Lecture 23:

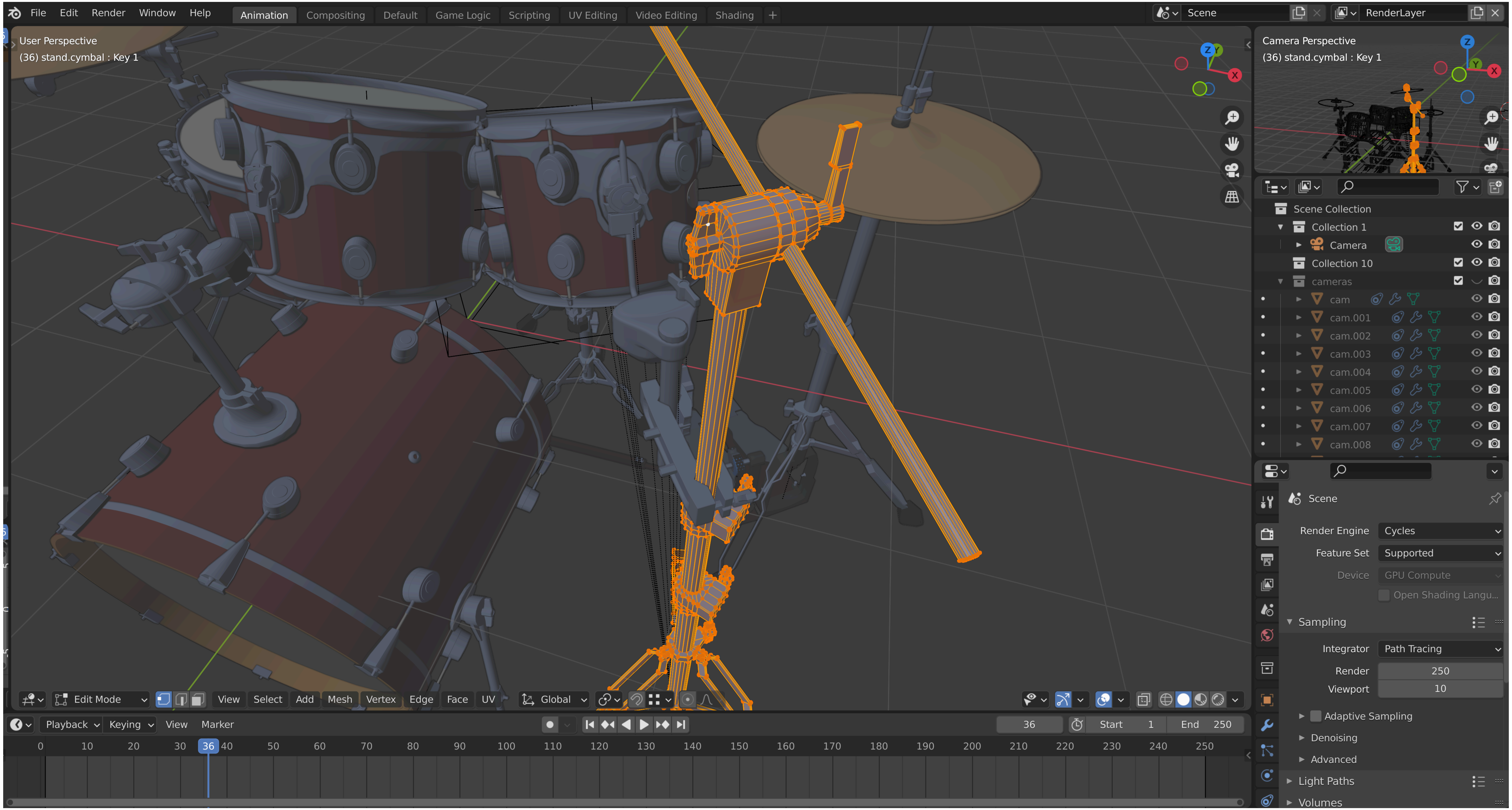
Special Topics: Neural Radiance Fields

Computer Graphics and Imaging
UC Berkeley CS184/284A

Where does the 3D model come from?

San Miguel Scene, 10.7M triangles

3D Modeling Software is Complex

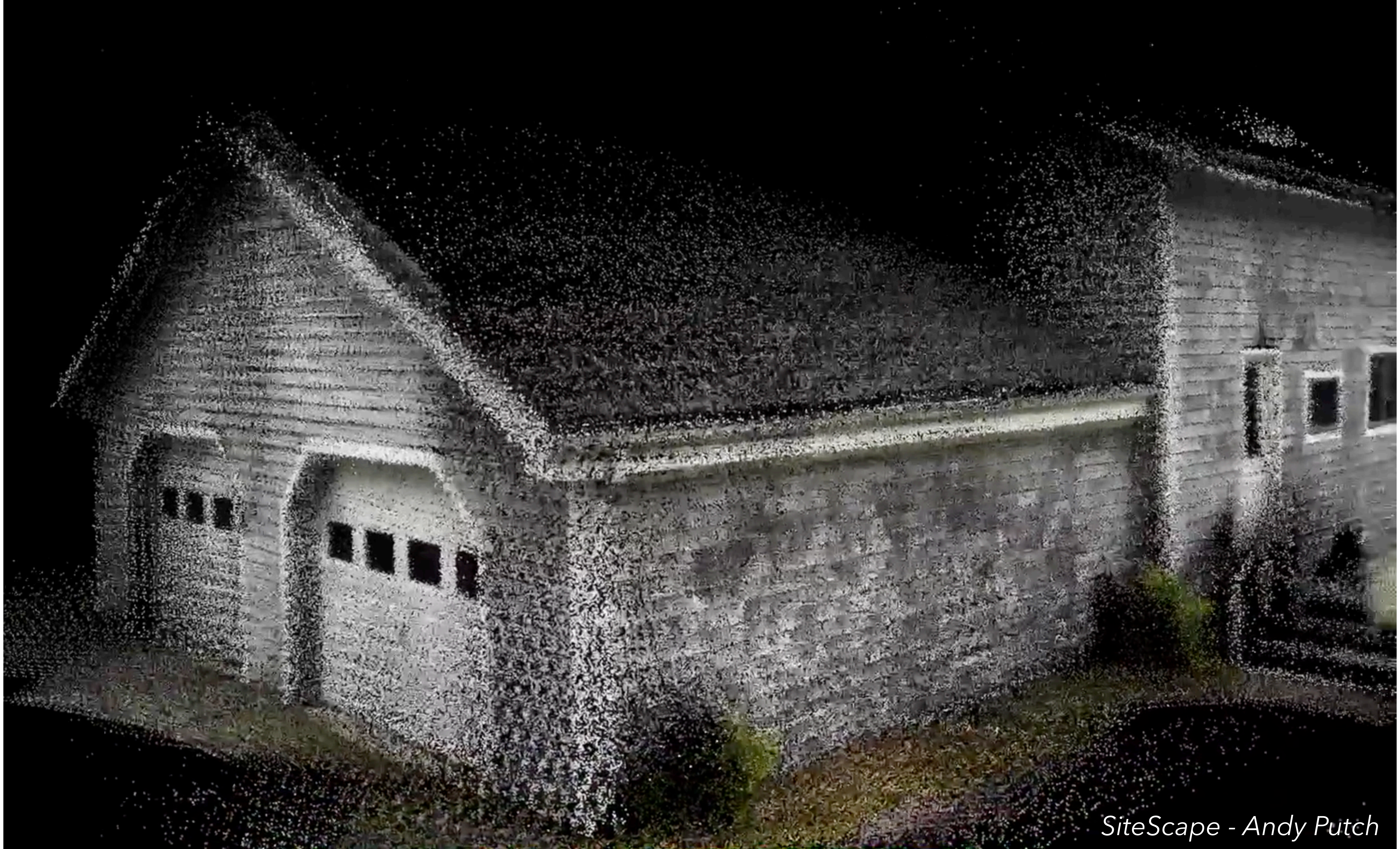


The Next Leap: How A.I. will change the 3D industry, Andrew Price

Depth Sensors

Microsoft Kinect

iPhone Depth Sensor



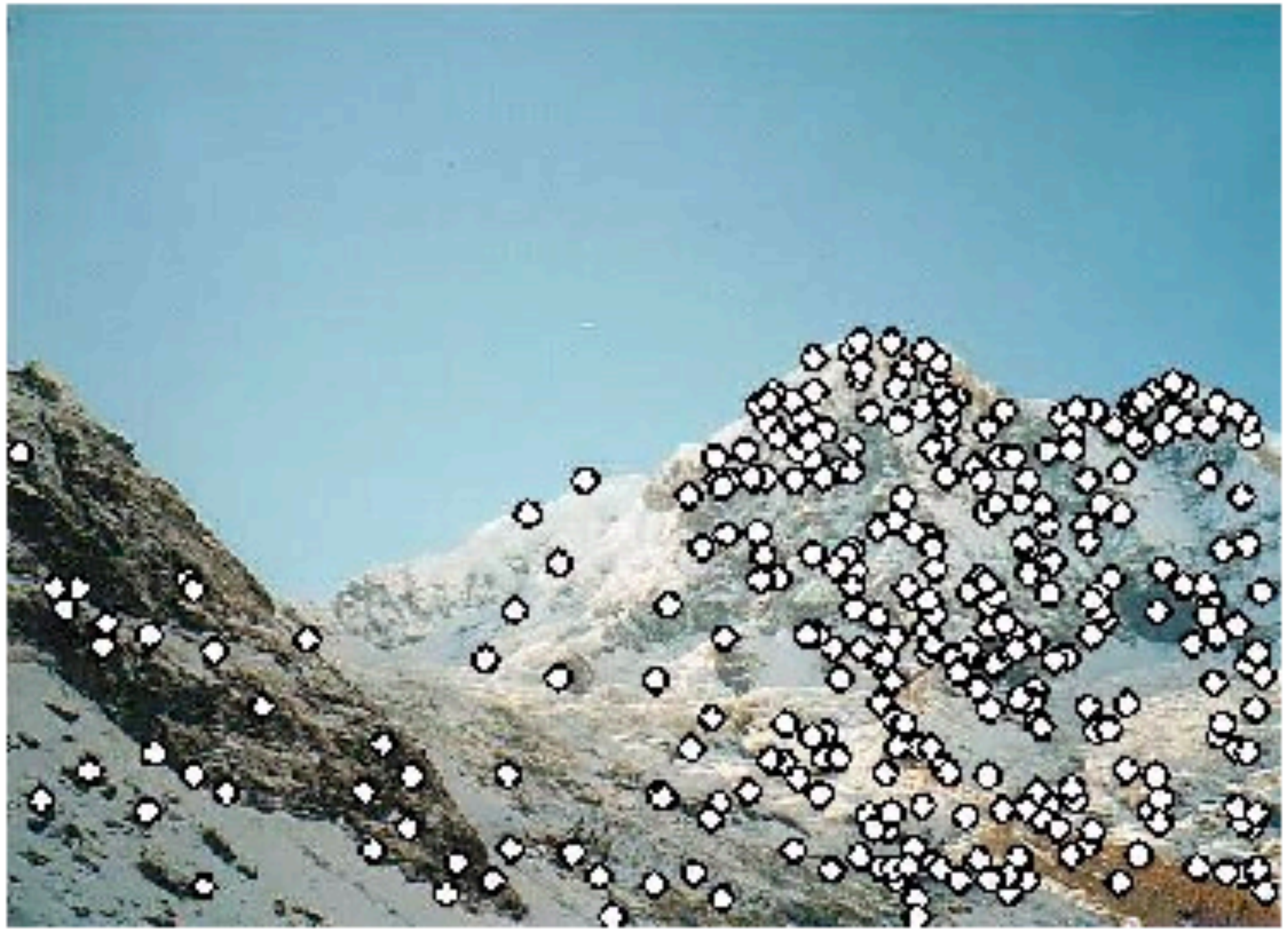
SiteScape - Andy Putch

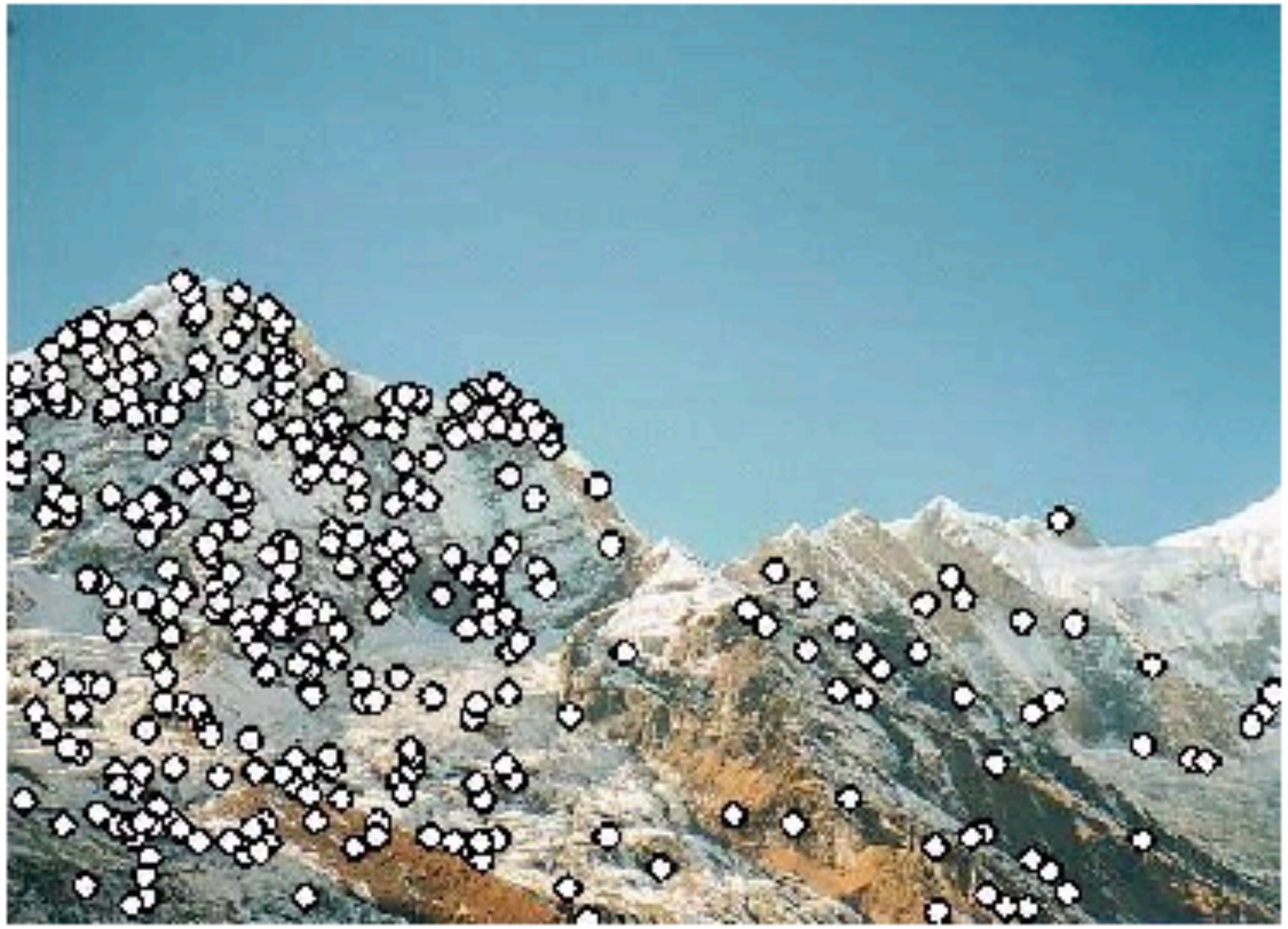
What if you don’t have depth?

Point Cloud from Images

Brown and Lowe (2007)

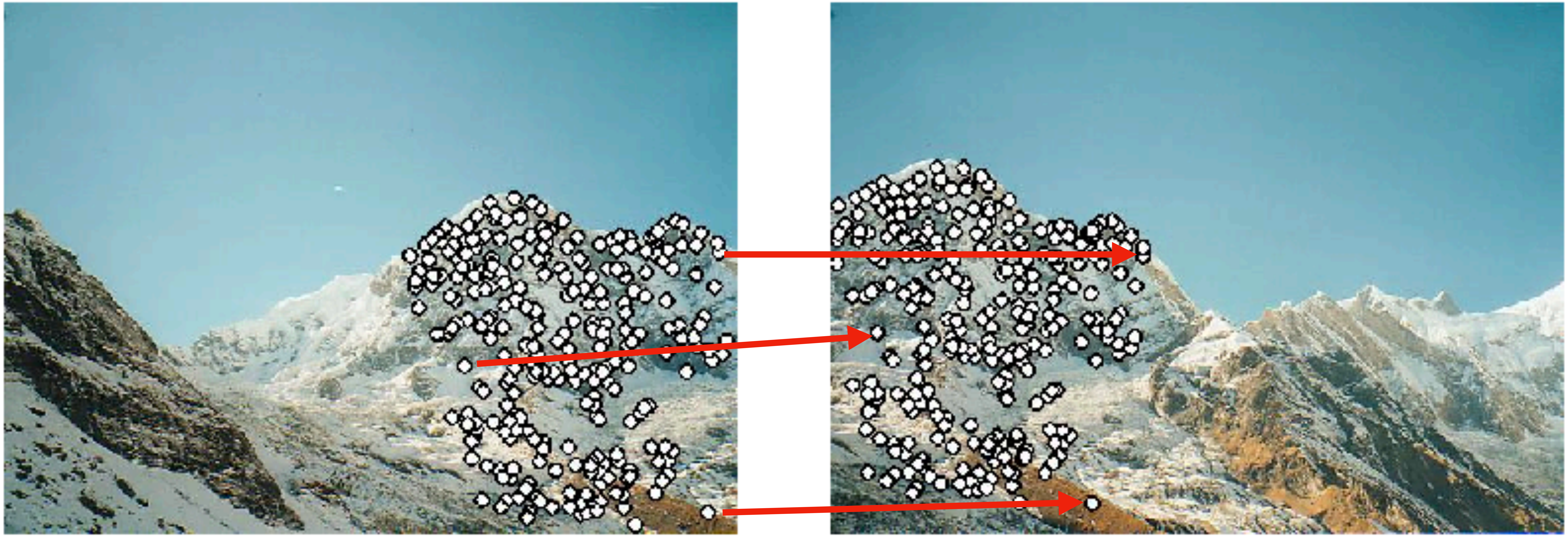
Point Cloud from Images





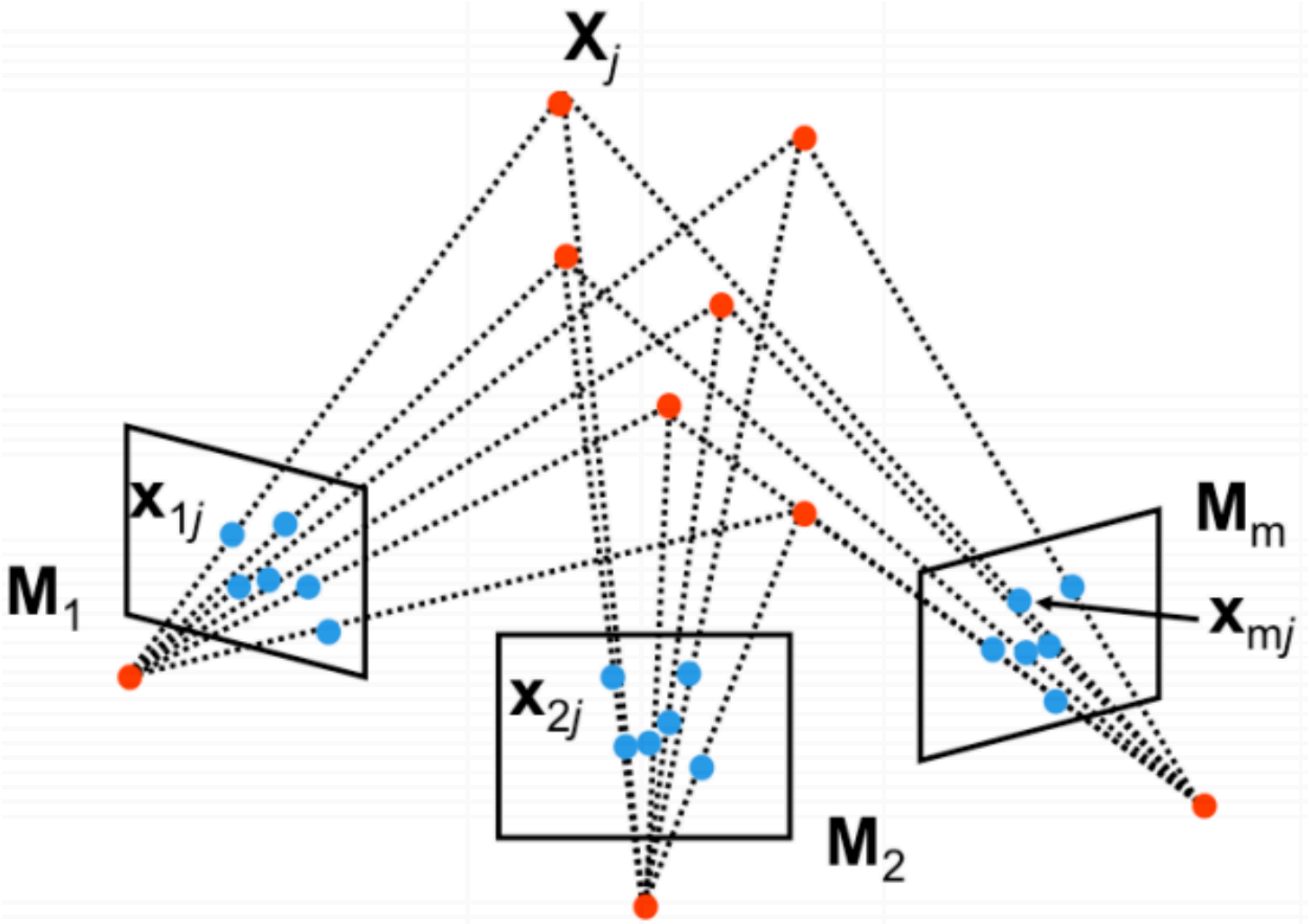
Brown and Lowe (2007)

Point Cloud from Images



Brown and Lowe (2007)

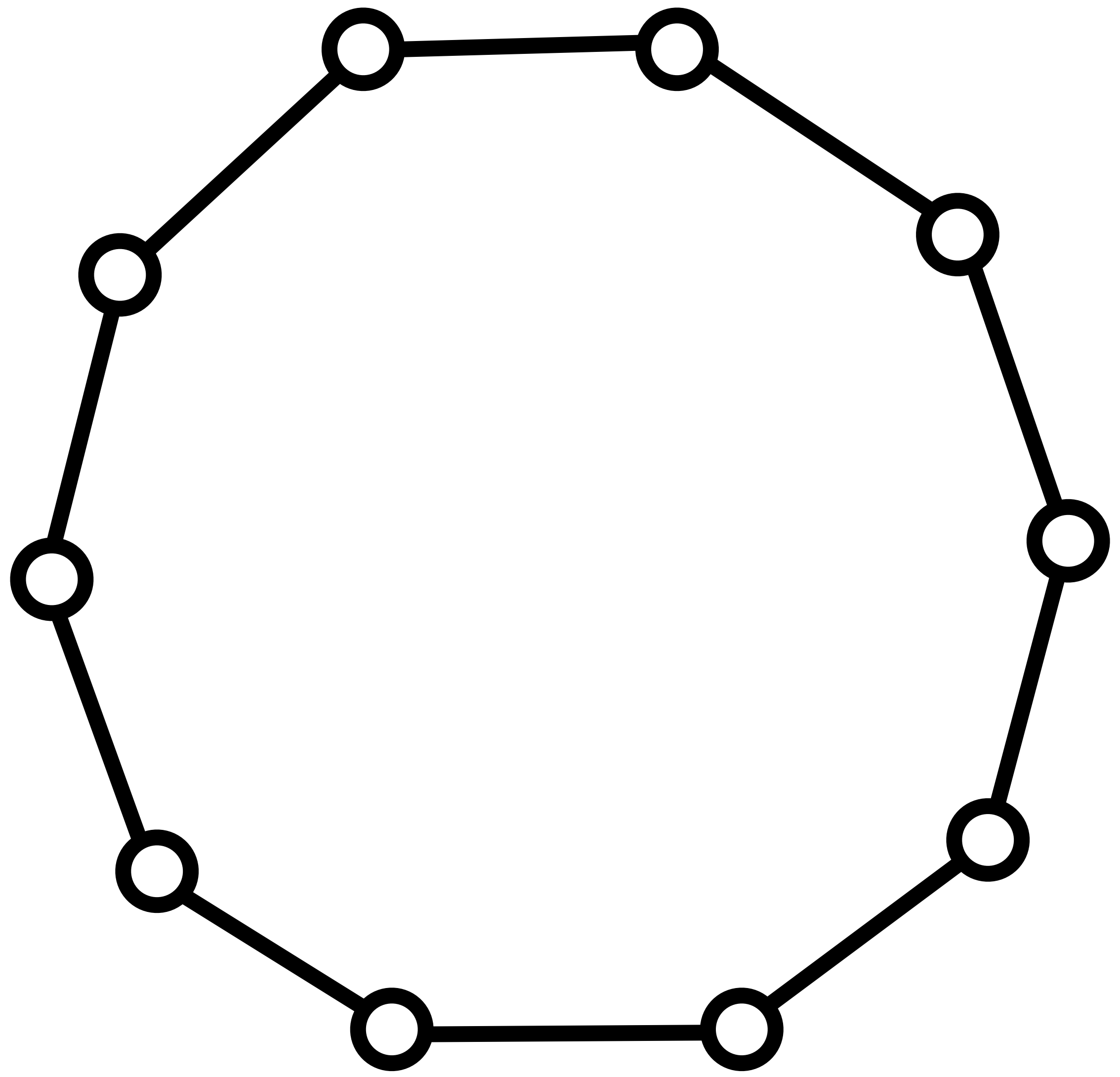
Point Cloud from Images

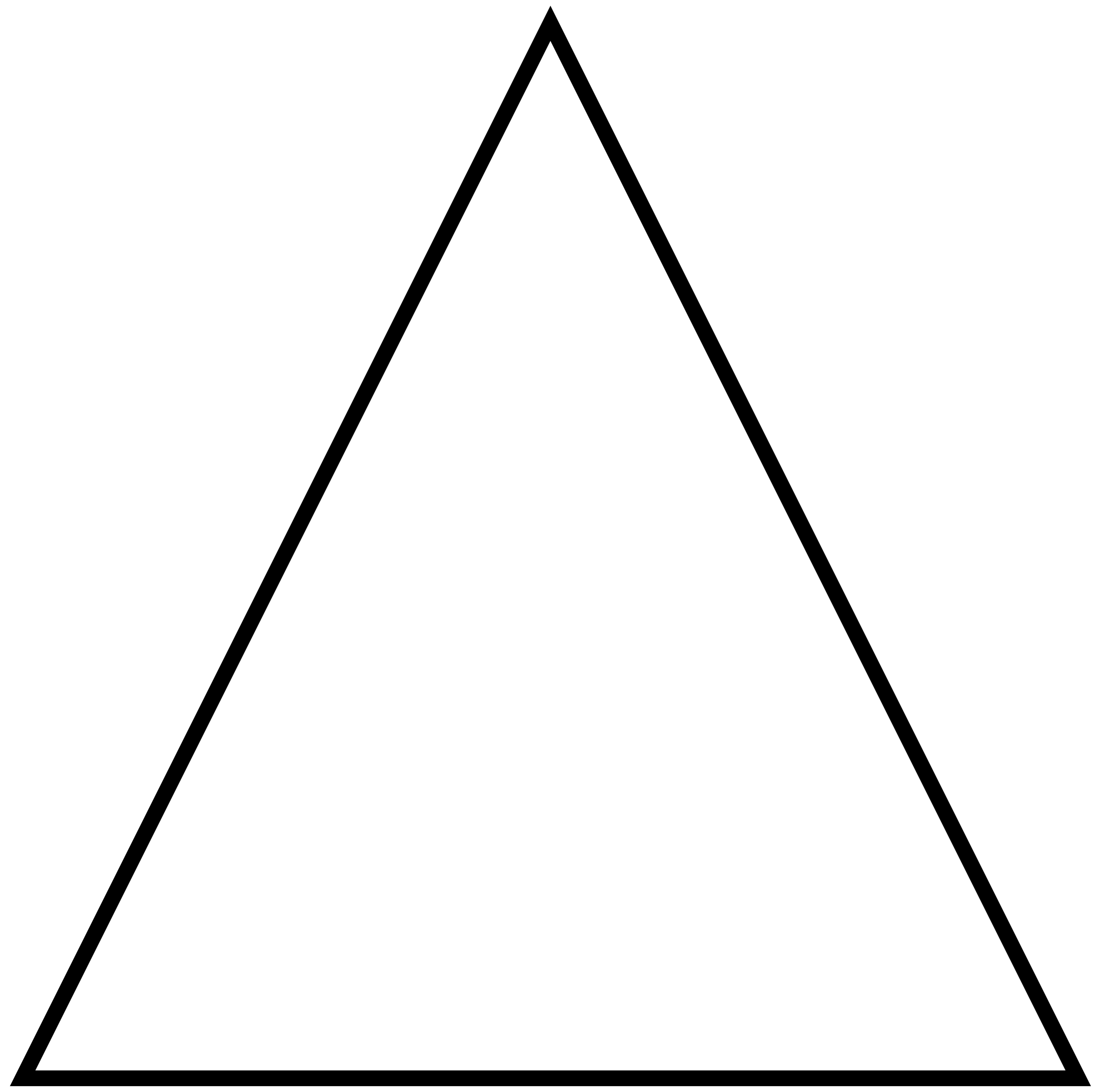


Hata and Savarese CS231a

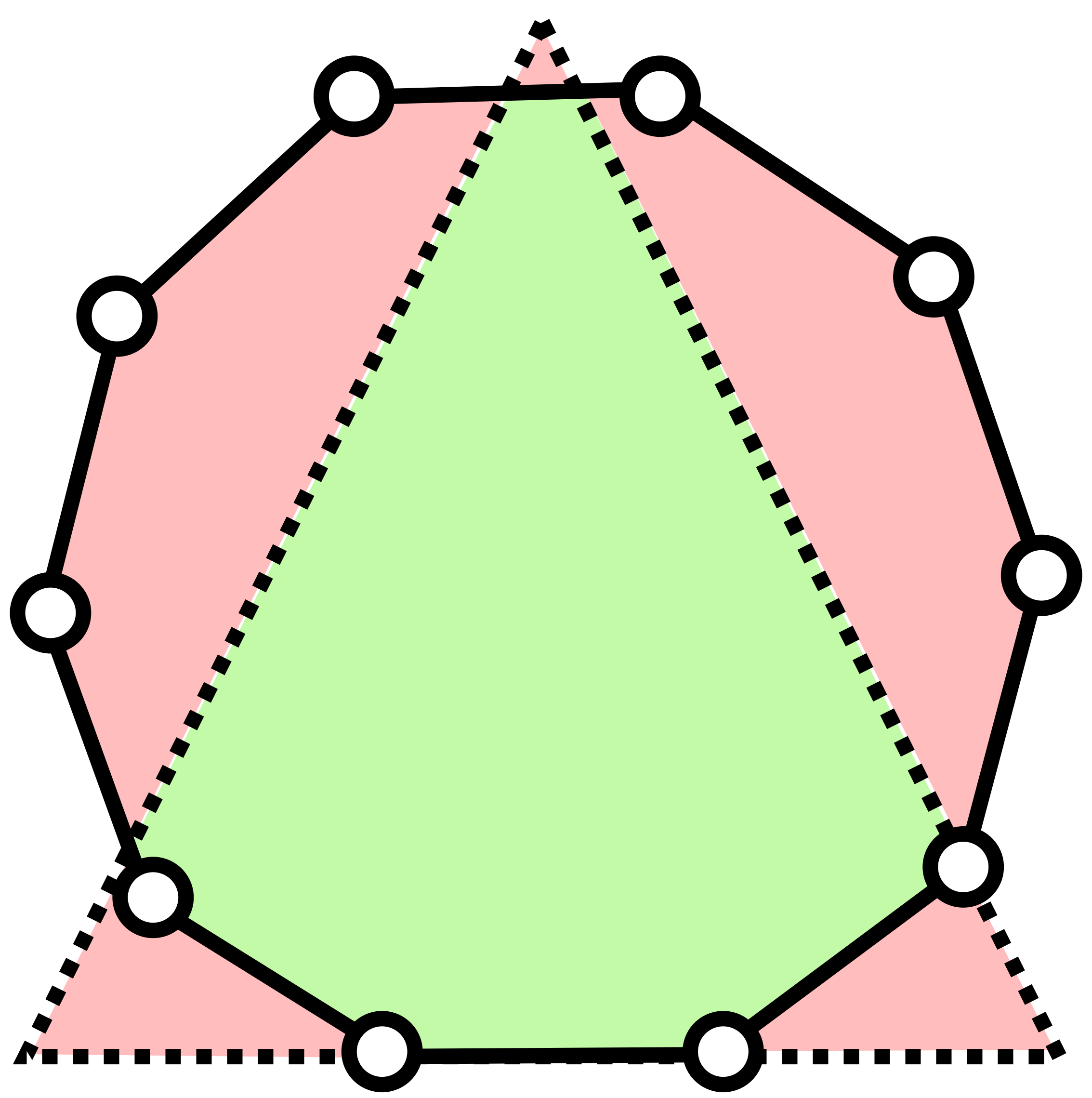
Can we operate on the geometry directly?

Gradient Based Optimization

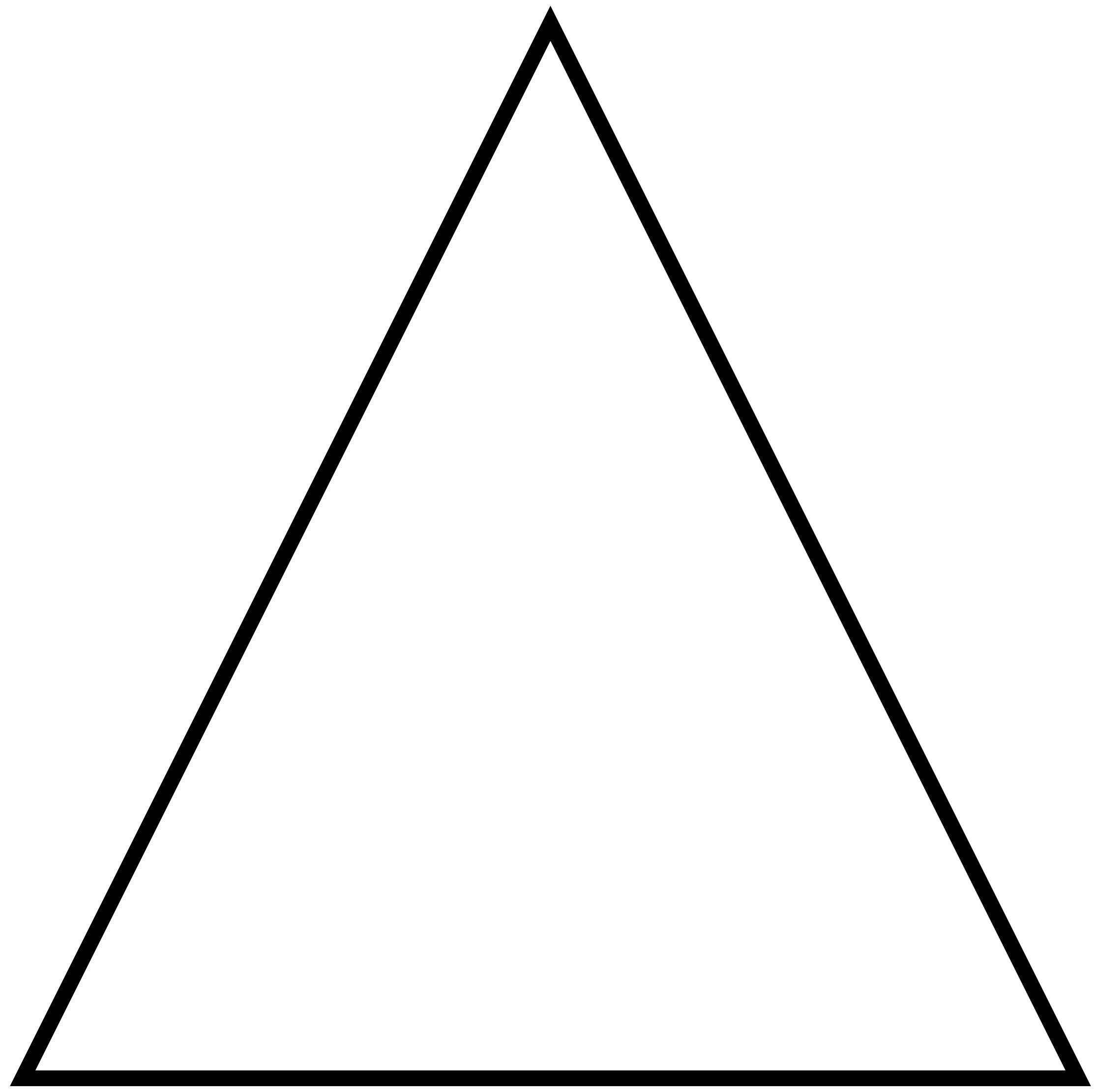

Initial Geometry


Target Geometry

Gradient Based Optimization

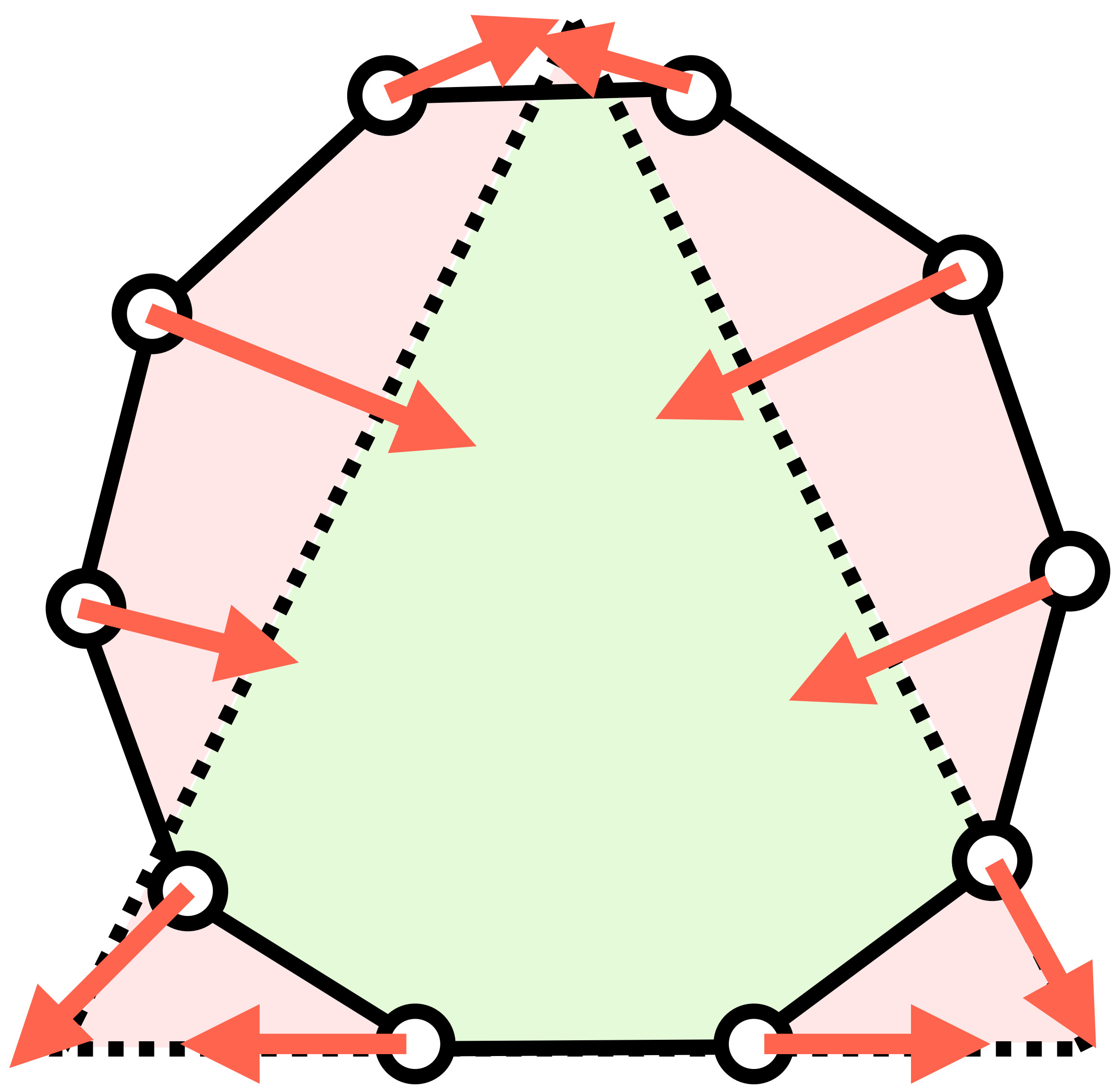


Compute Gradients

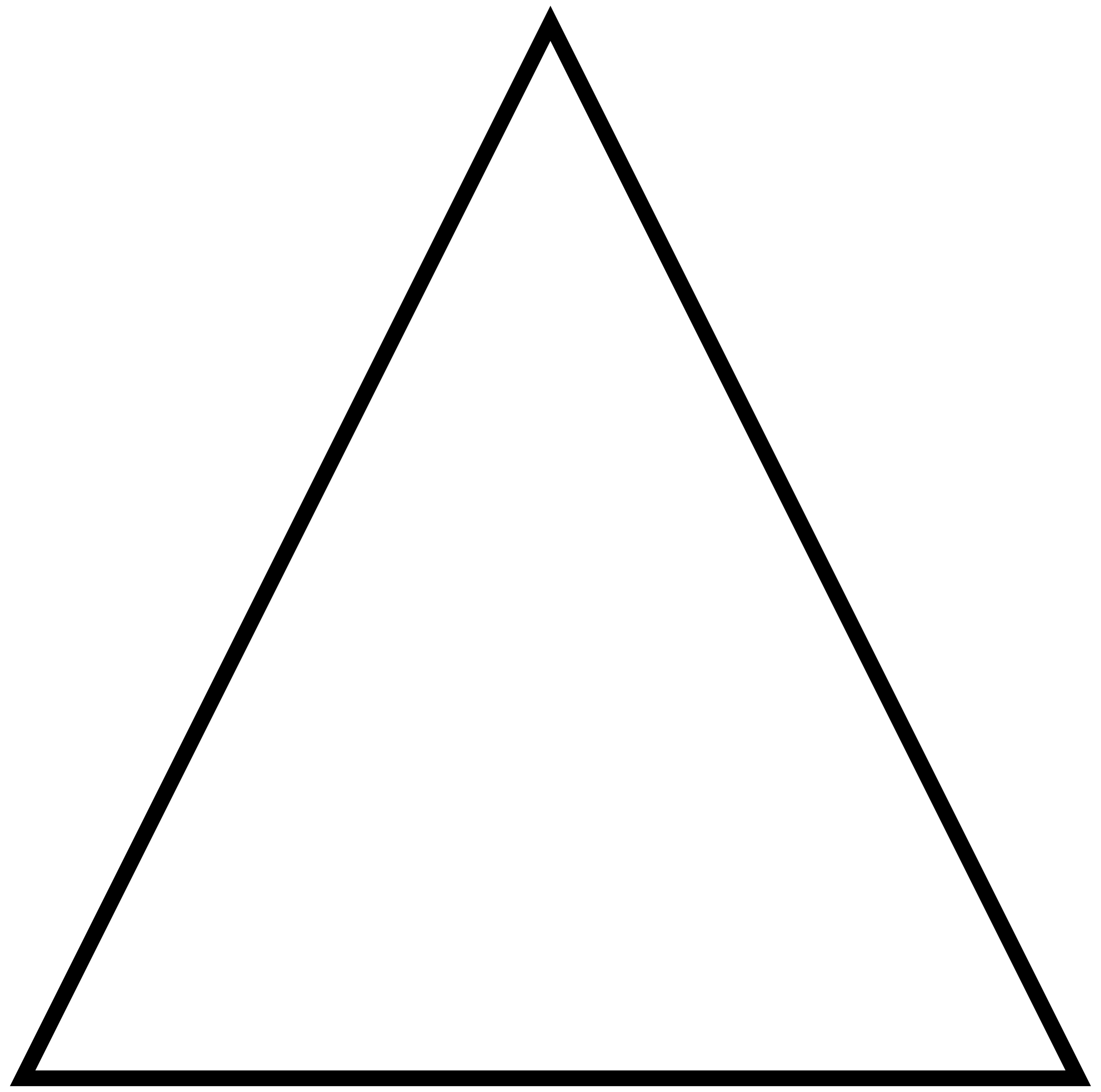


Target Geometry

Gradient Based Optimization

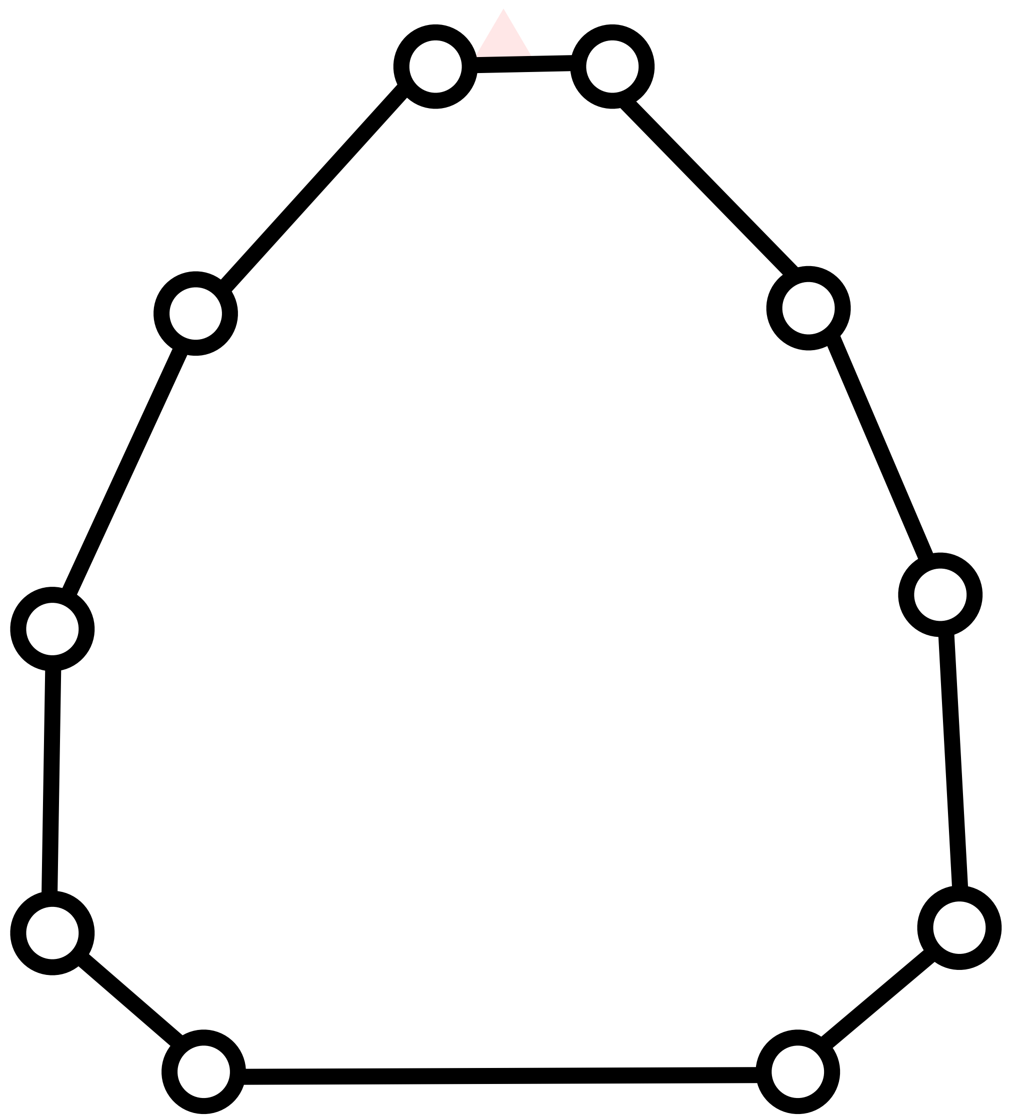


Compute Gradients

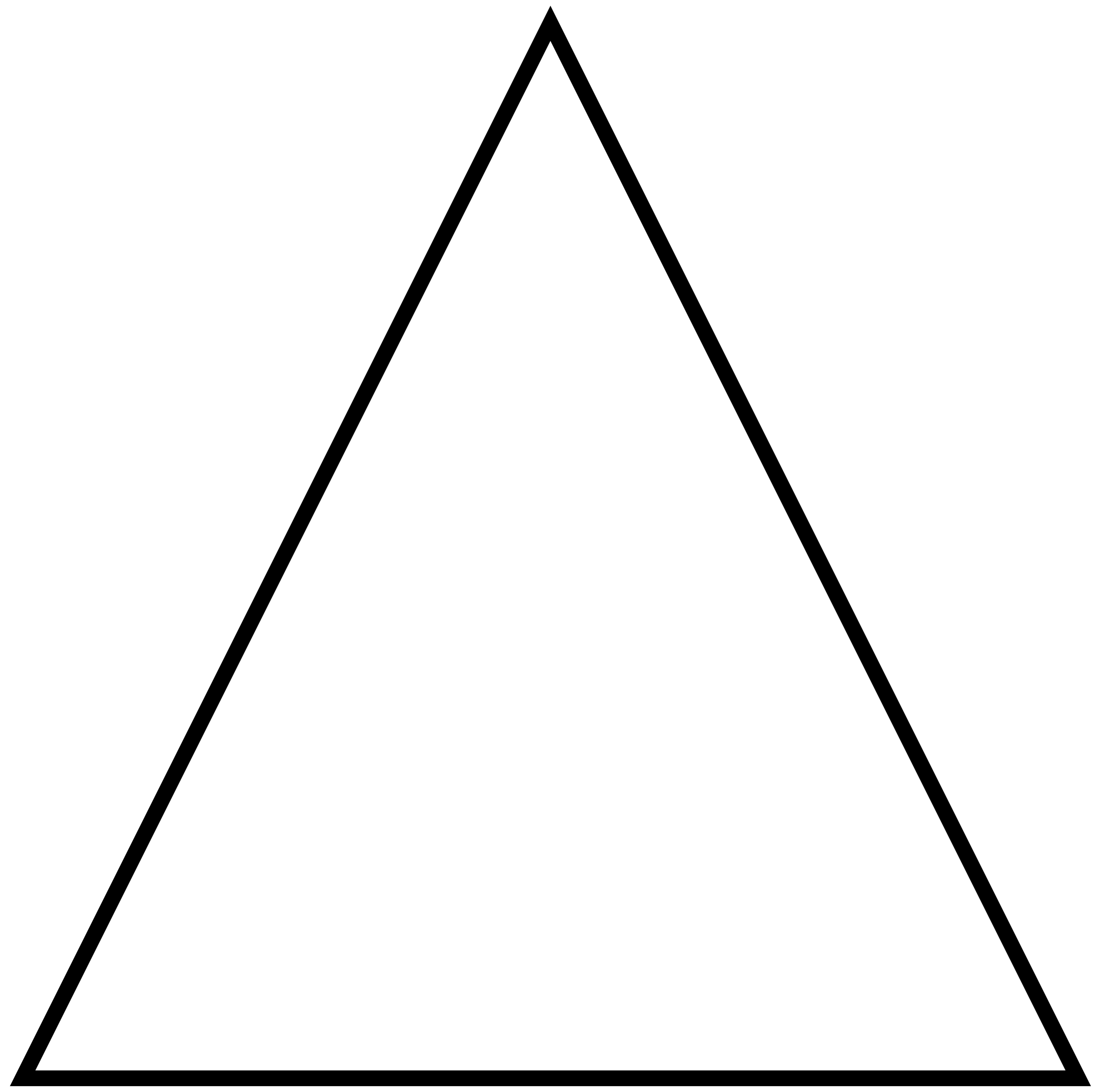


Target Geometry

Gradient Based Optimization

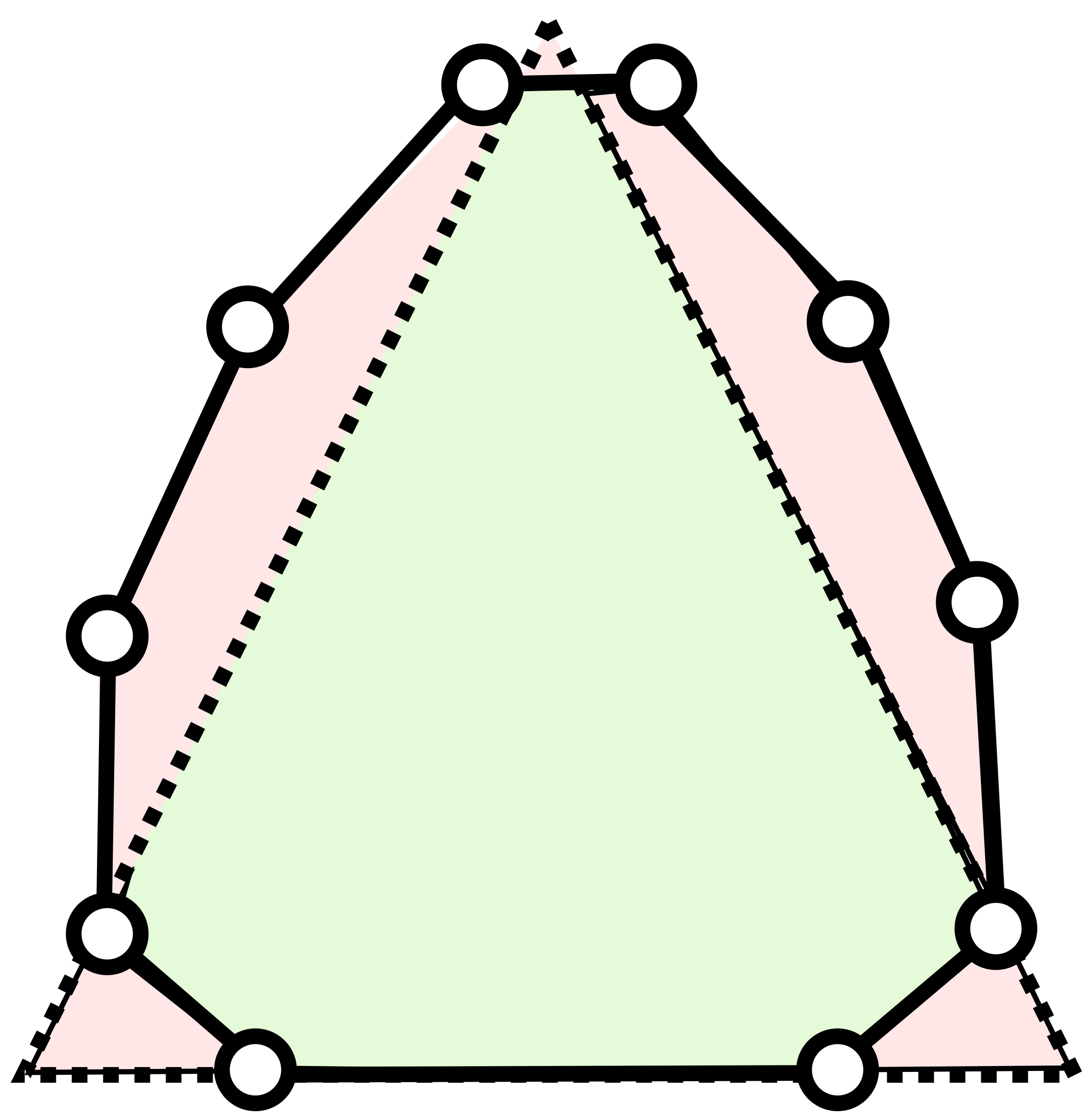


Update positions

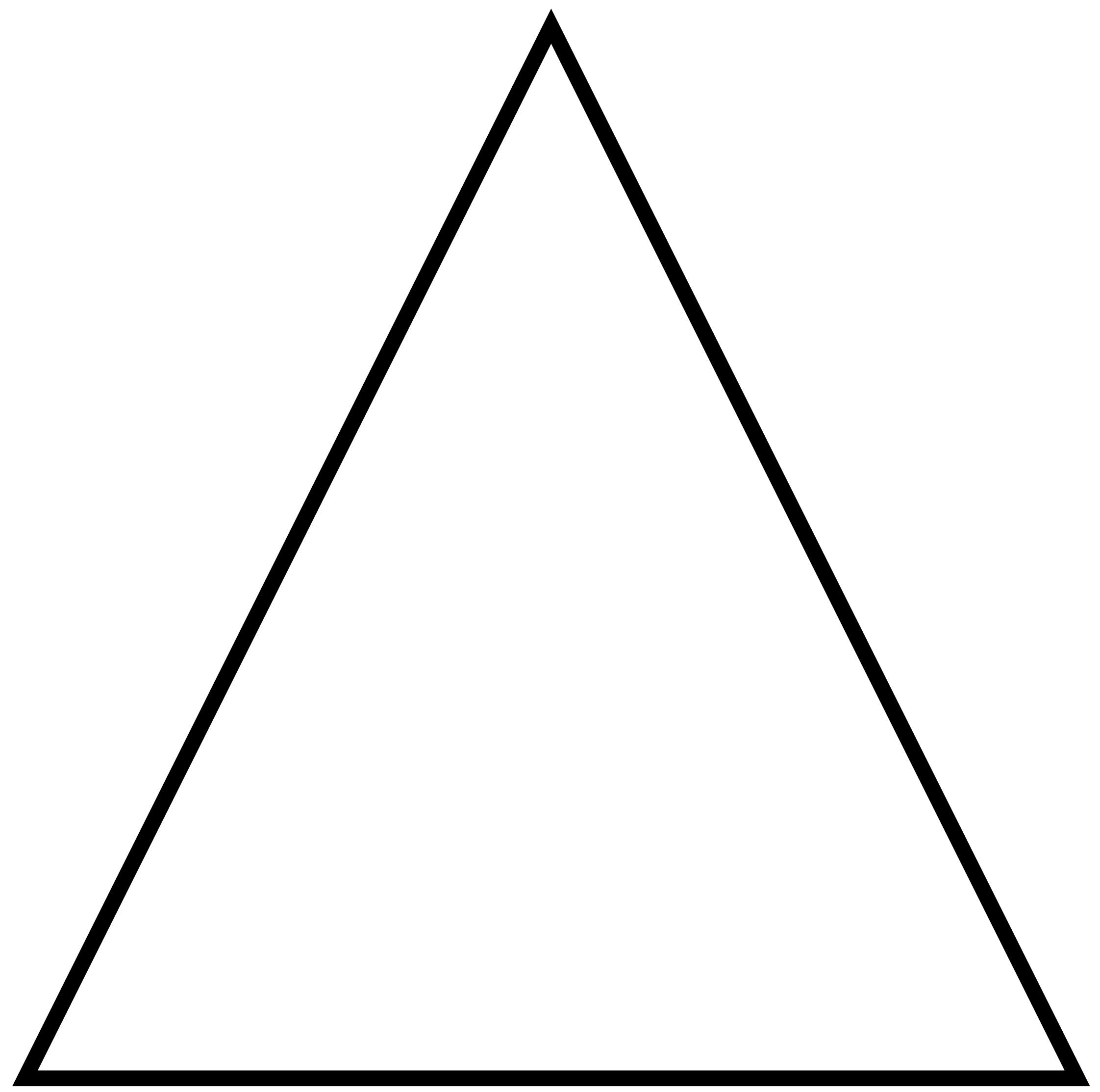


Target Geometry

Gradient Based Optimization

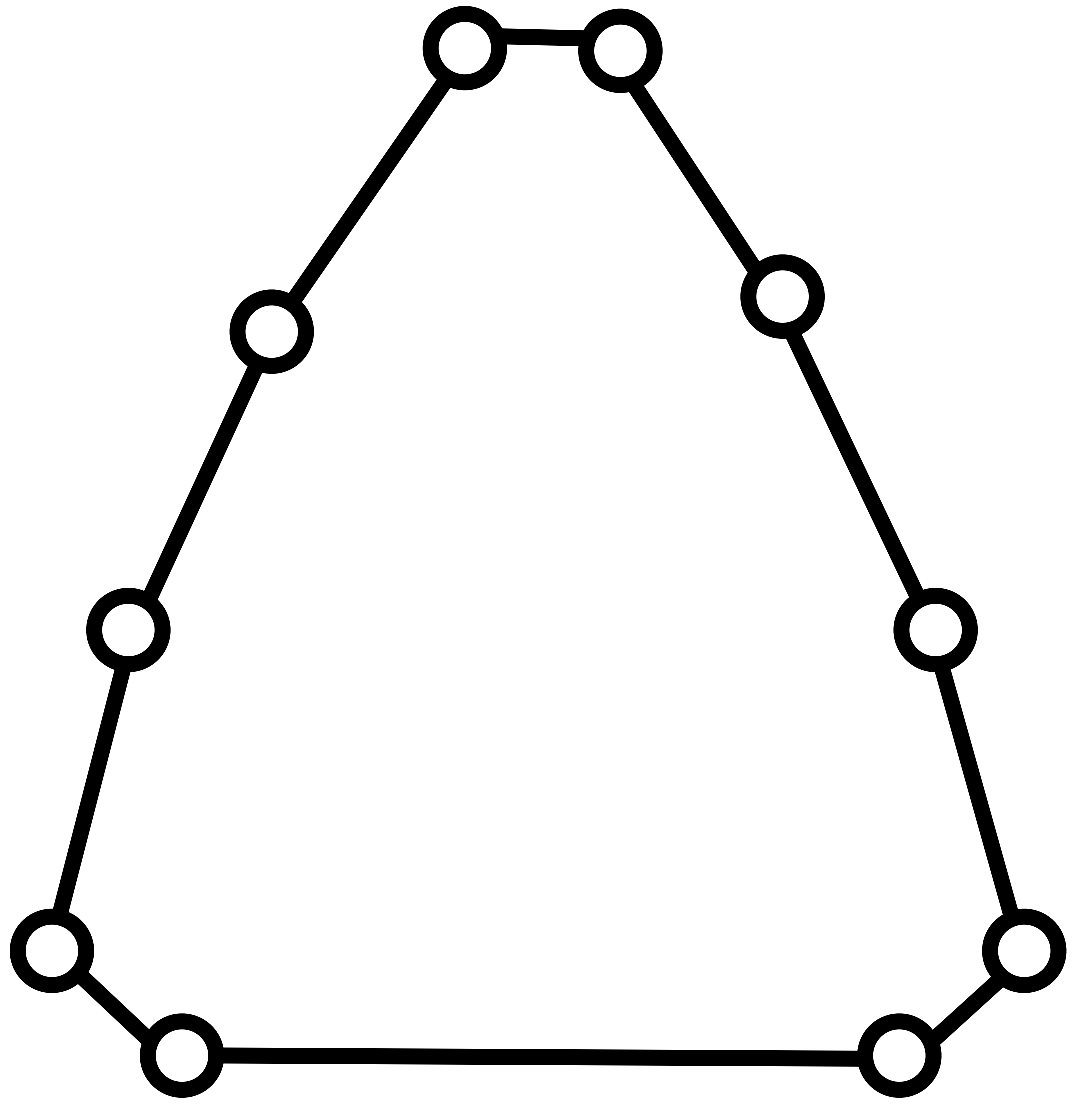


Compute New Error

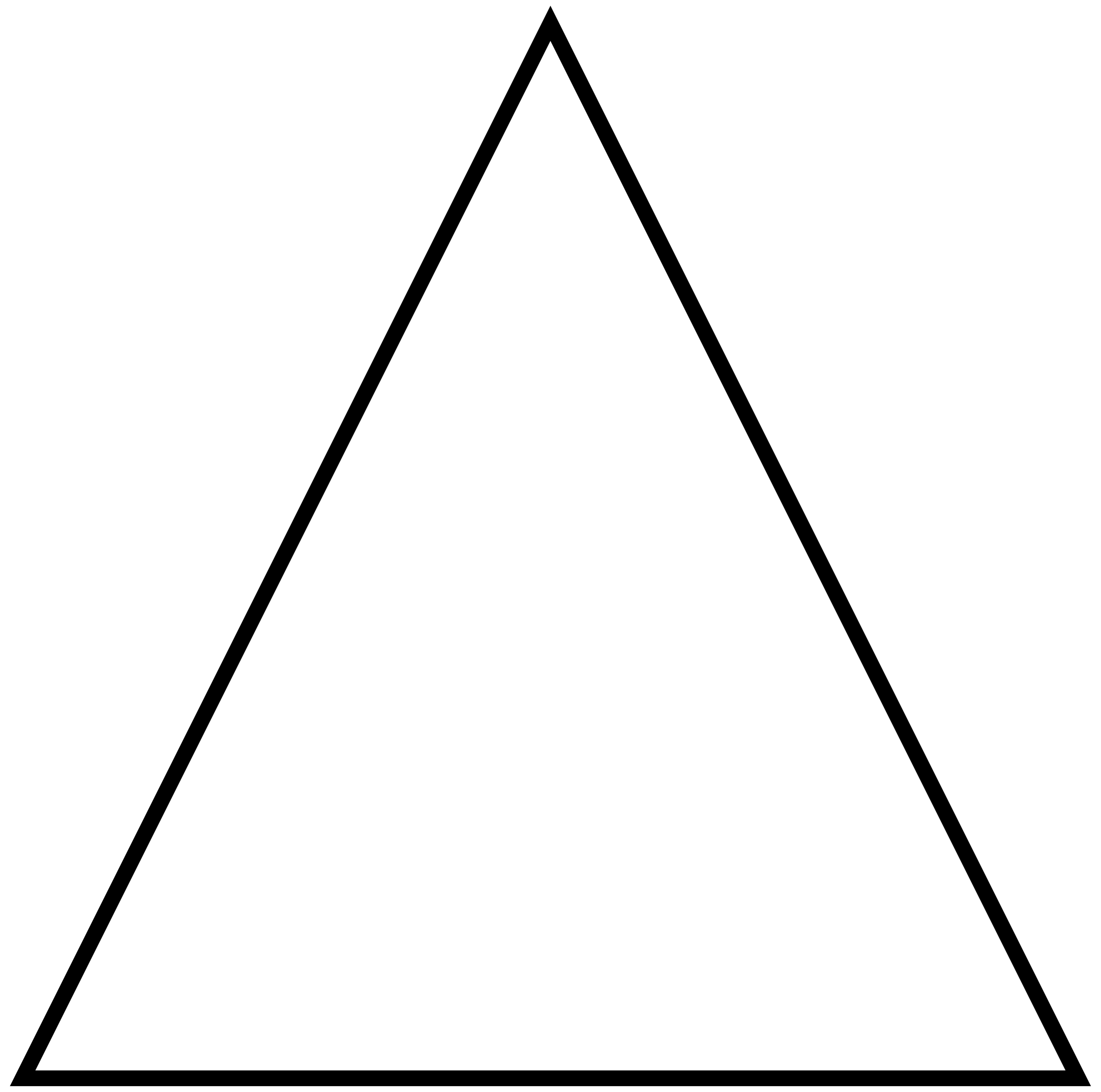


Target Geometry

Gradient Based Optimization

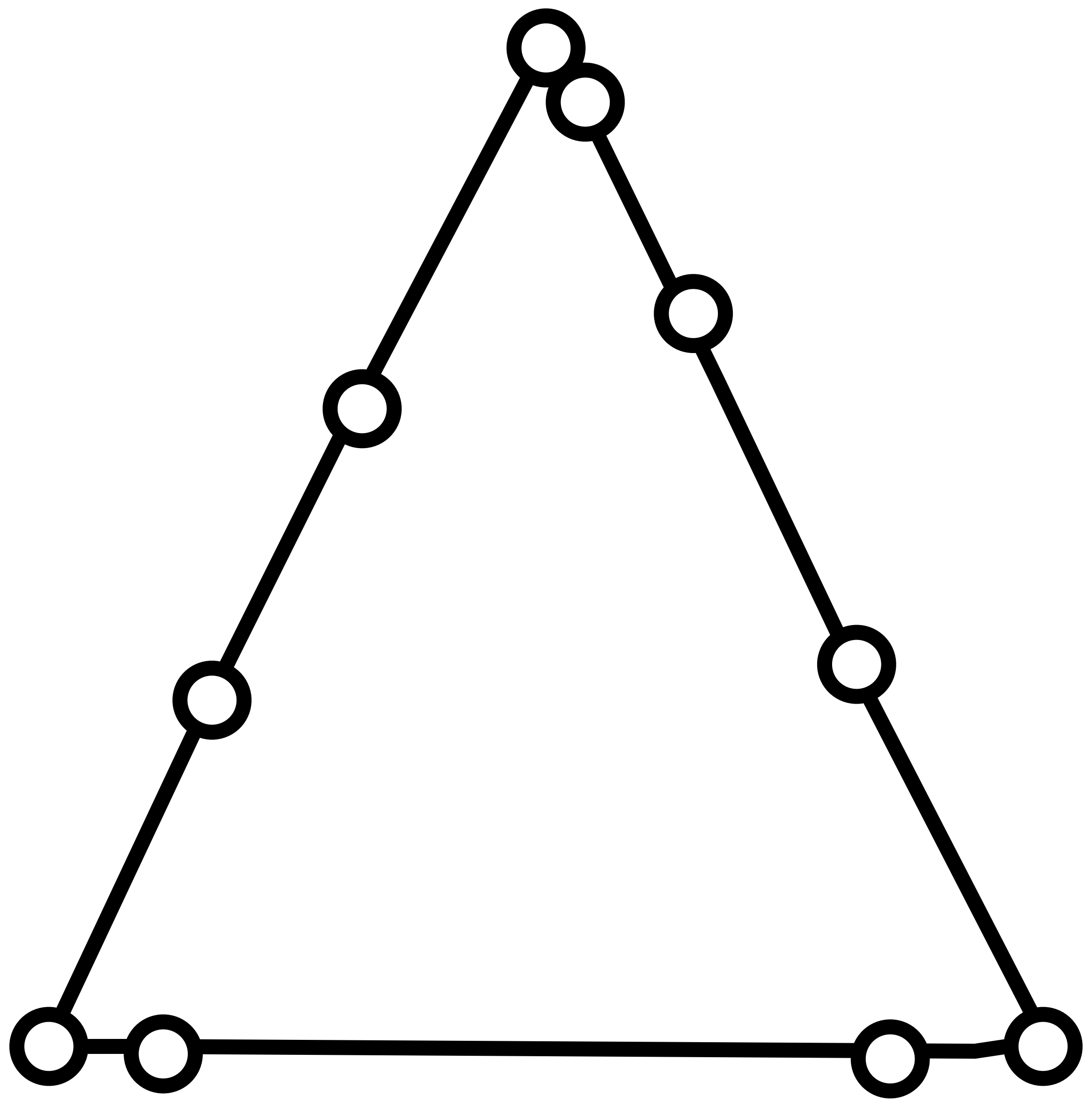


Repeat

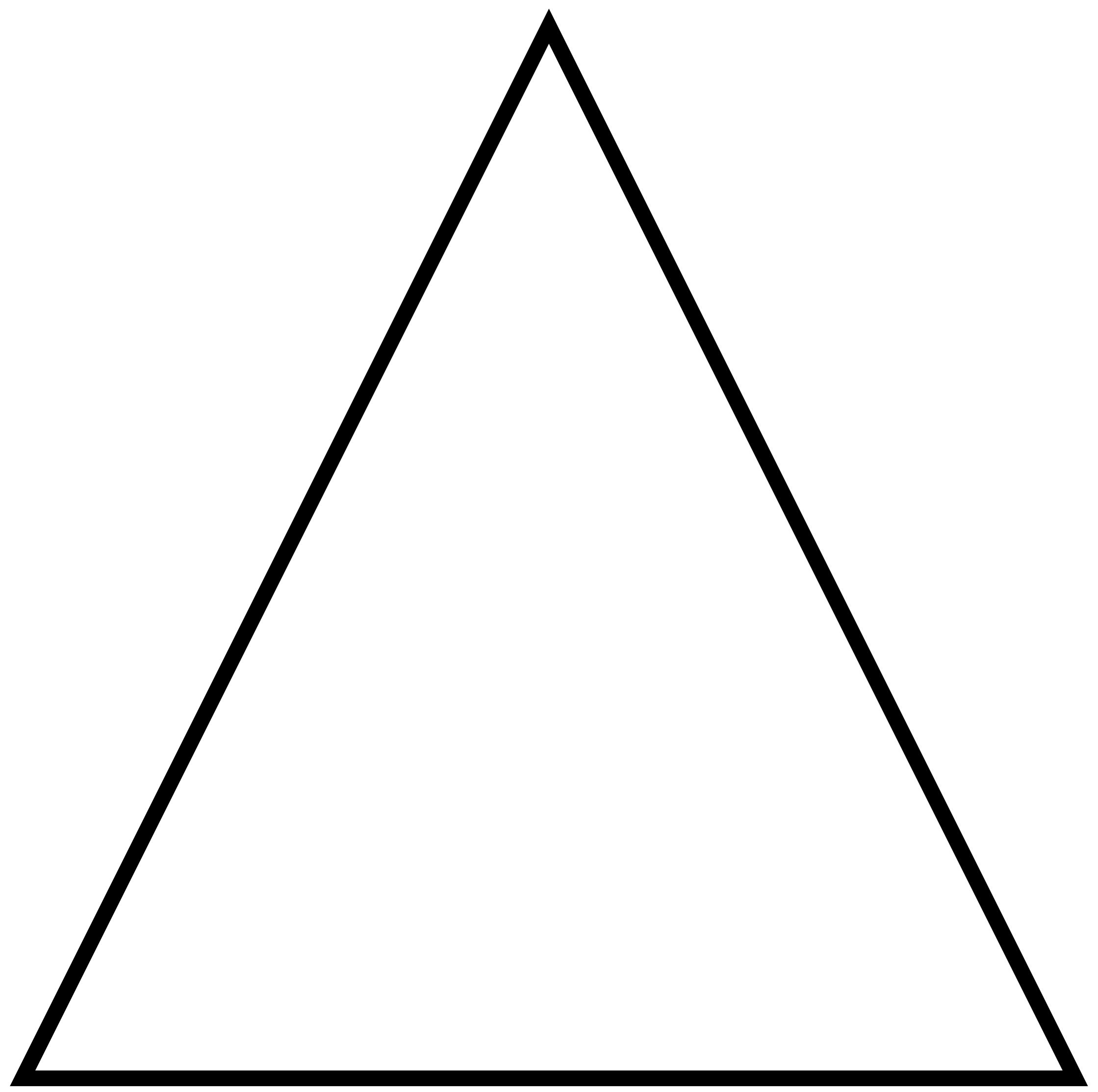


Target Geometry

Gradient Based Optimization

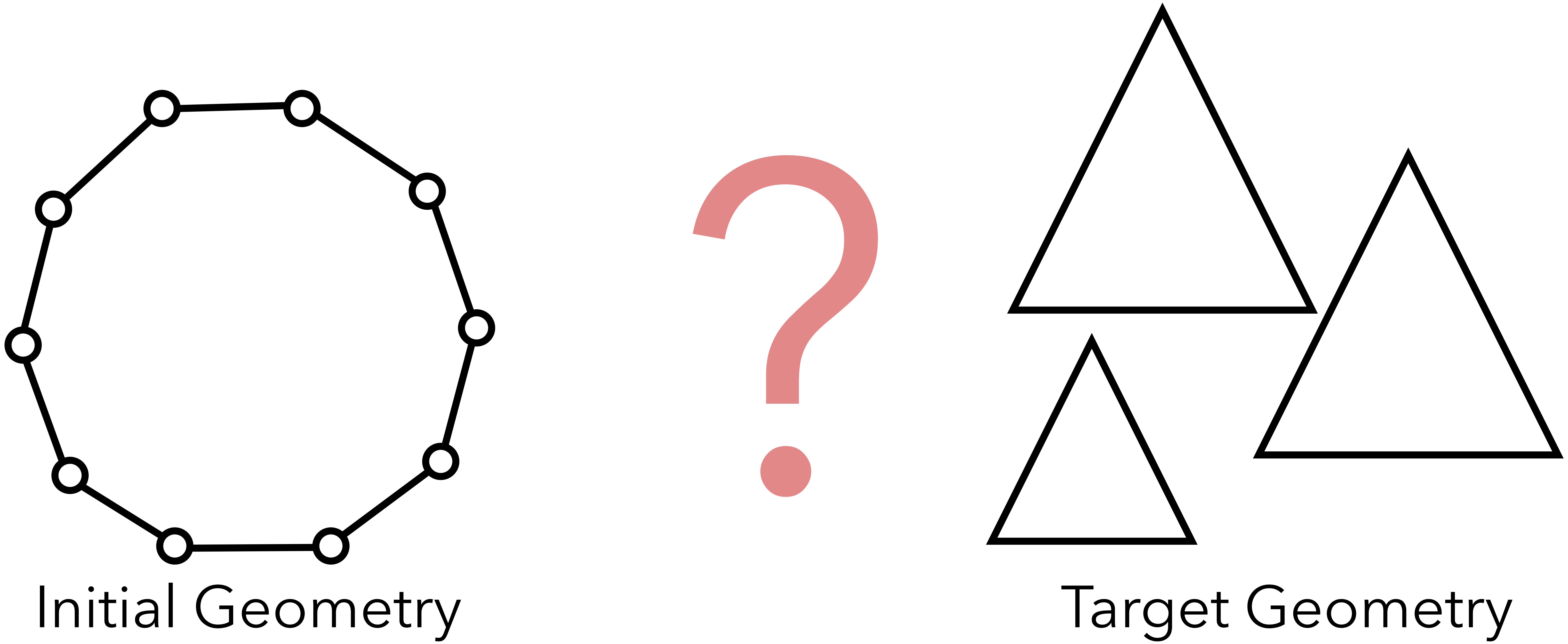


Repeat

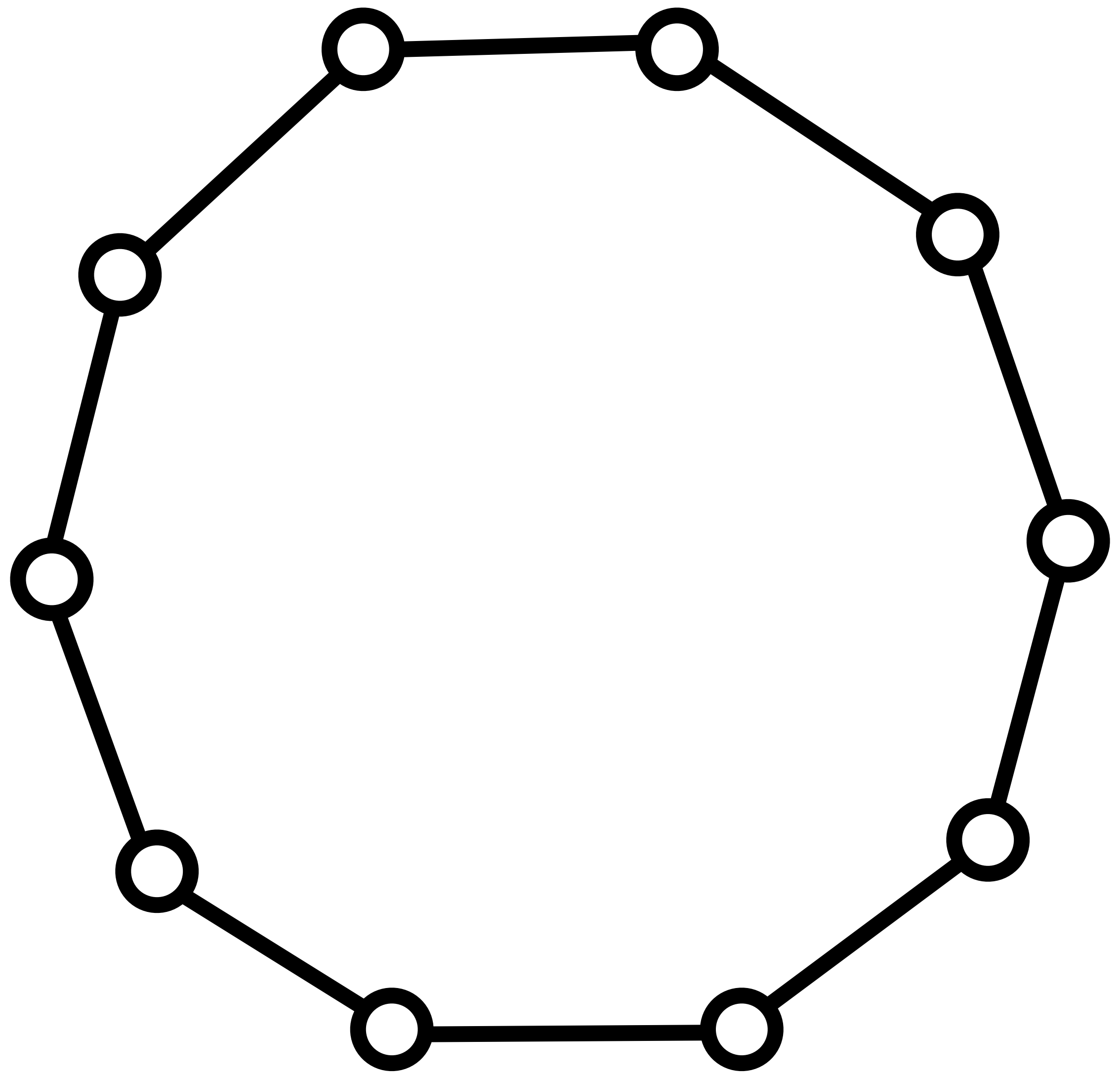


Target Geometry

Gradient Based Optimization

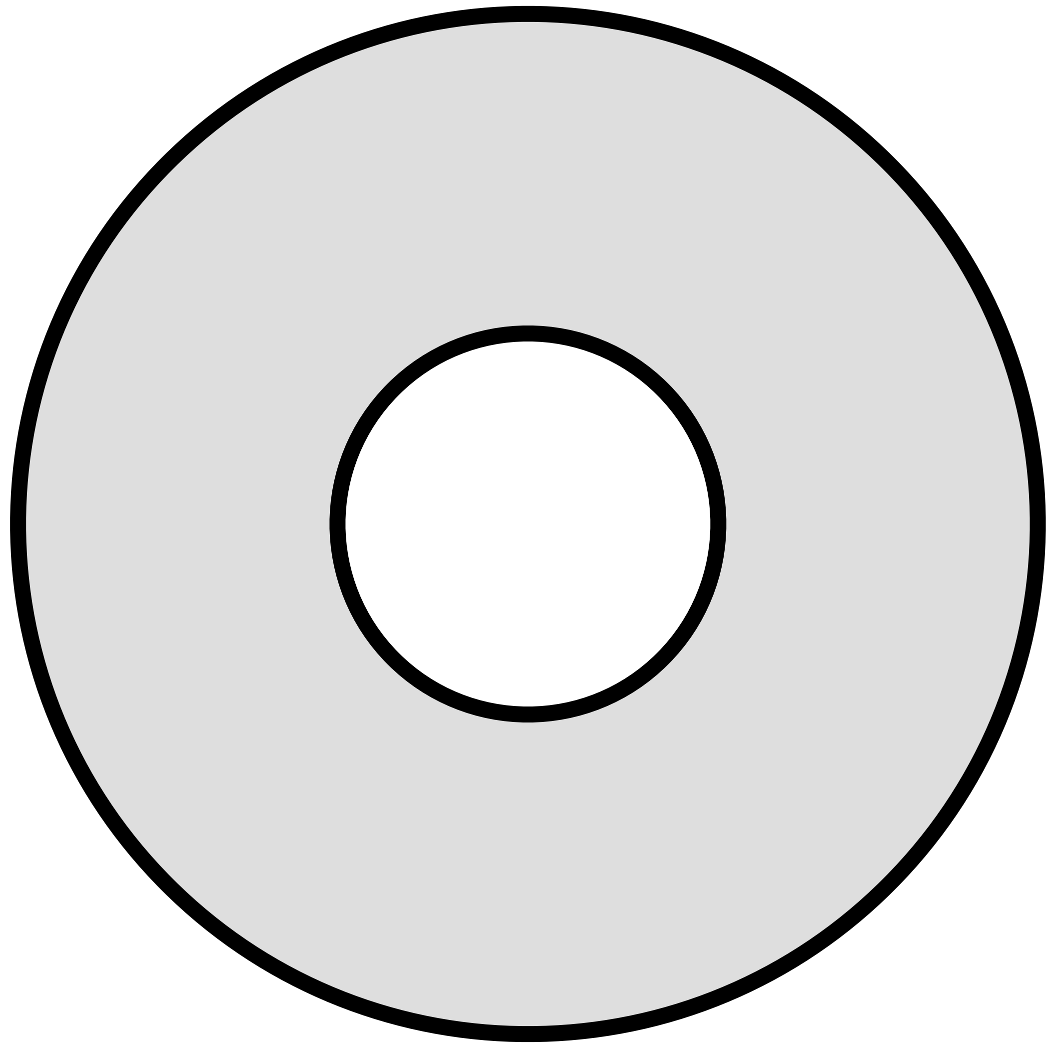


Gradient Based Optimization



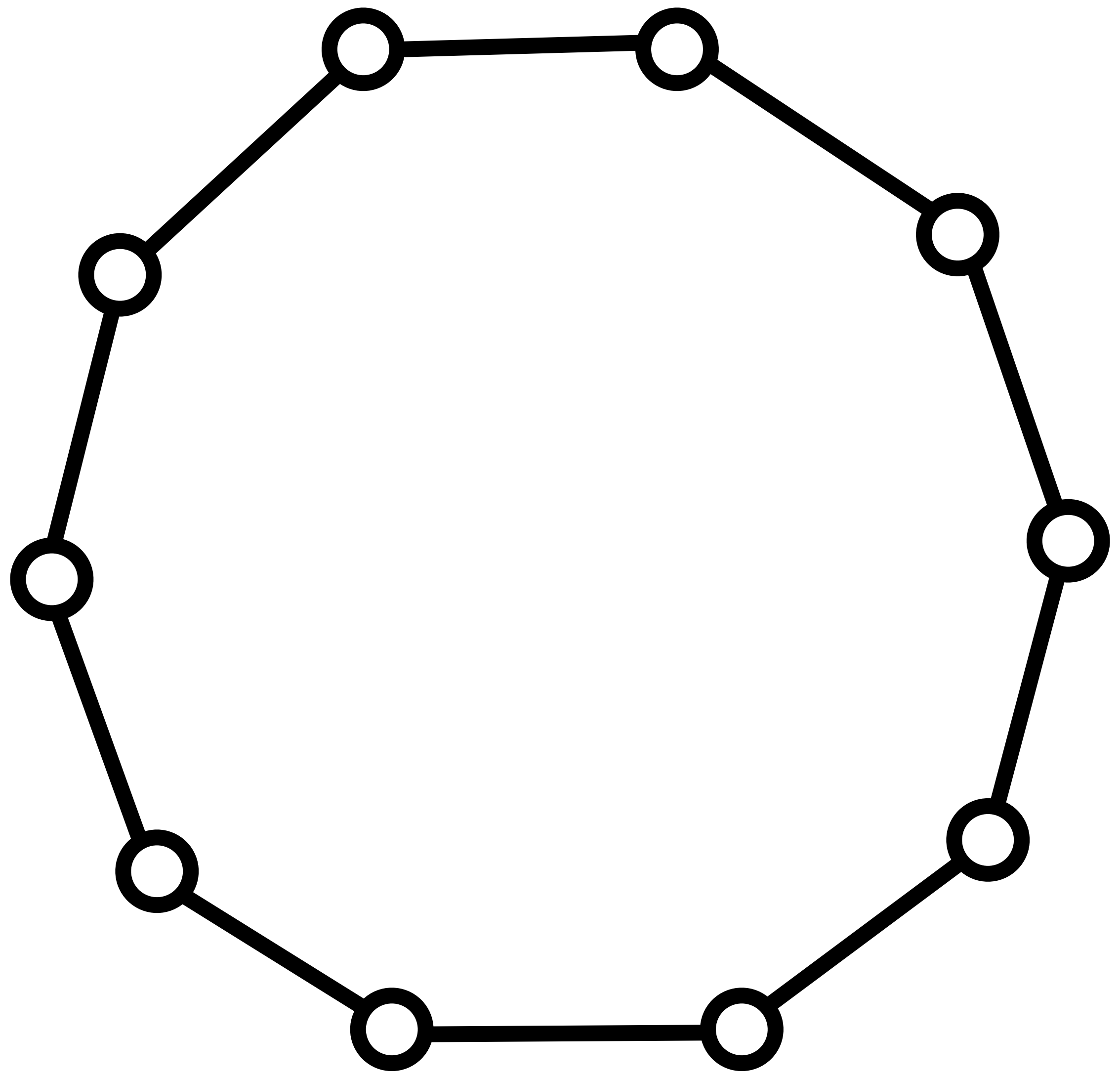
Initial Geometry

?



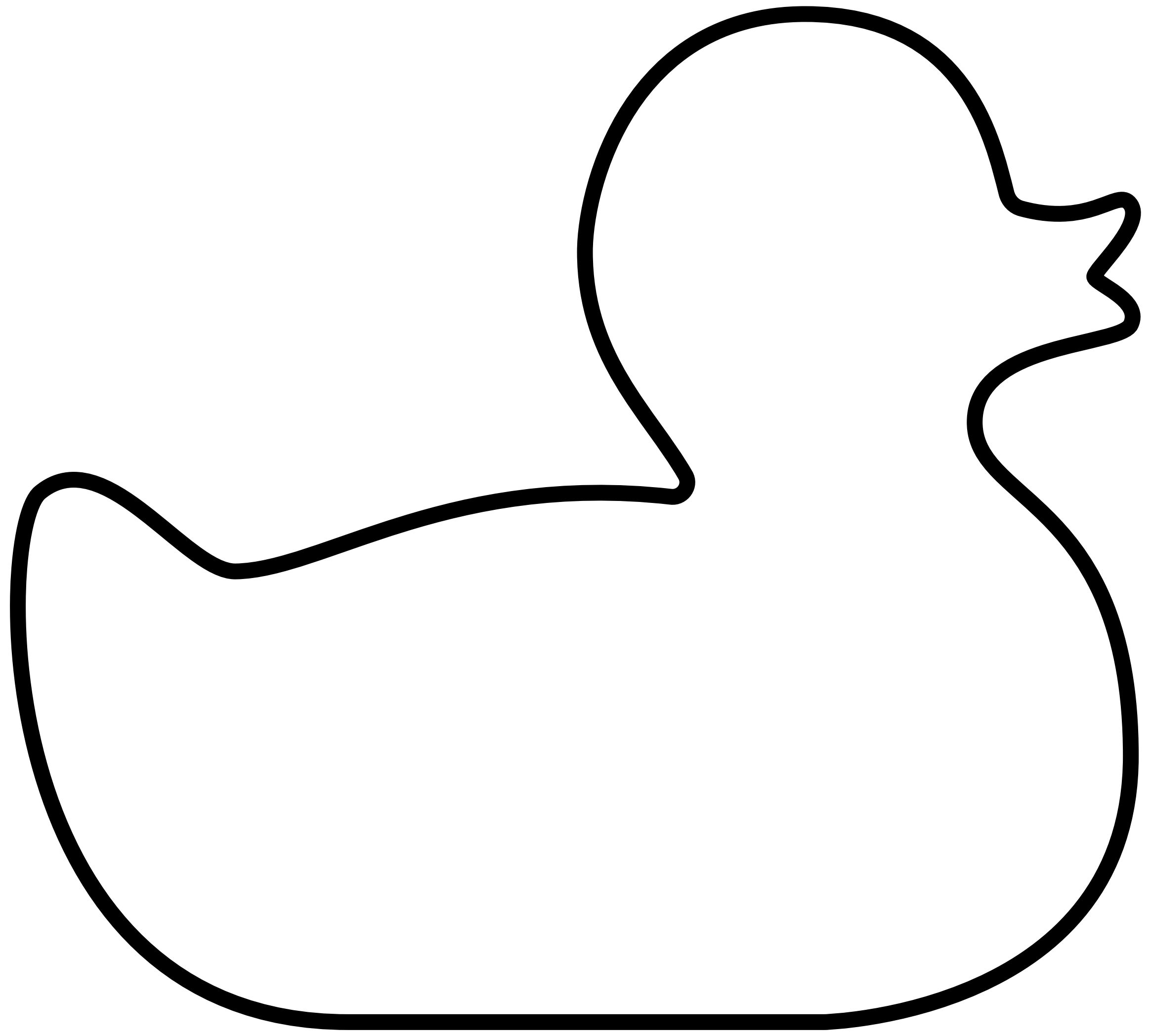
Target Geometry

Gradient Based Optimization



Initial Geometry

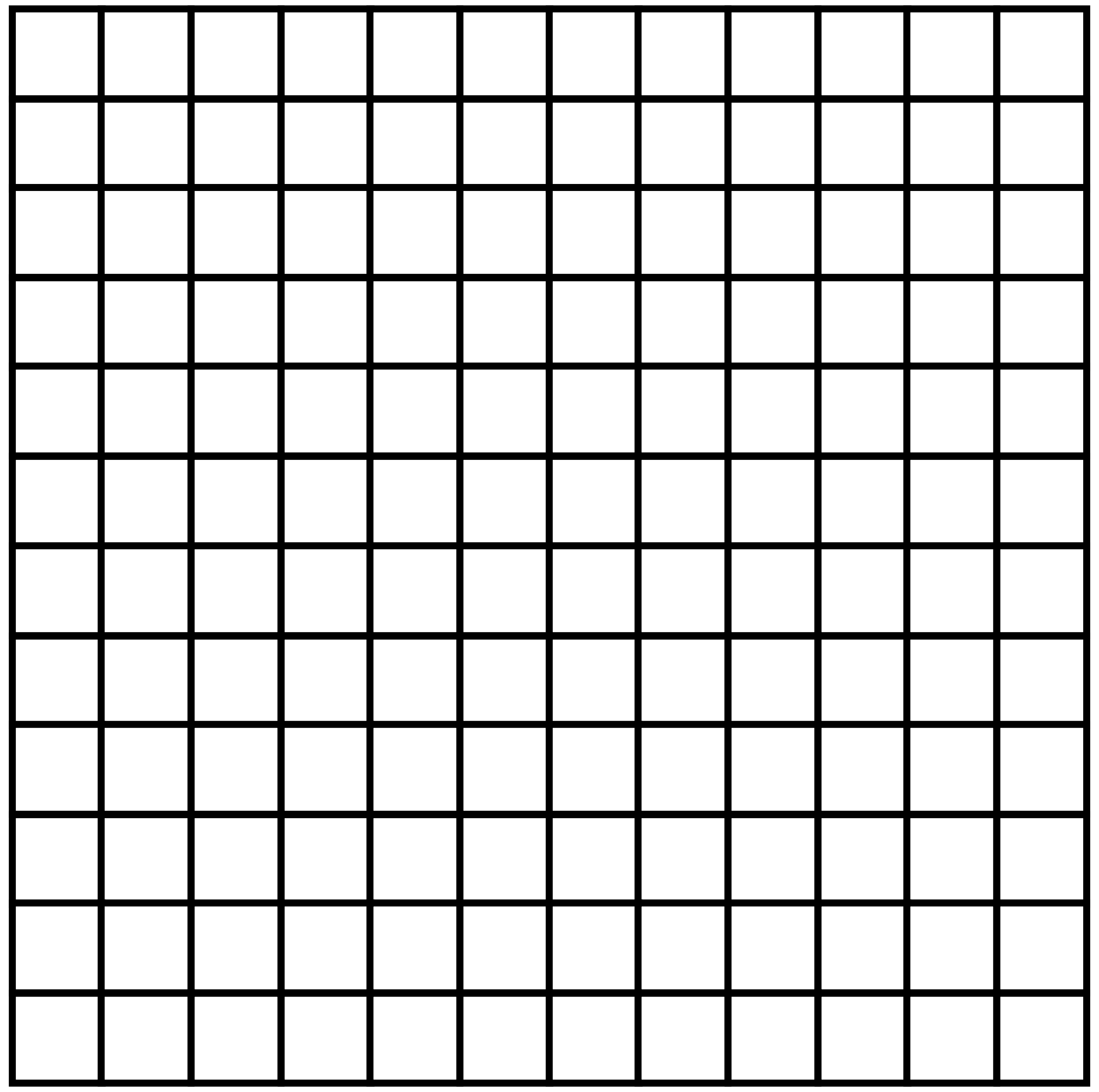
?



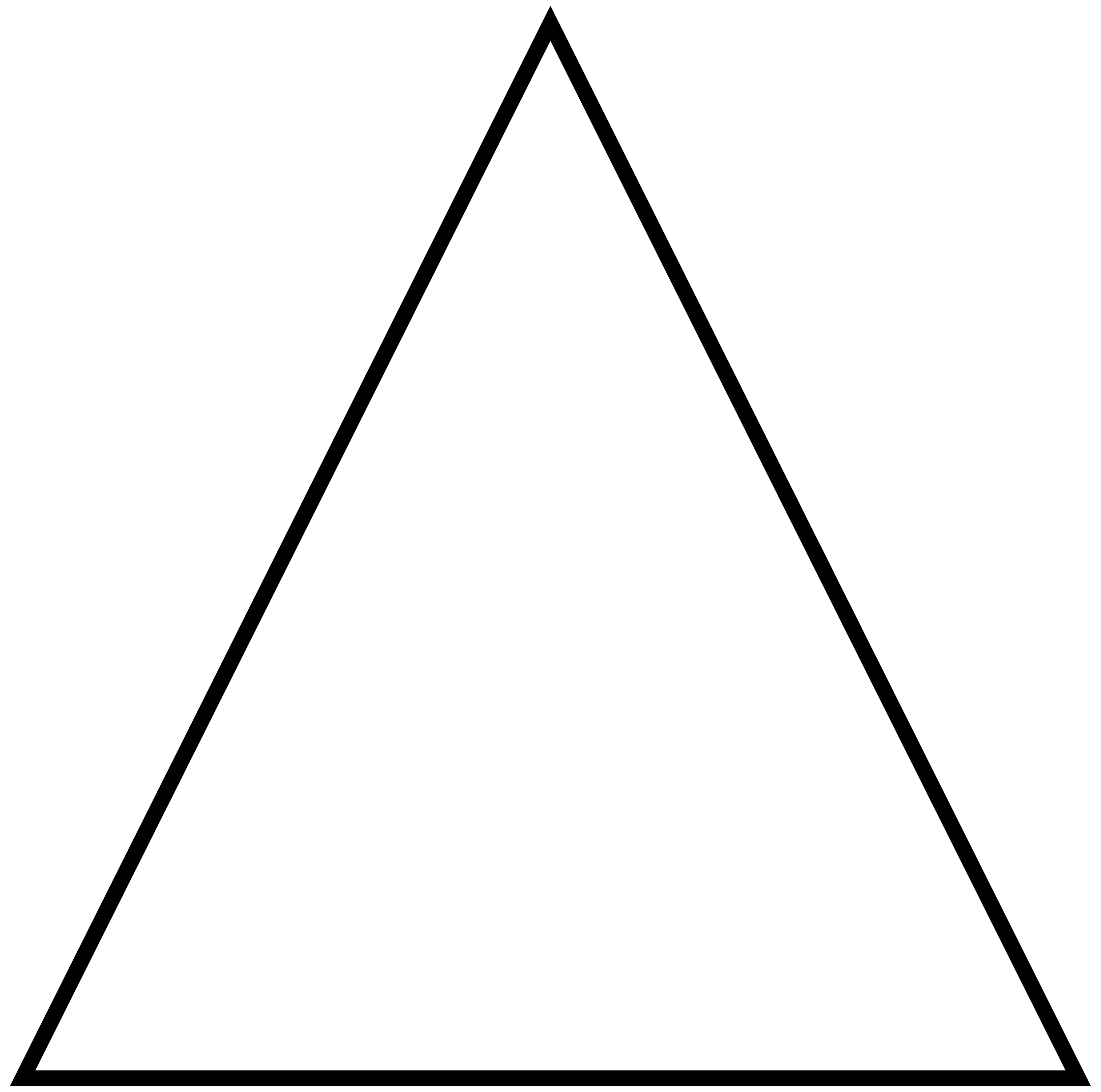
Target Geometry

Voxel Representation

Gradient Based Optimization

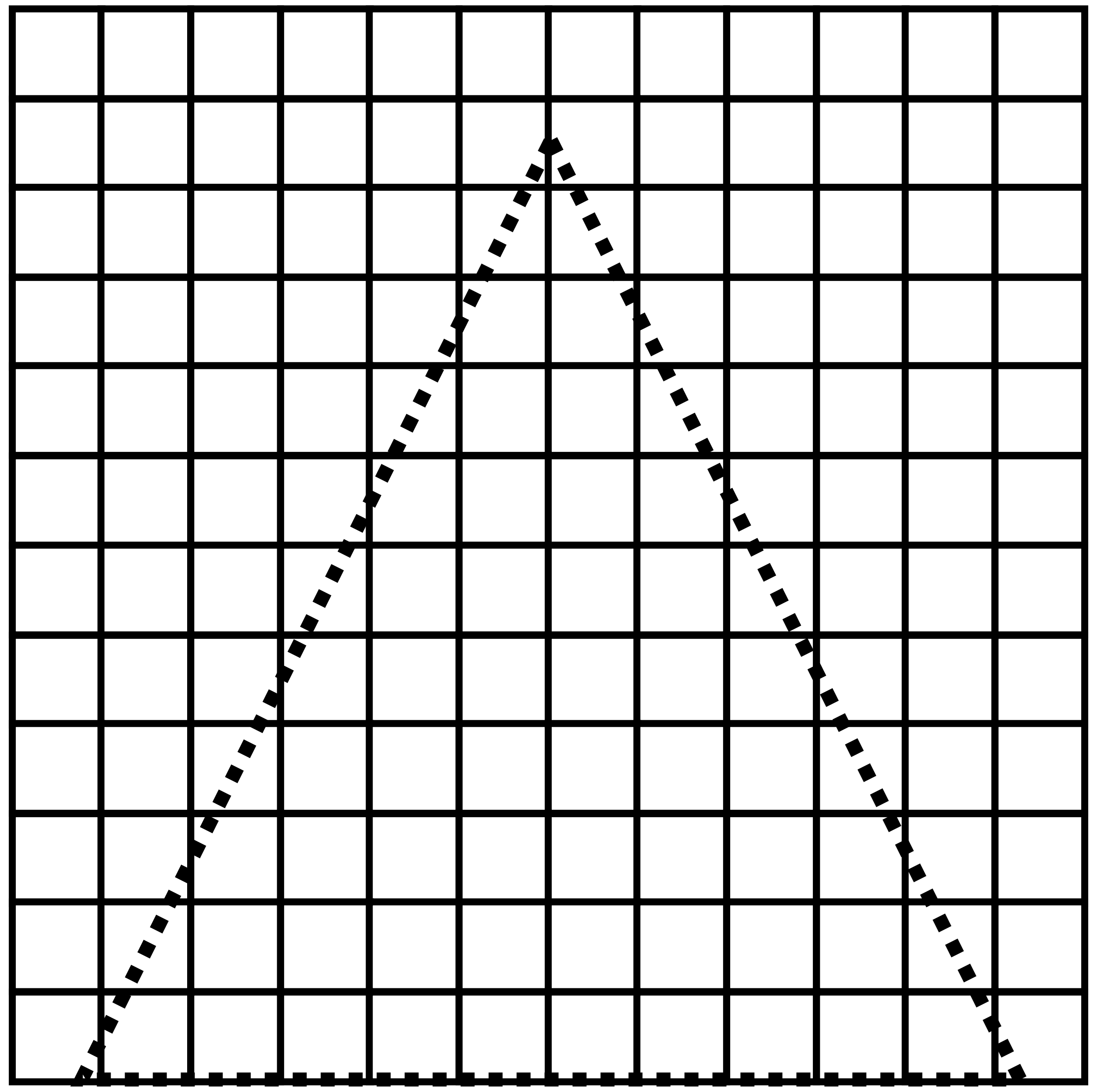


Initialized Grid

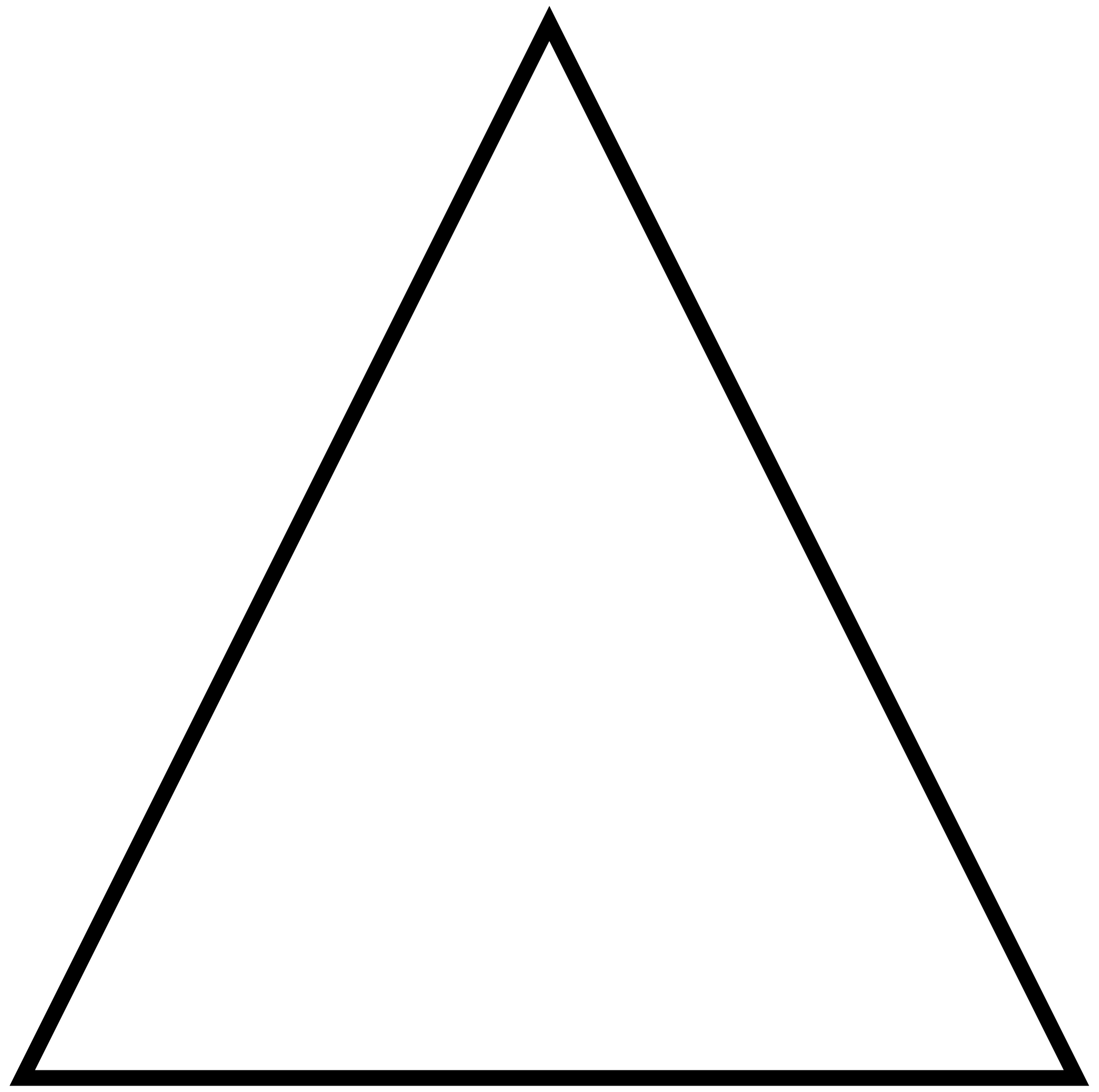


Target Geometry

Gradient Based Optimization

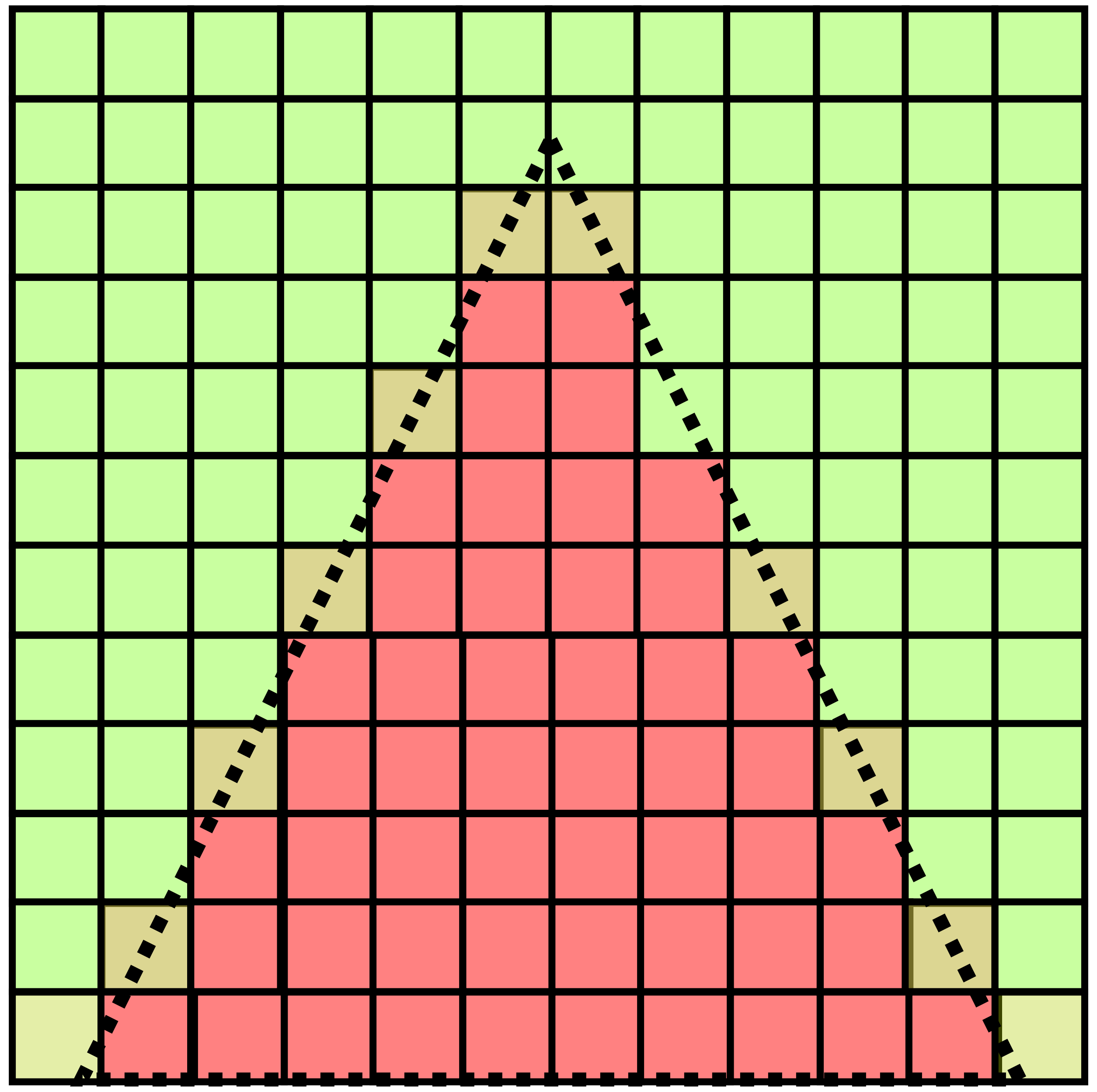


Initialized Grid

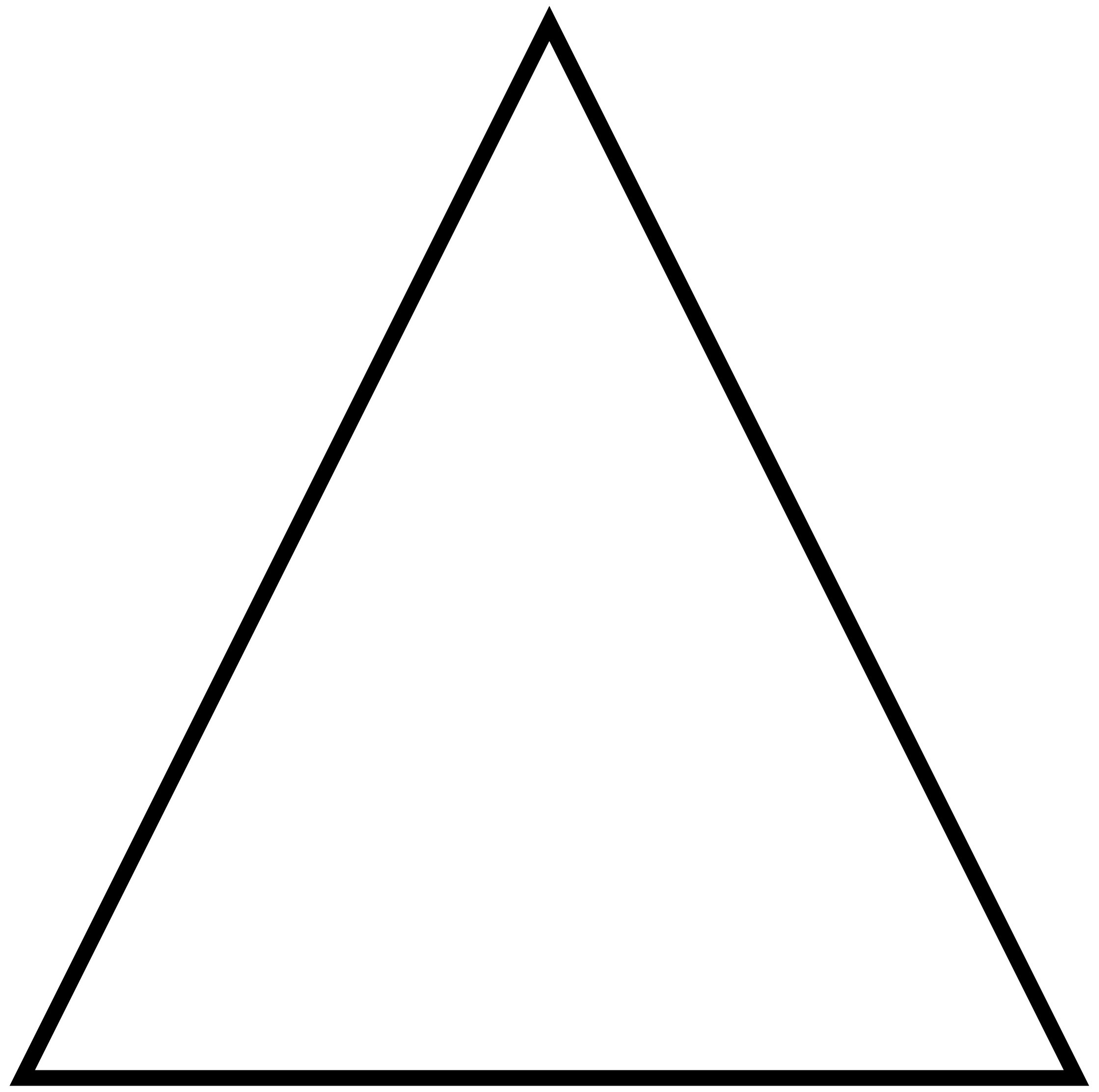


Target Geometry

Gradient Based Optimization

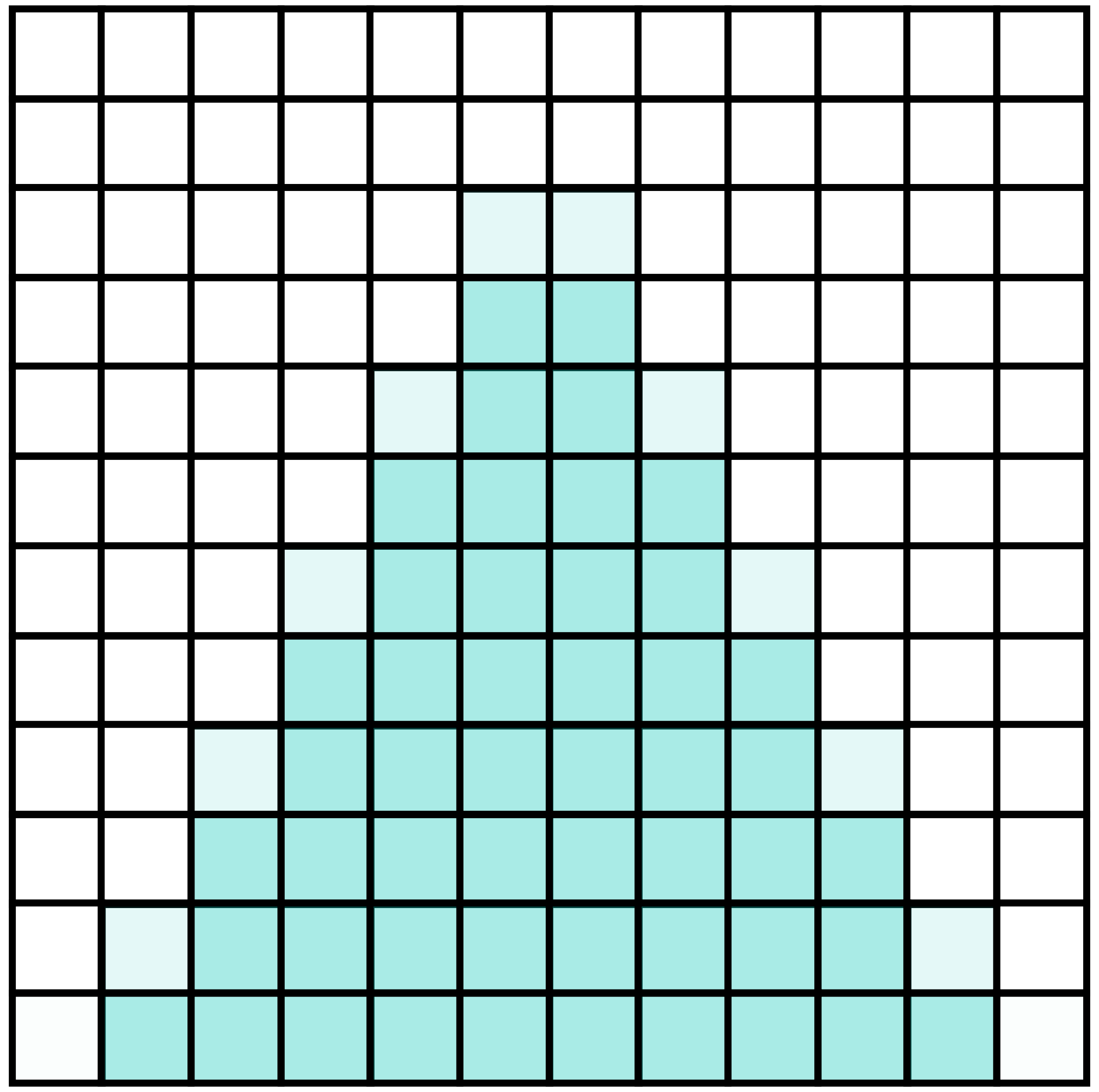


Loss

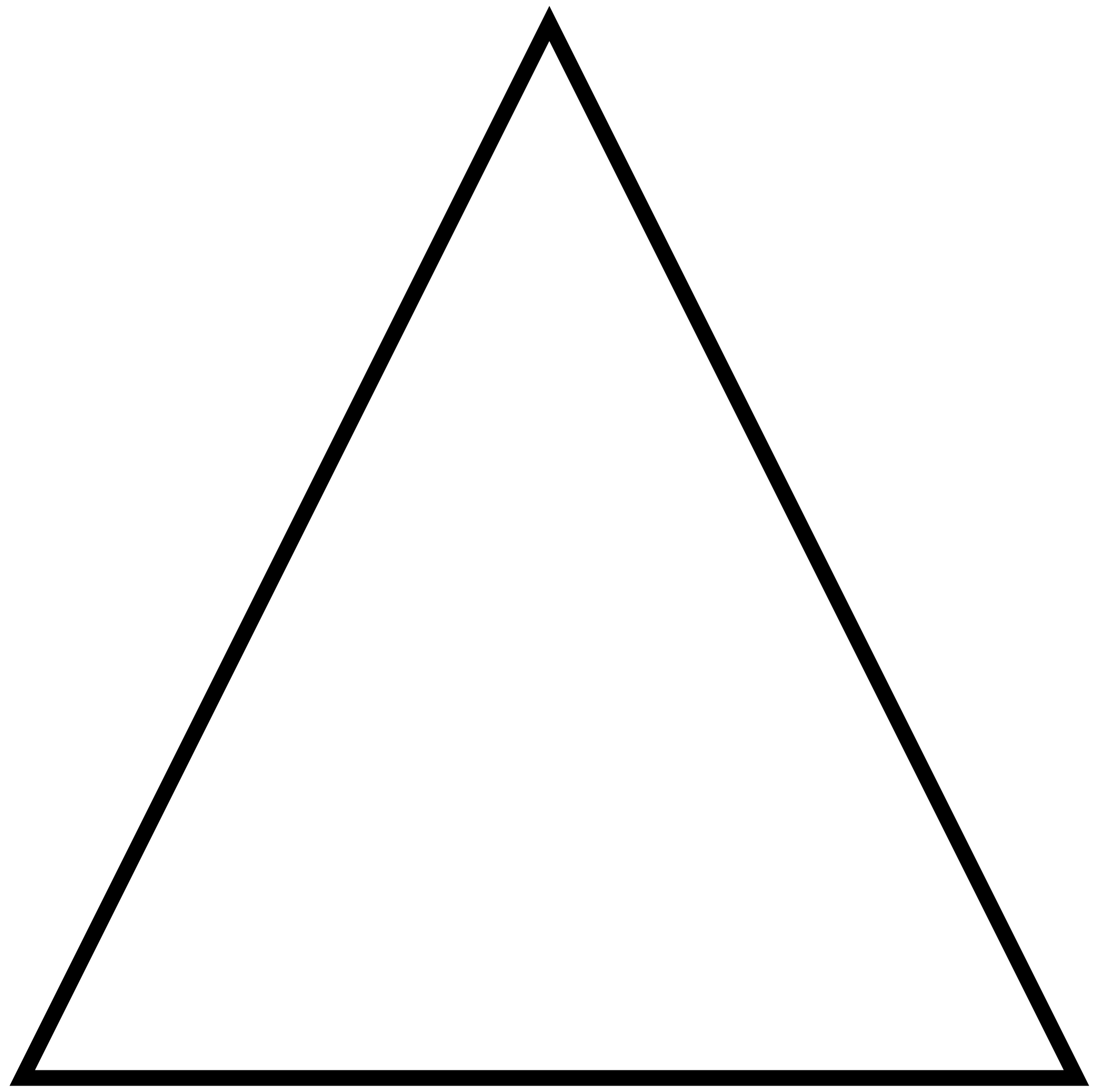


Target Geometry

Gradient Based Optimization

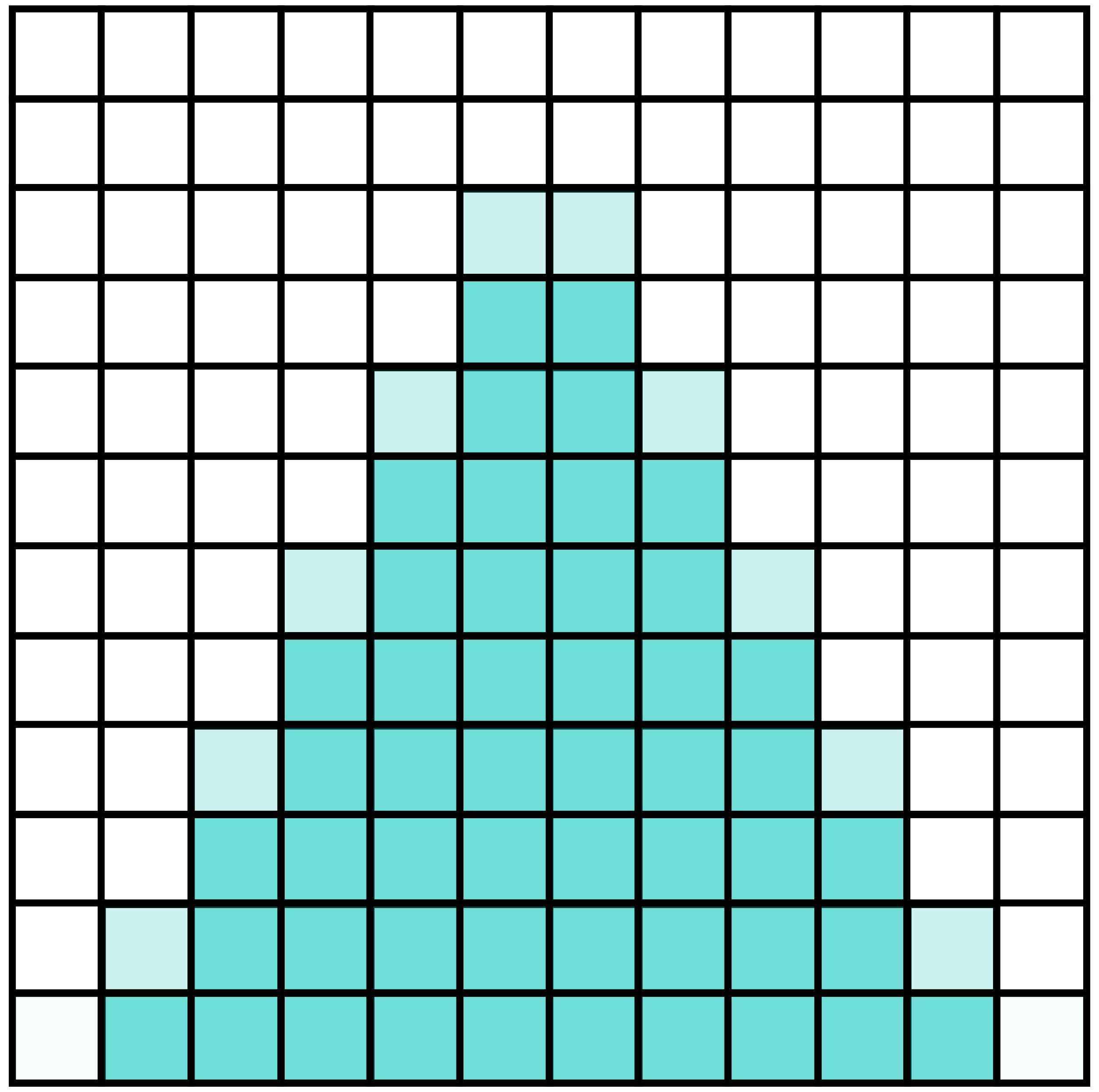


Gradient Step

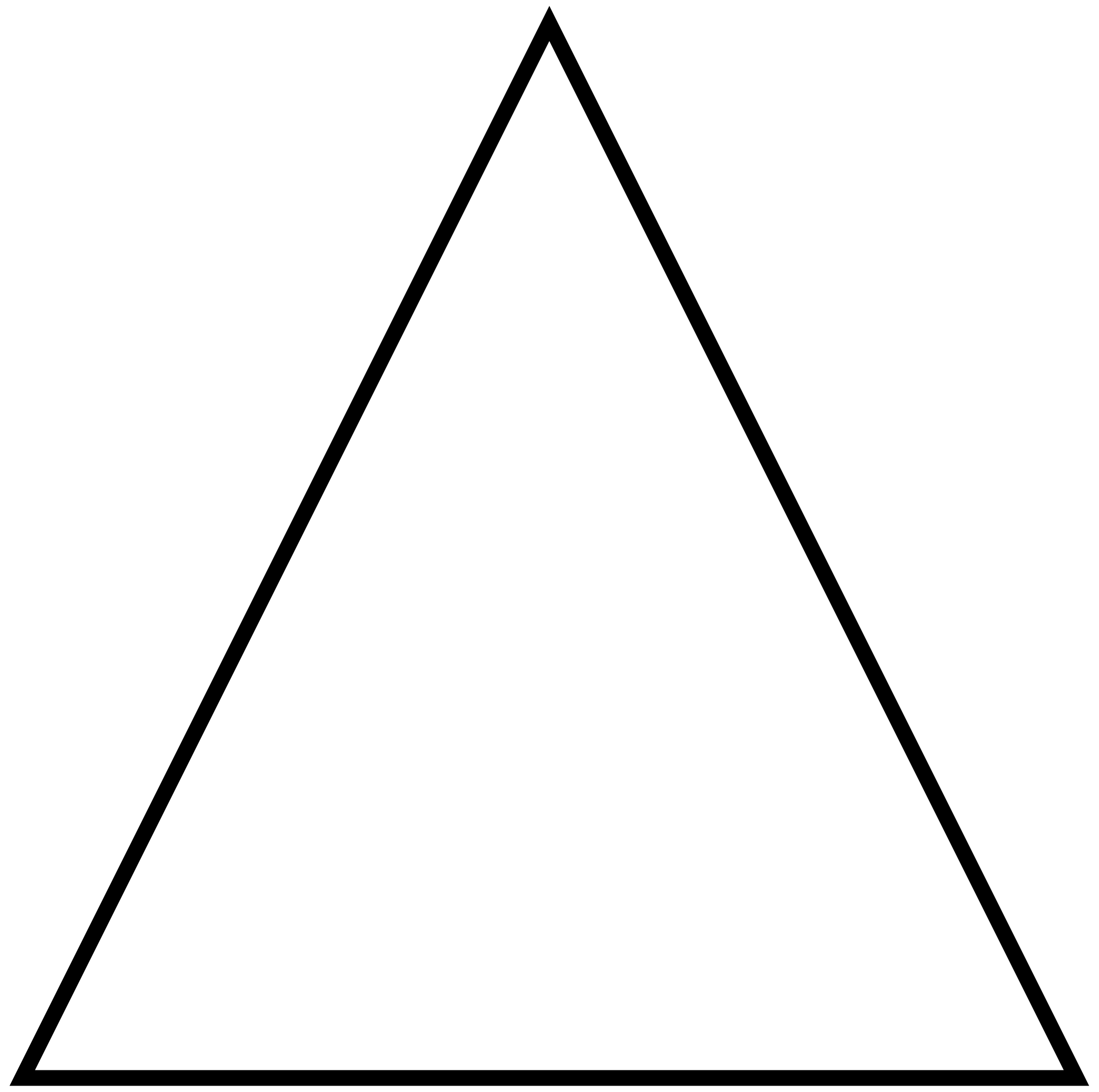


Target Geometry

Gradient Based Optimization

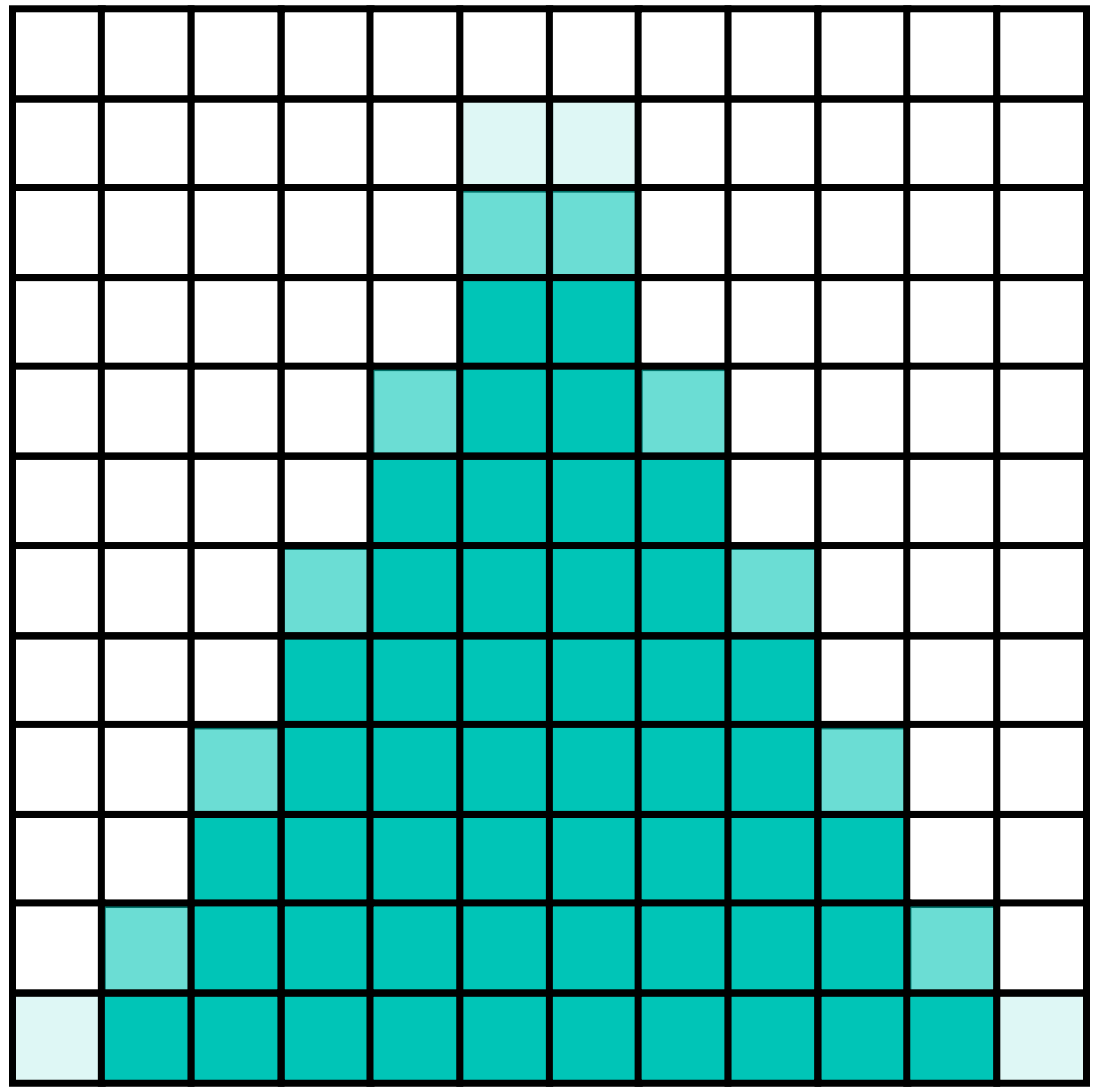


Repeat

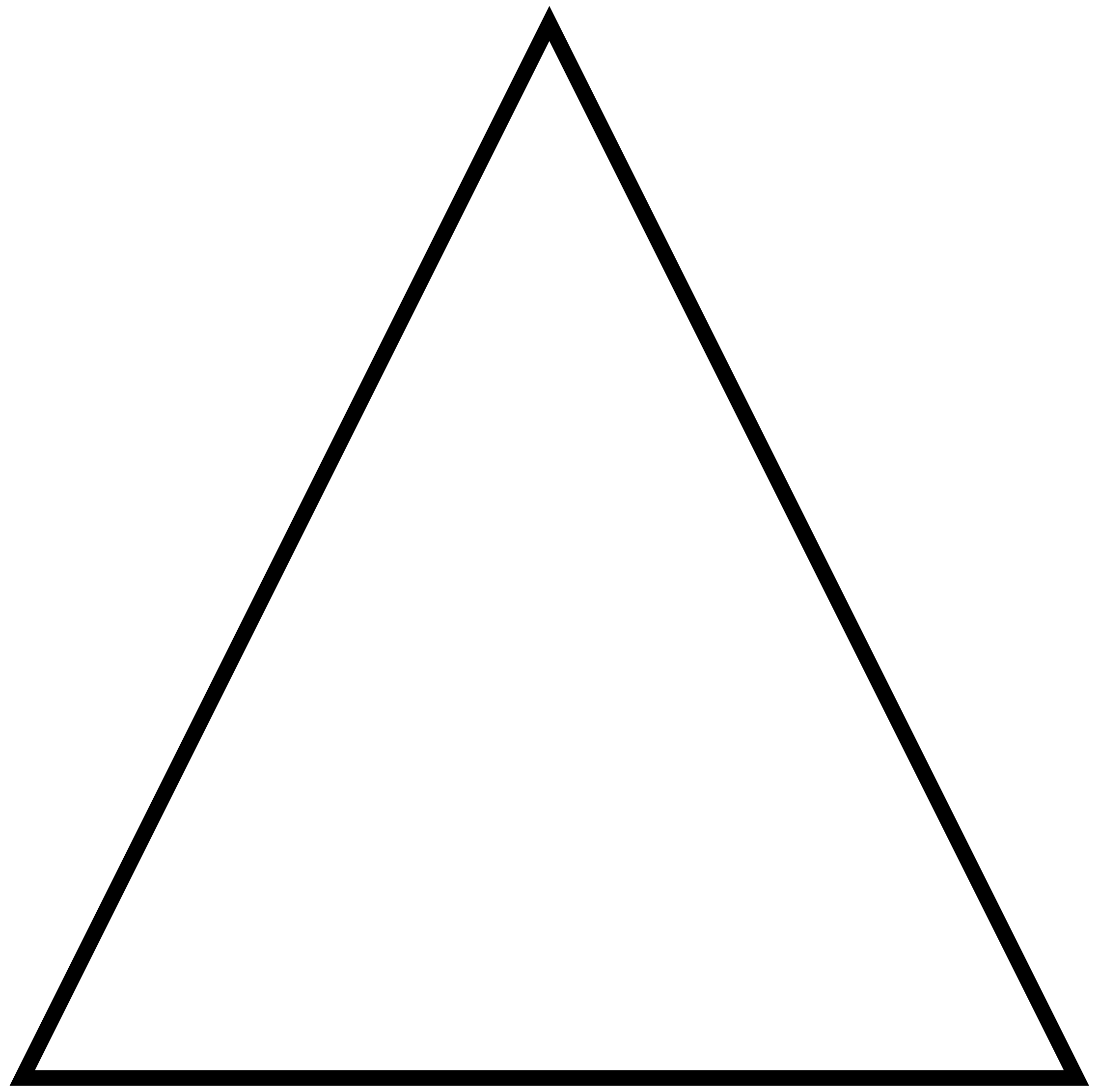


Target Geometry

Gradient Based Optimization

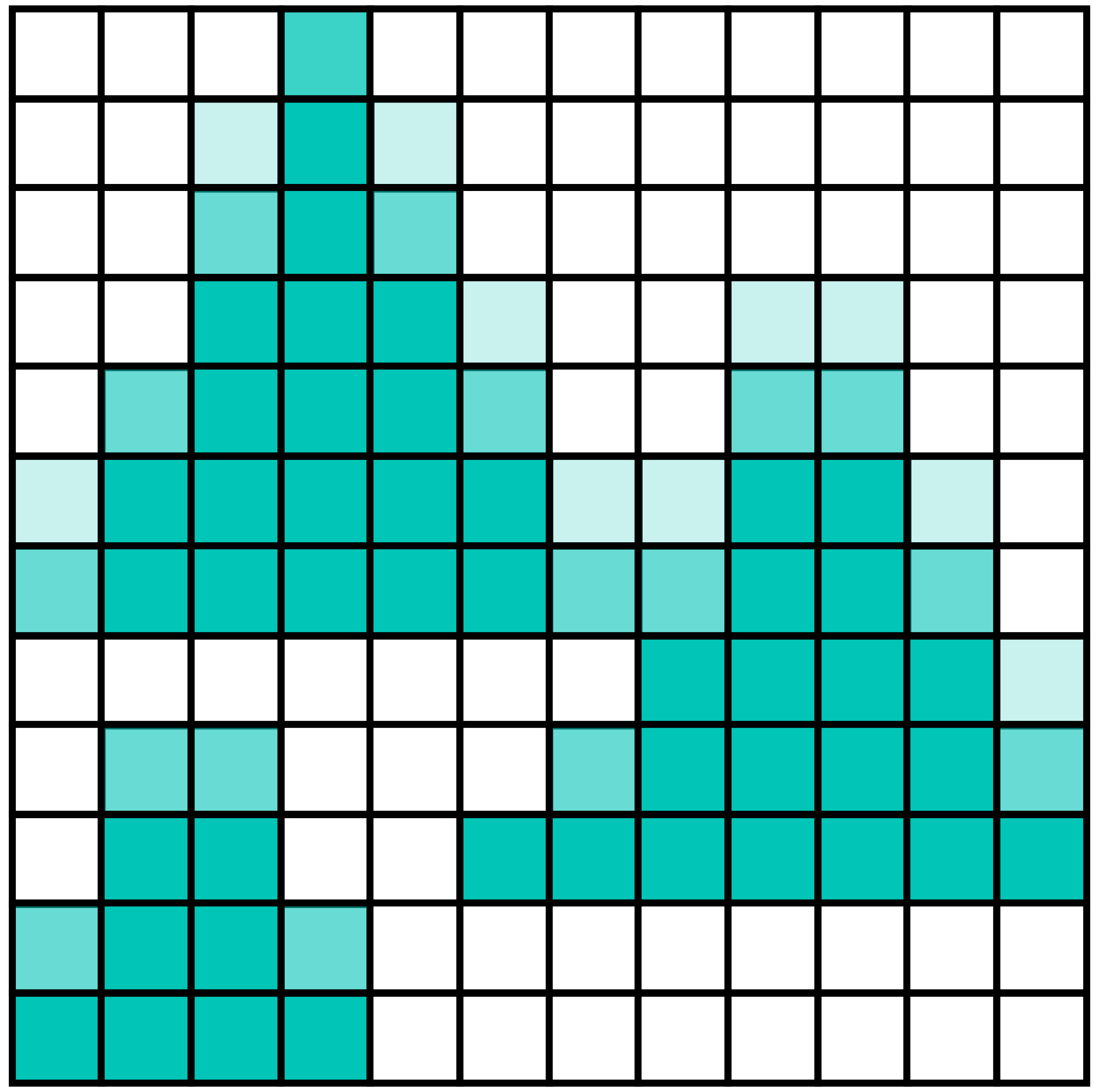


Repeat

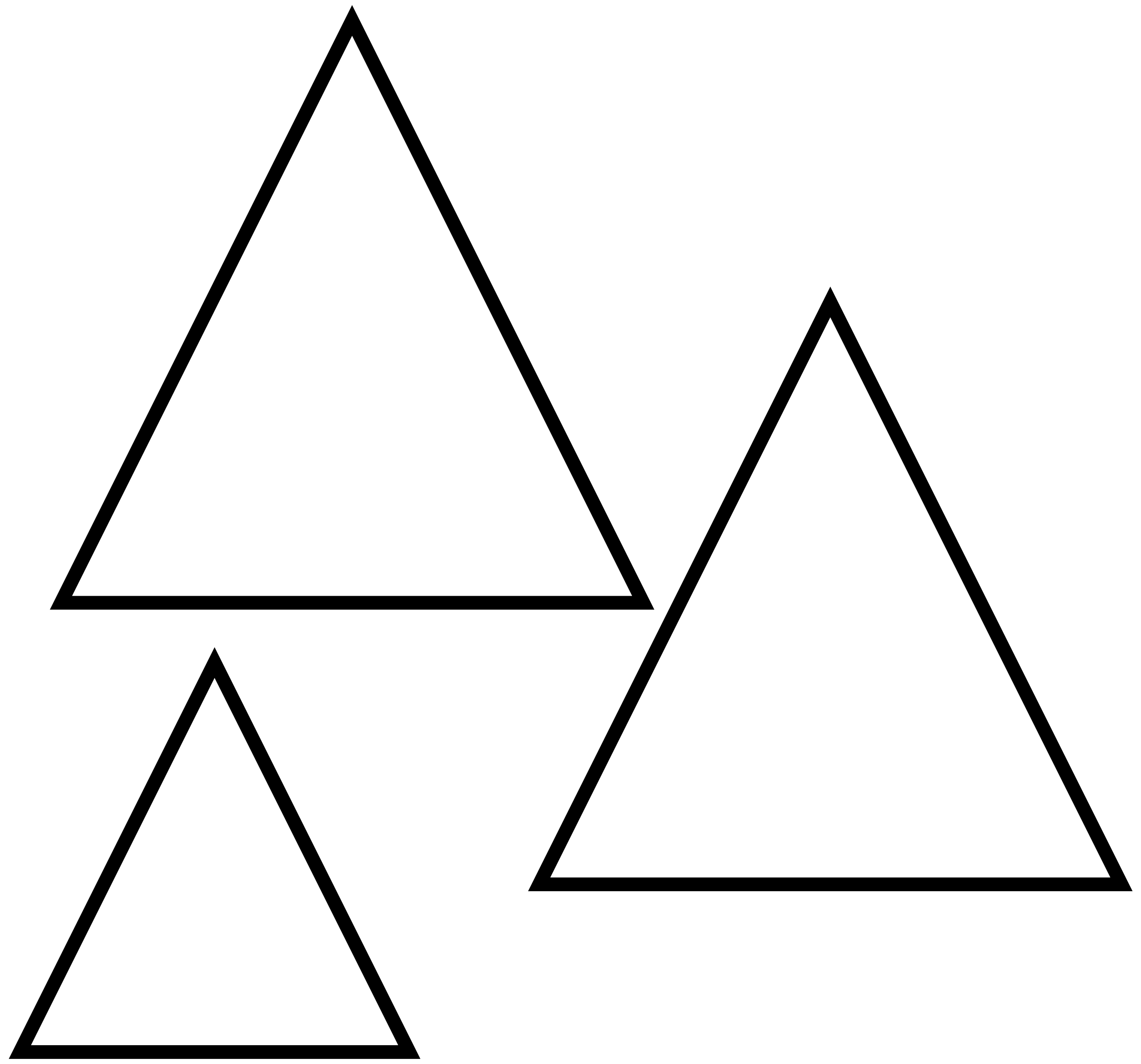


Target Geometry

Gradient Based Optimization

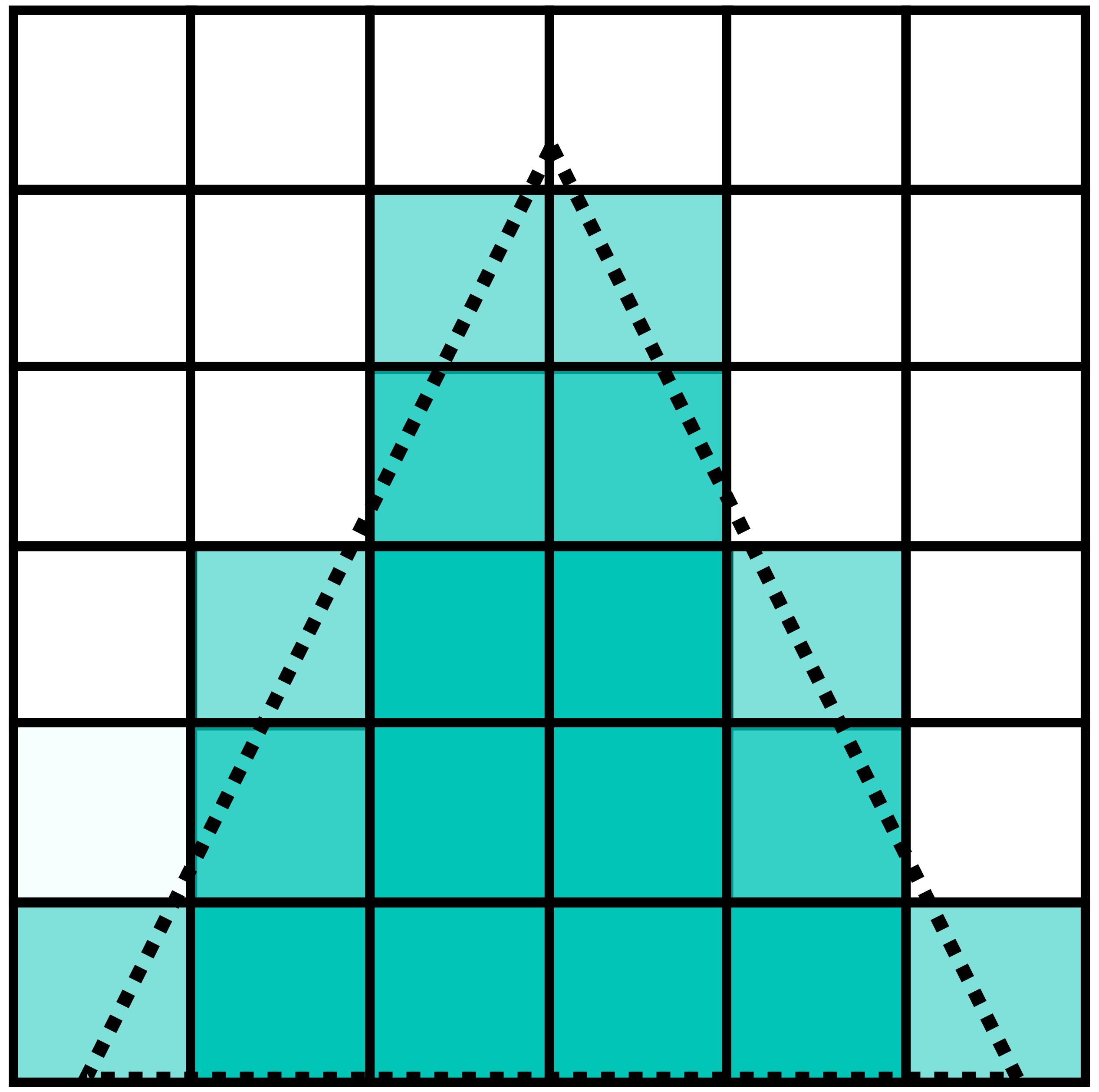


Reconstruction

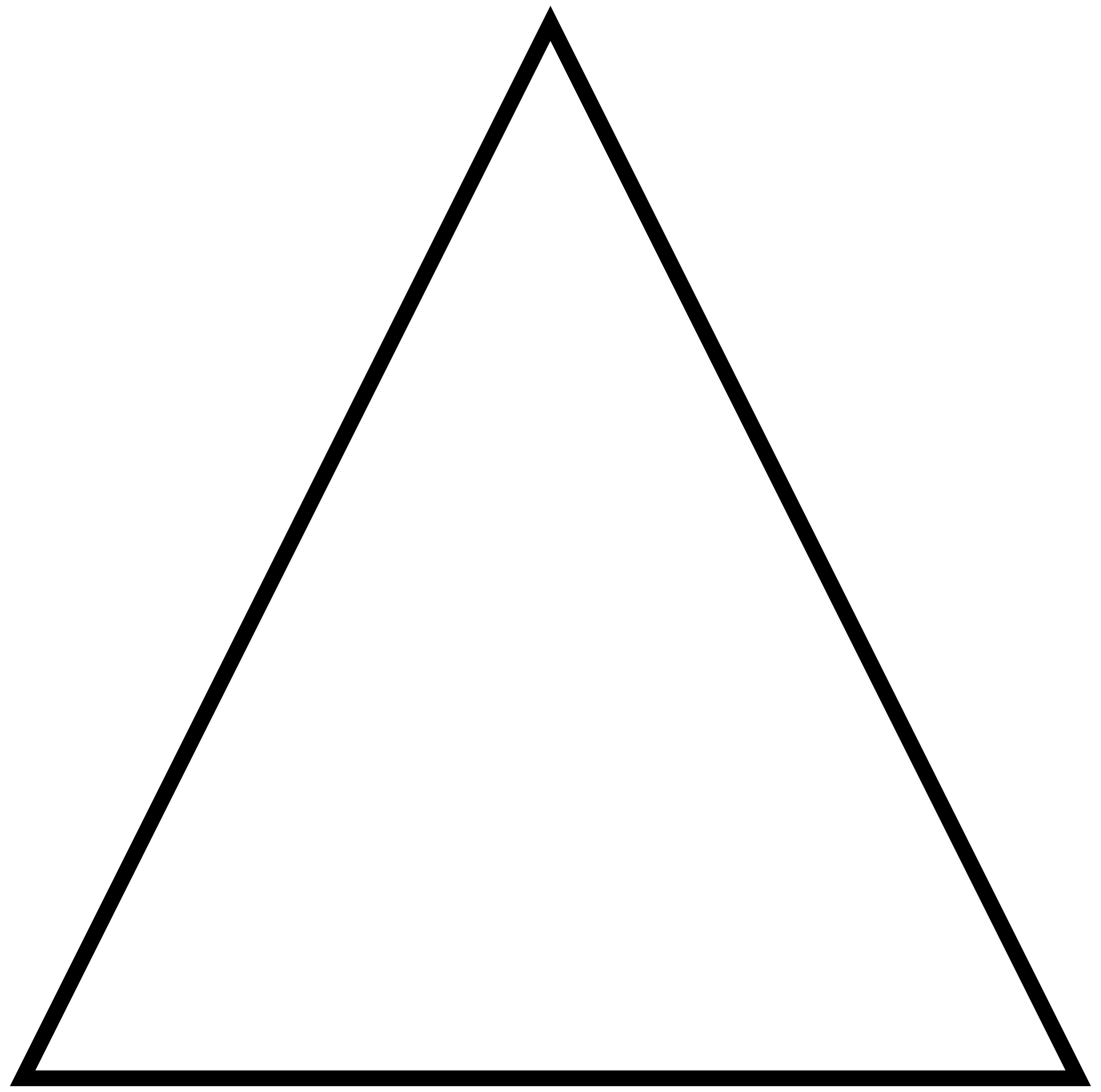


Target Geometry

Gradient Based Optimization

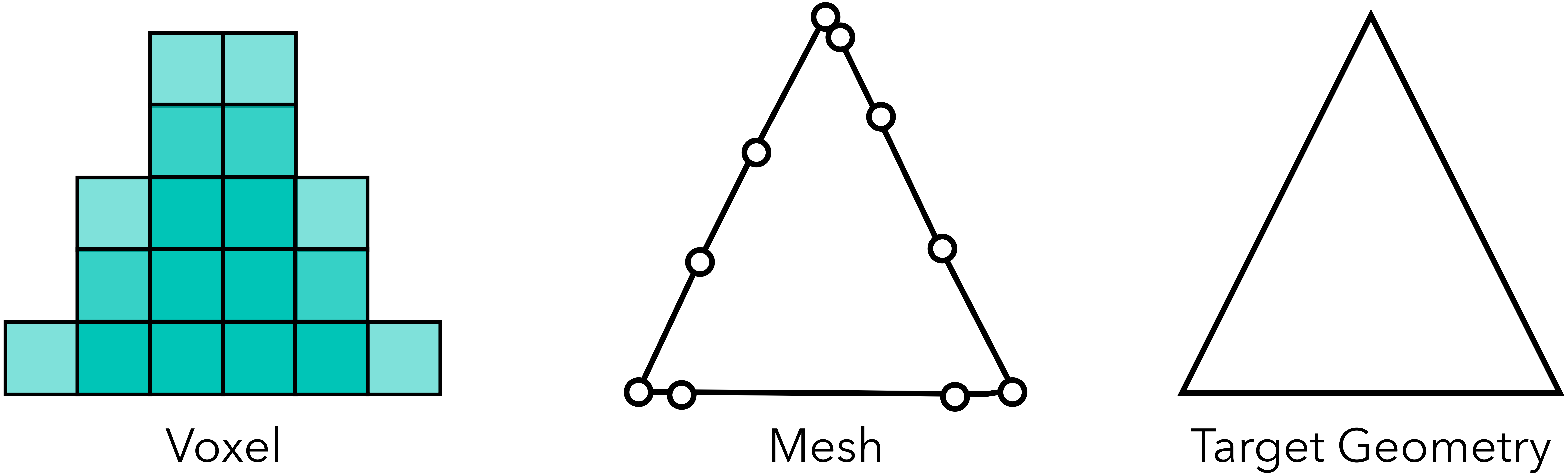


Reconstruction

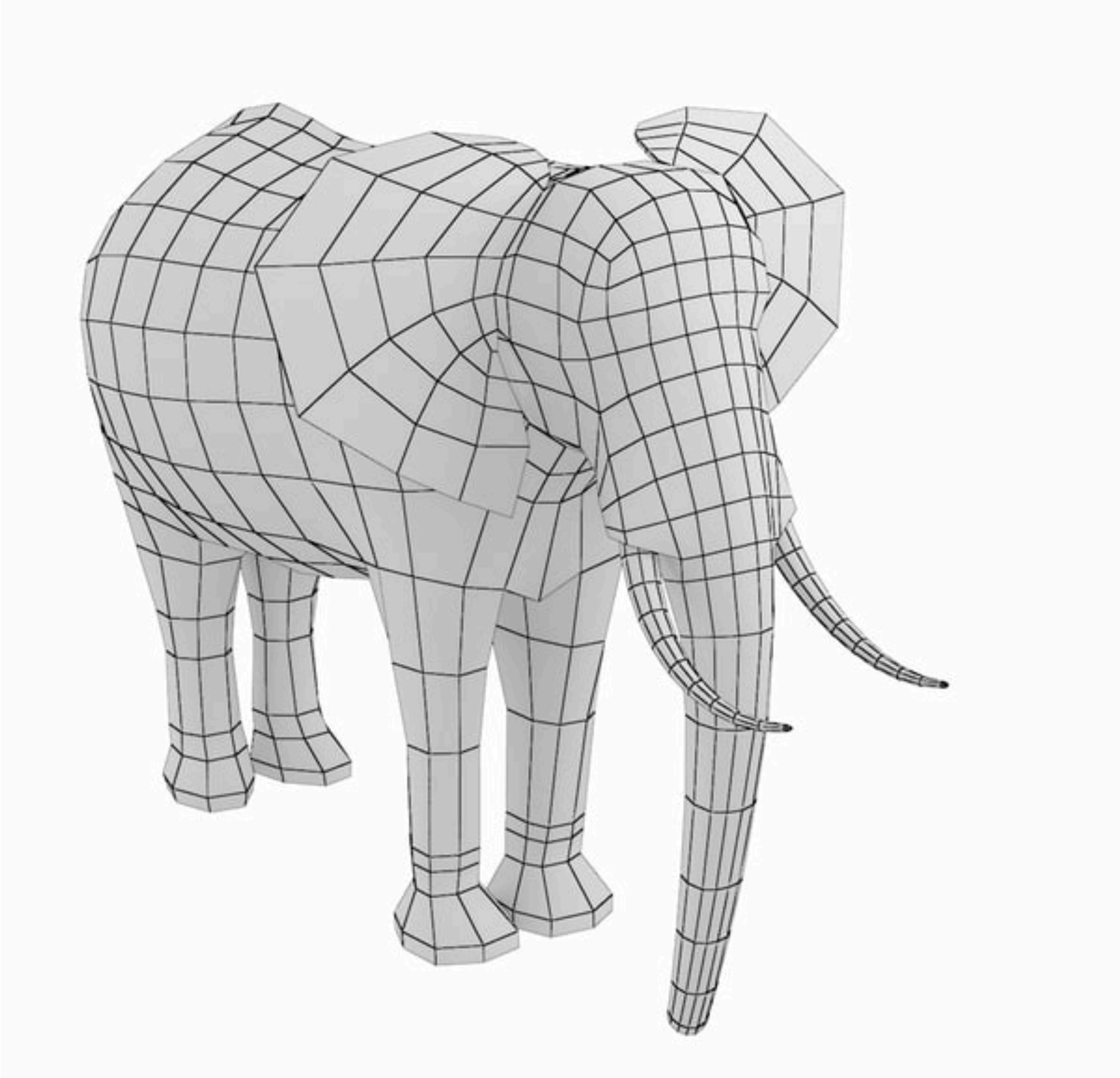


Target Geometry

Gradient Based Optimization



Geometry Representations



Mesh Representation

Voxel Representation

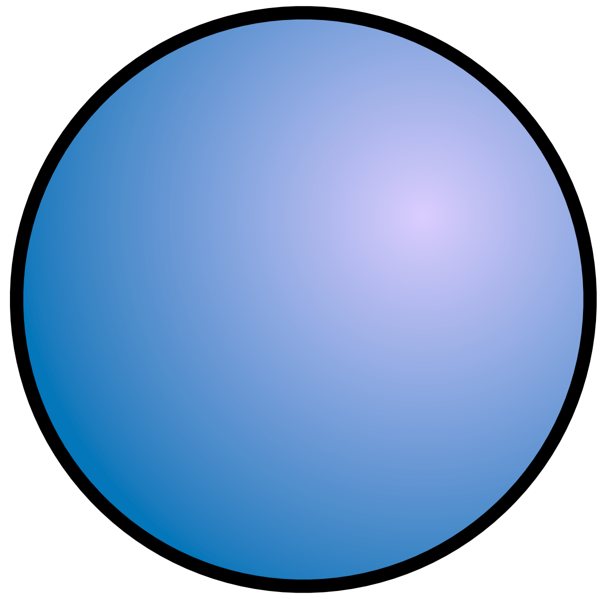
Small memory footprint

Hard to optimize

Easy to optimize

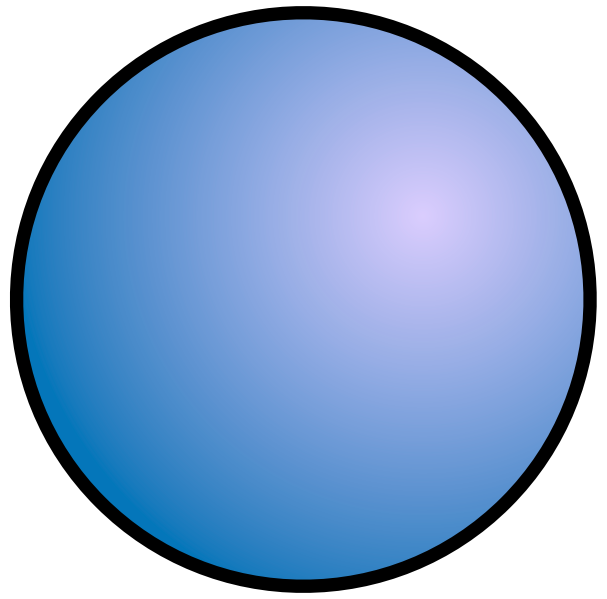
Large memory footprint

Implicit Functions

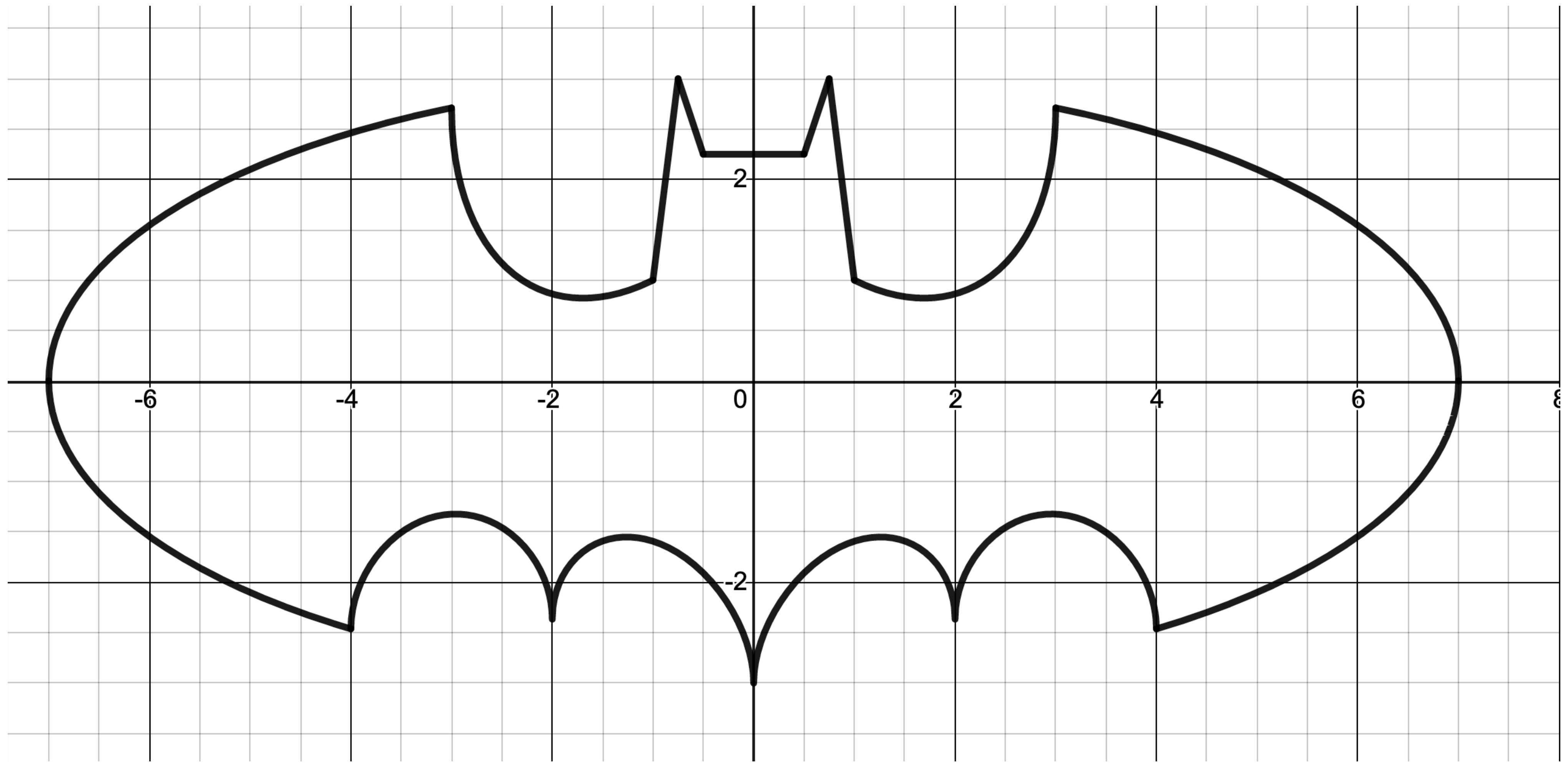


$x^2 + y^2 + z^2 = 1$

Implicit Functions



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



$$\left\{ |x| > 3{:}3\sqrt{-\left(\frac{x}{7}\right)^2 + 1} \right\}$$

$$\left\{ |x| > 4{:}-3\sqrt{-\left(\frac{x}{7}\right)^2 + 1} \right\}$$

$$\left|\frac{x}{2}\right| - \frac{3\sqrt{33} - 7}{112}x^2 + \sqrt{1 - \left(\mathrm{abs}\left(|x| - 2\right) - 1\right.}$$

$$\left\{ .75 < |x| < 1 : 9 - 8|x| \right\}$$

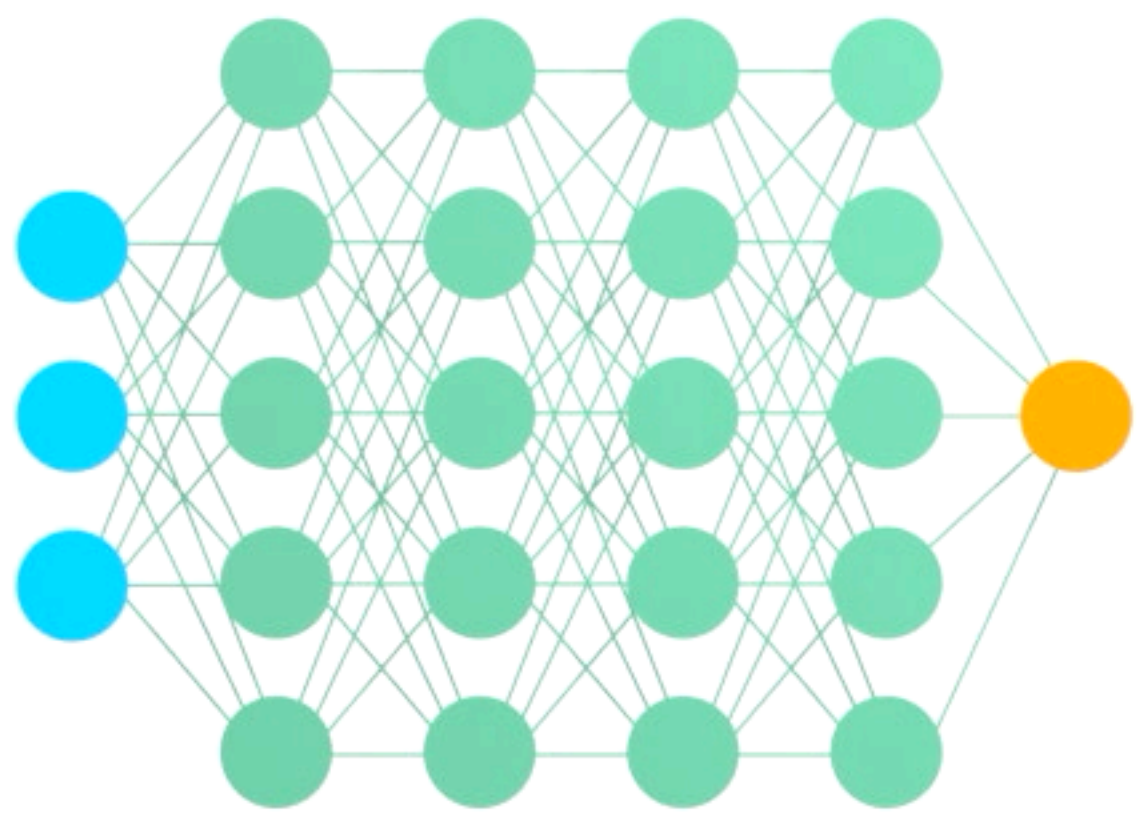
$$\left\{ .5 < |x| < .75{:}\ 3|x| + .75 \right\}$$

$$\left\{ |x| < .5 : 2.25 \right\}$$

$$\left\{ |x| > 1{:}1.5 - .5|x| - \frac{6\sqrt{10}}{14}\left(\sqrt{3 - x^2 + 2|x|}\right.\right.$$

Coordinate Based Neural Network

Input
Coordinate



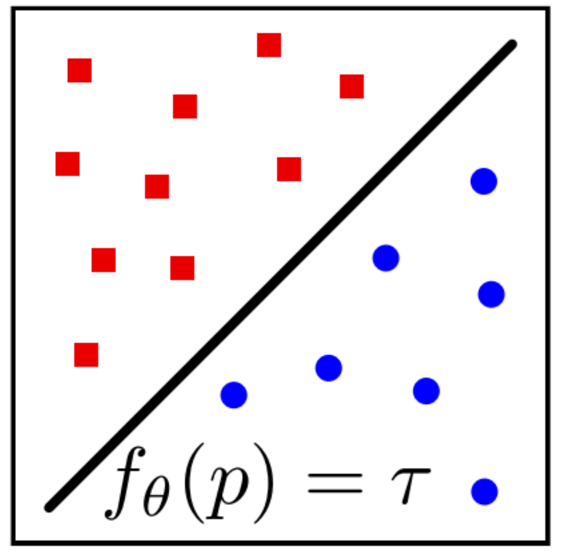
Value at
Coordinate

Multi Layer Perceptron

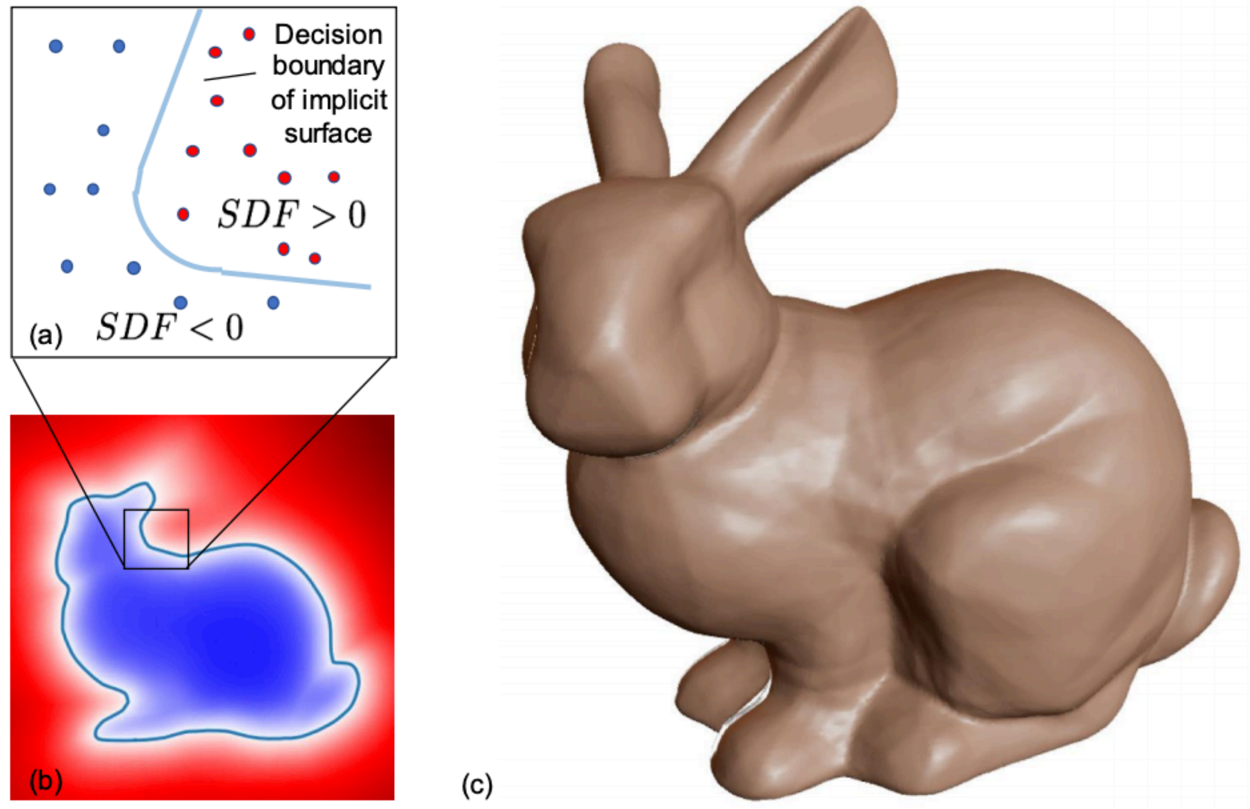
MLP

Neural networks as a continuous shape representation

Occupancy Networks
(Mescheder et al. 2019)
$(x, y, z) \rightarrow$ occupancy

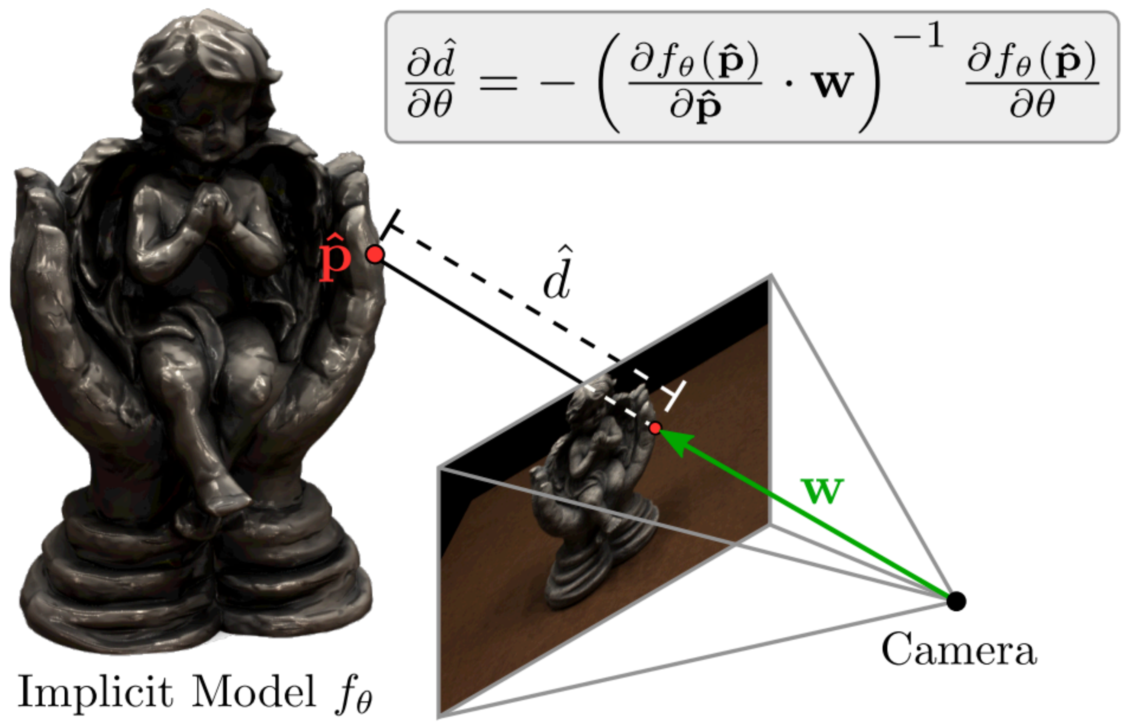


DeepSDF
(Park et al. 2019)
$(x, y, z) \rightarrow$ distance



Scene Representation Networks
(Sitzmann et al. 2019)
$(x, y, z) \rightarrow$ latent vec. (color, dist.)

Differentiable Volumetric Rendering
(Niemeyer et al. 2020)
$(x, y, z) \rightarrow$ color, occ.



NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis

Matthew Tancik*[1] Pratul P. Srinivasan*[1,3] Ben Mildenhall*[1,3]

Jonathan T. Barron[3] Ravi Ramamoorthi[2] Ren Ng[1]

[1] [2] UC San Diego [3] Google

* Denotes Equal Contribution

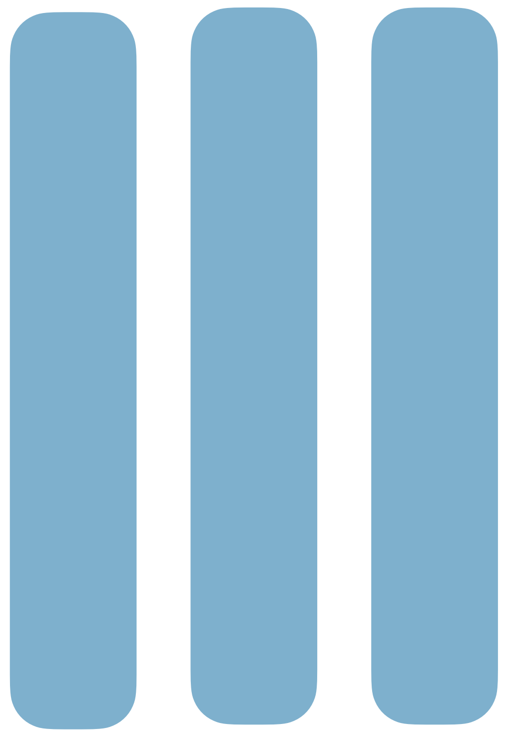
Radiance

Light Traveling Along A Ray

1. Radiance is the fundamental field quantity that describes the distribution of light in an environment
    - Radiance is the quantity associated with a ray
    - Rendering is all about computing radiance
2. Radiance is invariant along a ray in a vacuum

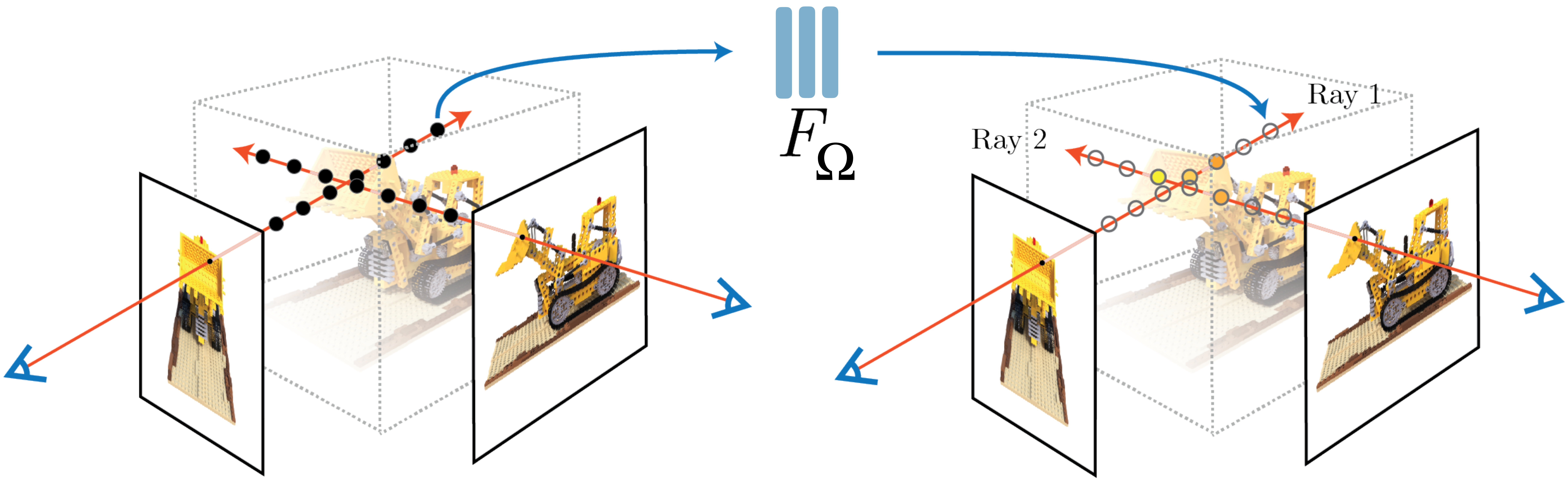
Representing a scene as a continuous 5D function

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



MLP
9 layers,
256 channels

Generate views with traditional volume rendering



Generate views with traditional volume rendering

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

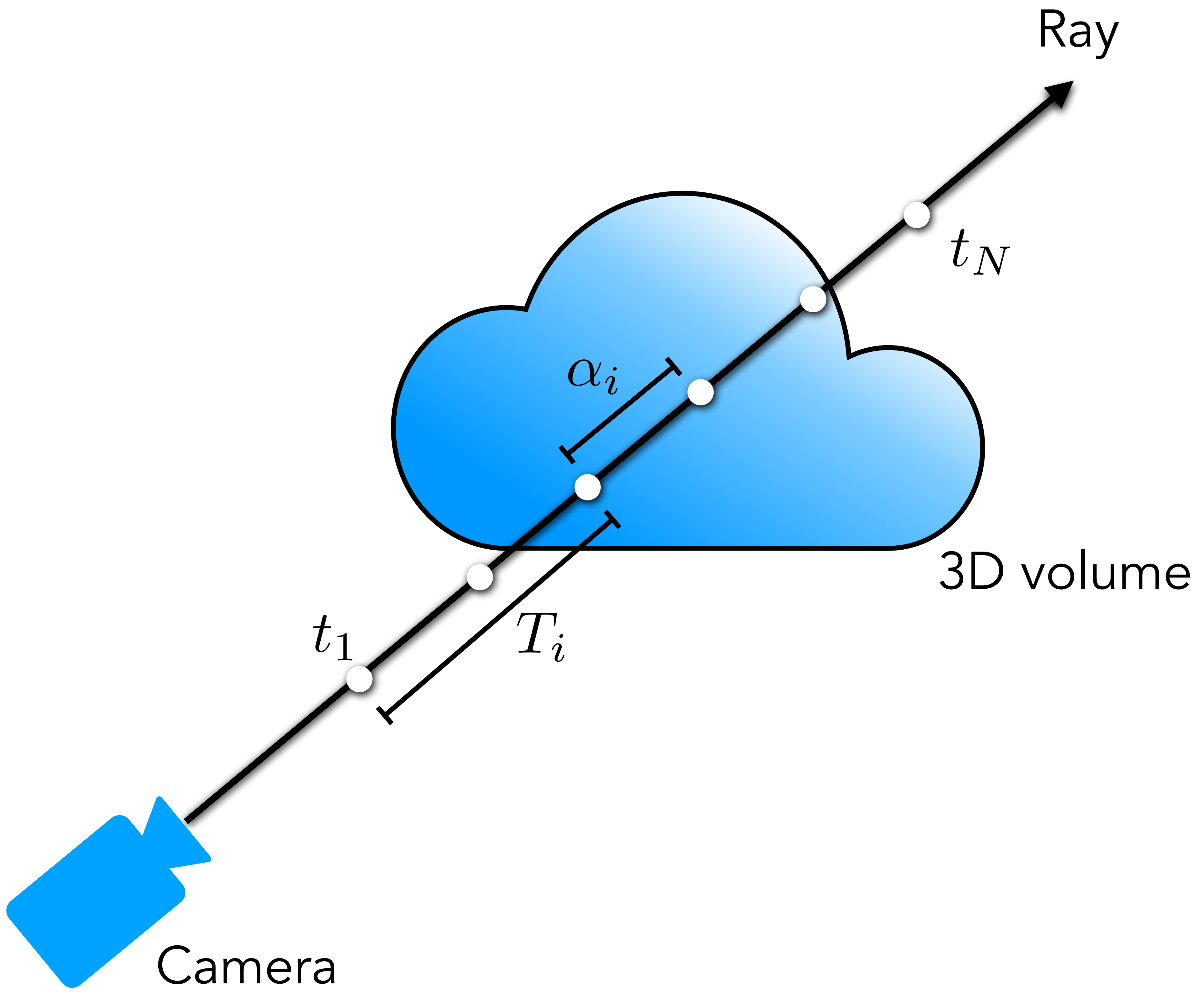
$$C \approx \sum_{i=1}^{N} \underbrace{T_i \alpha_i}_{\text{weights}} \underbrace{c_i}_{\text{colors}}$$

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*:

$$\alpha_i = 1 - e^{-\sigma_i \delta t_i}$$



Effective resolution is tied to distance between samples

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

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

weights

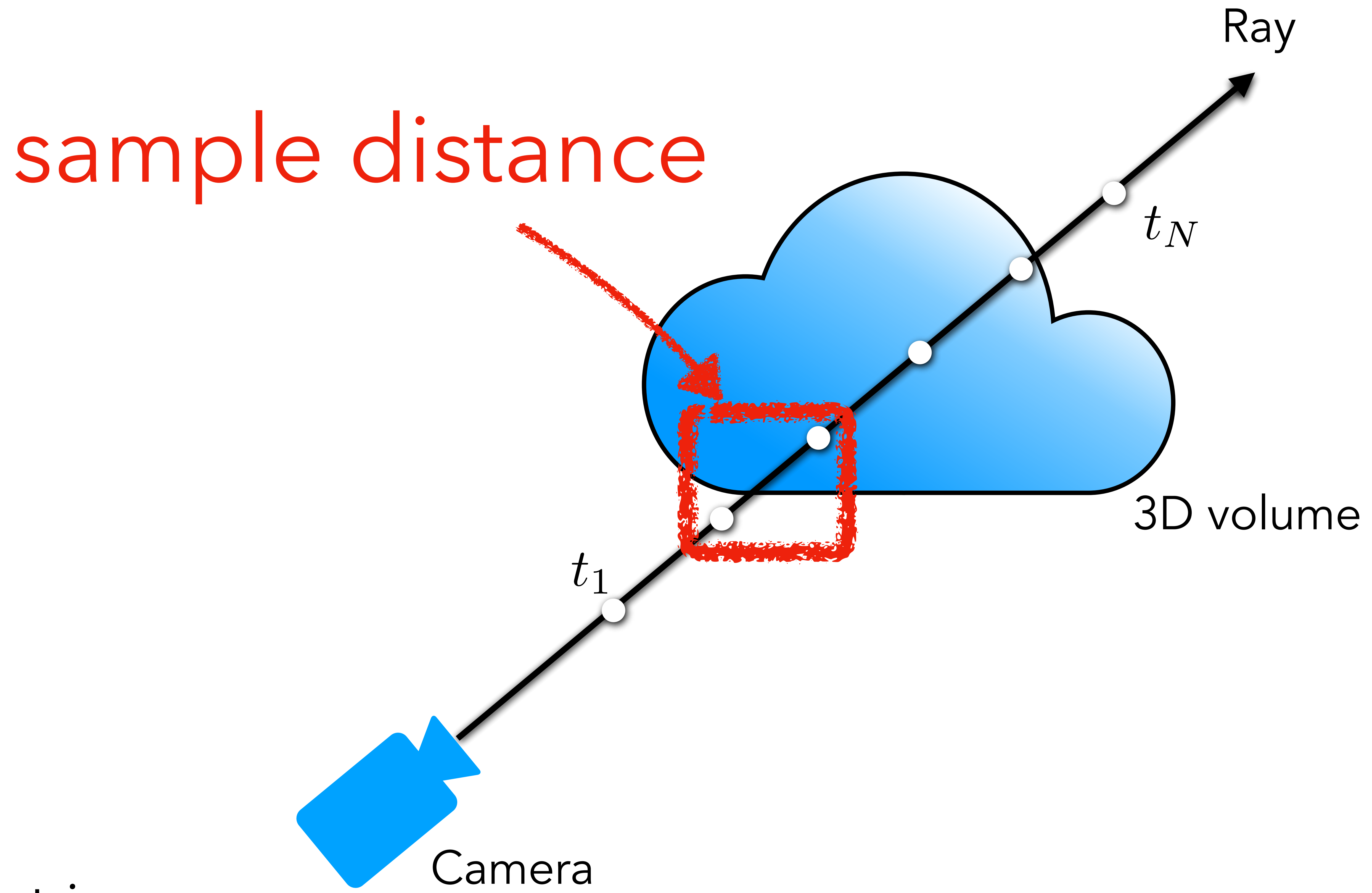
colors

How much light is blocked earlier along ray:

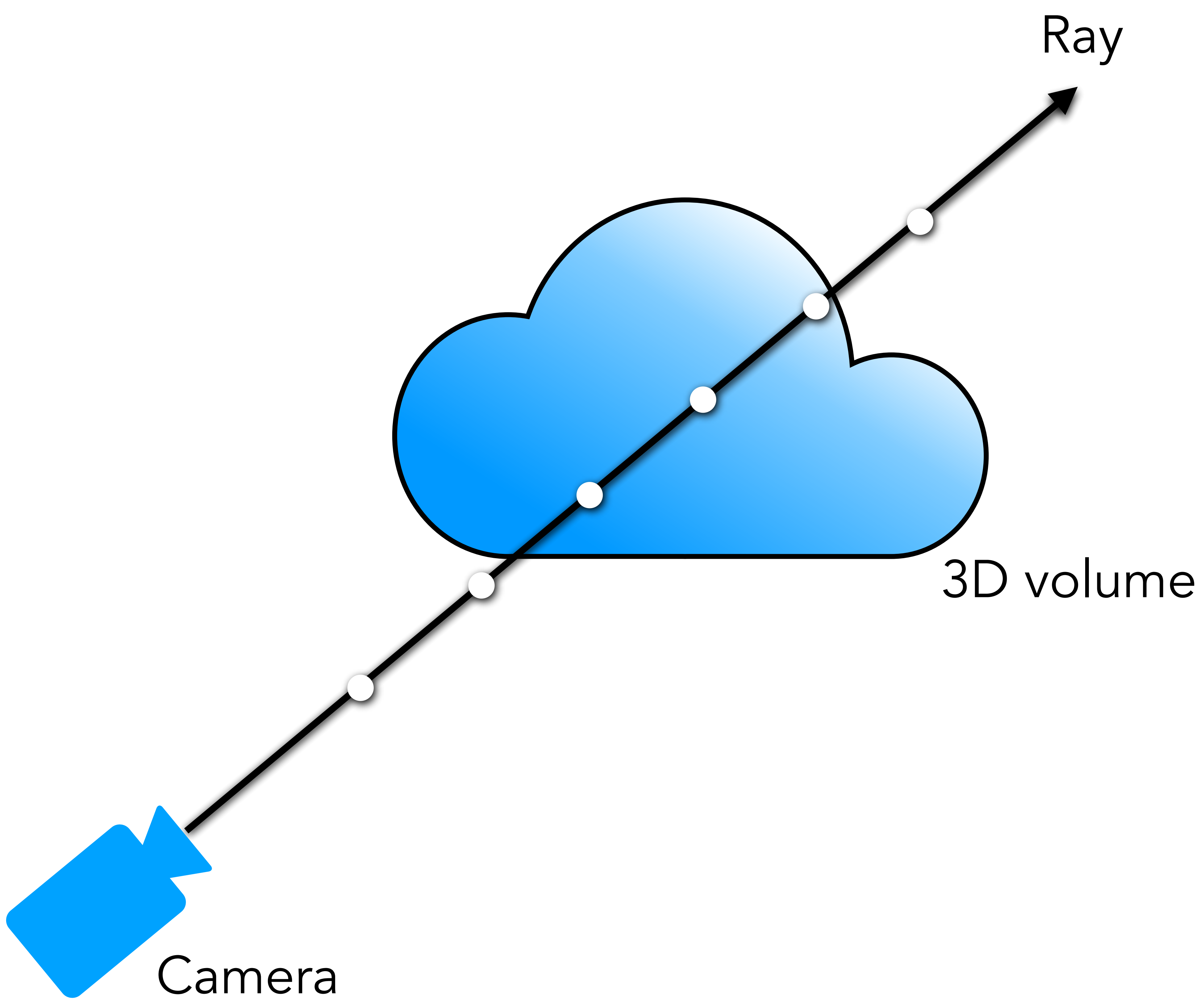
$$T_i = \prod_{j=1}^{i-1} (1 - \alpha_j)$$

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

$$\alpha_i = 1 - e^{-\sigma_i \delta t_i}$$



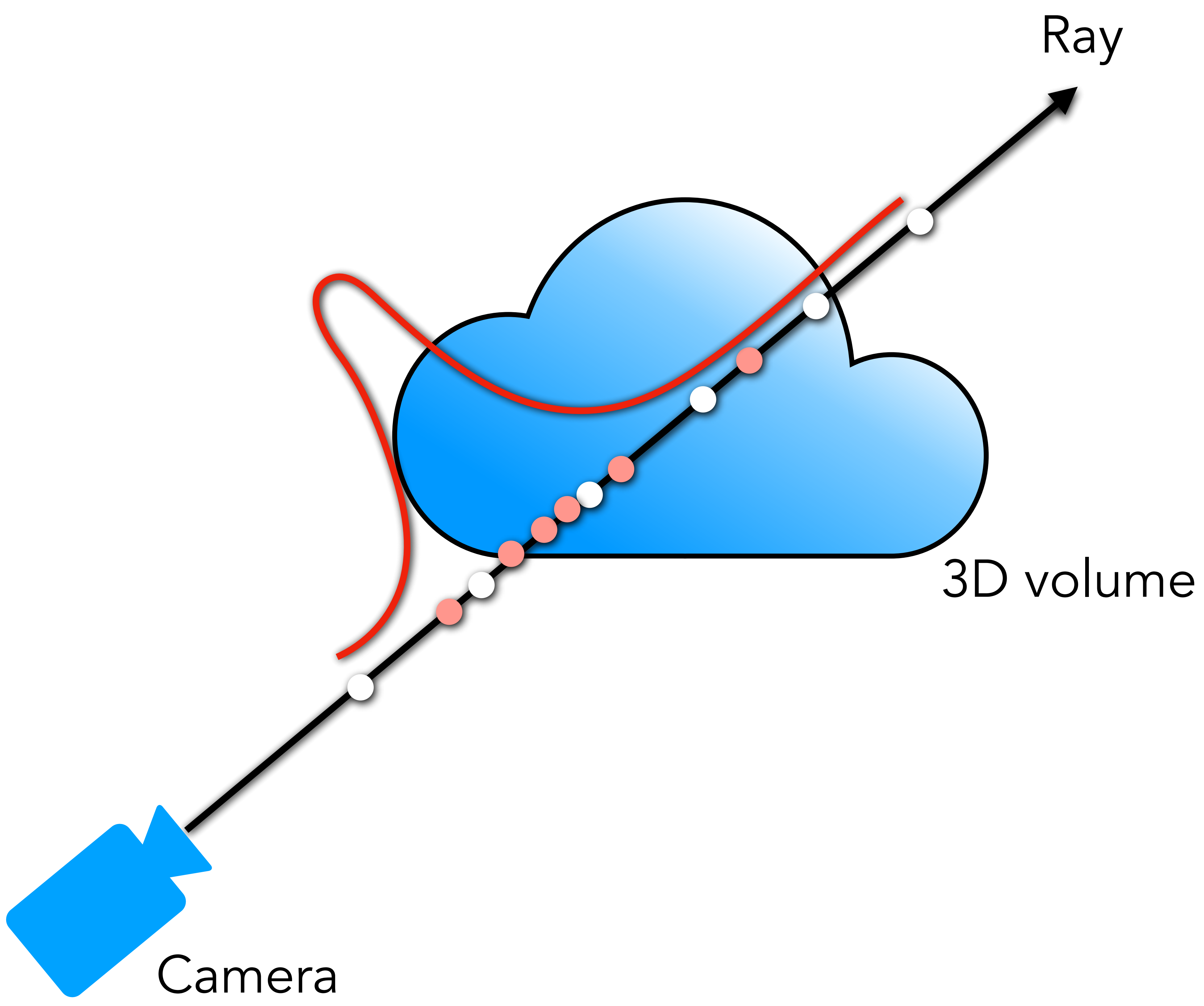
Can we allocate samples more efficiently? Two pass rendering



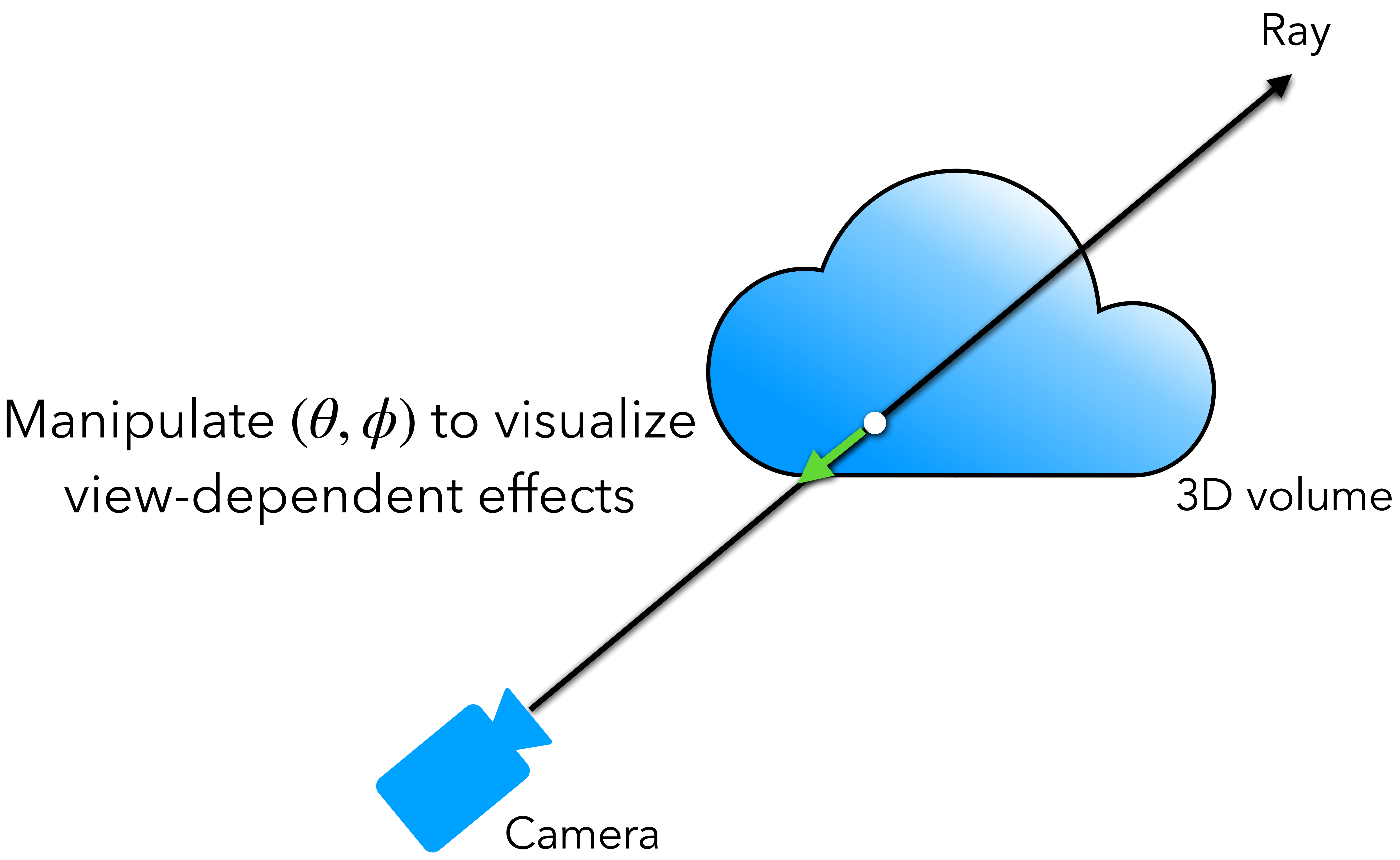
Two pass rendering: fine

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

treat weights as probability
distribution for new samples



Viewing directions as input



Volume rendering is differentiable

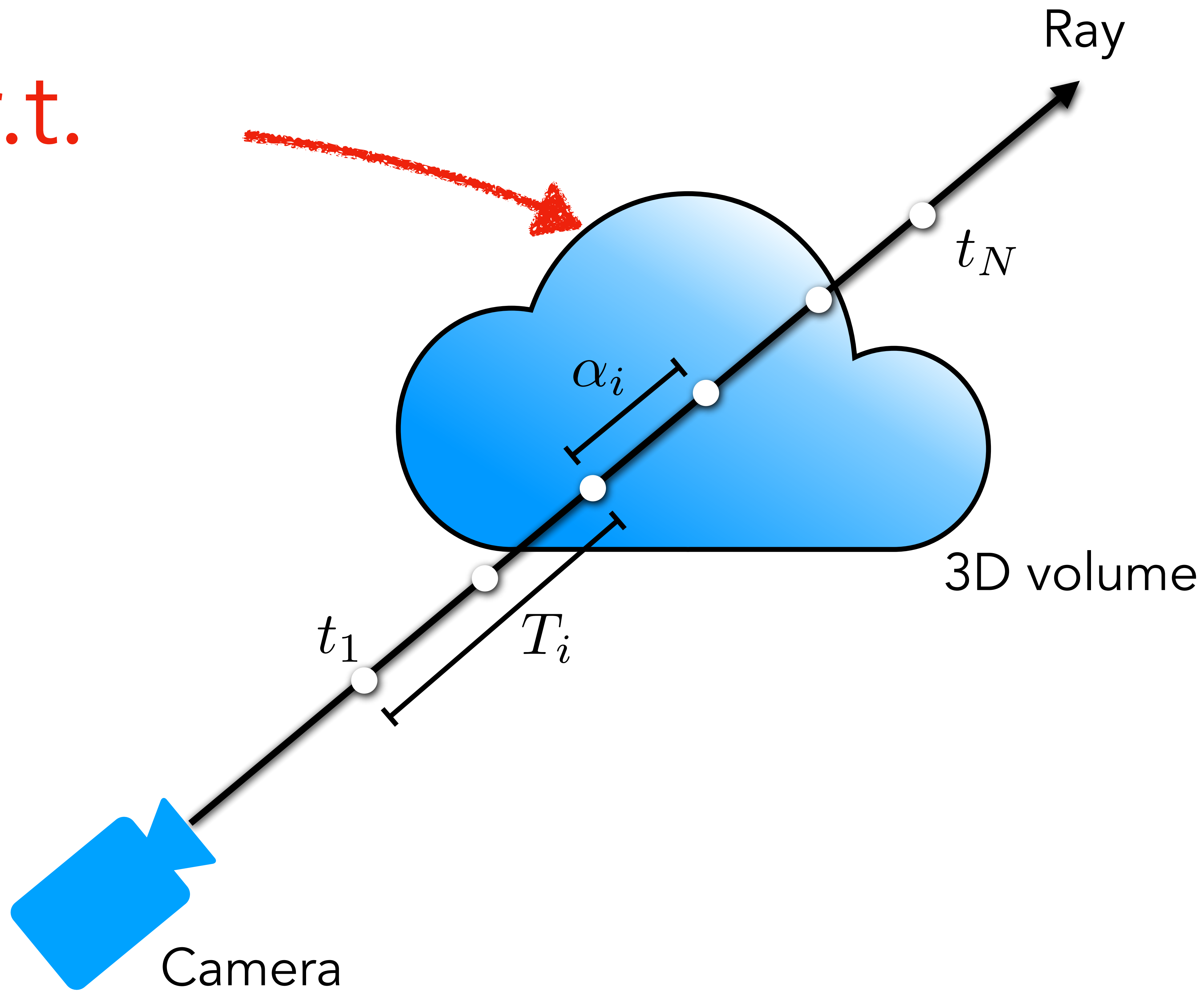
Rendering model for ray r(t) = o + td:

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

differentiable w.r.t.

colors

weights



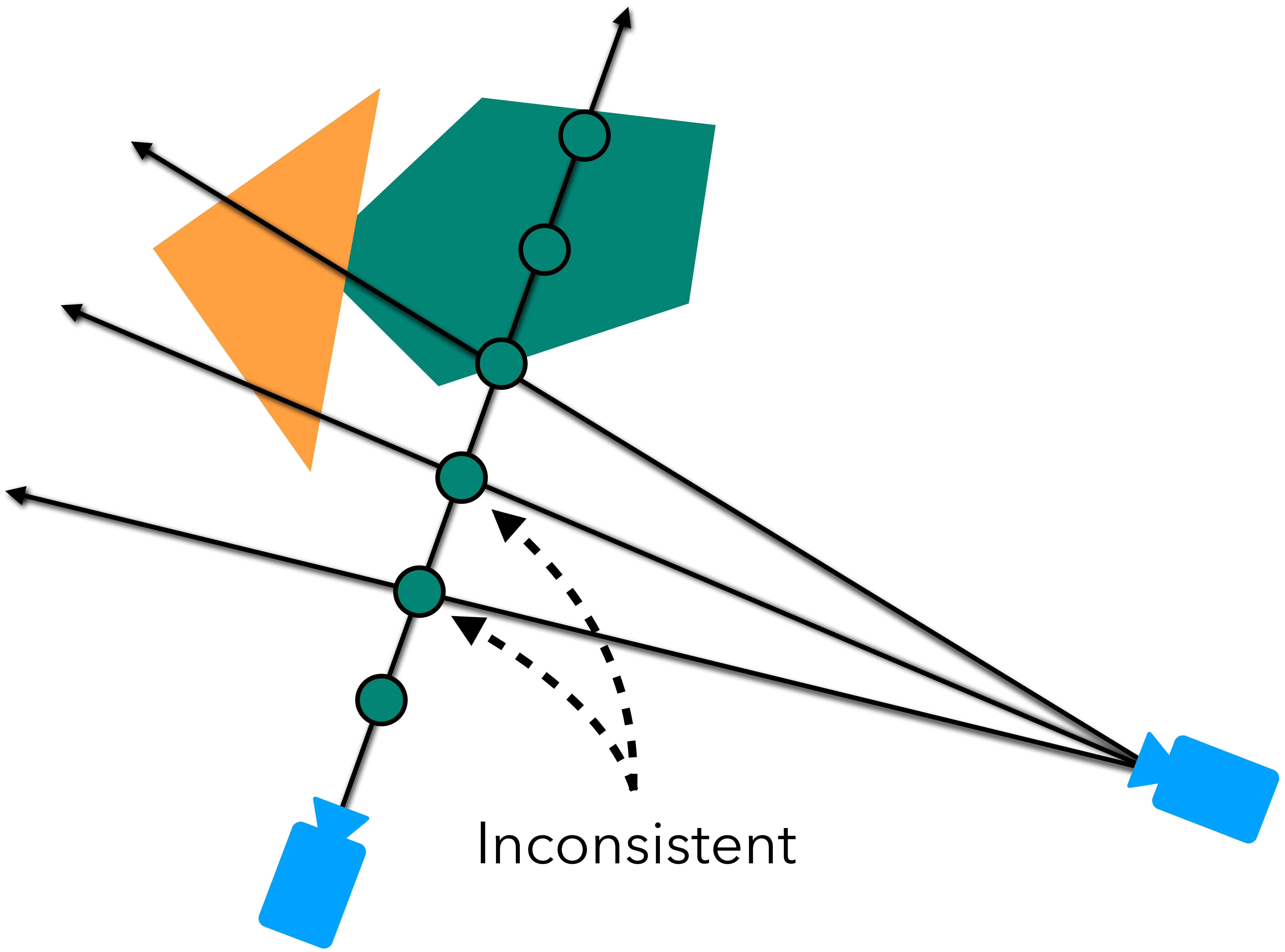
How much light is blocked earlier along ray:

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

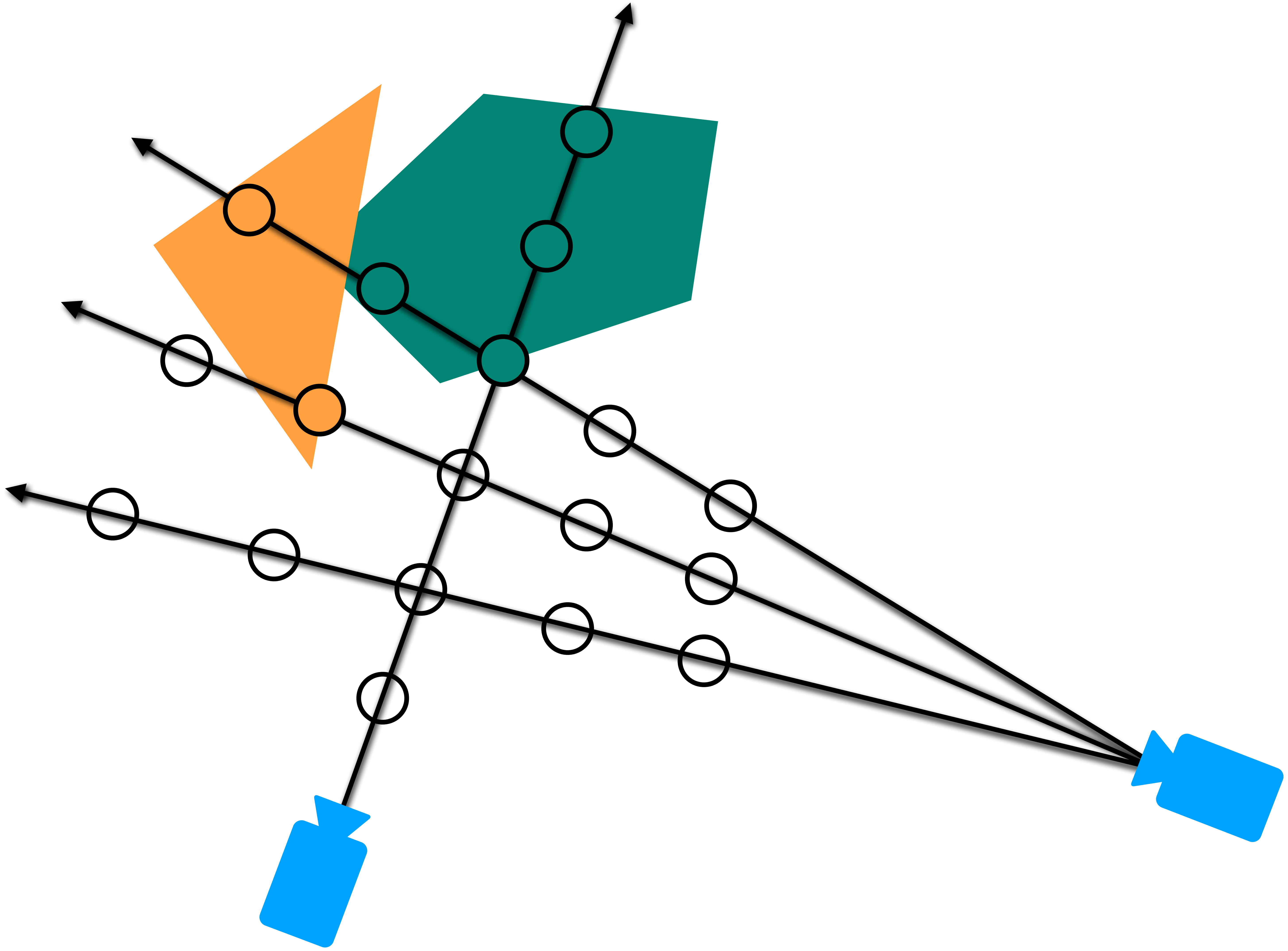
How much light is contributed by ray segment *i*:

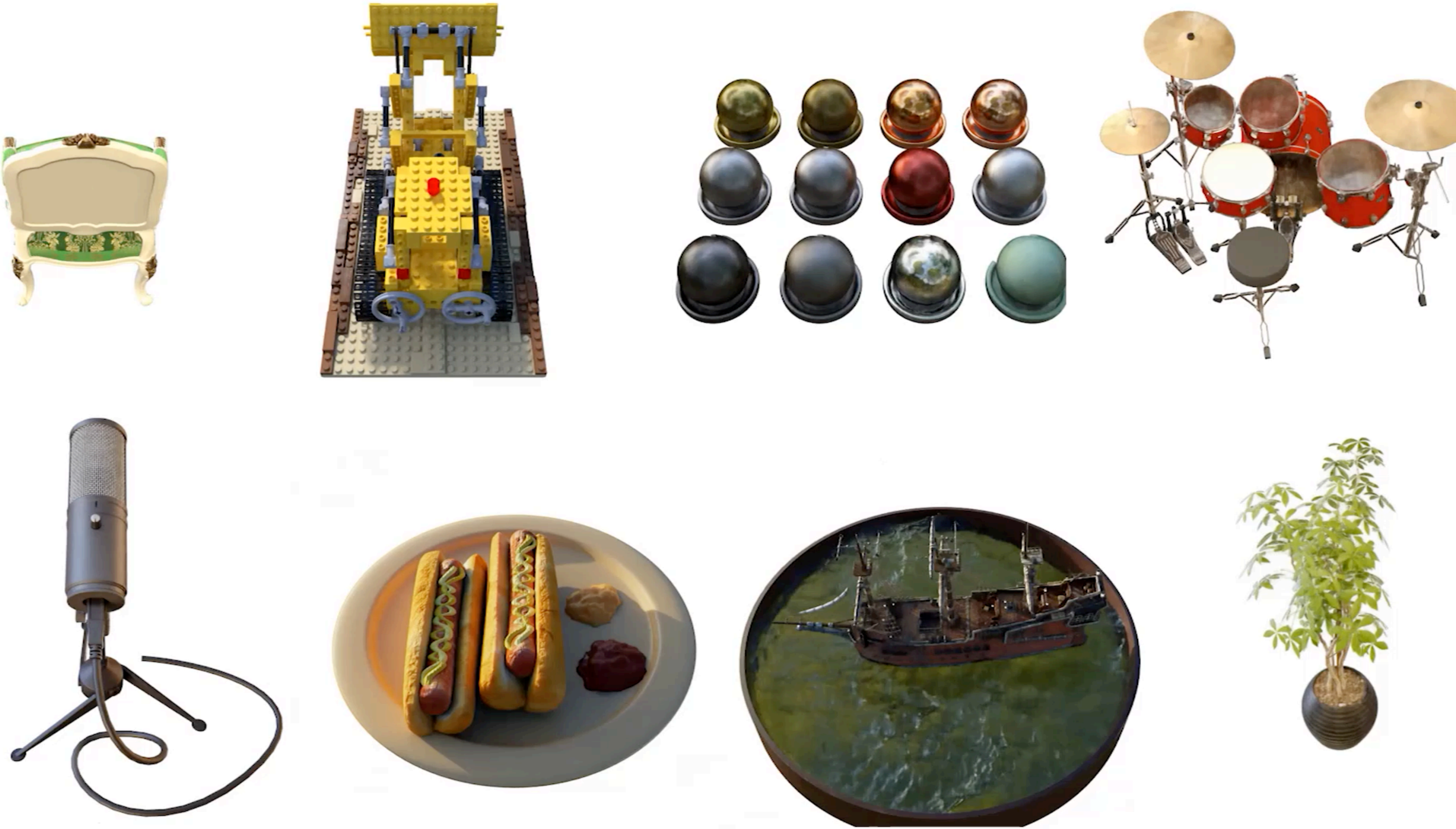
$$\alpha_i = 1 - e^{-\sigma_i \delta t_i}$$

Multiview Consistency as Supervision



Multiview Consistency as Supervision

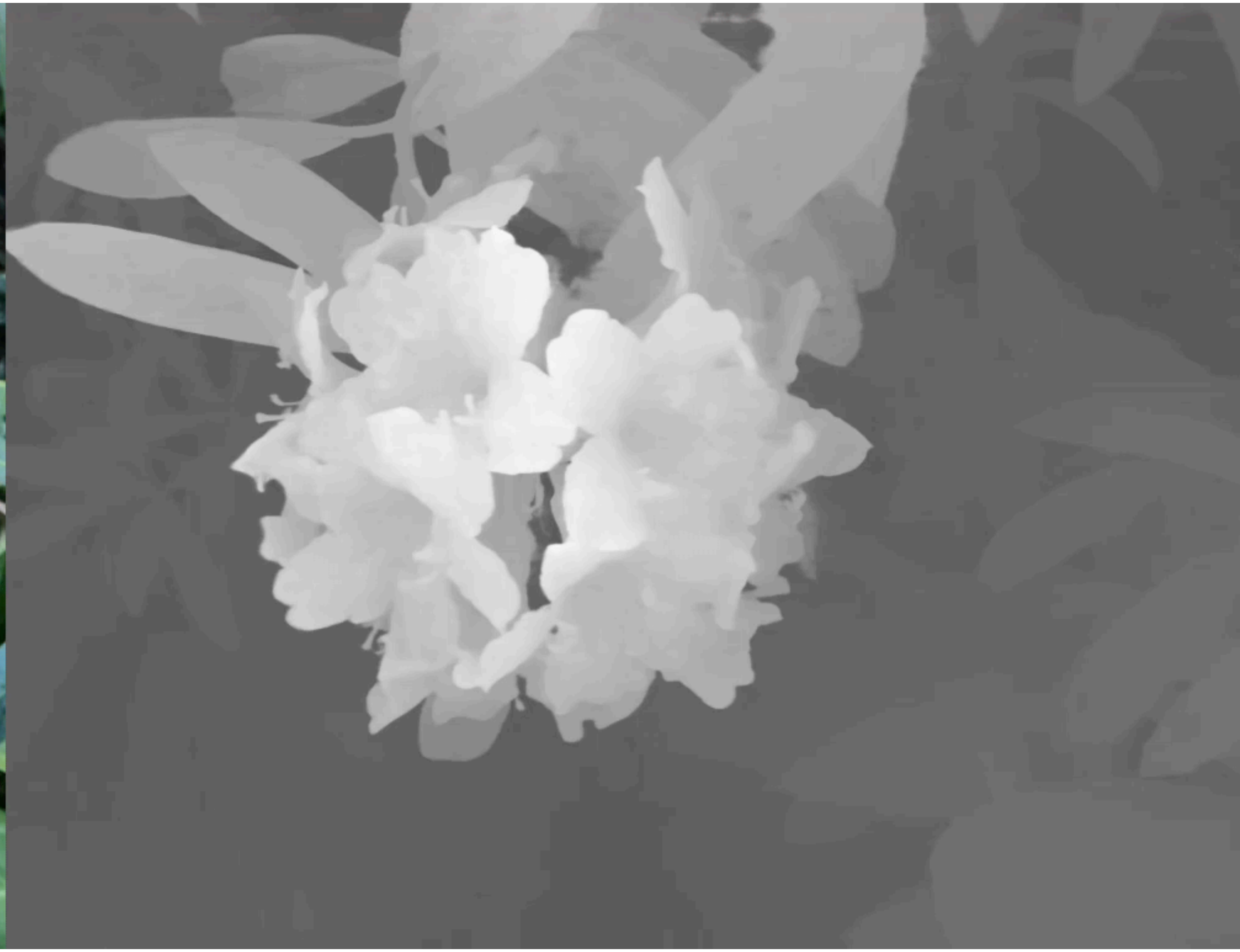




NeRF encodes convincing view-dependent effects using directional dependence

NeRF encodes convincing view-dependent effects using directional dependence

NeRF encodes detailed scene geometry with occlusion effects

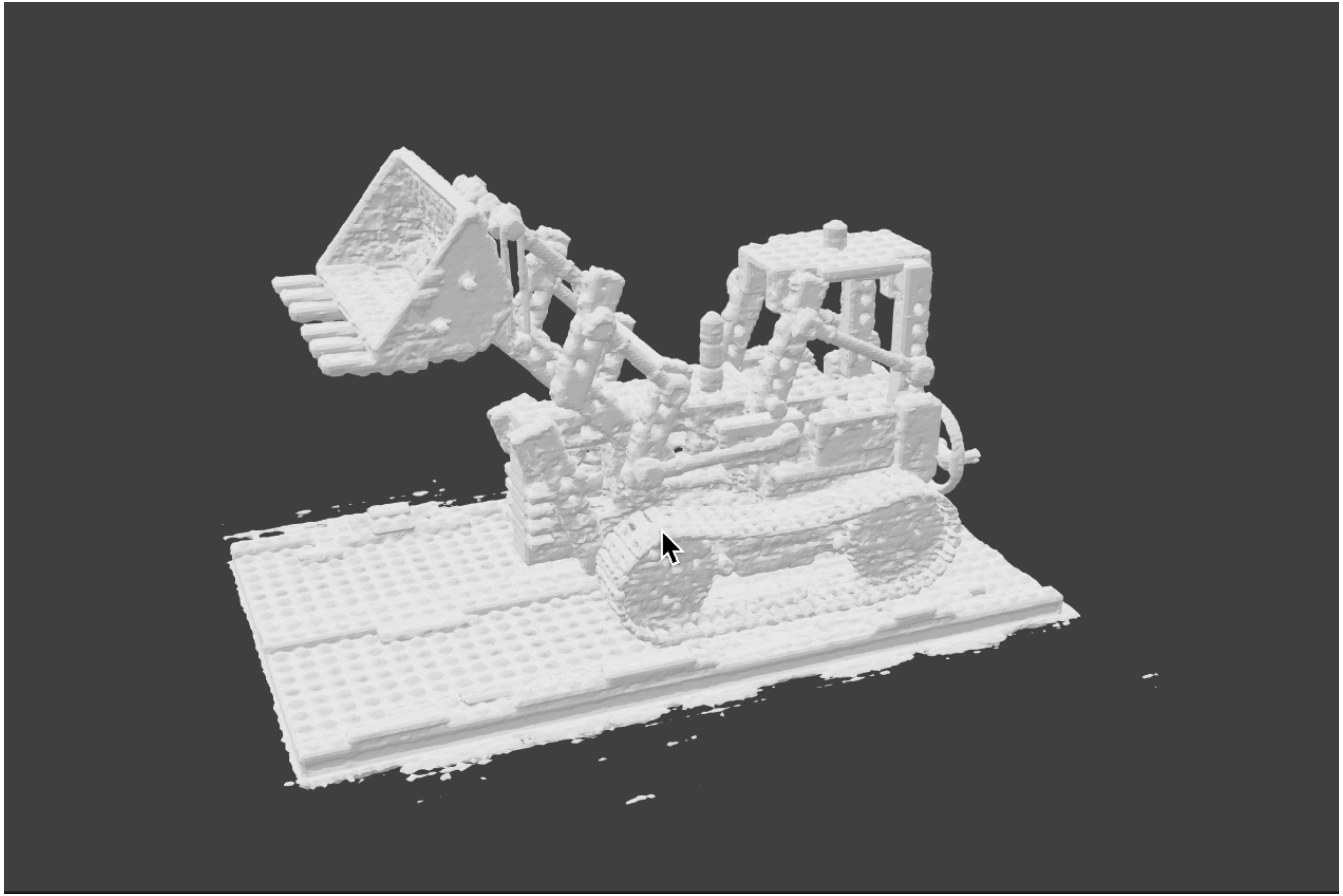


NeRF encodes detailed scene geometry with occlusion effects

NeRF encodes detailed scene geometry with occlusion effects

NeRF encodes detailed scene geometry



Naive implementation produces blurry results

NeRF (Naive)

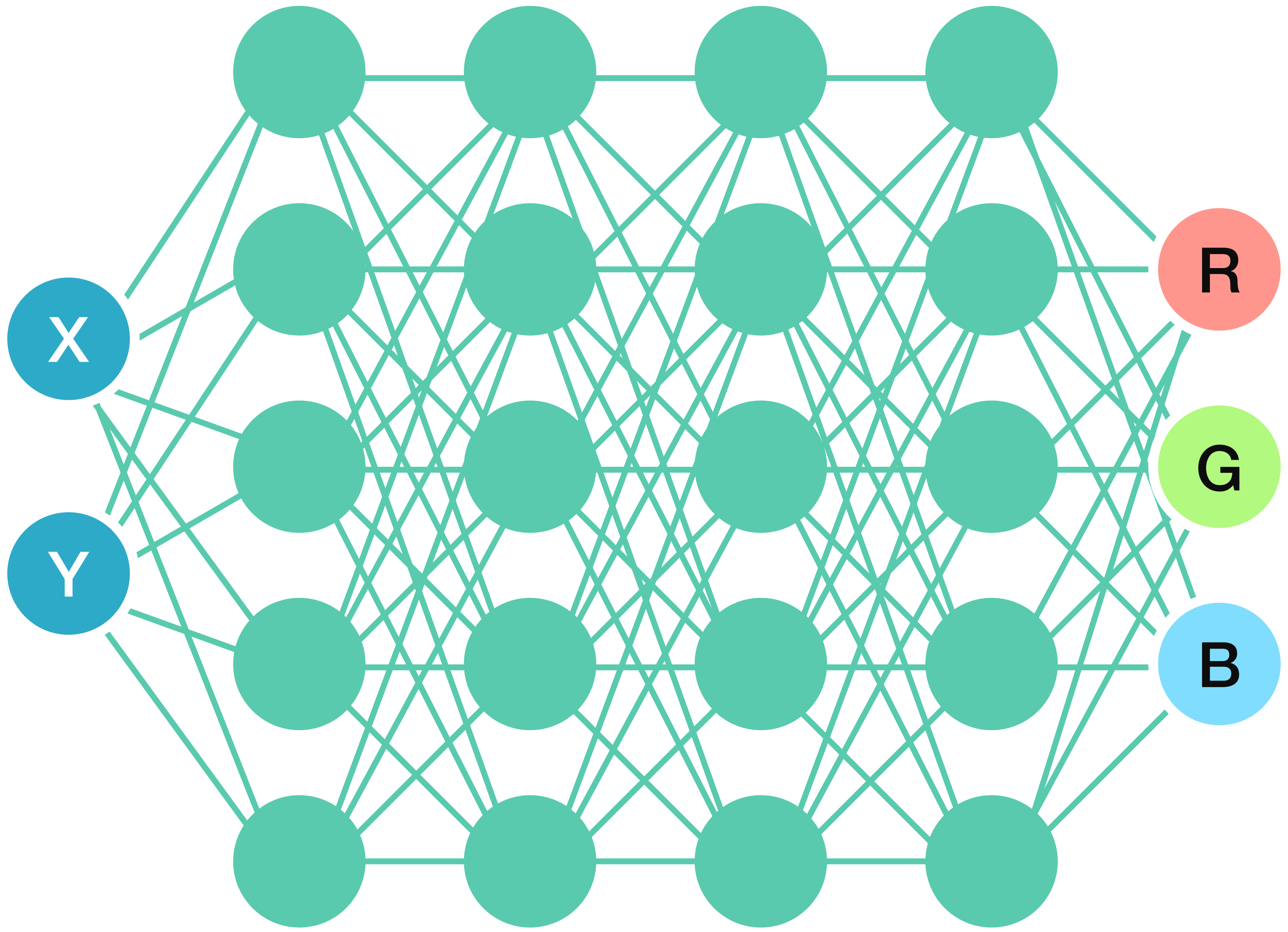
NeRF (with positional encoding)

Challenge:

How to get MLPs to represent higher frequency functions?

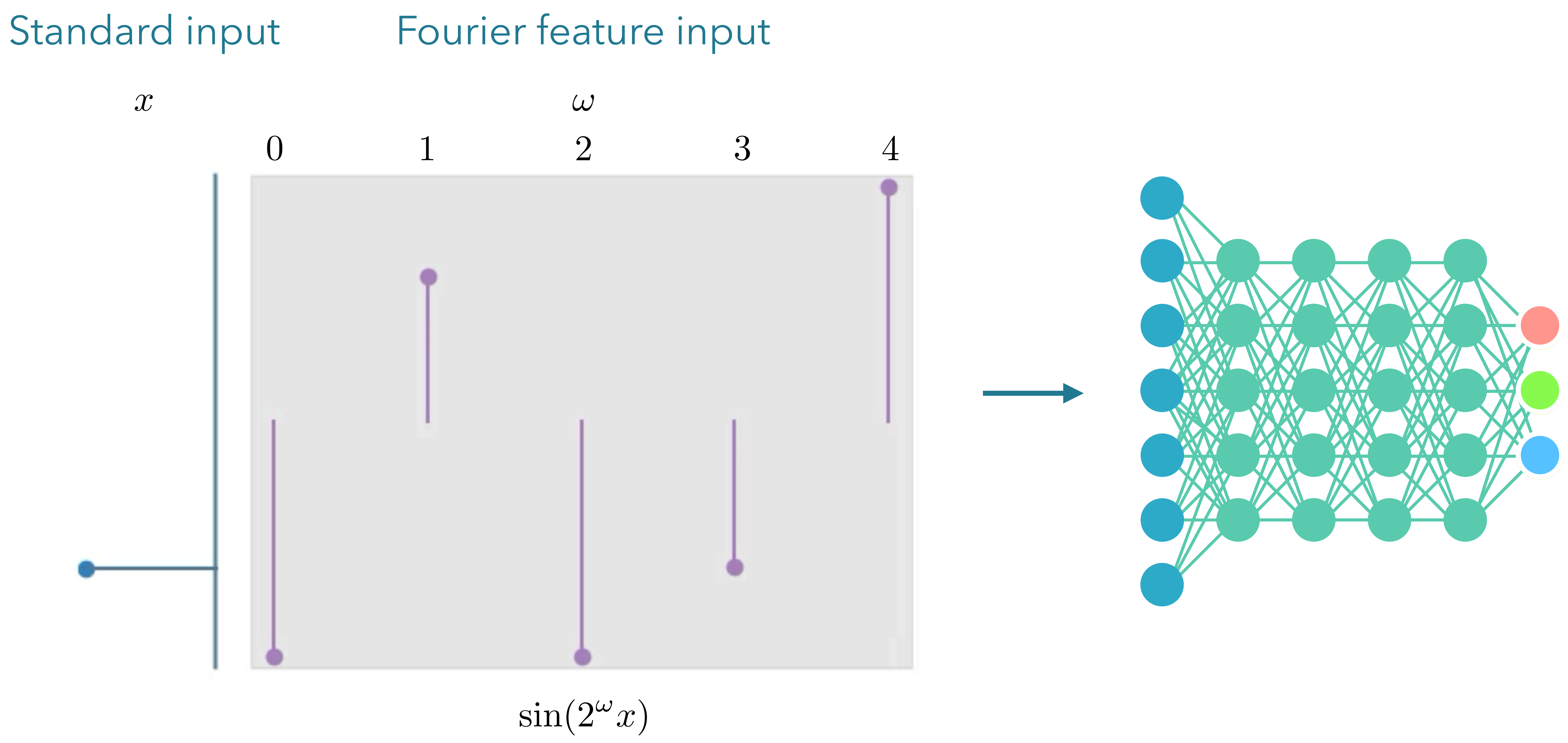
Rahaman et al. 2019, Basri et al. 2020

Image Representation

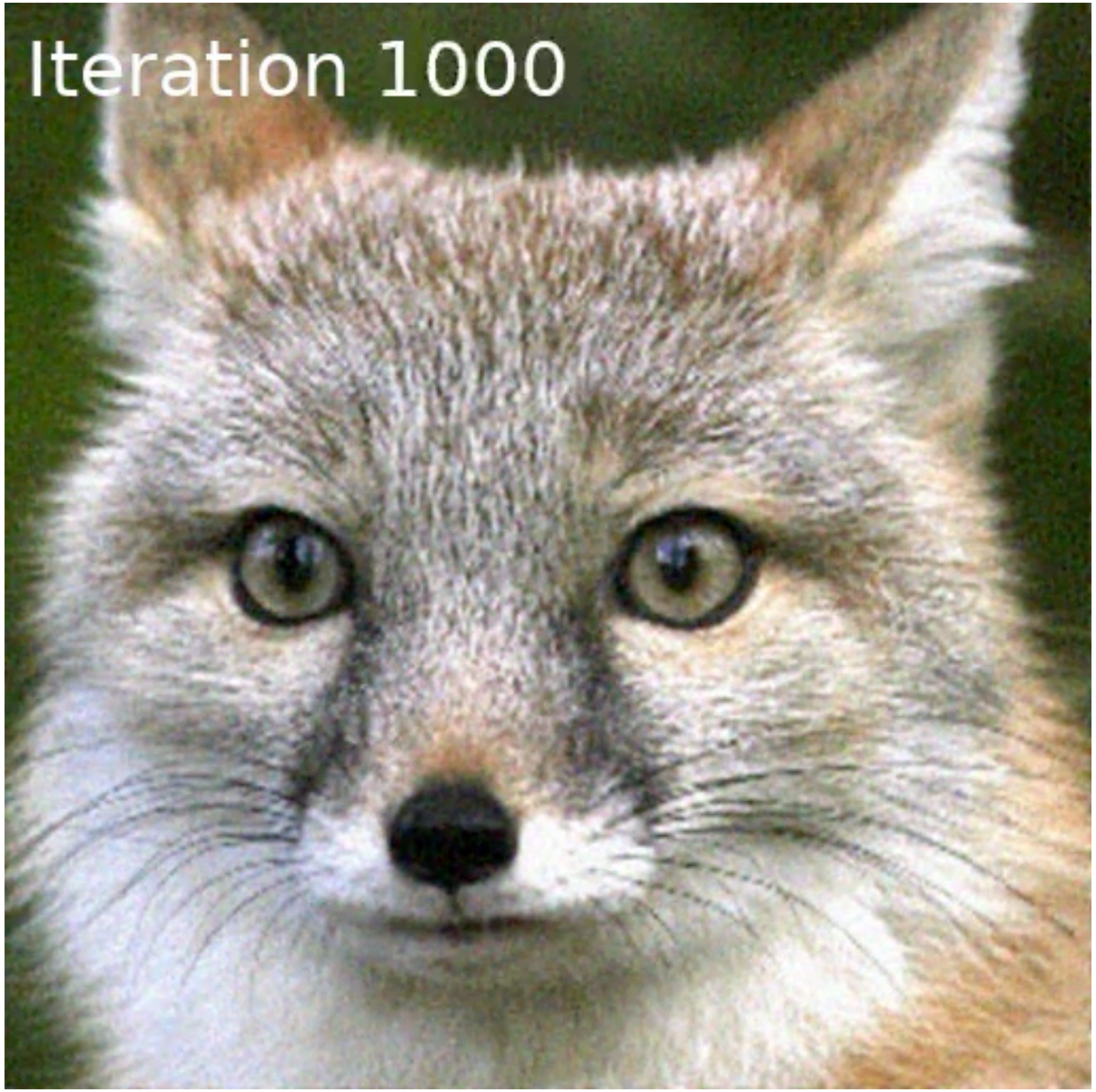


MLP output

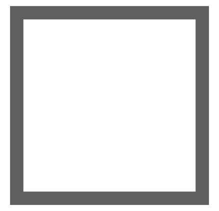
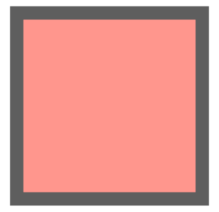
Supervision image

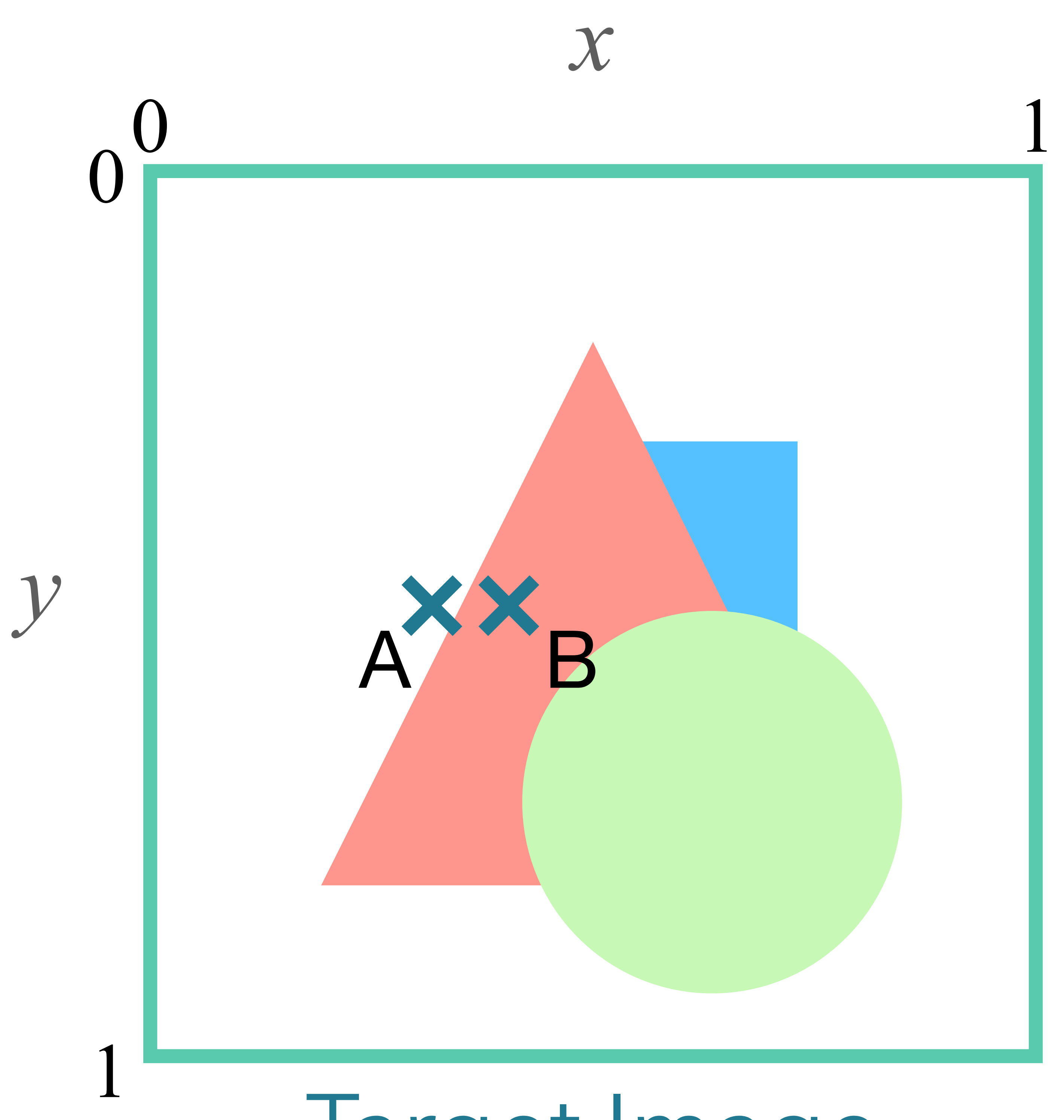

Standard input
Fourier feature input
x
ω
0
1
2
3
4
$\sin(2^{\omega}x)$

Standard MLP

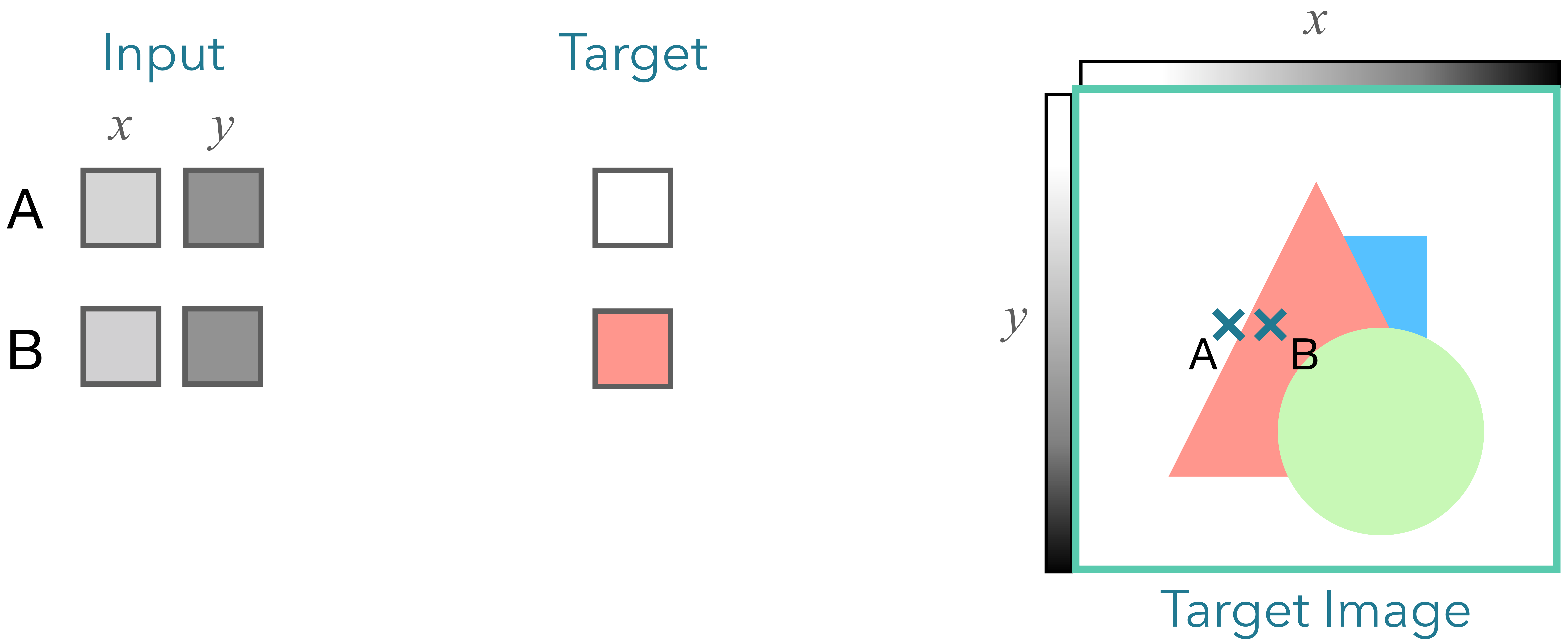


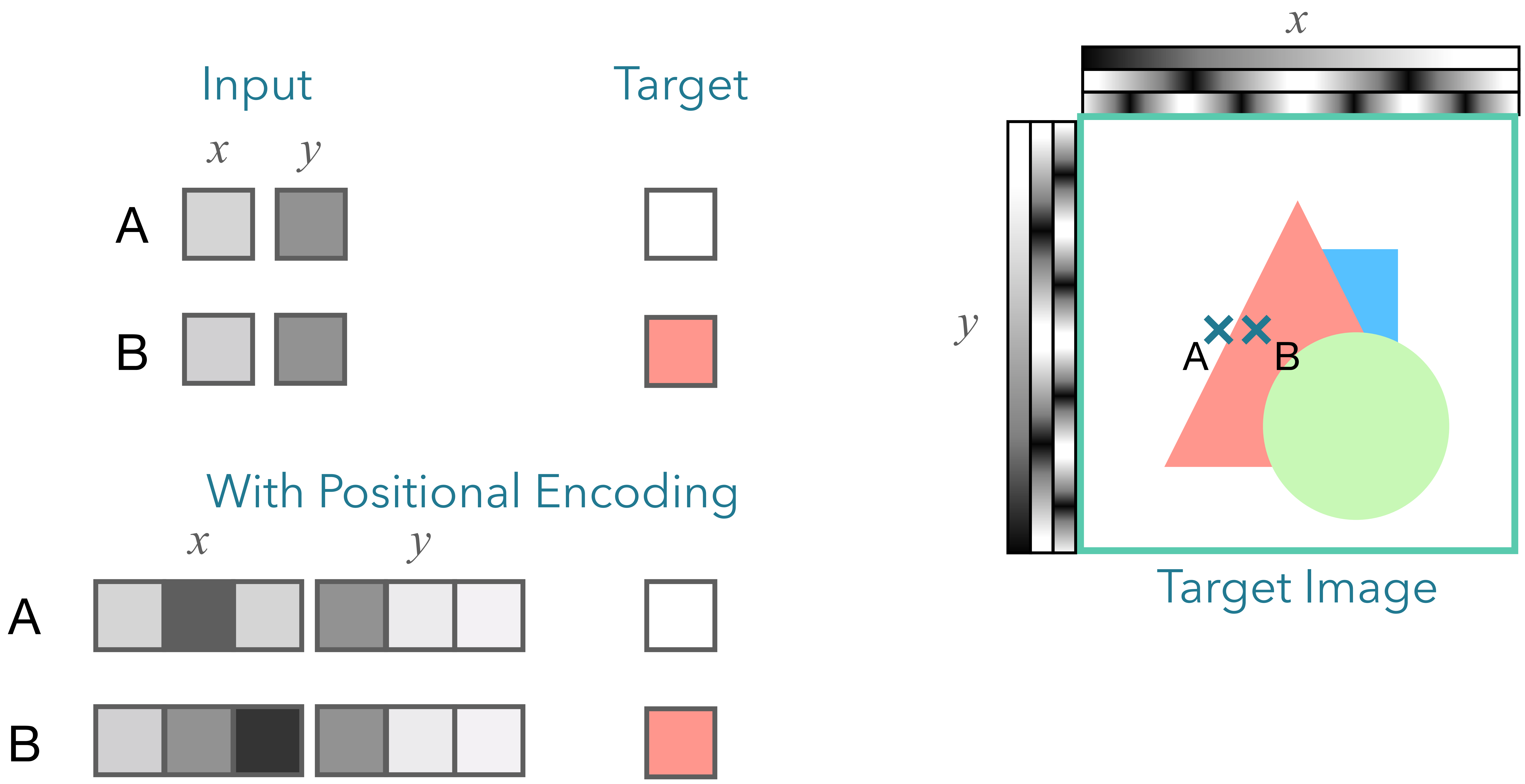
MLP with Fourier features

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



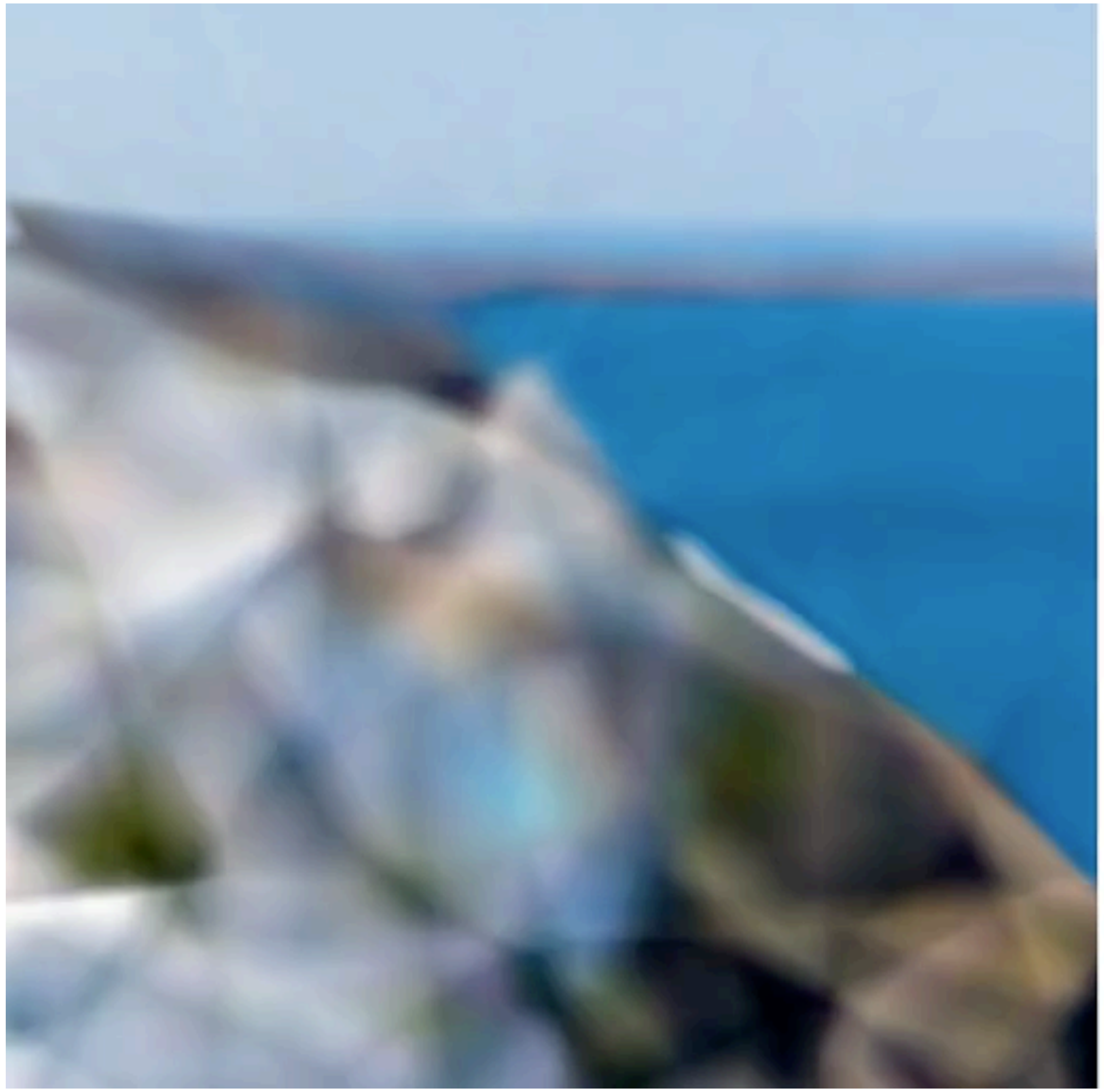
Target Image


Input
x
y
A
B
Target
x
y
A
B
Target Image

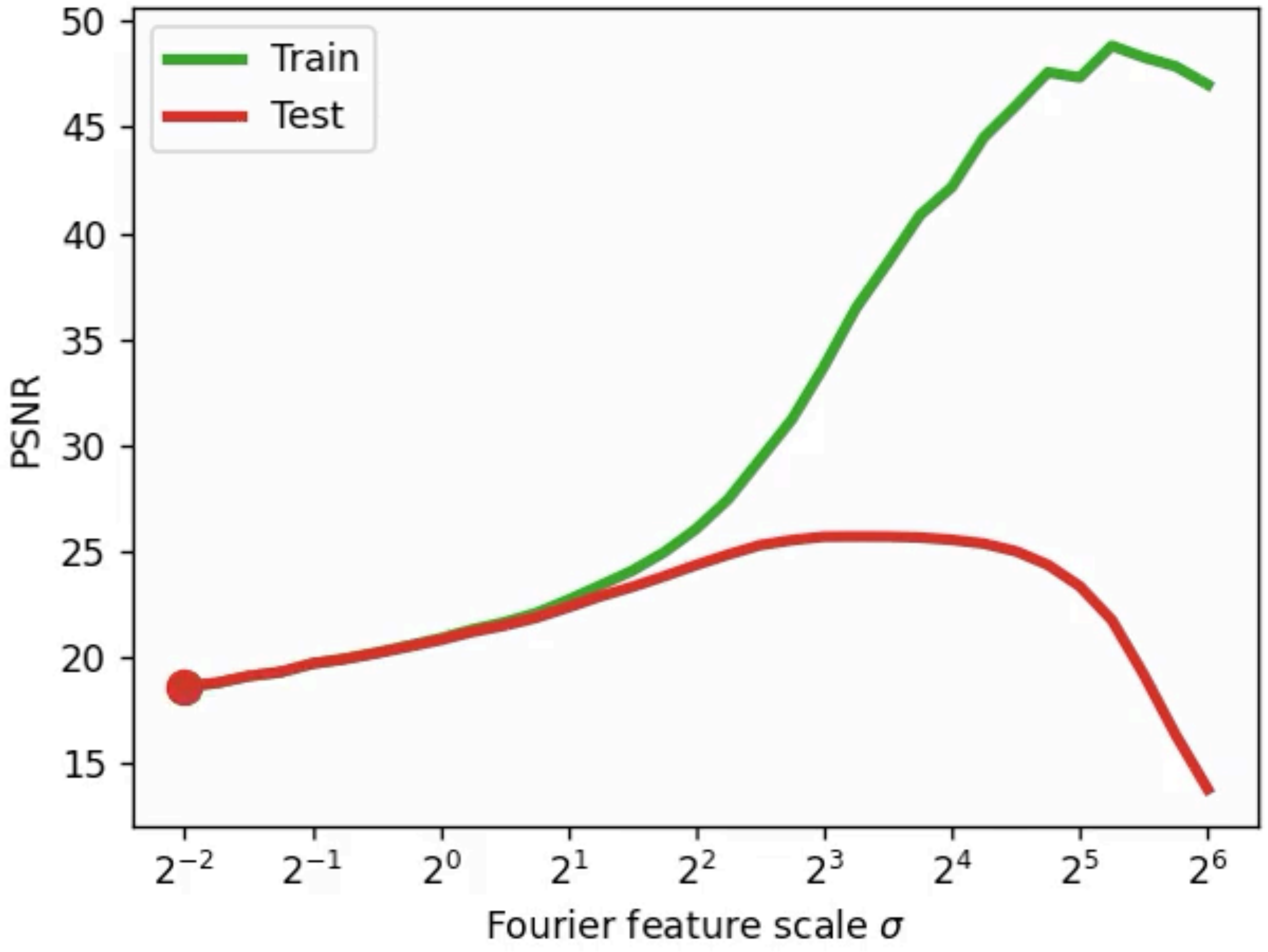


Input
Target
x
y
A
B
With Positional Encoding
x
y
A
B
x
y
A
B
Target Image

Performance depends on max encoding frequency



Network output



Performance vs. scale value

Mapping Code

Original

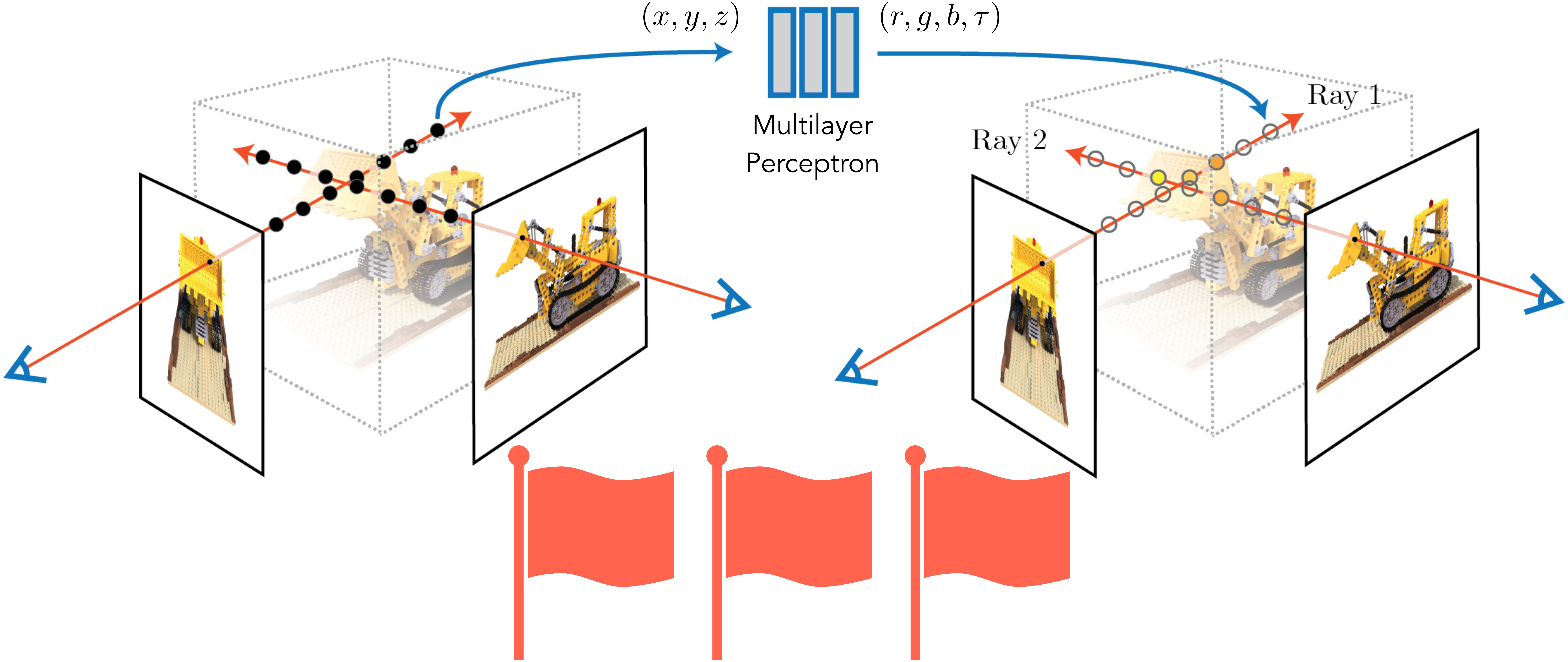
```
x = input_coordinate
x = nn.Dense(x, features=256)
              ⋮
```

With Positional Encoding

```
x = input_coordinate
B = SCALE * np.random.normal(shape=(input_dims, NUM_FEATURES))
rff_input = np.concatenate([np.sin((2*np.pi*x) @ B), np.cos((2*np.pi*x) @ B)], axis=-1)
x = nn.Dense(rff_input, features=256)
              ⋮
```

NeRF Extensions

Point Samples of a Continuous Function



Solution: Prefiltering with a mipmap

Pyramidal Parametrics, Lance Williams, SIGGRAPH 1983 mipmaps reduce image aliasing

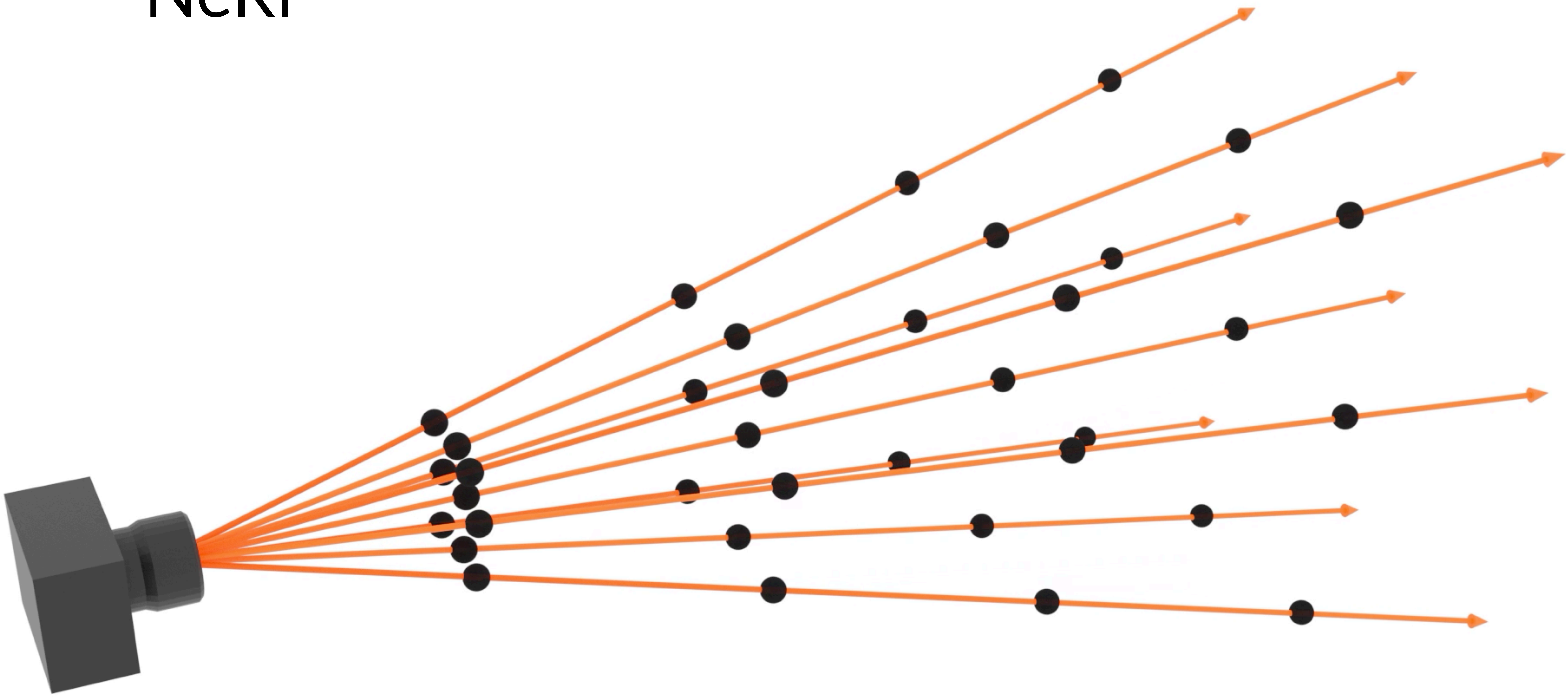
mip-NeRF reduces NeRF aliasing

How can we prefilter a neural network?

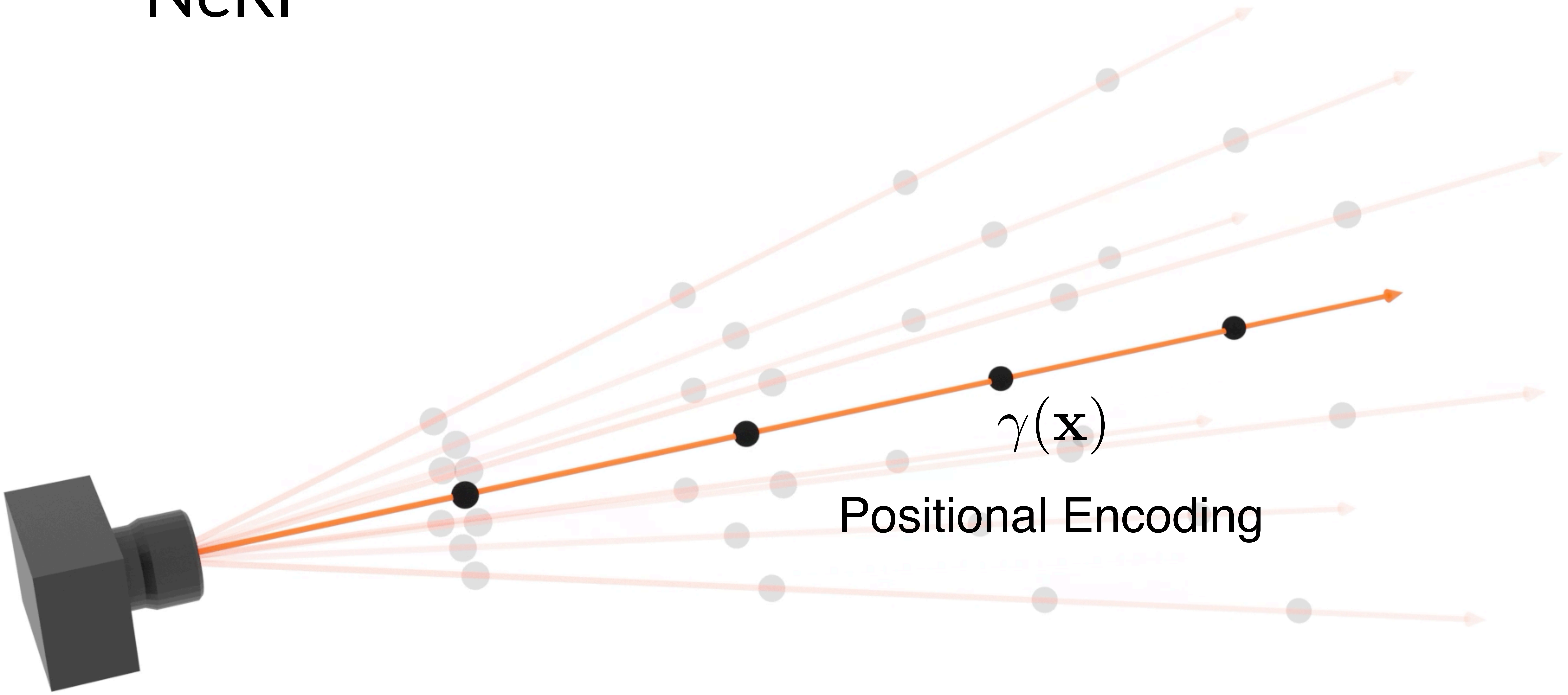
mip = “*multum in parvo*”, Latin for “much in little”

Barron et al. (2021)

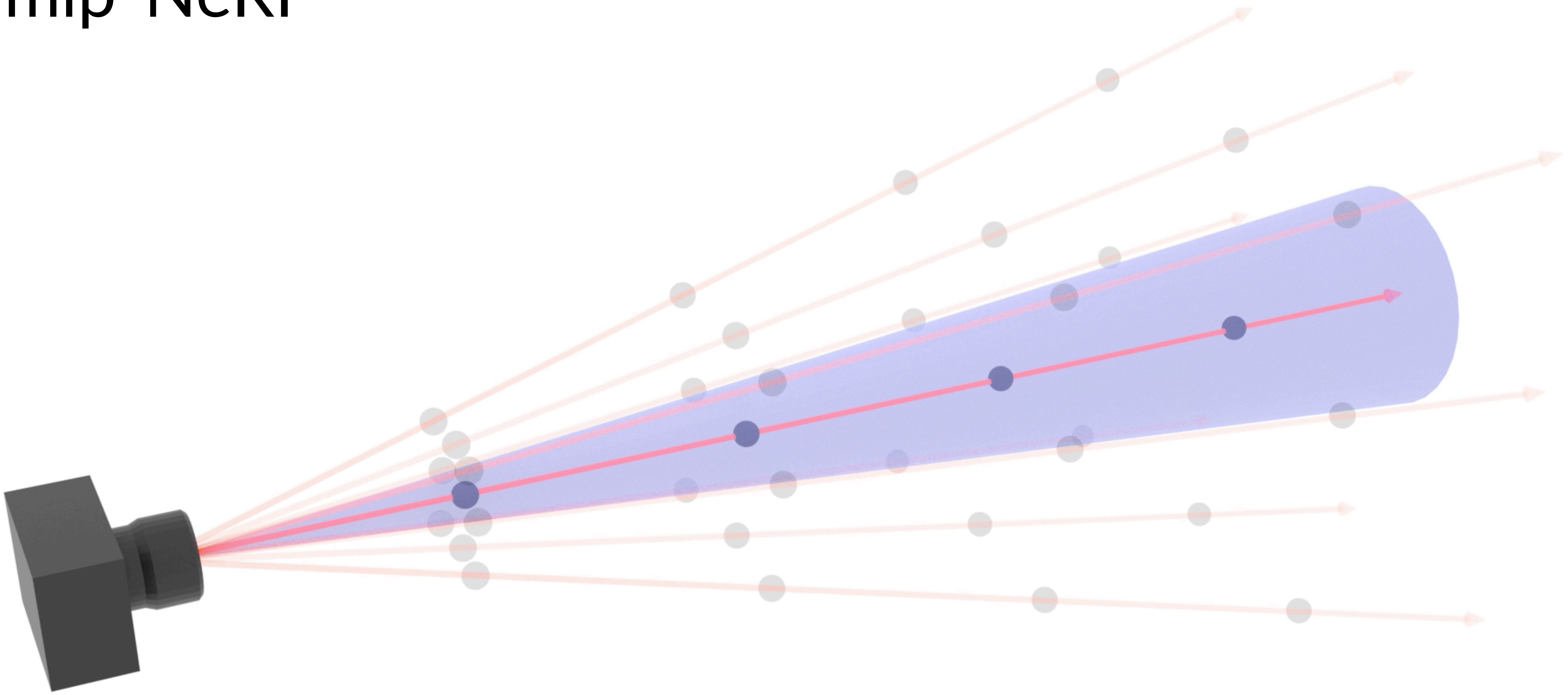
NeRF



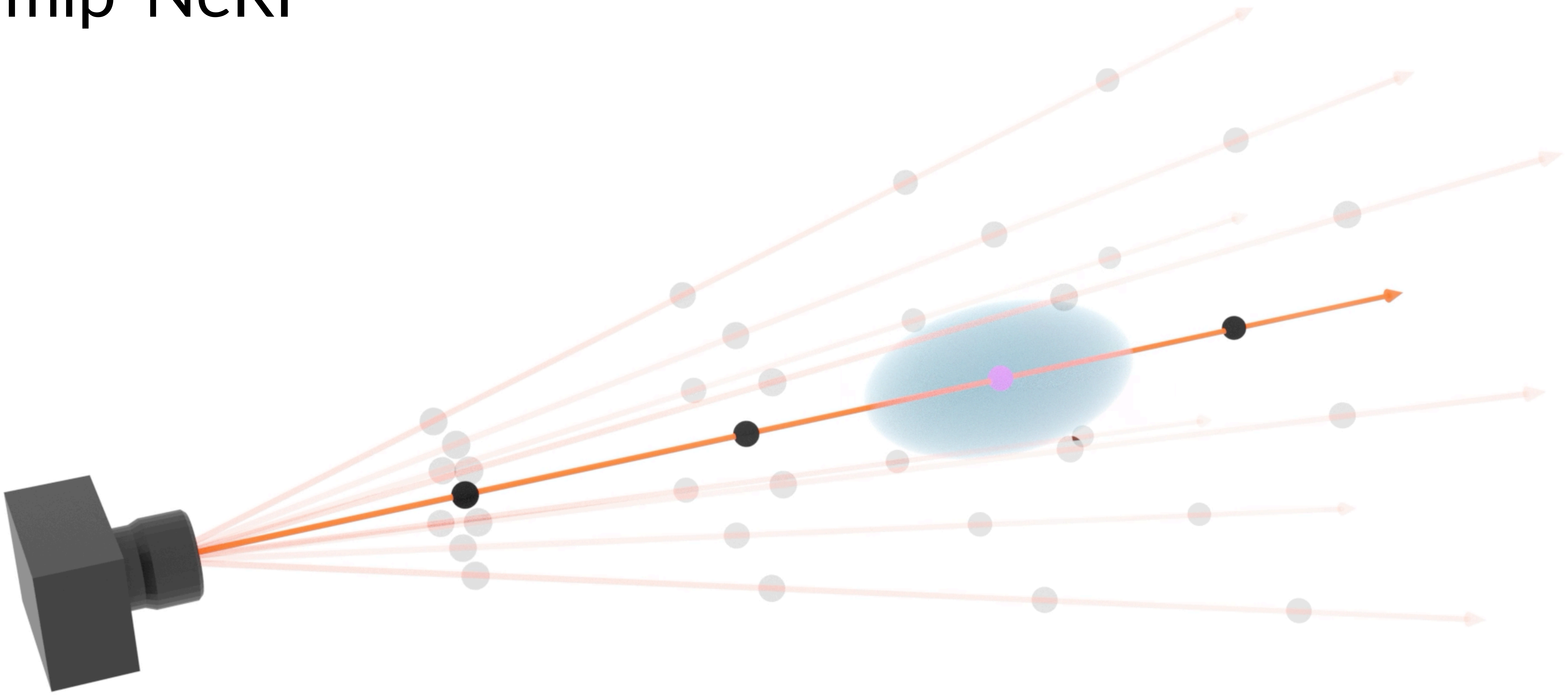
NeRF



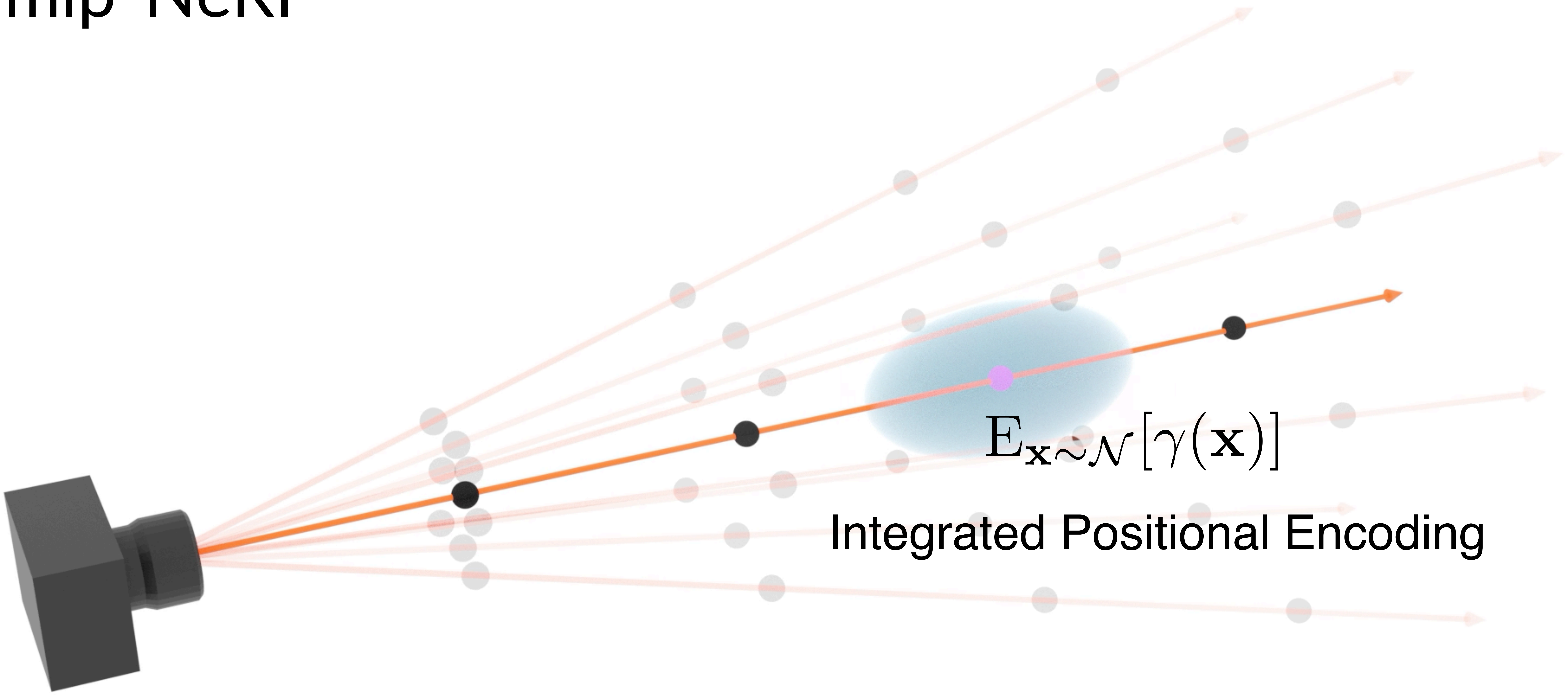
mip-NeRF



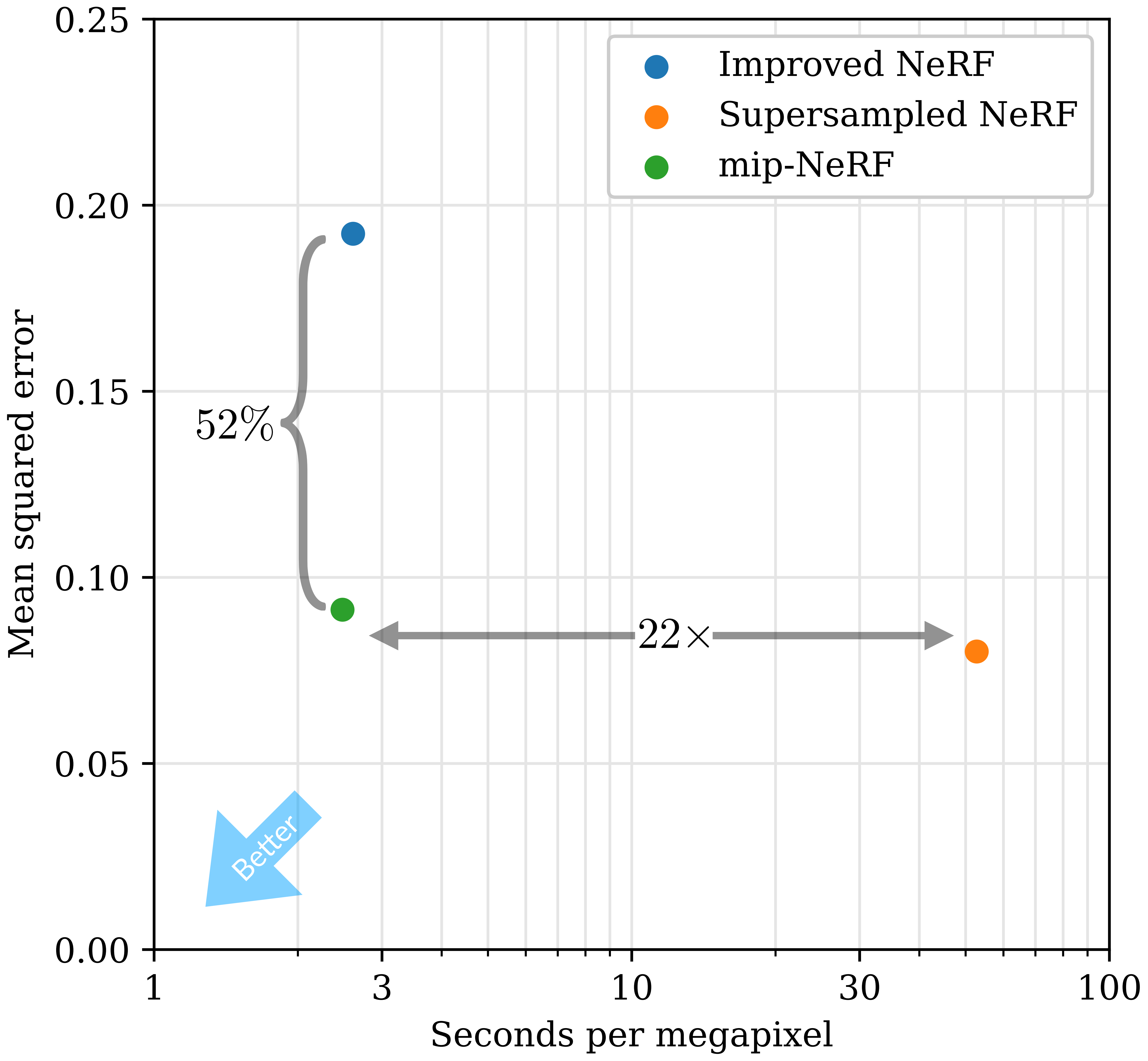
mip-NeRF



mip-NeRF



$\mathrm{E}_{\mathbf{x}\sim\mathcal{N}}[\gamma(\mathbf{x})]$
Integrated Positional Encoding



7% faster training

50% smaller model

Improved NeRF

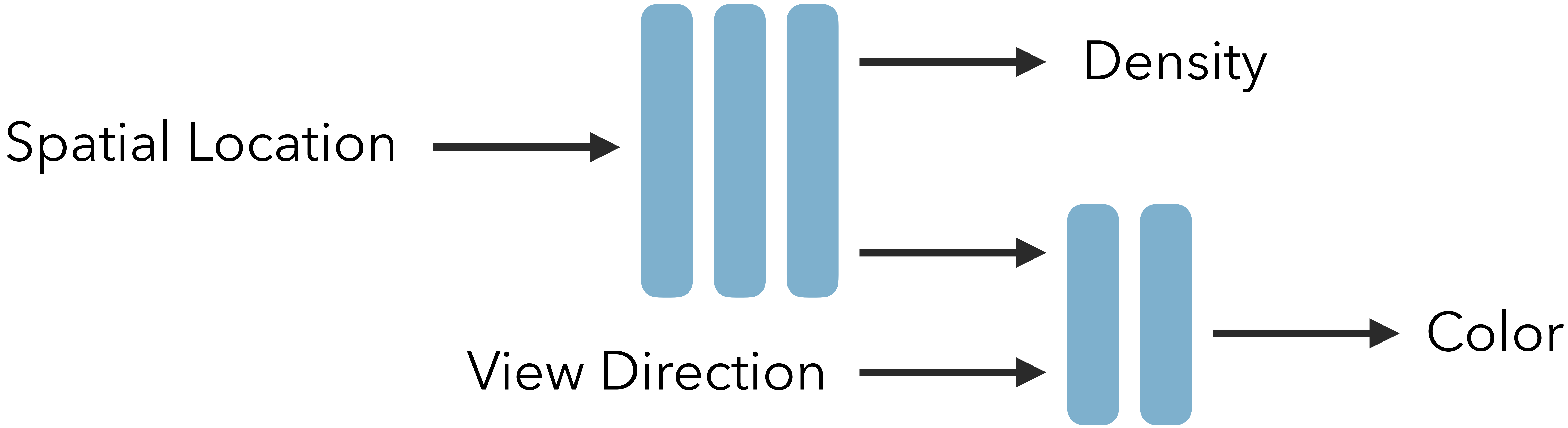
Mip-NeRF

NeRF in the Wild

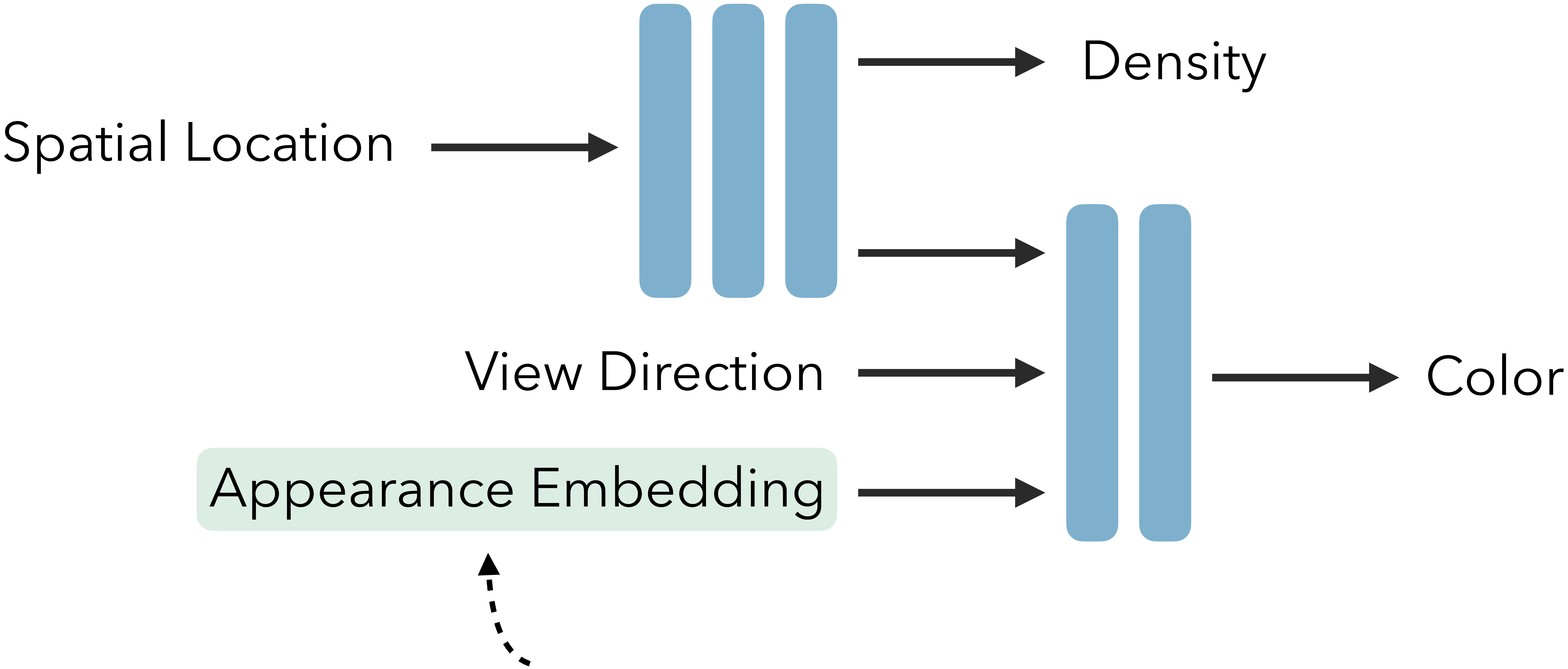
"In the wild" photographs

NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections, Martin-Brualla et al. CVPR 2021

Modifications to NeRF



Modifications to NeRF



Learnable higher dimensional input that is unique for each input value.


Old Town Square
Prague, Czech Republic

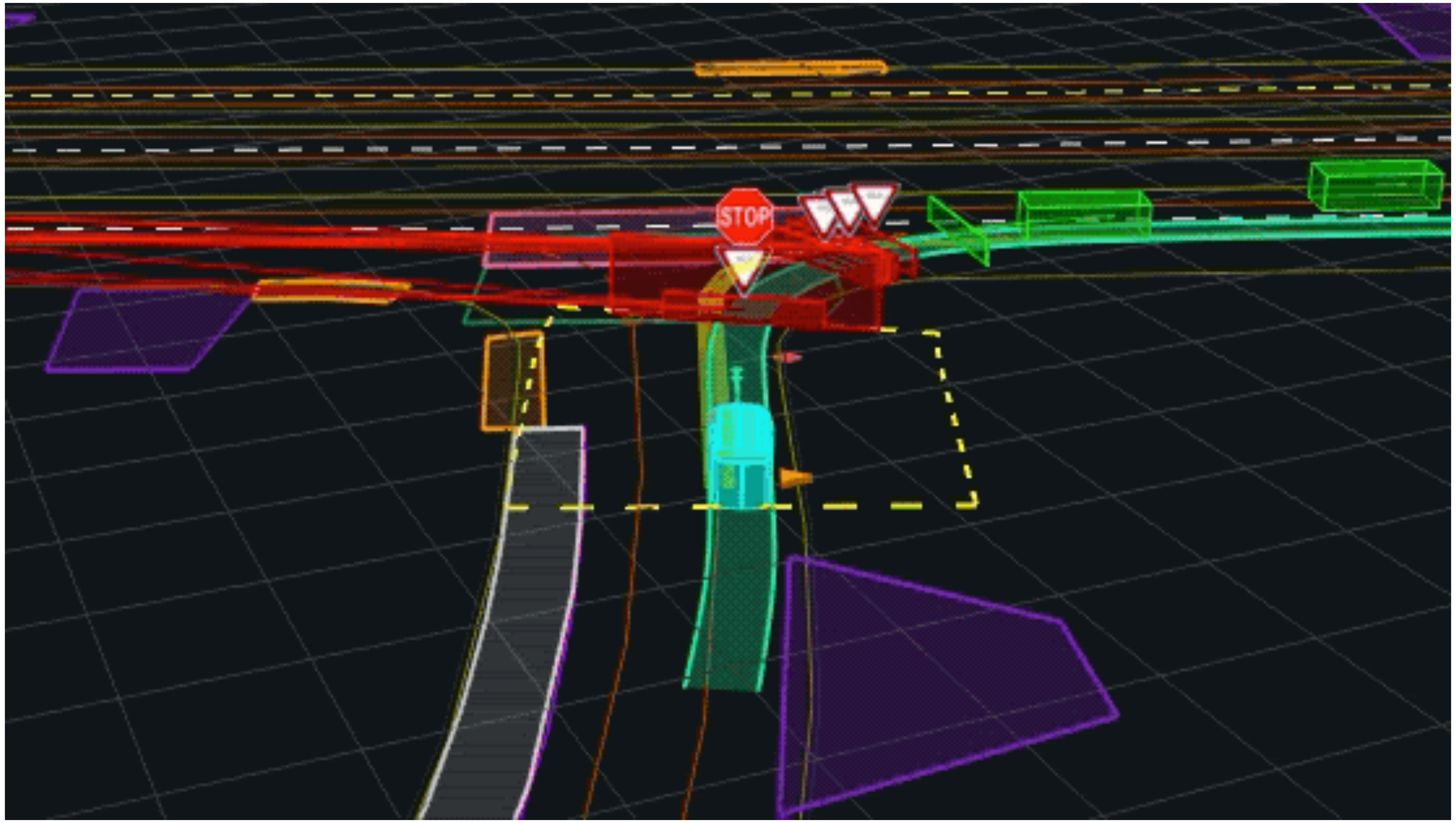
Block-NeRF: Scalable Large Scene Neural View Synthesis

Matthew Tancik[1] Vincent Casser[2] Xinchen Yan[2] Sabeek Pradhan[2]
Ben Mildenhall[3] Pratul P. Srinivasan[3] Jonathan T. Barron[3] Henrik Kretzschmar[2]

1 2 WAYMO 3 Google

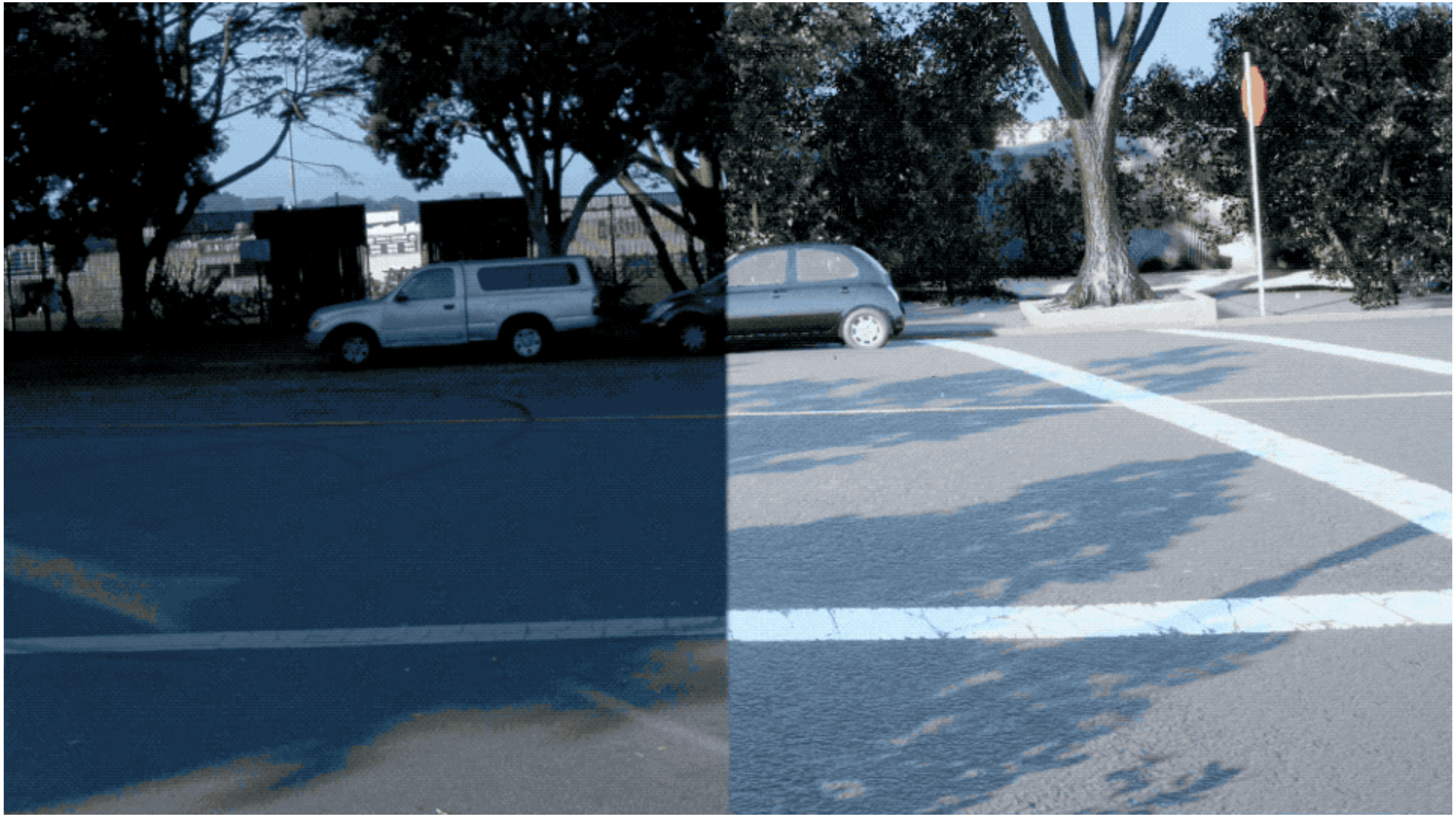
Safety Testing

Simulation is Essential for Scale



Lidar Simulation

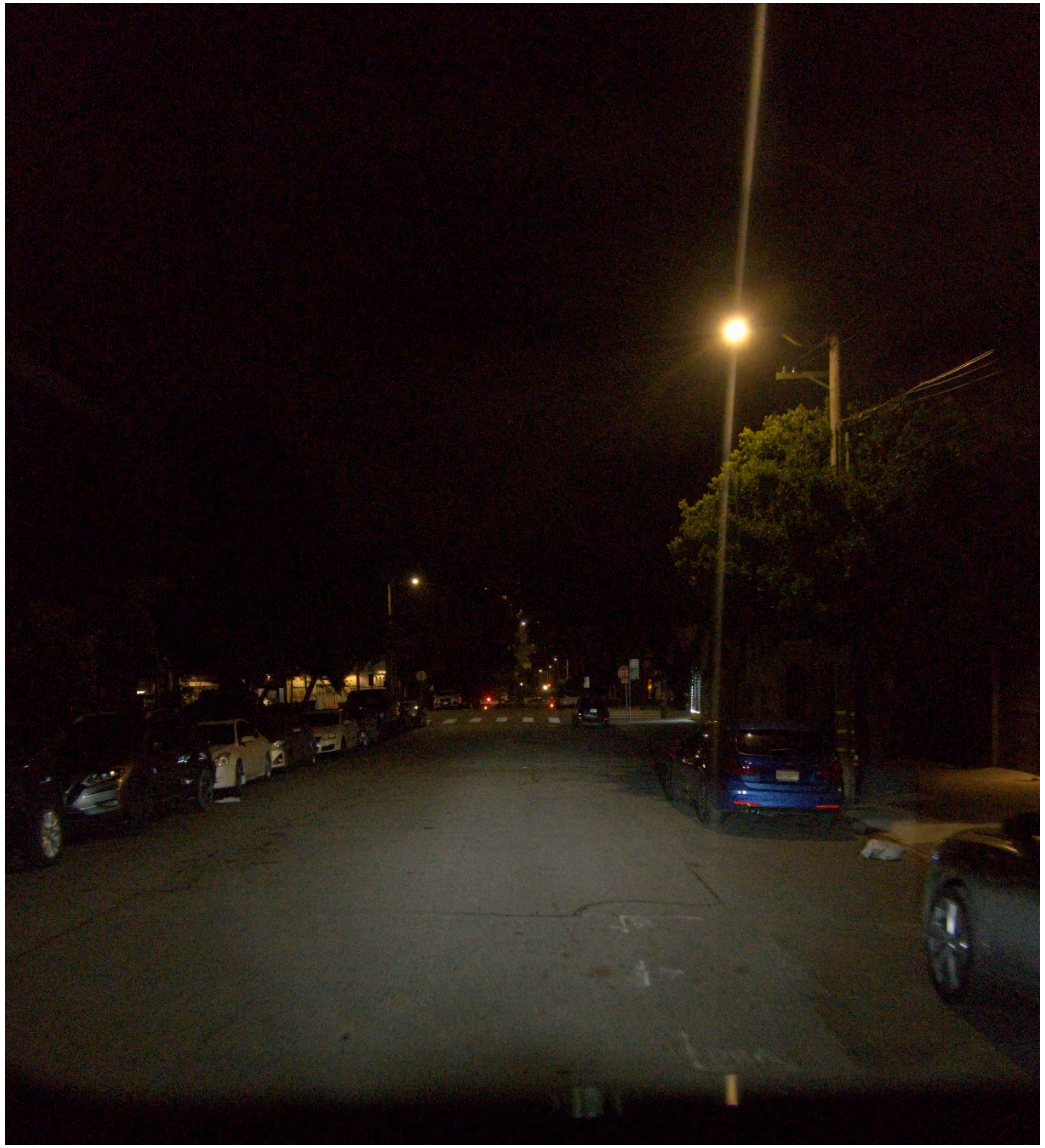
Camera Simulation using Models



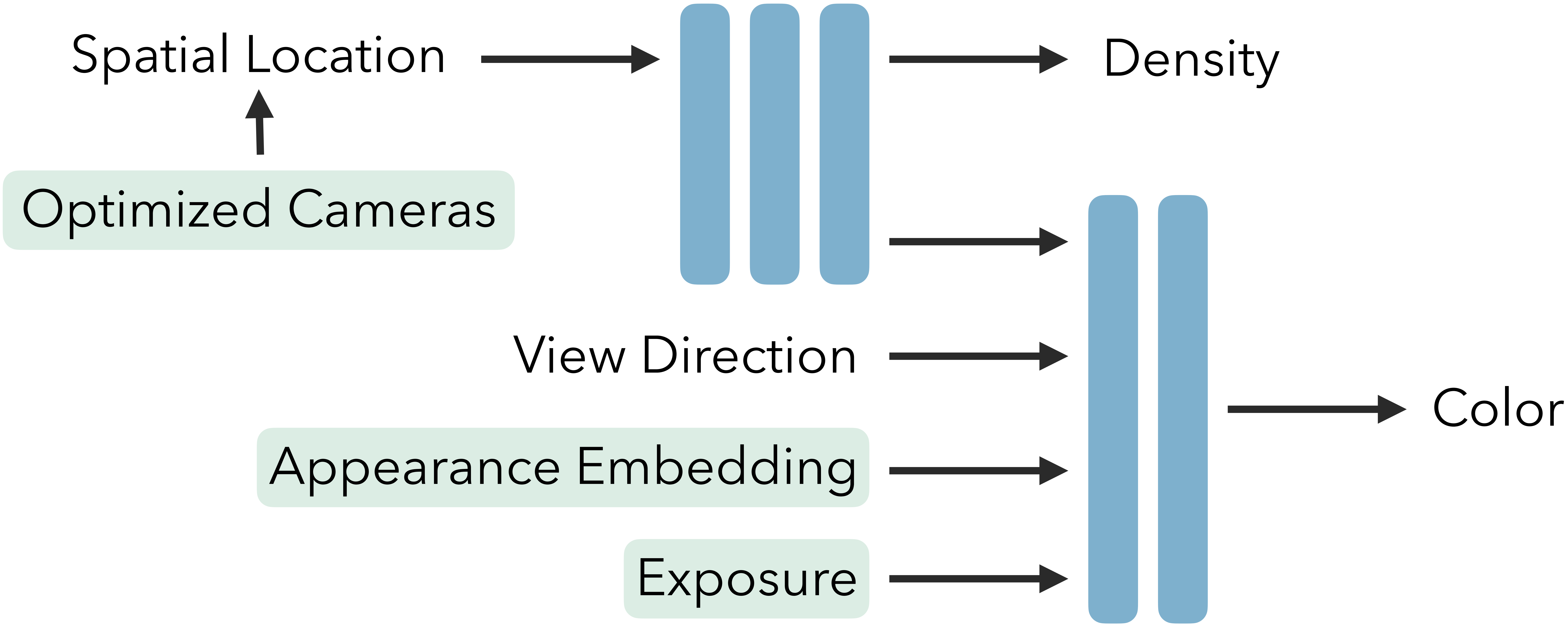
Access to lots of data

+ 360 degree view
+ Camera poses

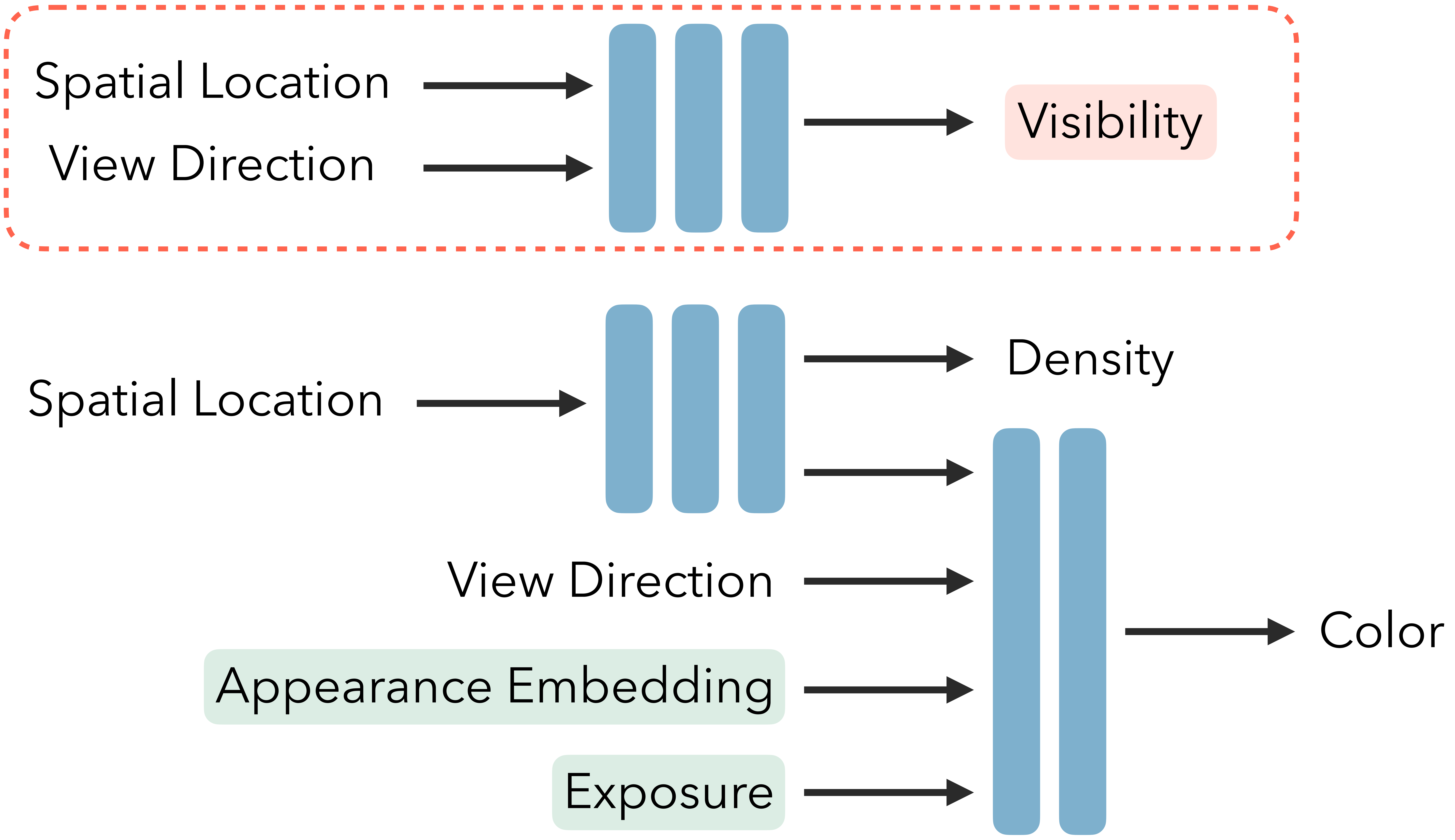
… but highly variable



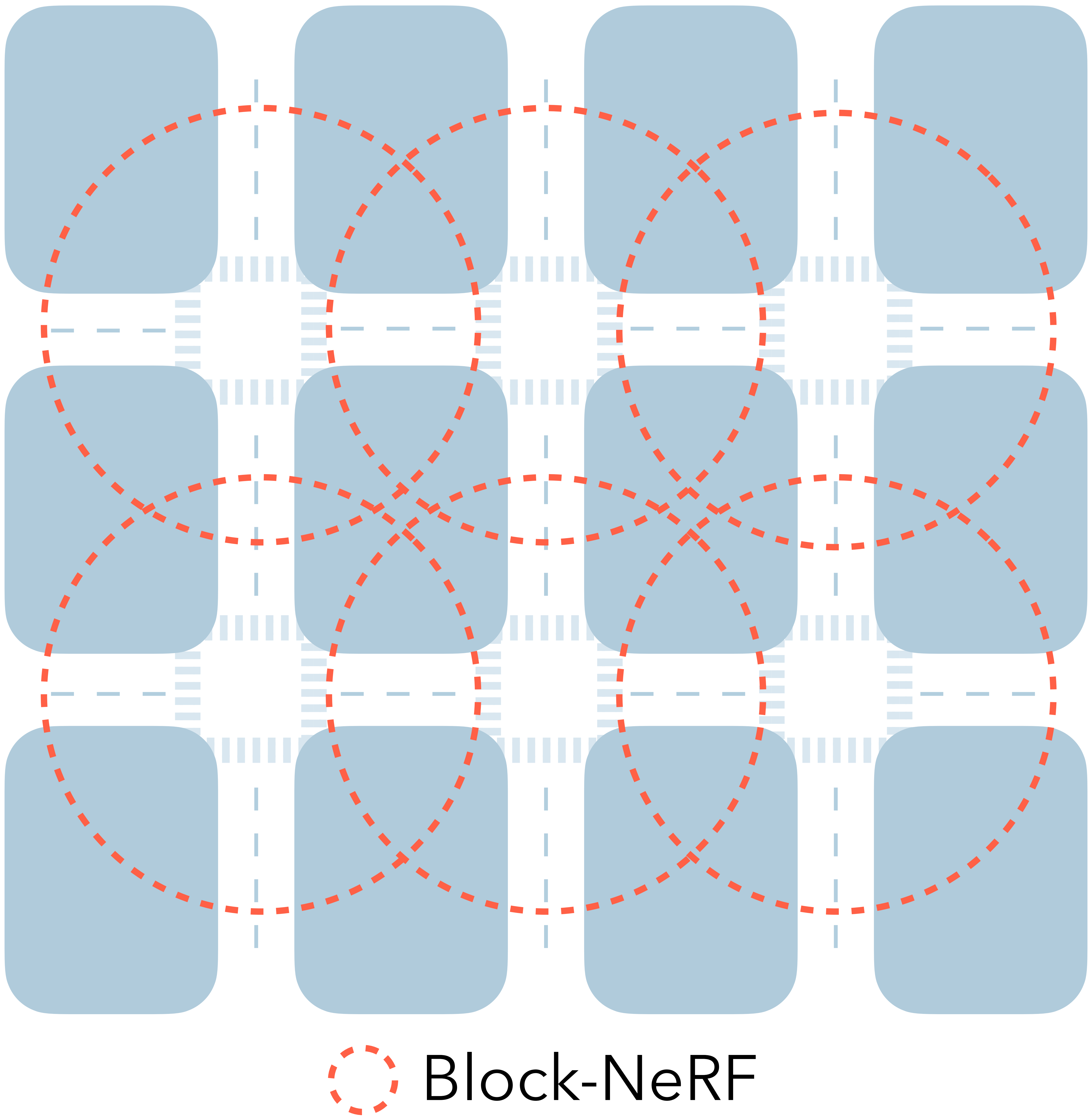
Modifications to NeRF



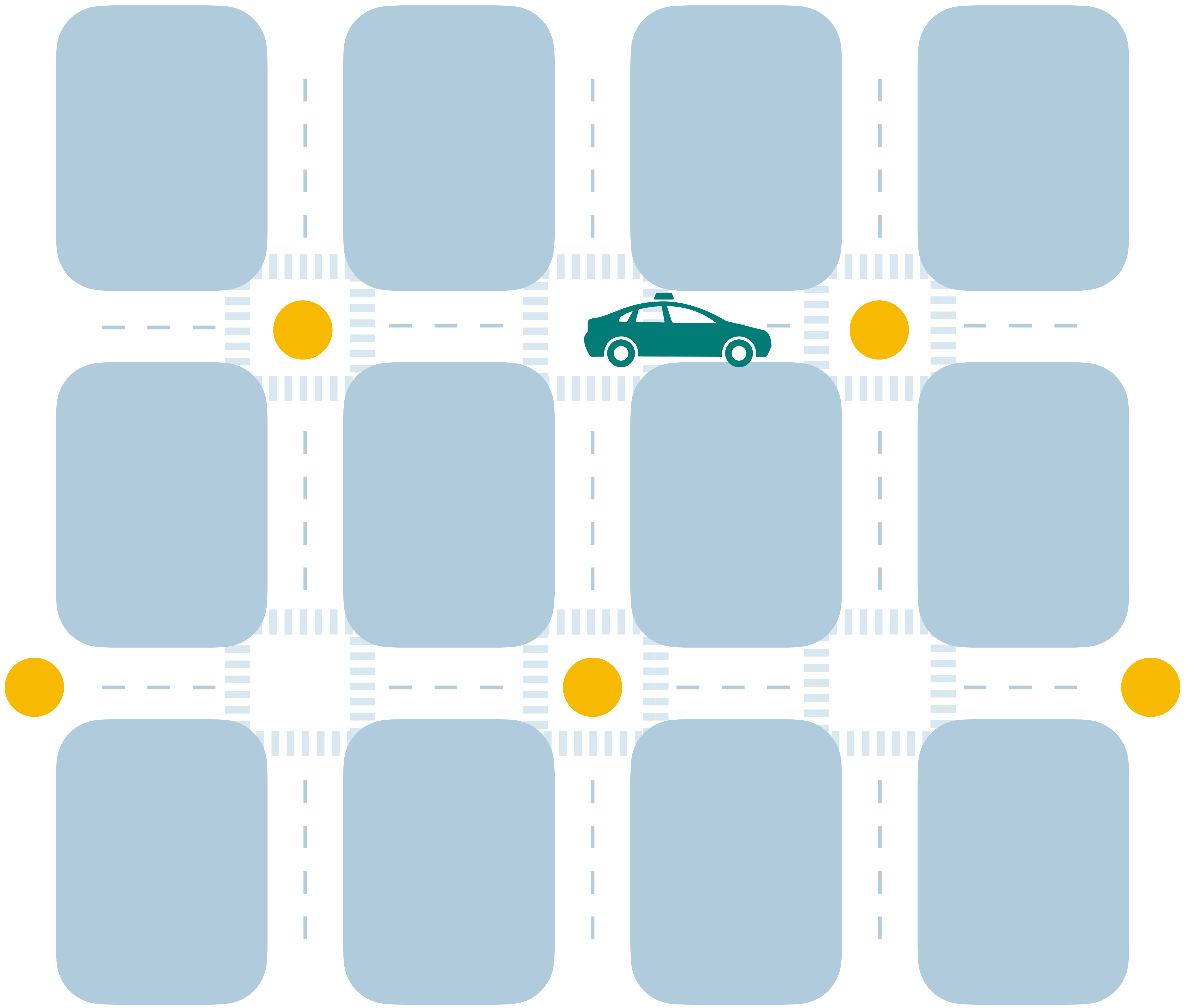
Modifications to NeRF



Scaling to Large Scenes

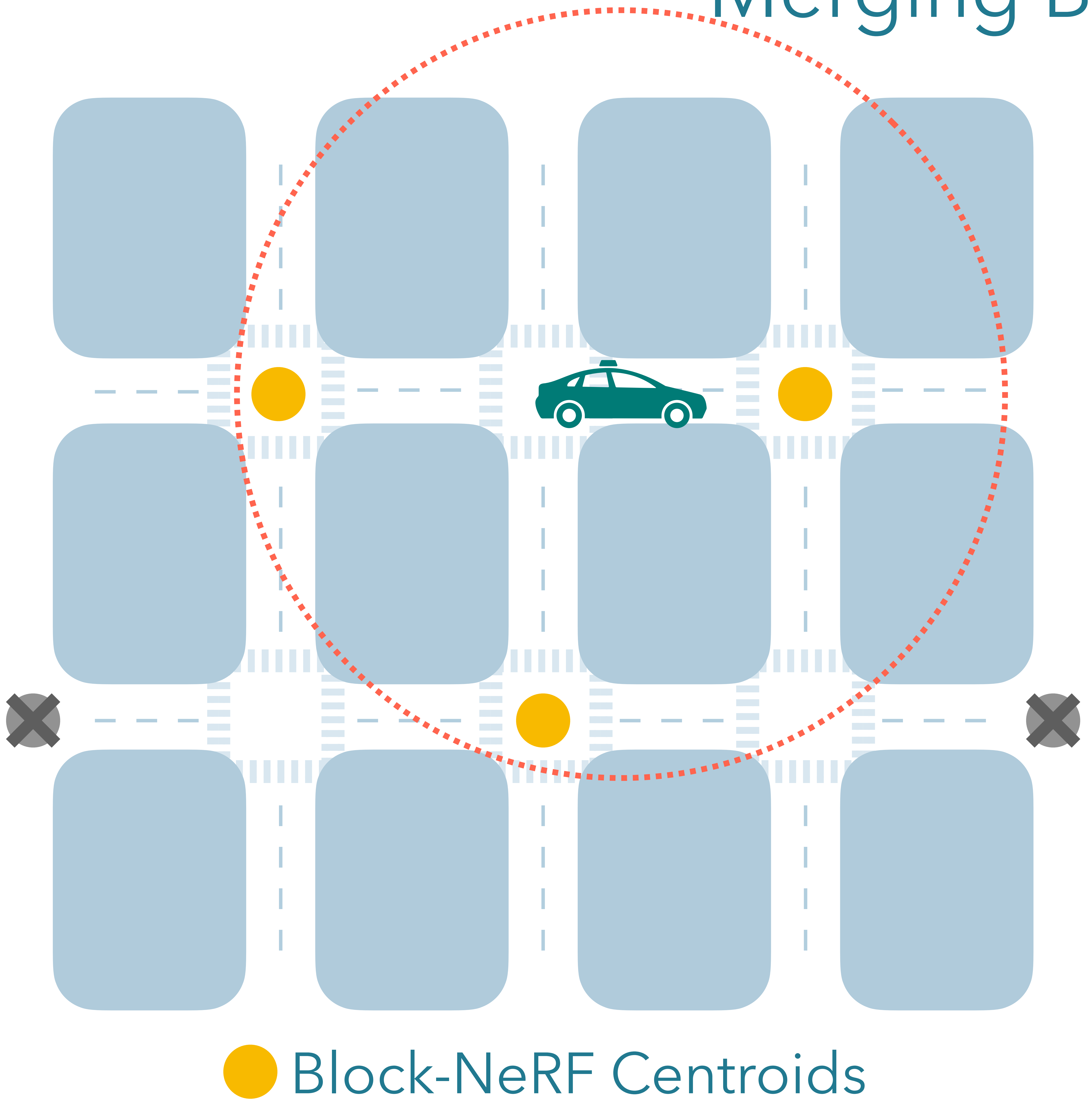


Merging Block-NeRFs

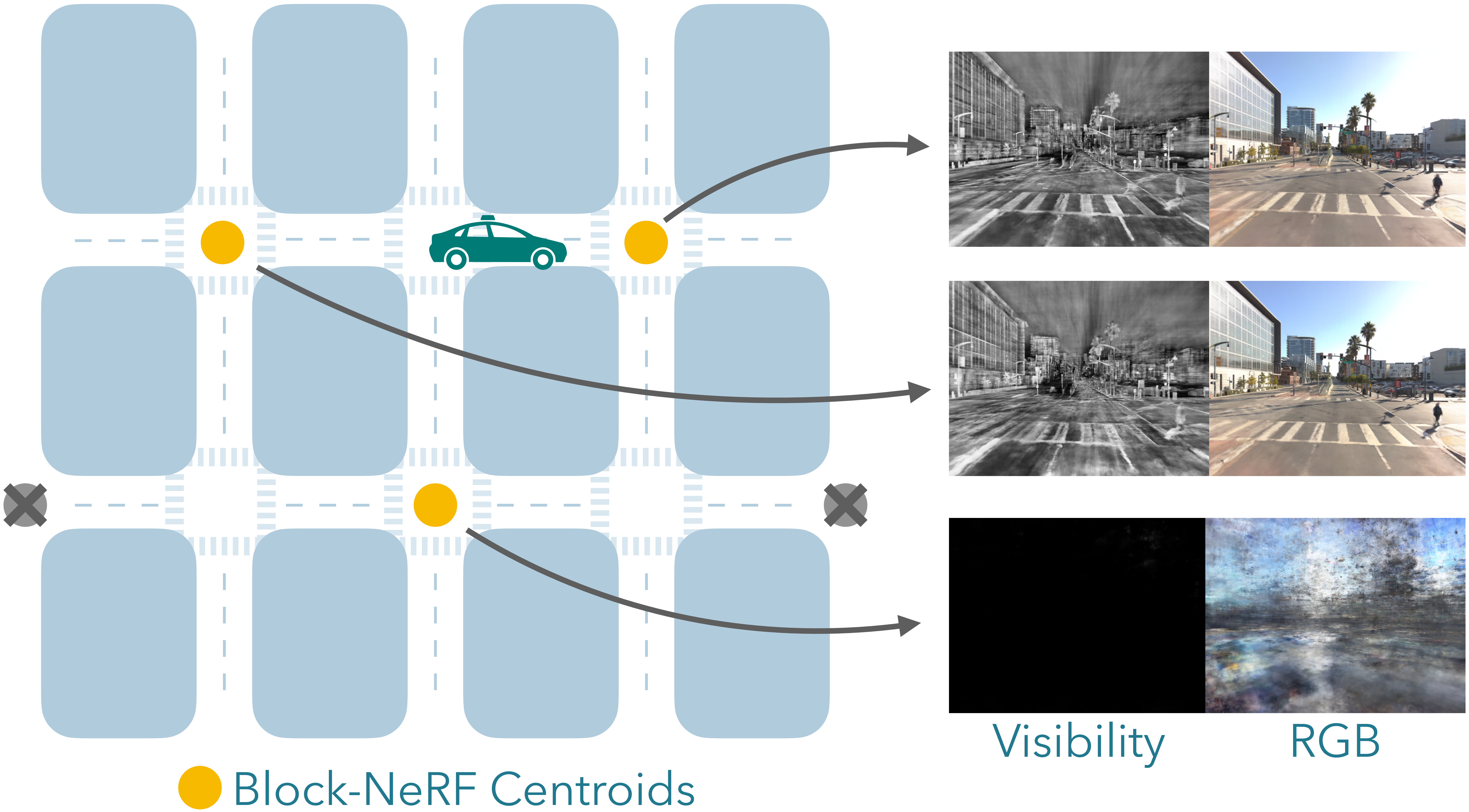


Block-NeRF Centroids

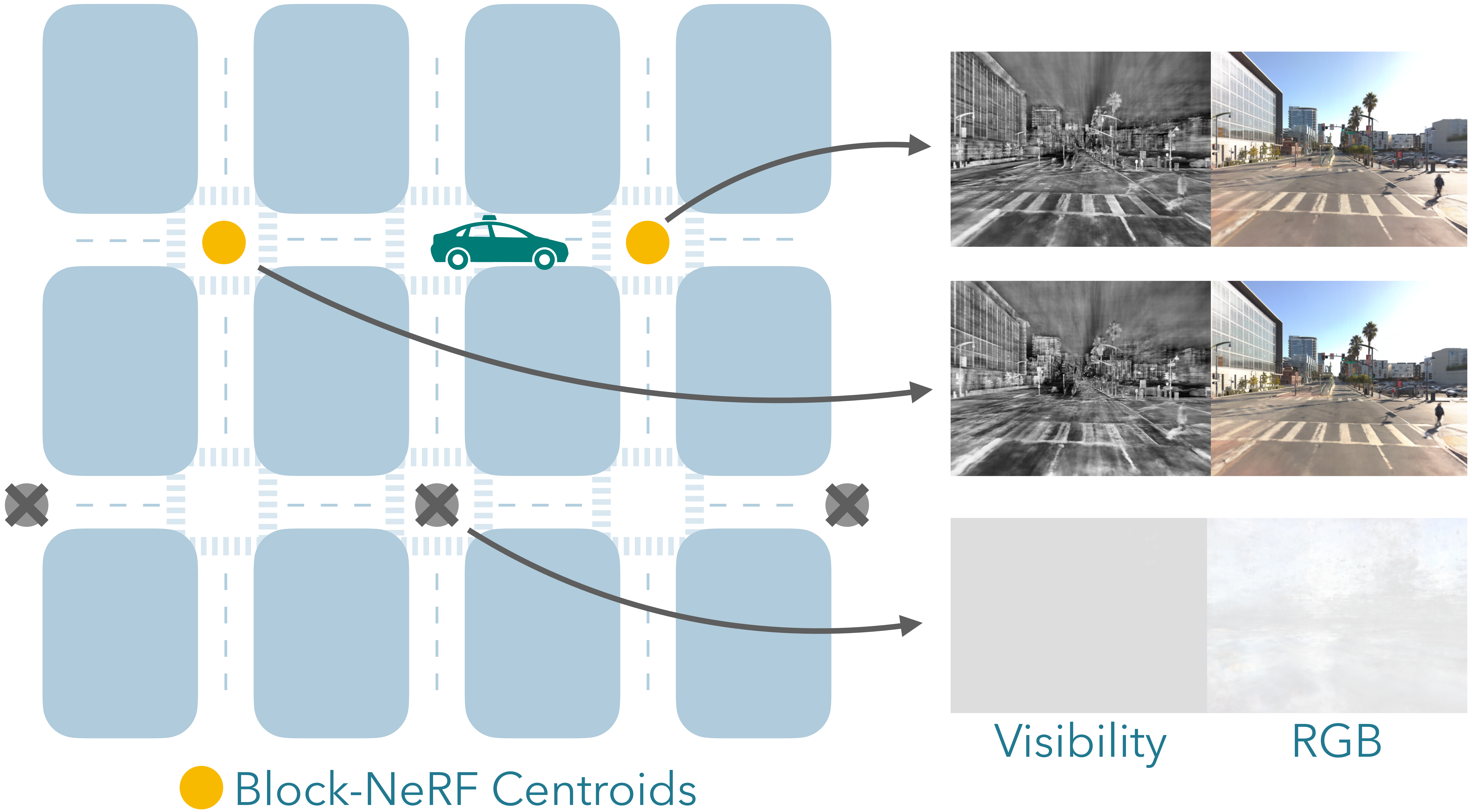
Merging Block-NeRFs



Merging Block-NeRFs

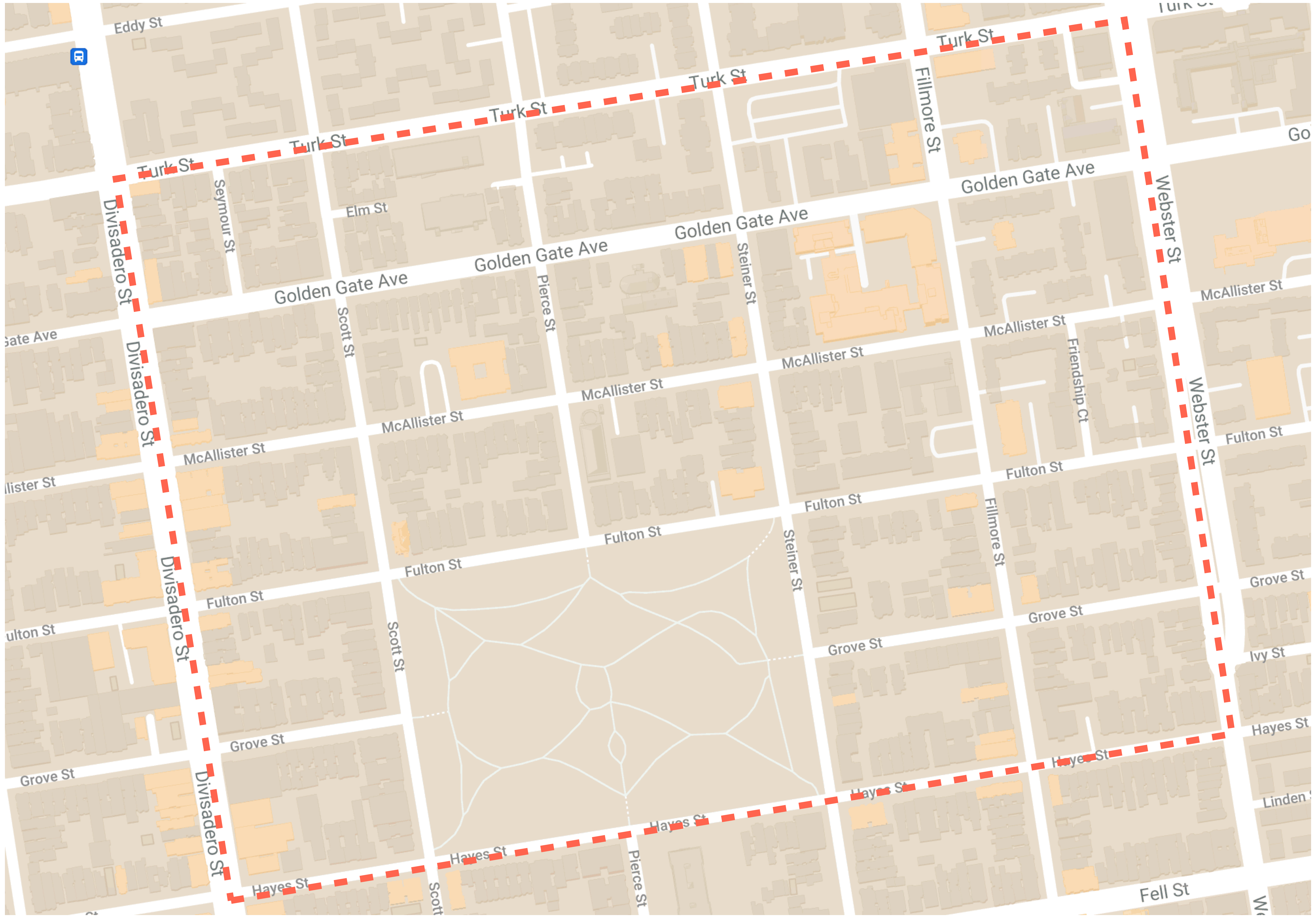


Merging Block-NeRFs



Merging Block-NeRFs

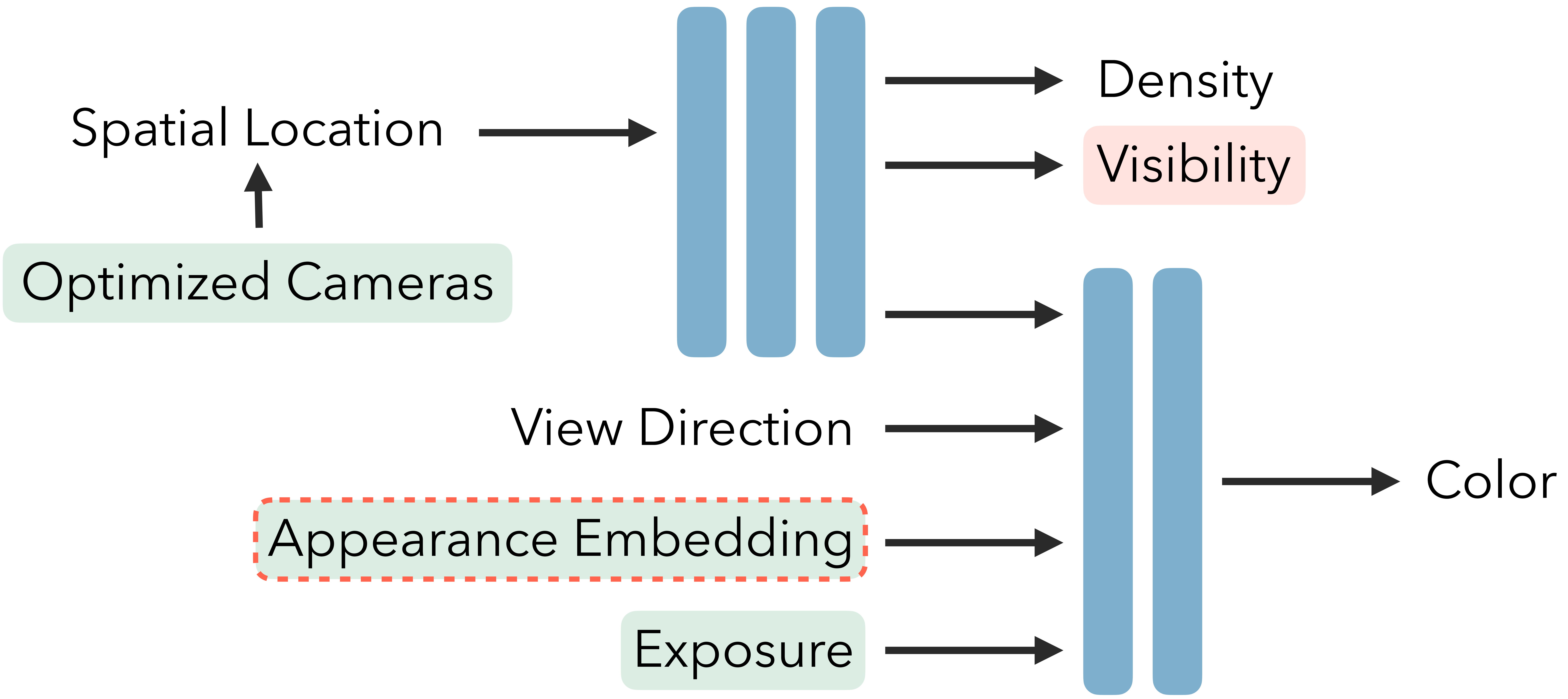
Large Scene Reconstruction



Alamo Square Neighborhood, San Francisco

- 2.8 Million Images
- Captured over 3 months
- 1330 Data collection runs
- 13.4 Hours of driving
- 35 Block-NeRFs

Modifications to NeRF



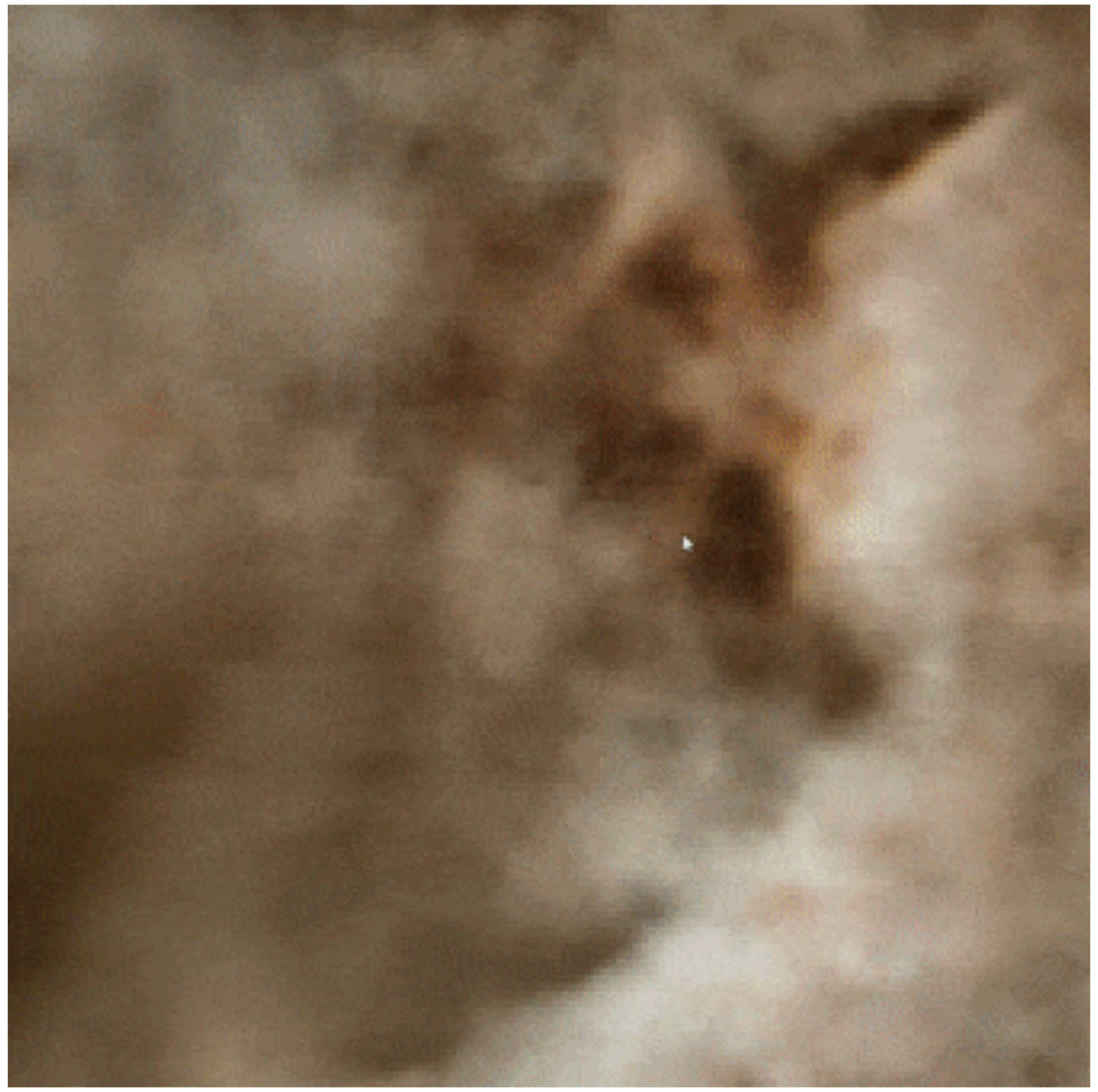
Appearance Modulation

Easy to locally update scene

During Construction

After Construction

Instant NGP (train NeRF in seconds)

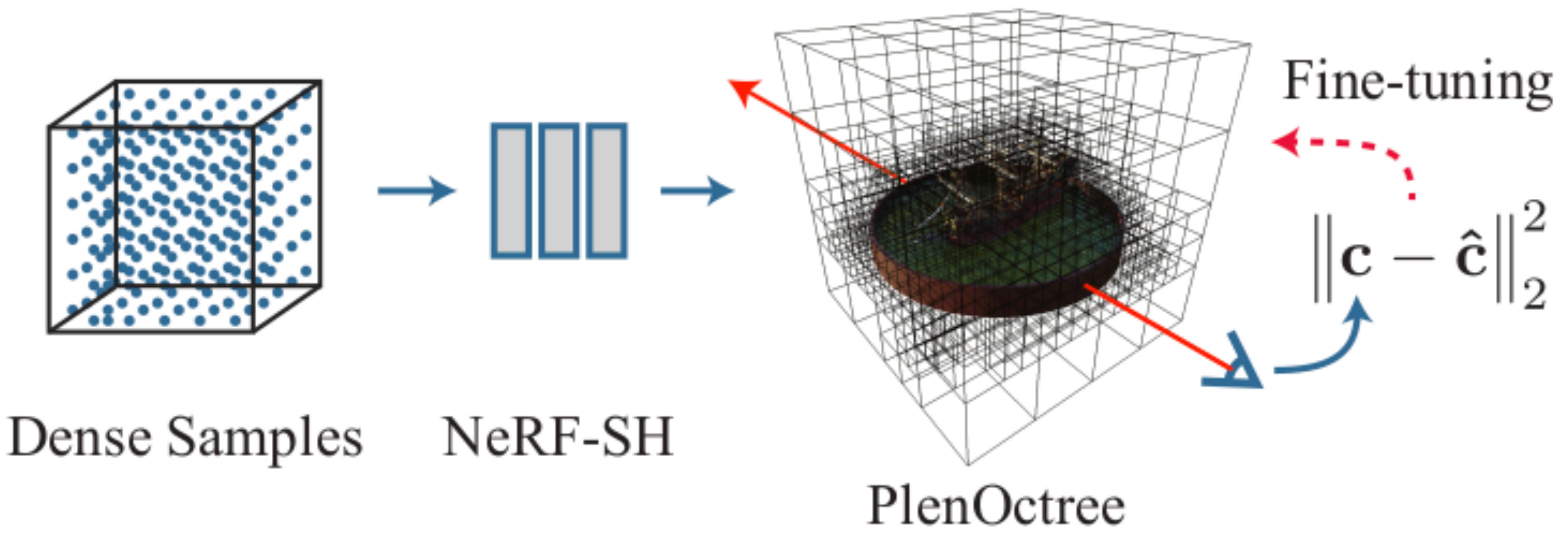


Instant Neural Graphics Primitives with a Multiresolution Hash Encoding, Müller et al. arXiv 2022

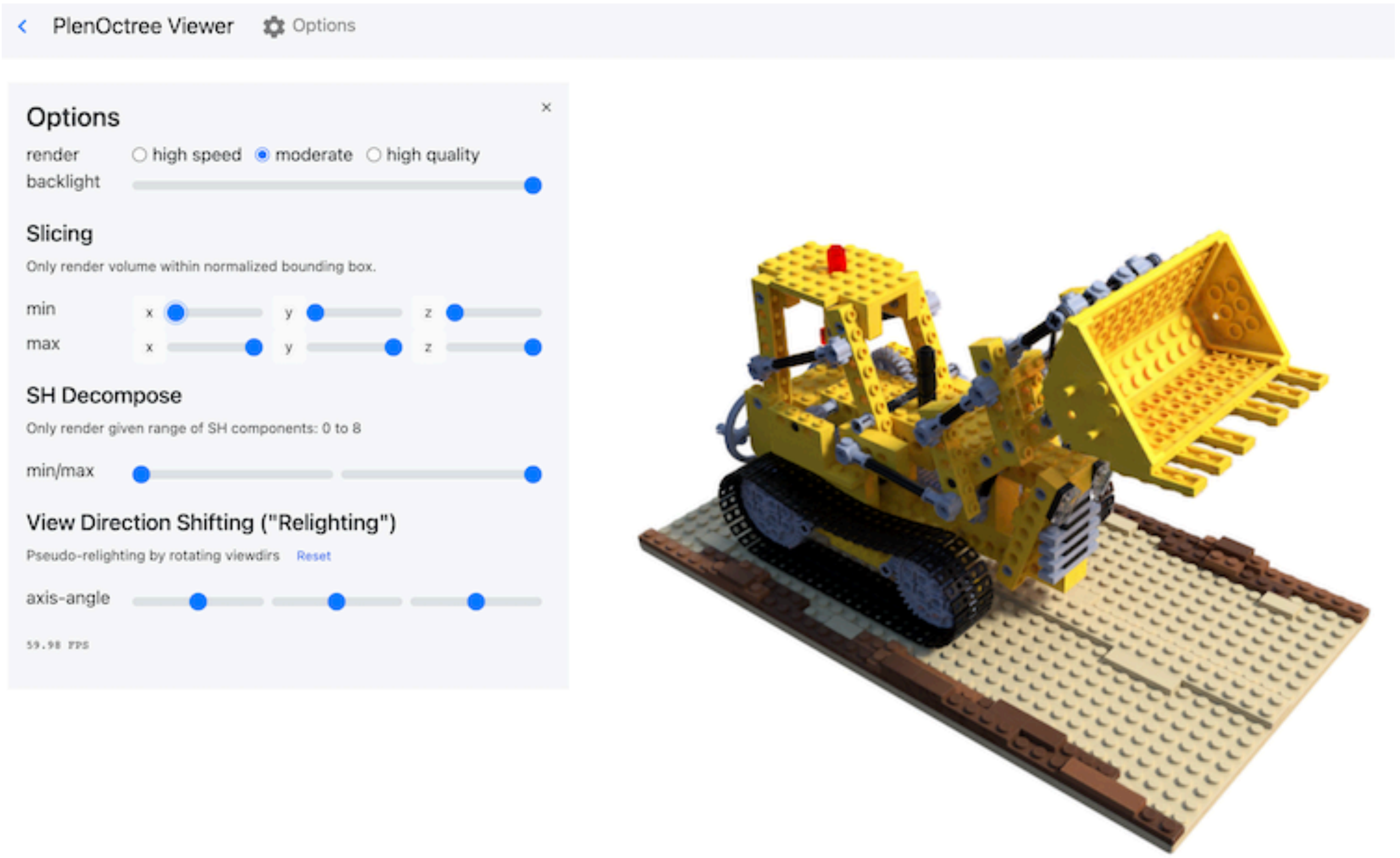
Instant NGP (train NeRF in seconds)

PlenOctrees

Render a NeRF in your browser - alexyu.net/plenoctrees/



Conversion to a PlenOctree



Real-time Rendering

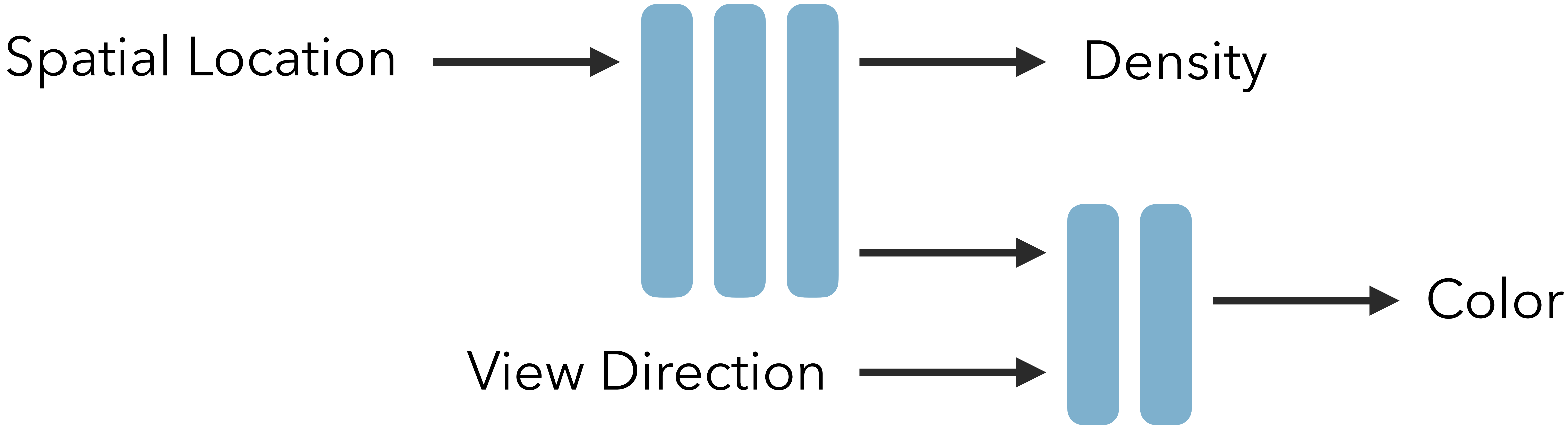
PlenOctrees: For Real-time Rendering of Neural Radiance Fields, Yu et al. arXiv 2021

NeRFies

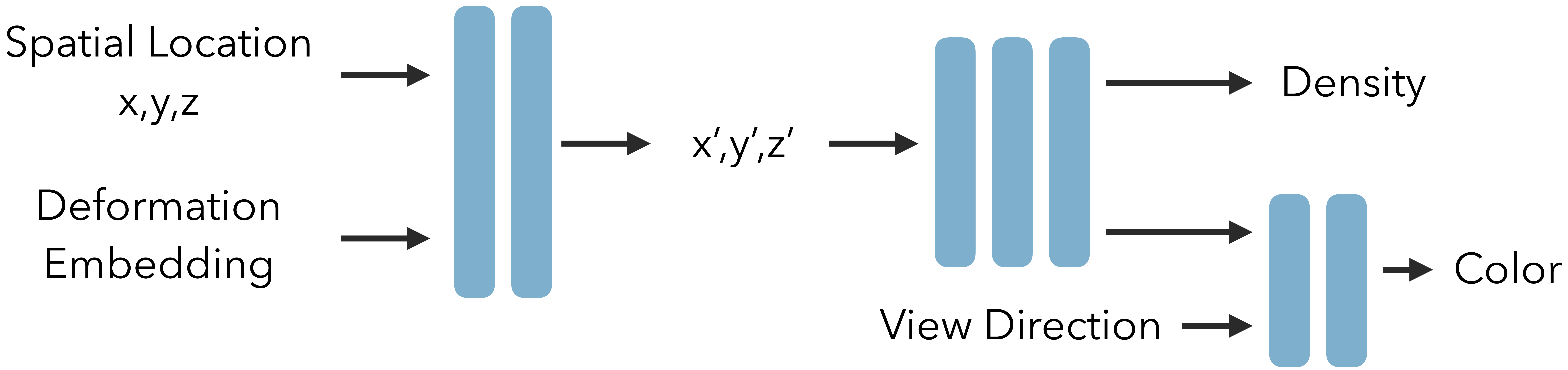
(a) Capture Process (b) Input (c) Nerfie (d) Nerfie Depth

Deformable Neural Radiance Fields, Park et al. arXiv 2020

Modifications to NeRF



Modifications to NeRF



Deformable Neural Radiance Fields, Park et al. arXiv 2020

Video NeRF

Input video

Fixed Time, View Interpolation

Fixed View, Time Interpolation

Neural Scene Flow Fields for Space-Time View Synthesis of Dynamic Scenes, Li et al. CVPR 2021

Video NeRF

Input

Neural Scene Flow Fields for Space-Time View Synthesis of Dynamic Scenes, Li et al. CVPR 2021

3D GANs

Efficient Geometry-aware 3D Generative Adversarial Networks, Chan et al. arXiv 2021

Dream Fields

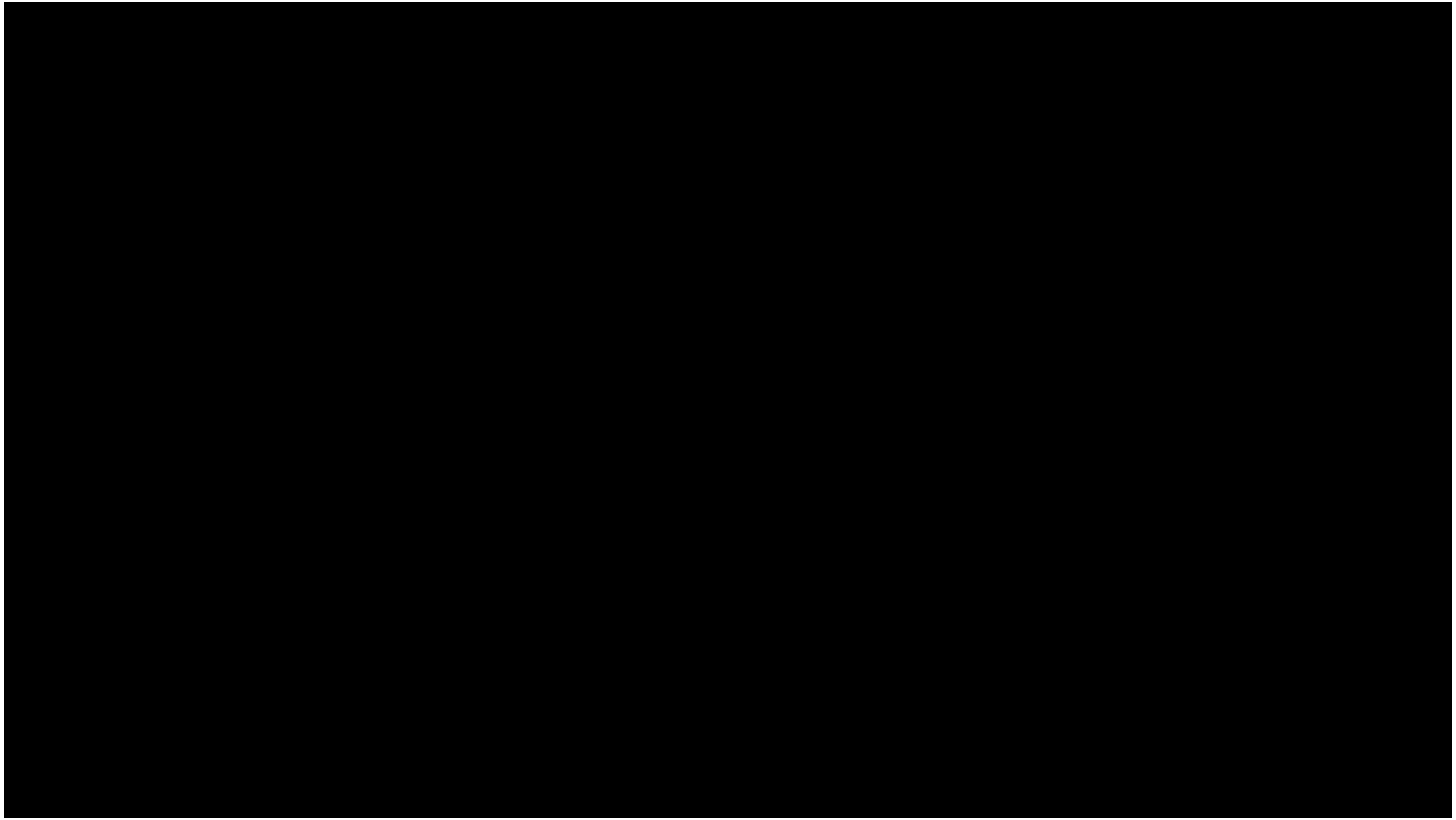
An armchair in the shape of a ___.
An armchair imitating a ___.

Zero-Shot Text-Guided Object Generation with Dream Fields, Jain et al. arXiv 2021

HumanNeRF

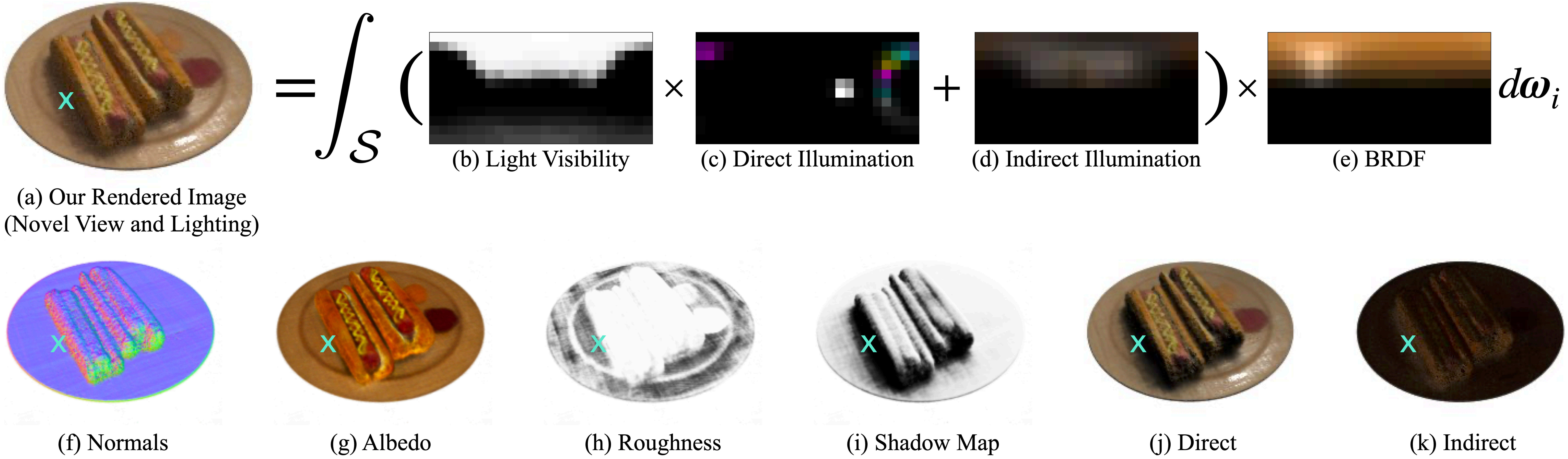
HumanNeRF: Free-viewpoint Rendering of Moving People from Monocular Video, Weng et al. arXiv 2022

NeRF in the Dark



NeRF in the Dark: High Dynamic Range View Synthesis from Noisy Raw Images, Mildenhall et al. arXiv 2021

Decompose Lighting and Materials

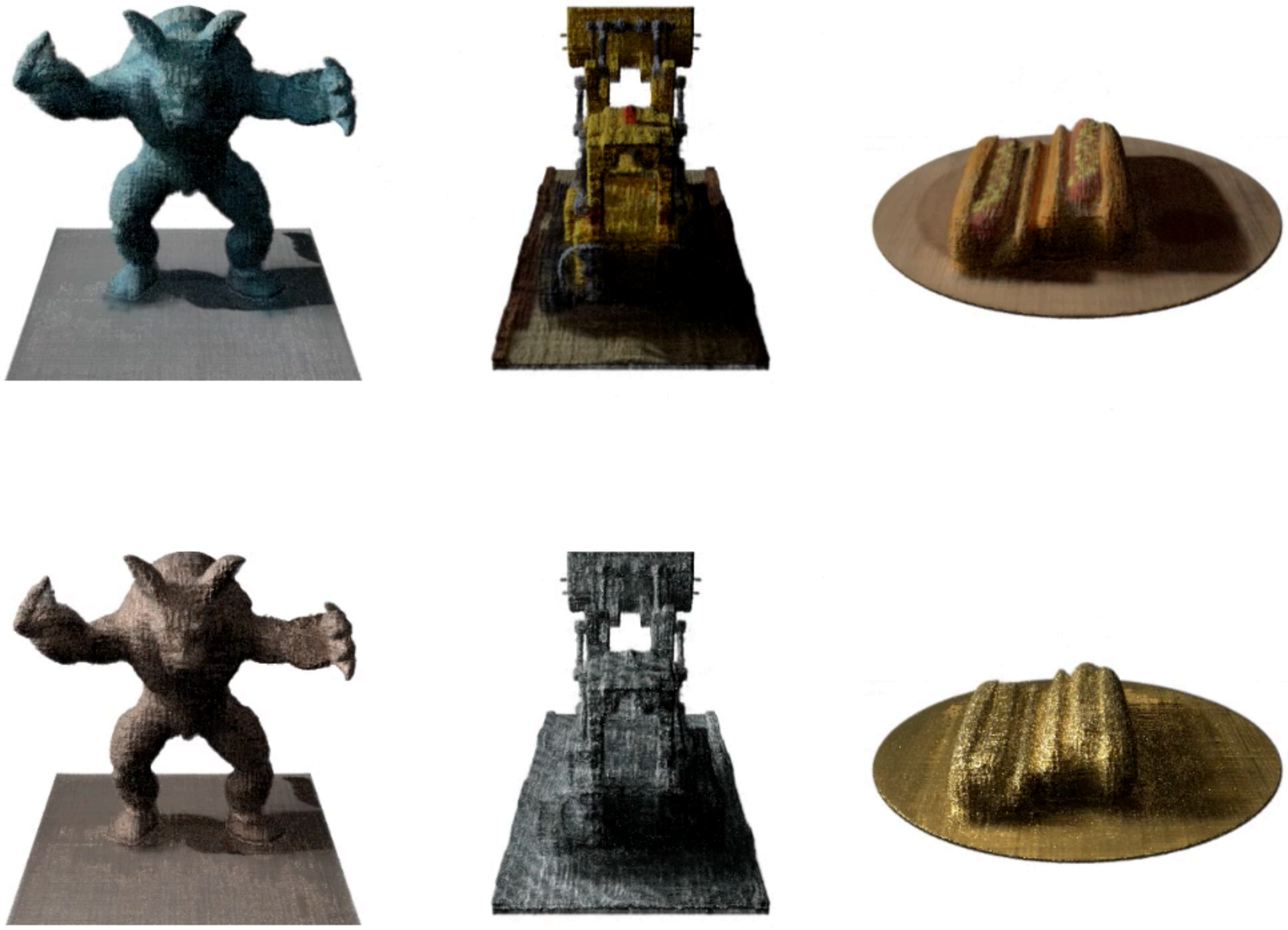


NeRV: Neural Reflectance and Visibility Fields for Relighting and View Synthesis, Srinivasan et al. CVPR 2021

Decompose Lighting and Materials

Edit Lighting

Edit Lighting and Materials

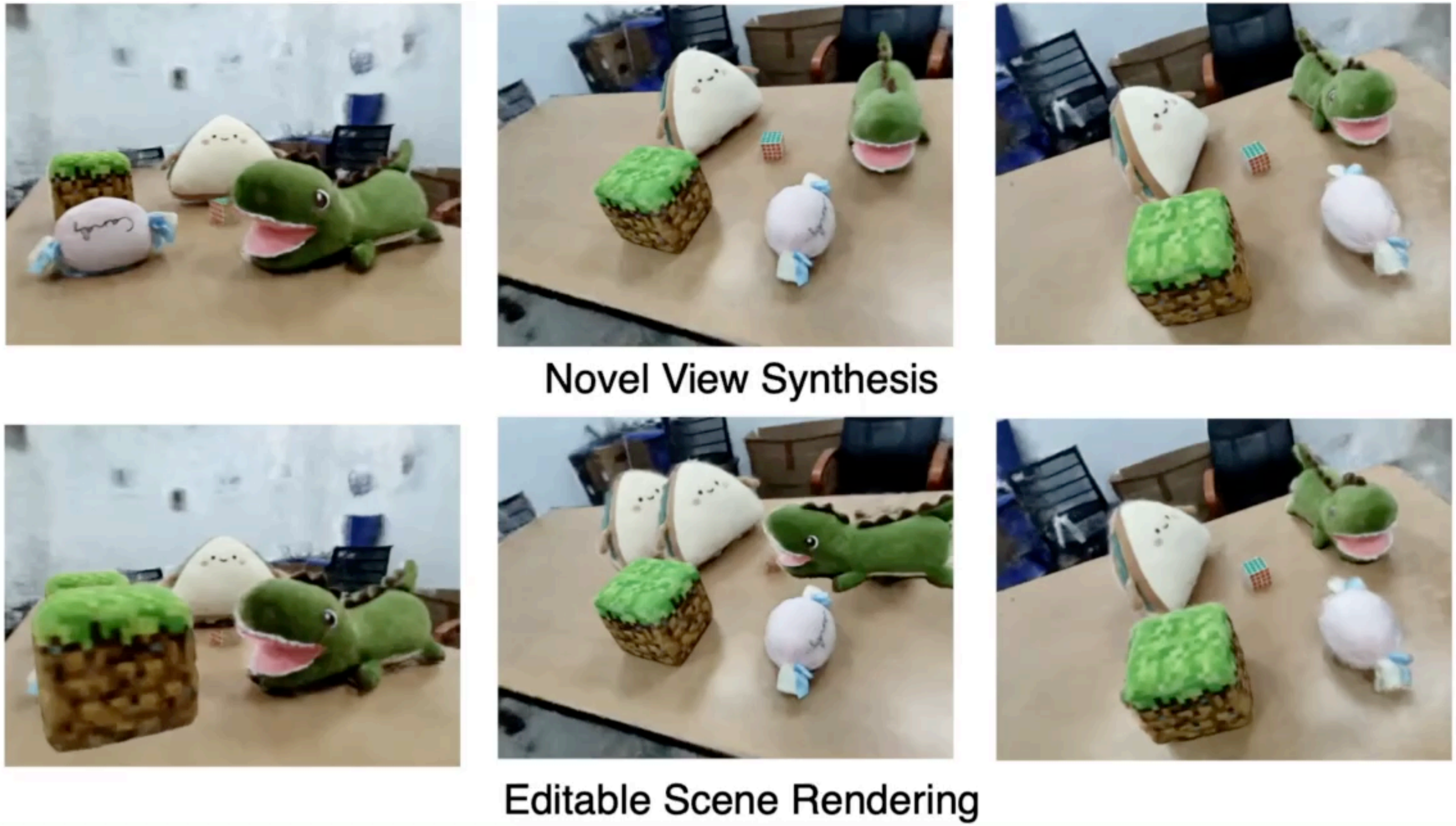


NeRV: Neural Reflectance and Visibility Fields for Relighting and View Synthesis, Srinivasan et al. CVPR 2021

iMap (NeRF SLAM)

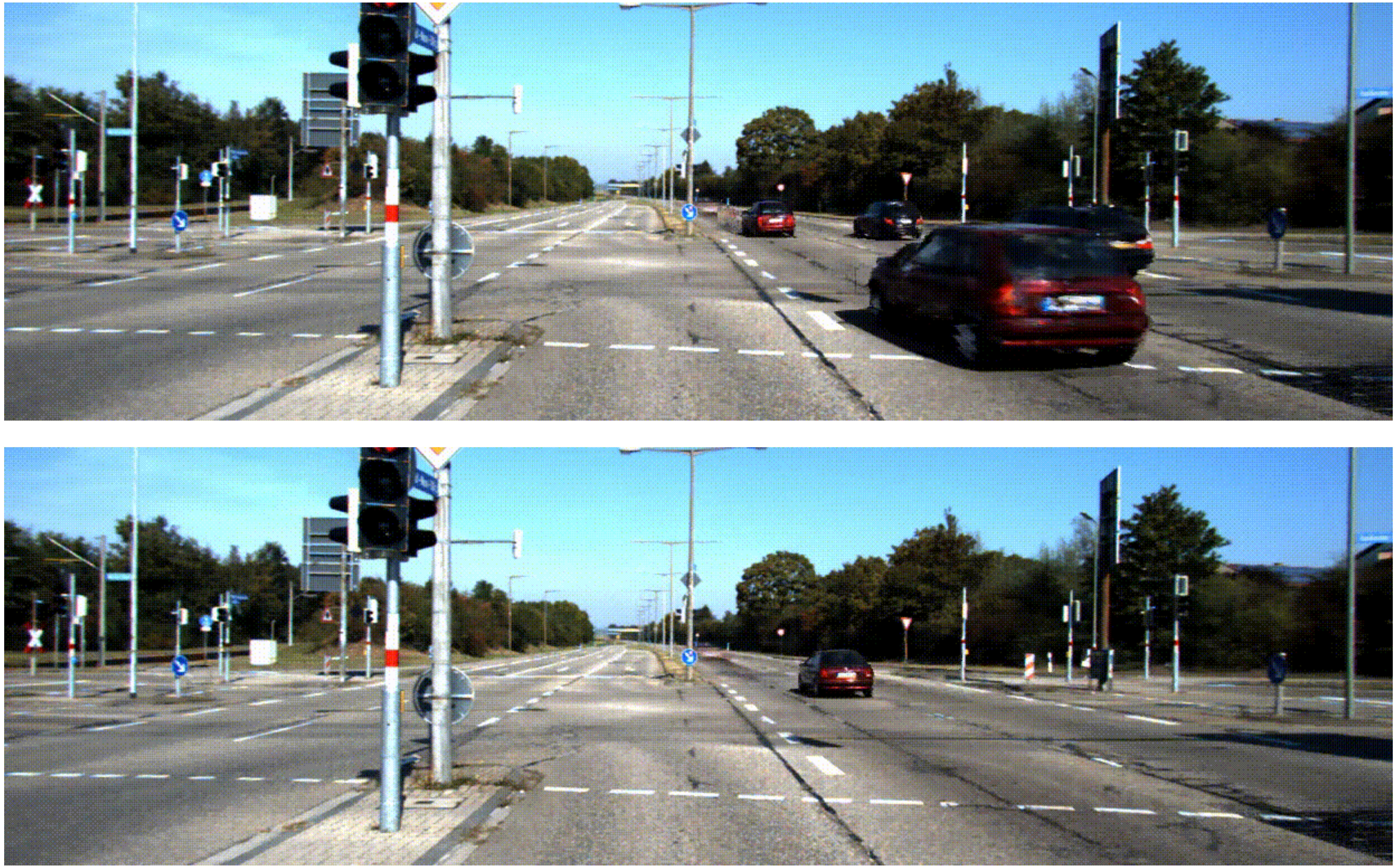
iMAP: Implicit Mapping and Positioning in Real-Time, Sucar et al. arXiv 2021

ObjectNeRF



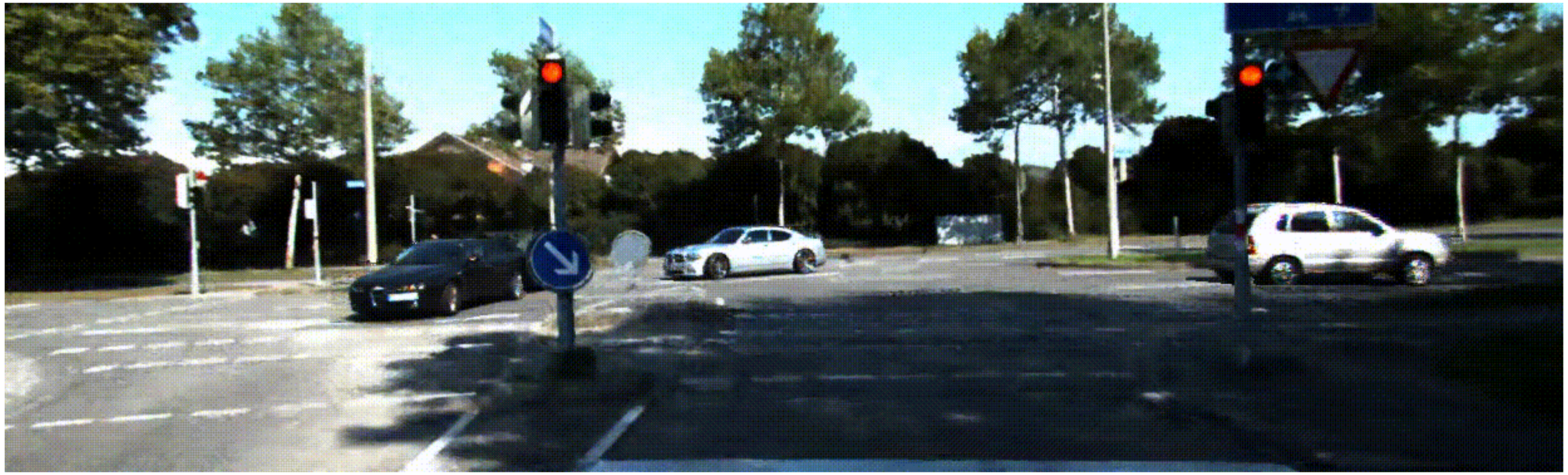
Learning Object-Compositional Neural Radiance Field for Editable Scene Rendering, Yang et al. ICCV 2022

Neural Scene Graphs for Dynamic Scenes



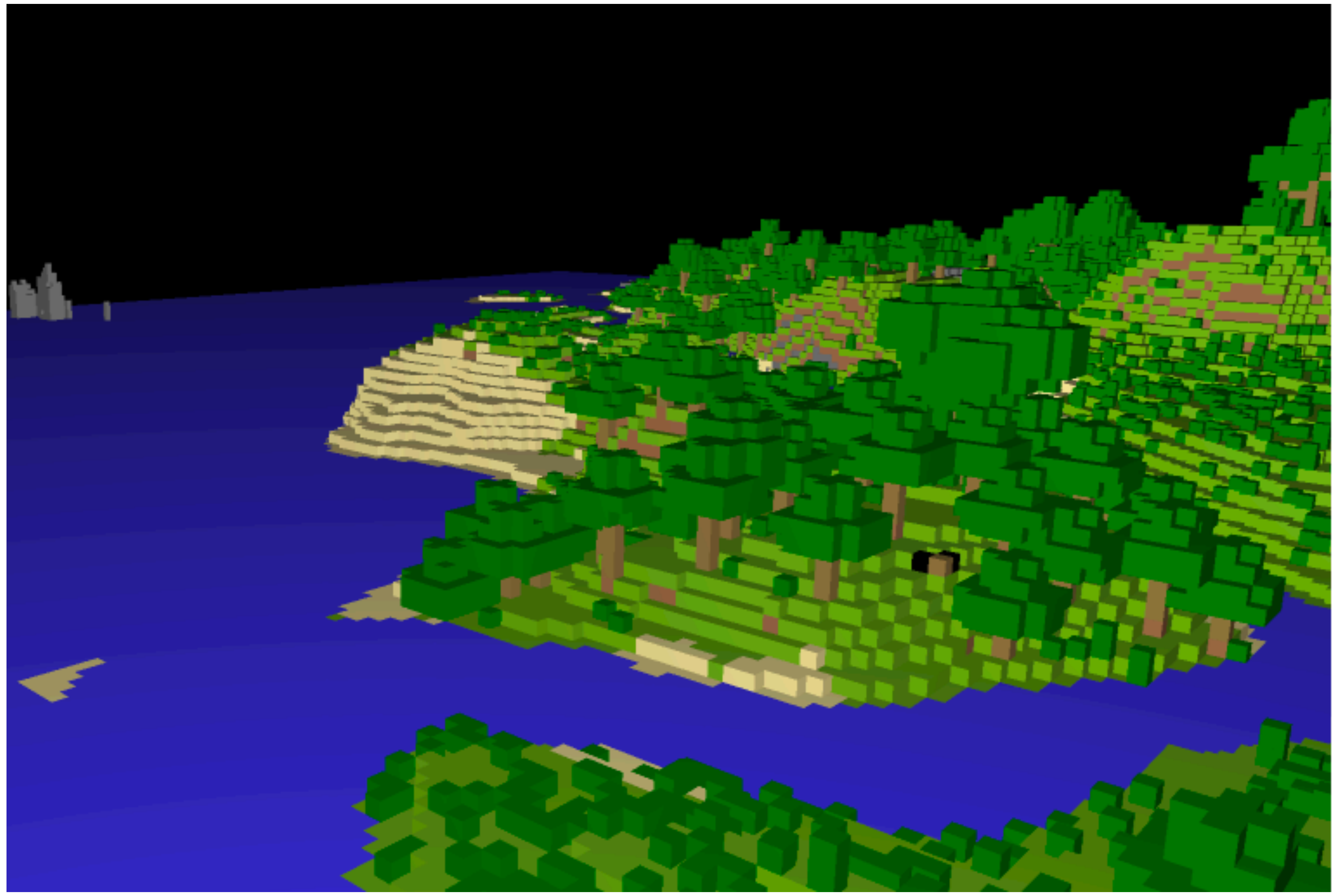
Neural Scene Graphs for Dynamic Scenes, Ost et al. CVPR 2021

Neural Scene Graphs for Dynamic Scenes



Neural Scene Graphs for Dynamic Scenes, Ost et al. CVPR 2021

GANcraft



Minecraft

GANcraft

GANcraft: Unsupervised 3D Neural Rendering of Minecraft Worlds, Hao et al. arXiv 2021

GANcraft

GANcraft: Unsupervised 3D Neural Rendering of Minecraft Worlds, Hao et al. arXiv 2021

GANcraft: Unsupervised 3D Neural Rendering of Minecraft Worlds, Hao et al. arXiv 2021