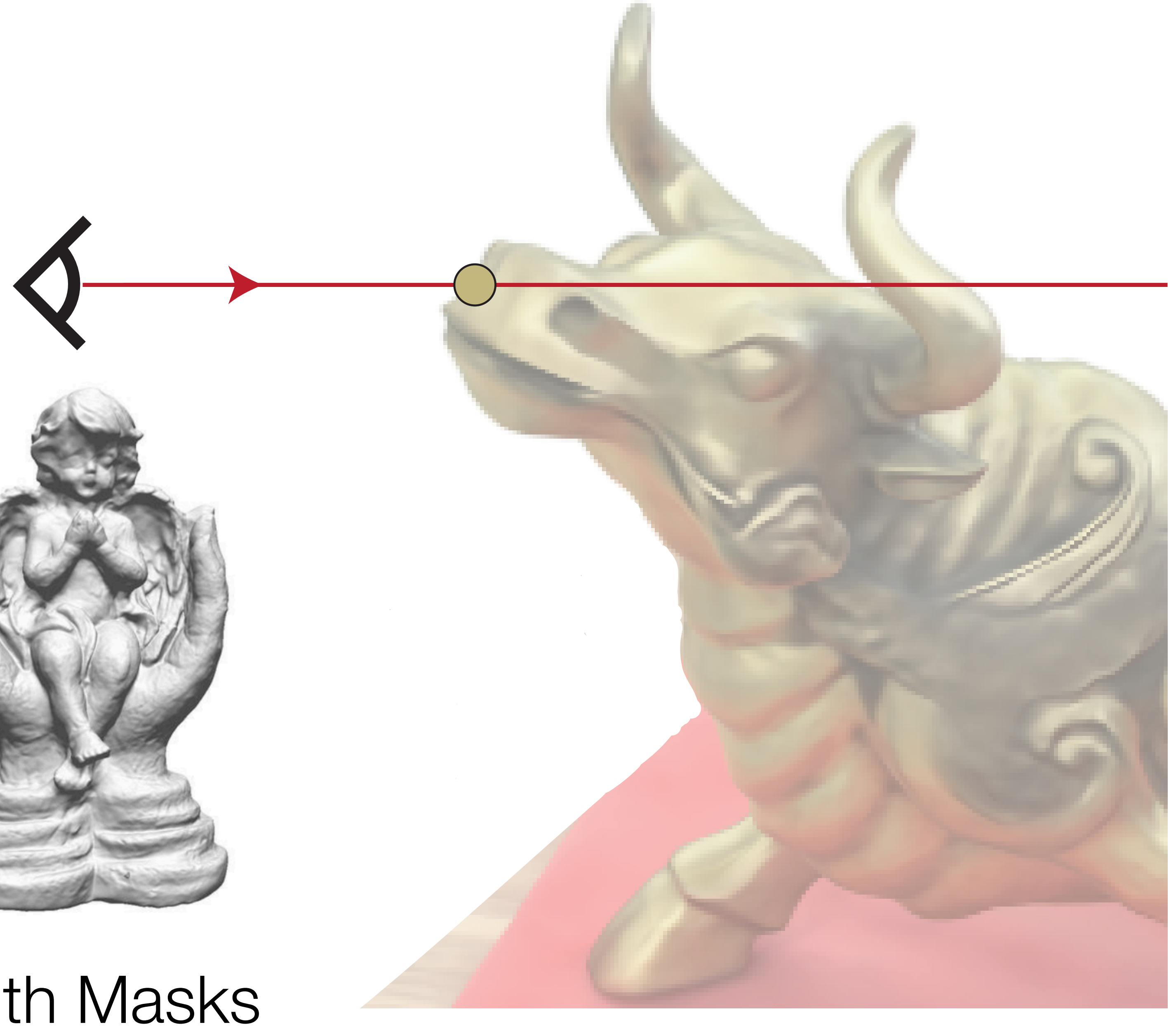
- Representation: Implicit surface
- Rendering: Find intersection
- Back-propagate: 1 sample

DVR [Niemeyer et al. '20]

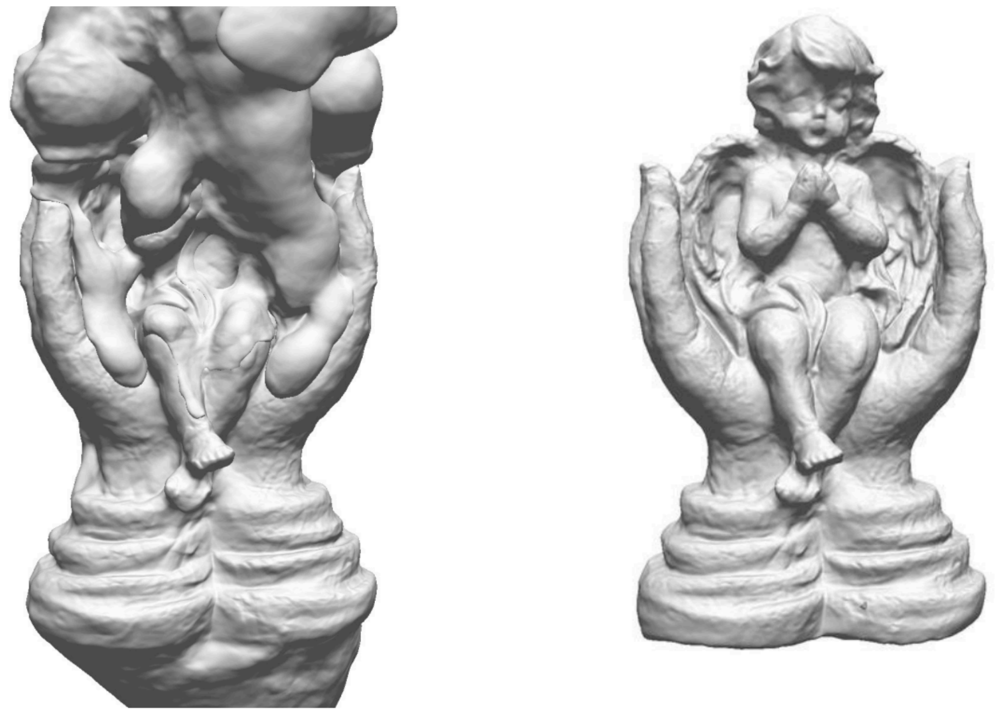
IDR [Yariv et al. '20]

- Limitation:

  Object masks



IDR [Yariv et al. '20]



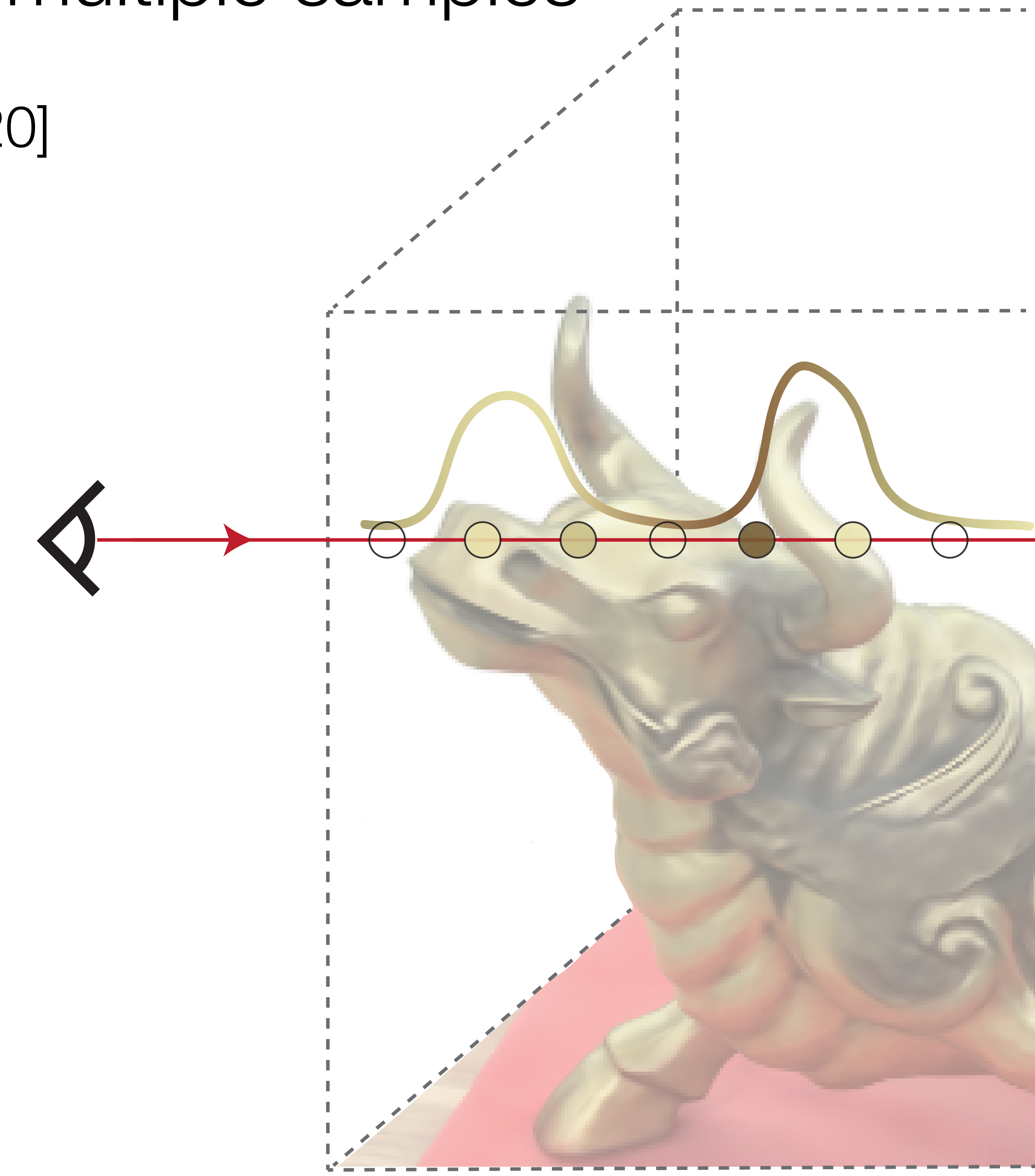
No Masks With Masks

<u>Volume Rendering:</u>

- Representation: Volume density
- Rendering: Integral approximation
- Back-propagate: multiple samples

NeRF [Mildenhall et al. '20]

…



Surface Rendering v.s. Volume Rendering

Surface Rendering:

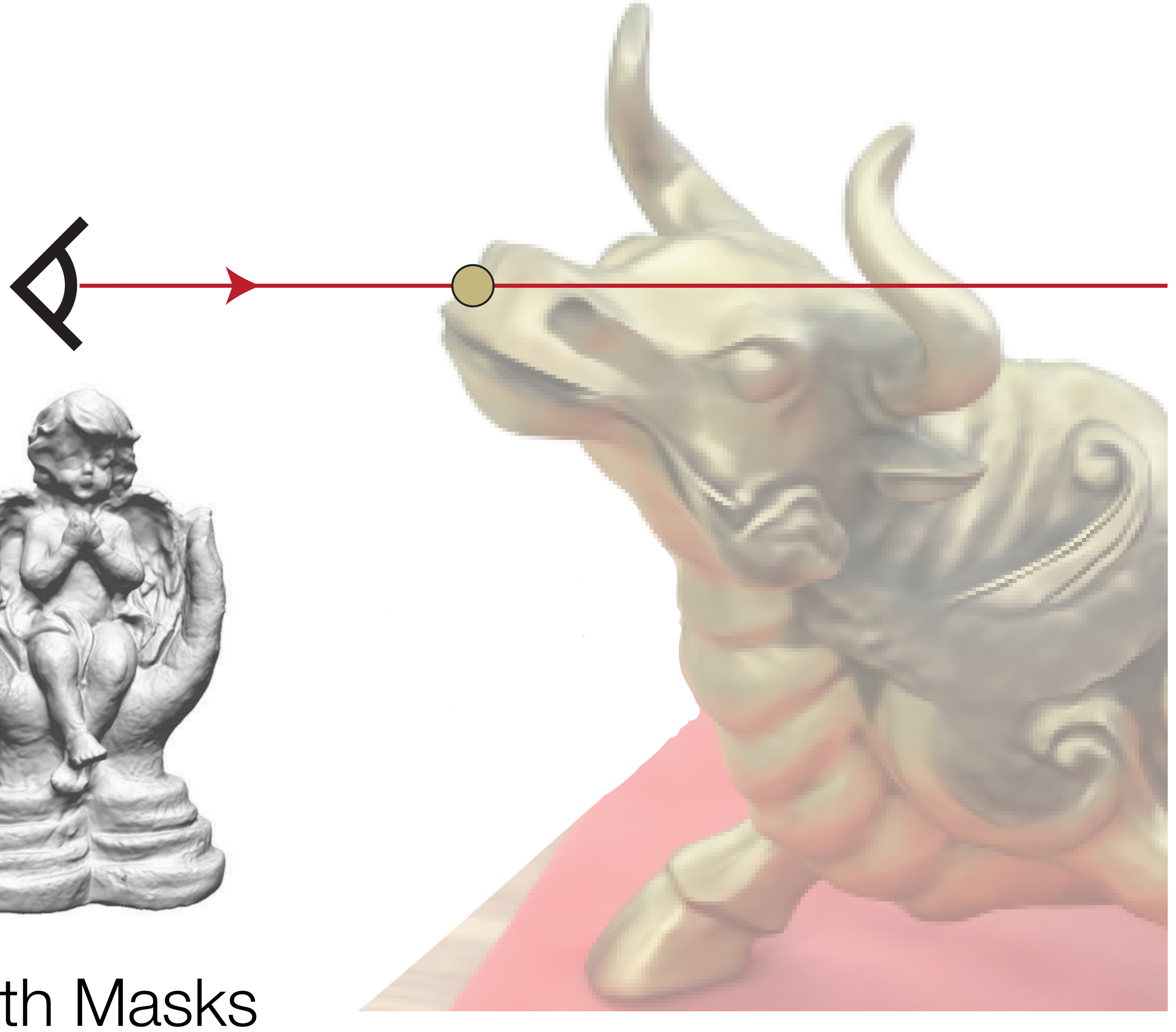
- Representation: Implicit surface
- Rendering: Find intersection
- Back-propagate: 1 sample

DVR [Niemeyer et al. '20]

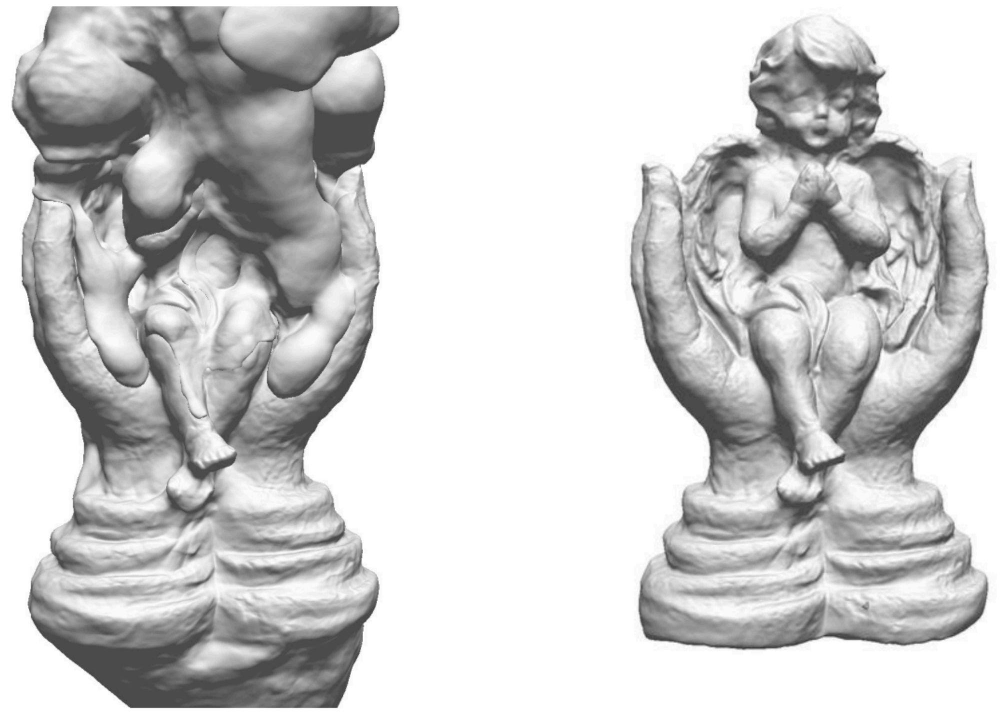
IDR [Yariv et al. '20]

- Limitation:

  Object masks



IDR [Yariv et al. '20]



No Masks | With Masks

Volume Rendering:

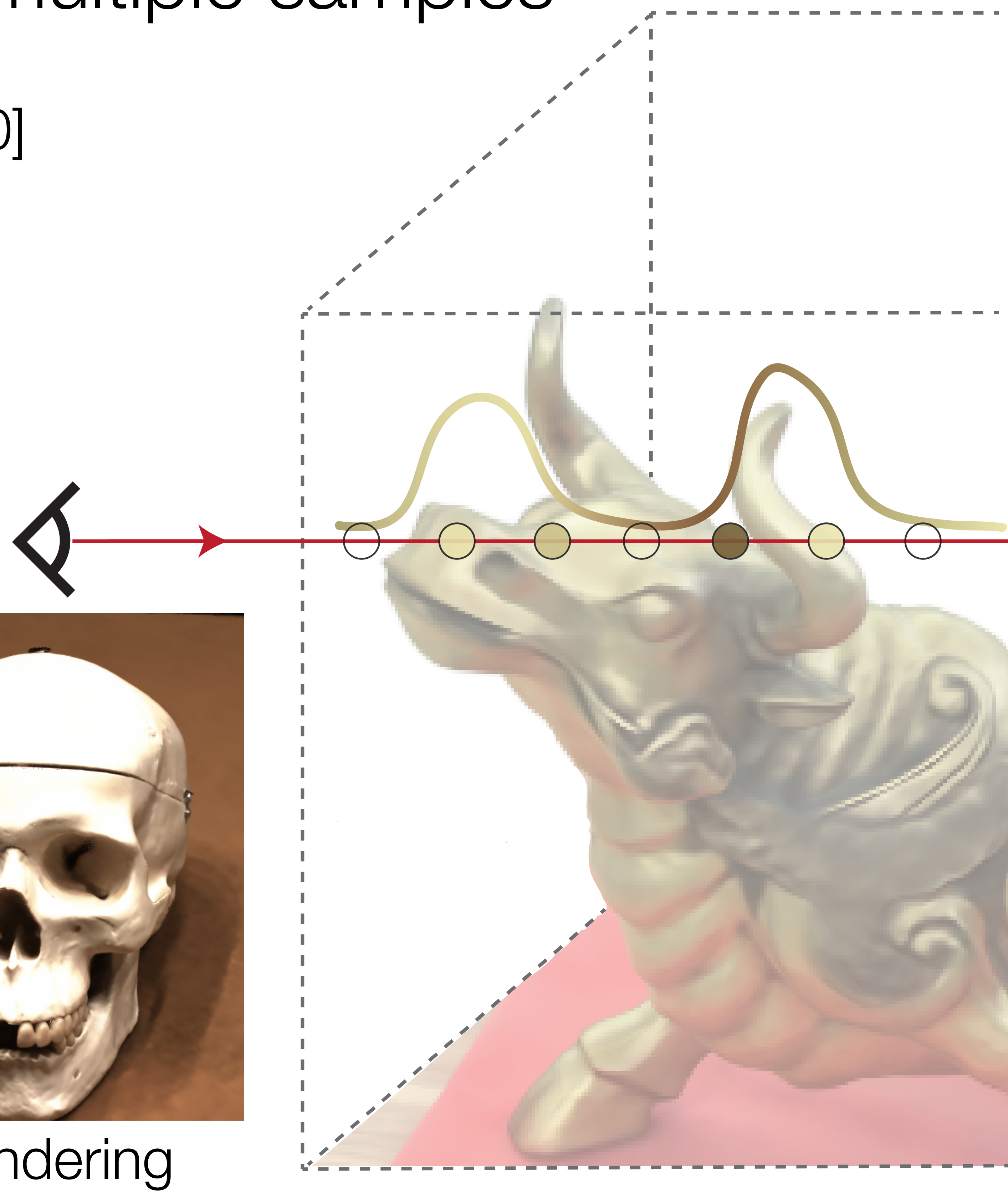
- Representation: Volume density
- Rendering: Integral approximation
- Back-propagate: multiple samples

NeRF [Mildenhall et al. '20]

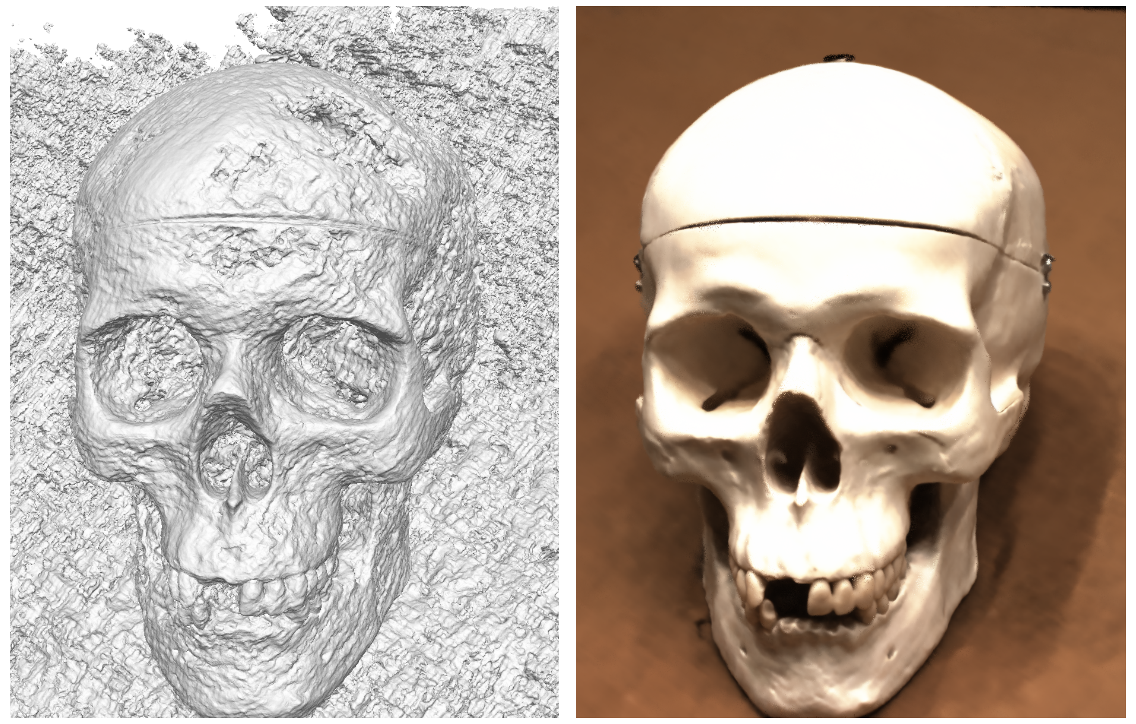
…

- Limitation:

  Noisy geometry



NeRF [Midenall et al. '20]



Density = 50 | Rendering

Surface Rendering v.s. Volume Rendering

Surface Rendering:

- Representation: Implicit surface
- Rend
- Back

DVR [Ni

IDR [Yar

- Limita

  Objec

Volume Rendering:

- Representation: Volume density
- n
- s

Can we get the best of both worlds?

IDR [Yariv et al. '20]

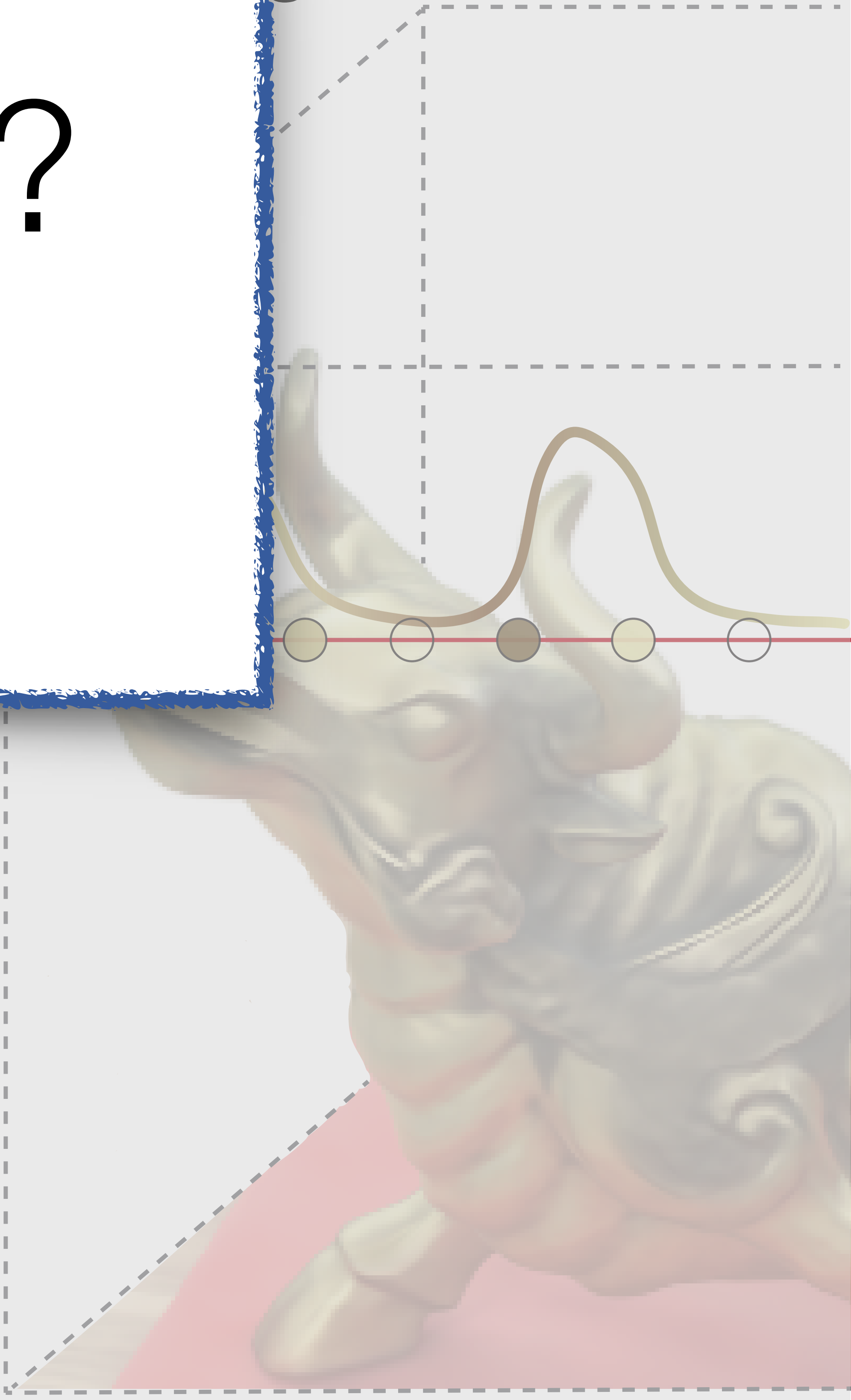
No Masks

With Masks

NeRF [Midenall et al. '20]

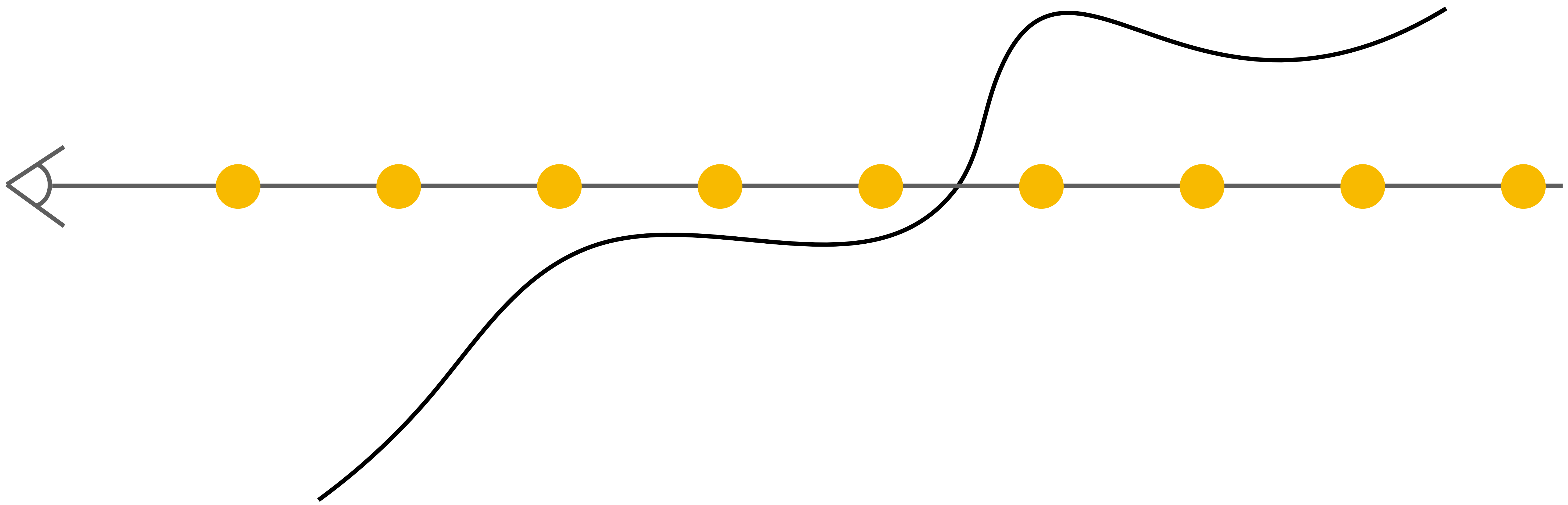
Density = 50

Rendering



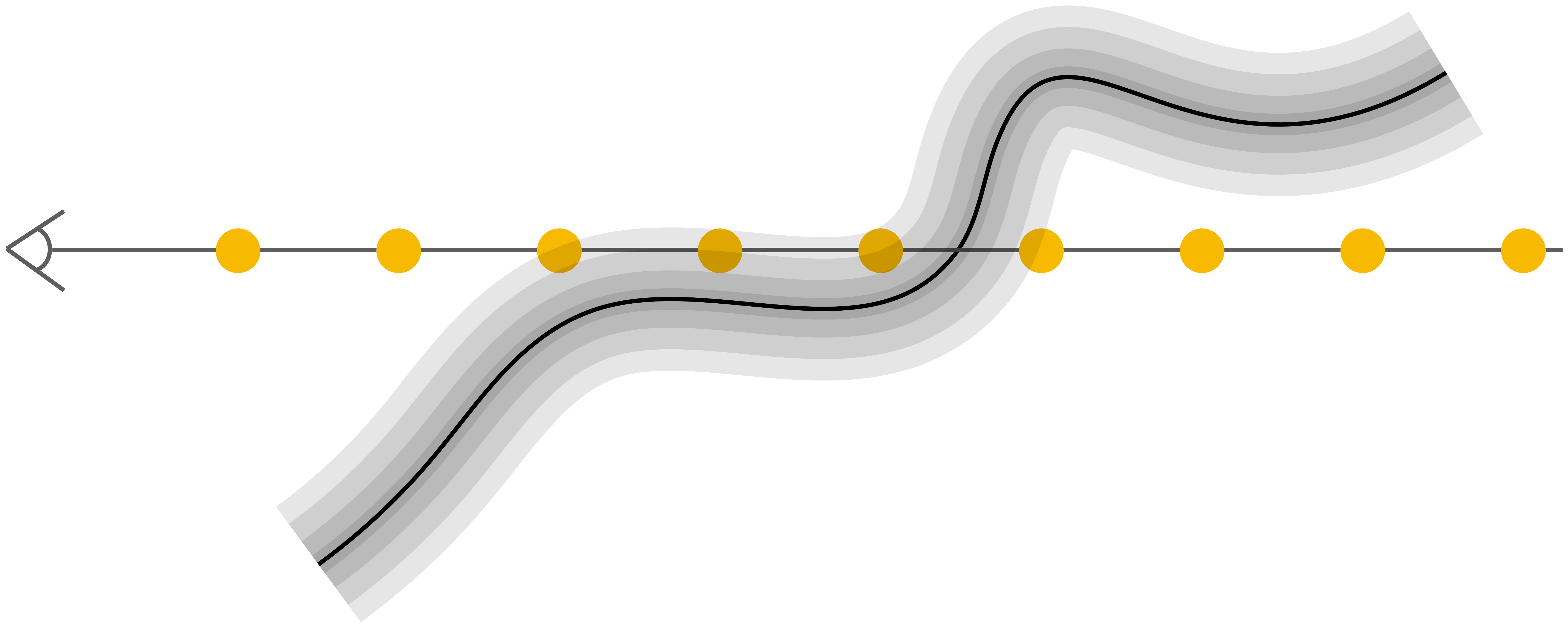
Surface reconstruction using volume rendering

How can we volume-render a surface?



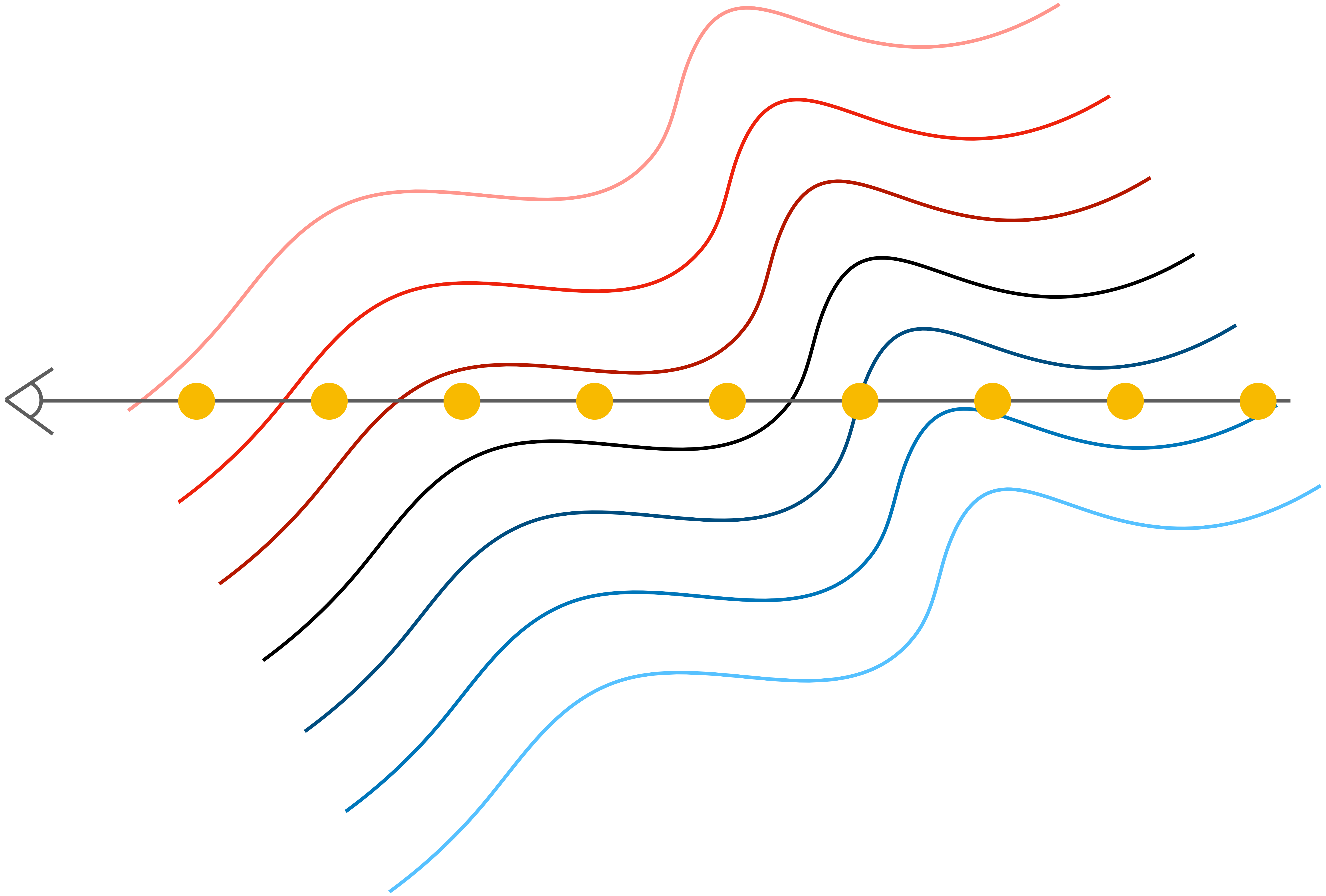
Surface reconstruction using volume rendering

⇒ Represent the scene as a “soft” surface



Surface reconstruction using volume rendering

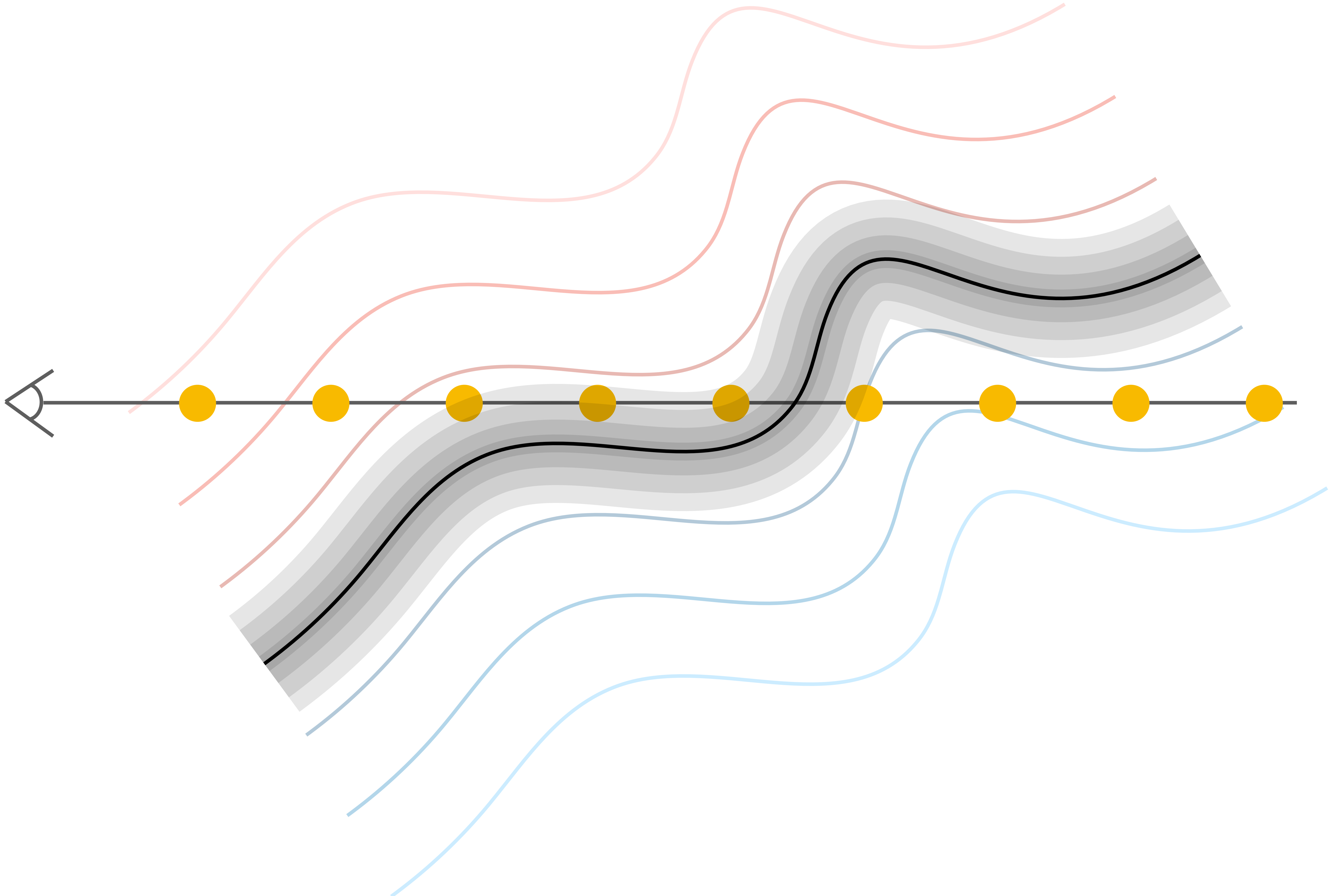
MLP-parameterized signed distance function $f(\mathbf{x})$



Volume rendering of neural implicit surfaces, Yariv et al., NeurIPS 2021

Surface reconstruction using volume rendering

Modeling density as: $\boldsymbol{\sigma}(\mathbf{x}) = \alpha\,\Psi_{\beta}(f(\mathbf{x}))$

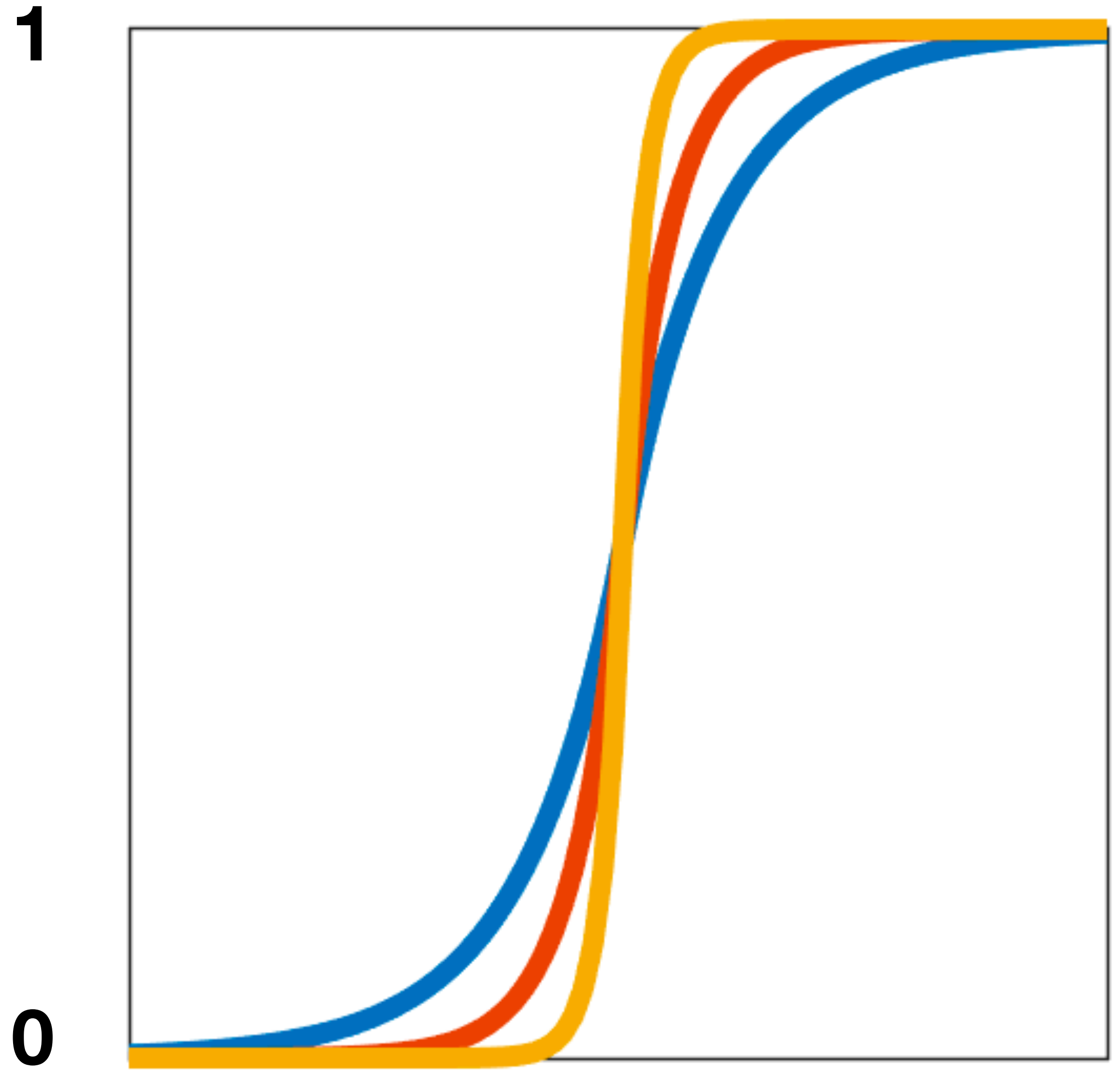


Volume rendering of neural implicit surfaces, Yariv et al., NeurIPS 2021

VolSDF: density $\boldsymbol{\sigma}$ as transformed SDF

Modeling density as: $\boldsymbol{\sigma}(\mathbf{x}) = \alpha\,\Psi_\beta(f(\mathbf{x}))$

Density



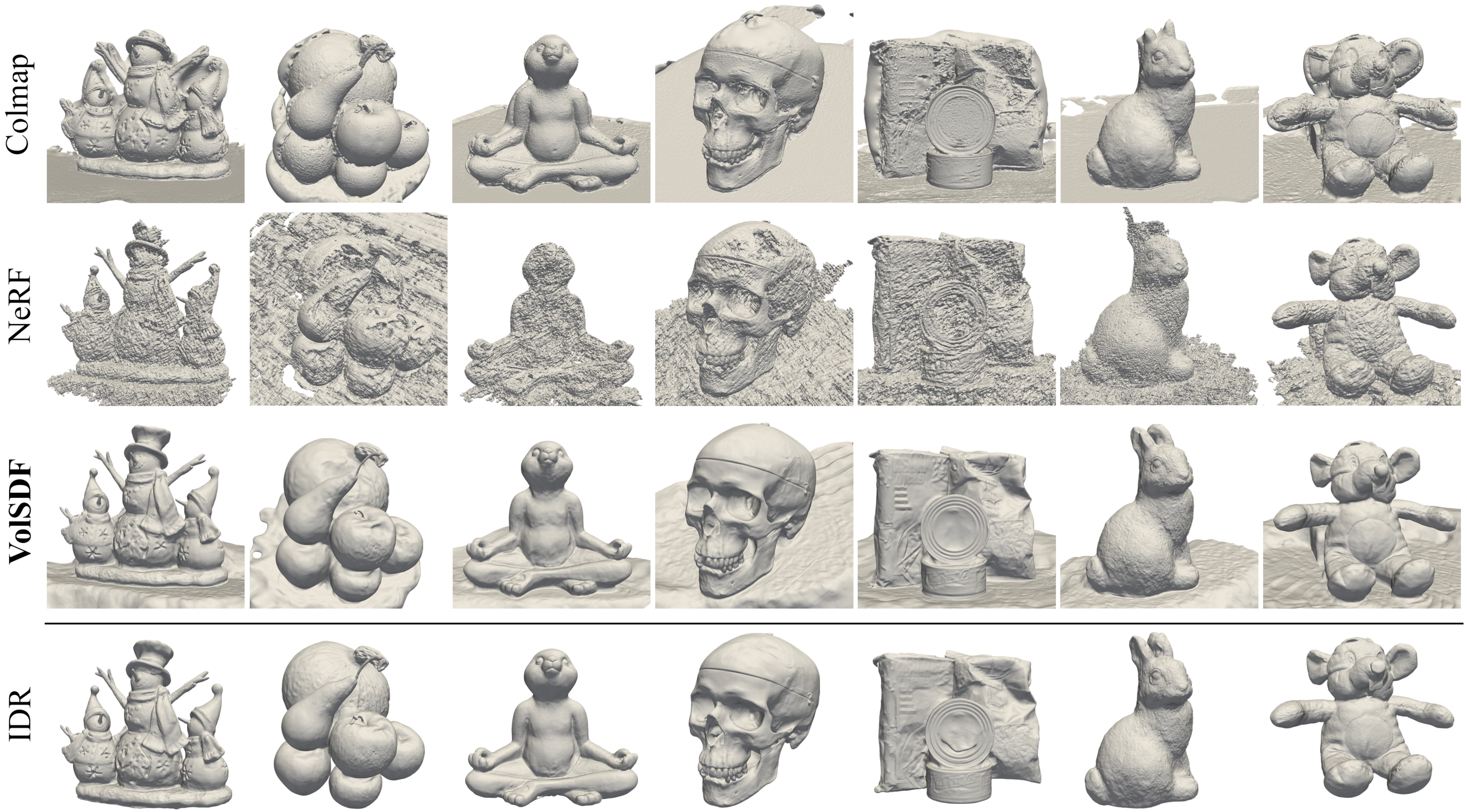
Laplace CDF

Signed Distance Function

*<u>Recall:</u> NeRF models Density $\boldsymbol{\sigma}(\boldsymbol{x}) : \mathbb{R}^3 \rightarrow \mathbb{R}^+$ is a general purpose MLP

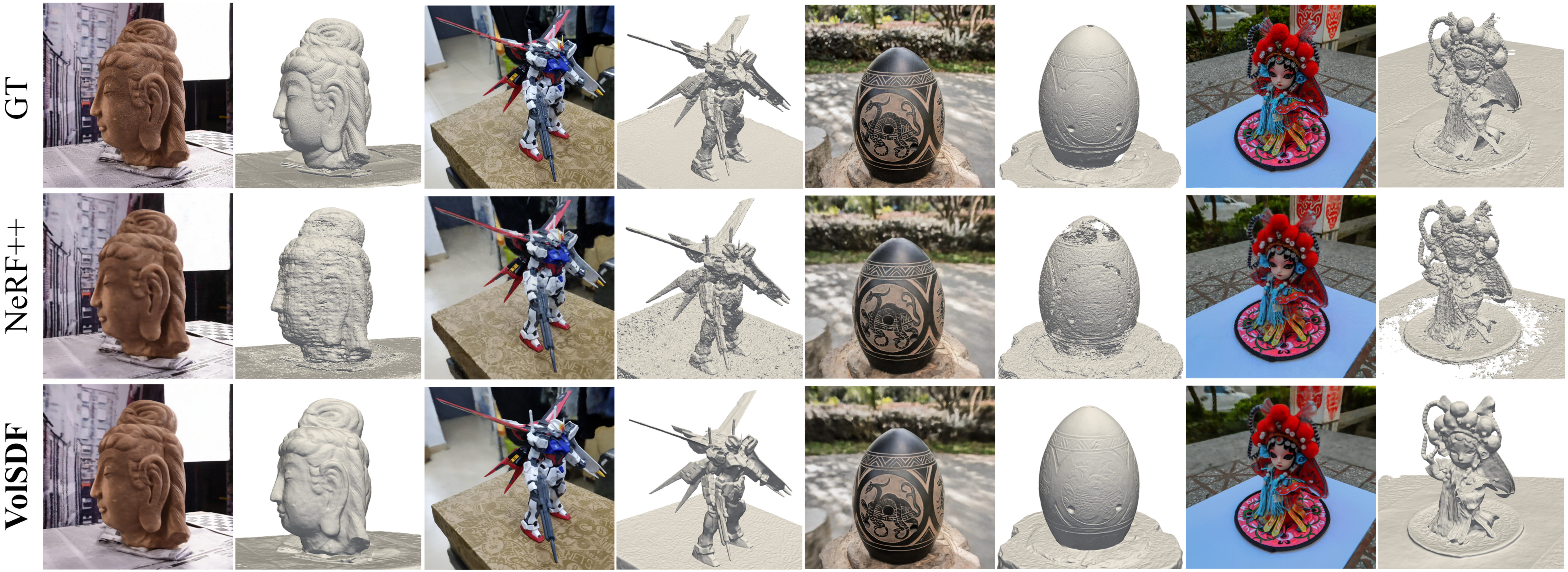
Volume rendering of neural implicit surfaces, Yariv et al., NeurIPS 2021

Results: comparisons



DTU: Large Scale Multi-view Stereopsis Evaluation [Jensen et al. 2014]

Results: comparisons



BlendedMVS: A Large-scale Dataset for Generalized Multi-view Stereo Networks [Yao et al. 2020]

Summary

Take-away messages:

- The underlying geometry, obtained by volume renderers is non-smooth and contains artifacts.
- Differentiable surface renderers produce highly accurate 3D reconstructions.
  However, they depend on object masks.
- Unifying surface and volume rendering is possible!

Summary

Take-away messages:

- The underlying geometry, obtained by volume renderers is non-smooth and contains artifacts.
- Differentiable surface renderers produce highly accurate 3D reconstructions.
  However, they depend on object masks.
- Unifying surface and volume rendering is possible!

Limitation:

- Per-scene overfit (no generalization)
- Computational expensive
- Assuming fully opaque surface

Summary

Take-away messages:

- The underlying geometry, obtained by volume renderers is non-smooth and contains artifacts.
- Differentiable surface renderers produce highly accurate 3D reconstructions.
  However, they depend on object masks.
- Unifying surface and volume rendering is possible!

Limitation:

- Per-scene overfit (no generalization)
- Computational expensive
- Assuming fully opaque surface

Other applications:

- 3D Generative models
- Dynamic scenes
- Camera optimization
- Material-Light decomposition
- So many more …

Summary

Take-away messages:

- The underlying geometry, obtained by volume renderers is non-smooth and contains artifacts.
- Differentiable surface renderers produce highly accurate 3D reconstructions.
  However, they depend on object masks.
- Unifying surface and volume rendering is possible!

Limitation:

- Per-scene overfit (no generalization)
- Computational expensive
- Assuming fully opaque surface

Other applications:

- 3D Generative models
- Dynamic scenes
- Camera optimization
- Material-Light decomposition
- So many more …

Survey papers/blogs:

- NeRF Explosion
  [Frank Dellaert ’20]
- State of the Art on Neural Rendering
  [Tewari et al. '20]
- Advances in Neural Rendering
  [Tewari et al. ’21]
- Neural Fields in Visual Computing and Beyond
  [Xie at al. ’21]

Questions?