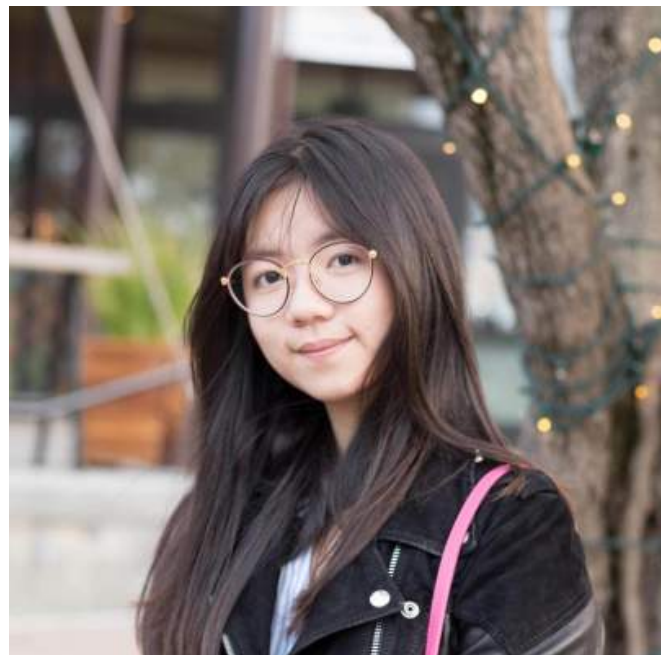# Space-time Neural Irradiance Fields for Free-Viewpoint Video
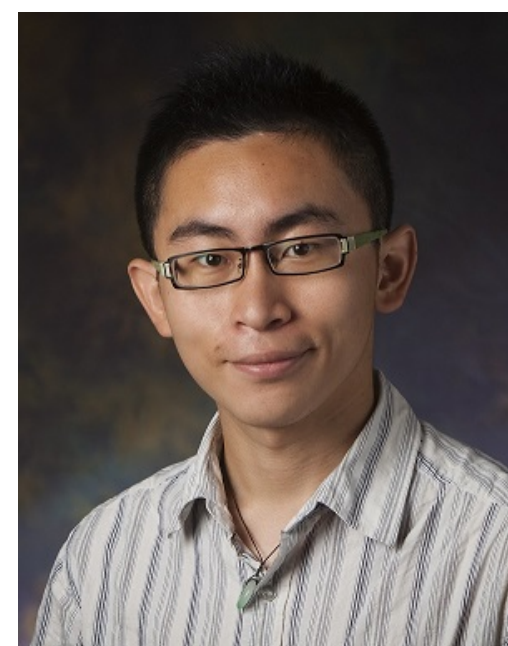
CVPR 2021

Wenqi Xian

Cornell Tech

Jia-Bin Huang

Virginia Tech

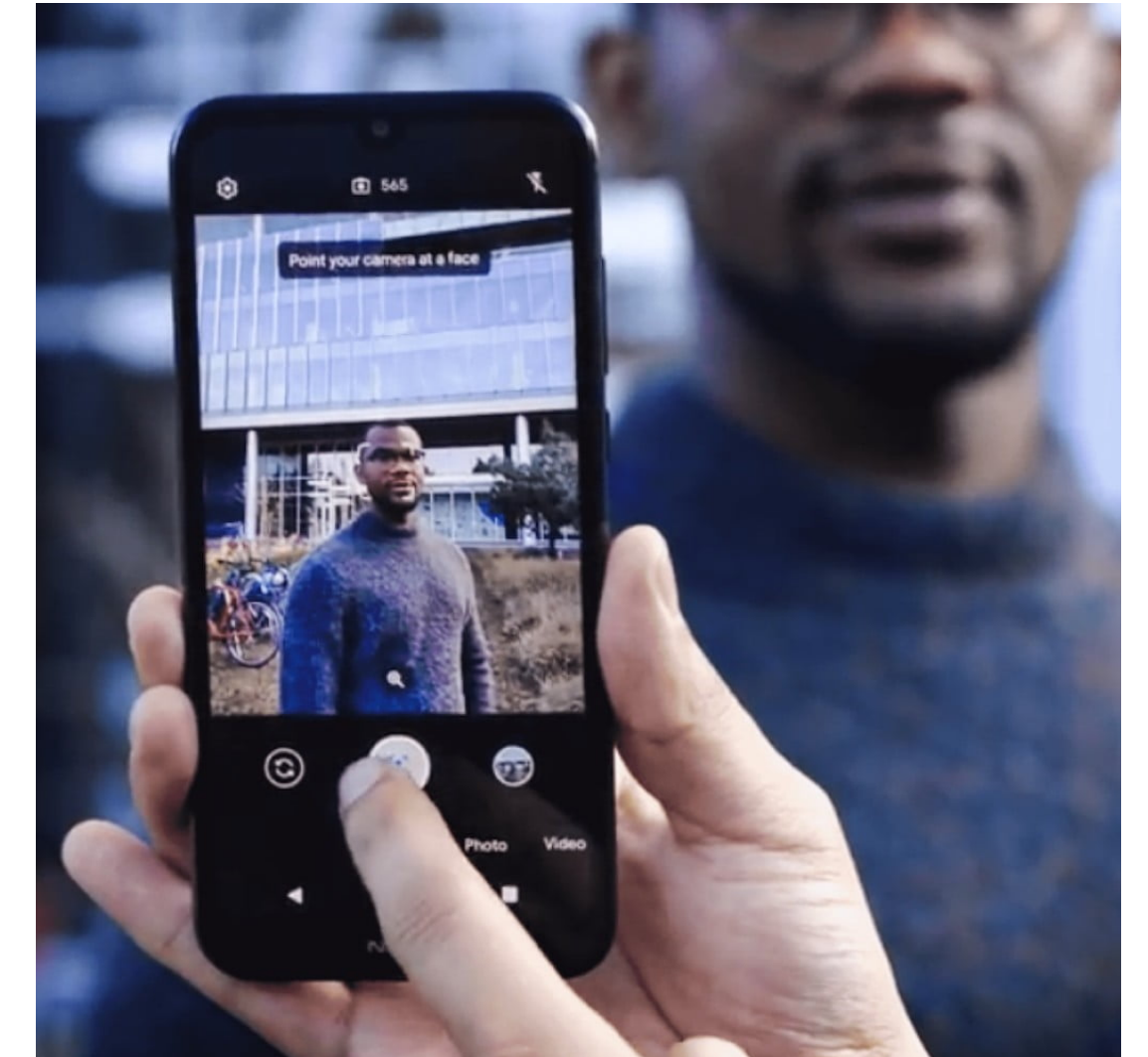**Johannes Kopf**
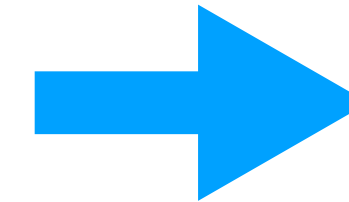
Facebook

Changil Kim

Facebook

Free-Viewpoint Video
Motivation

Multi-camera Rig

3D video on smartphones

How to achieve that from casual videos shot by **smartphones**?

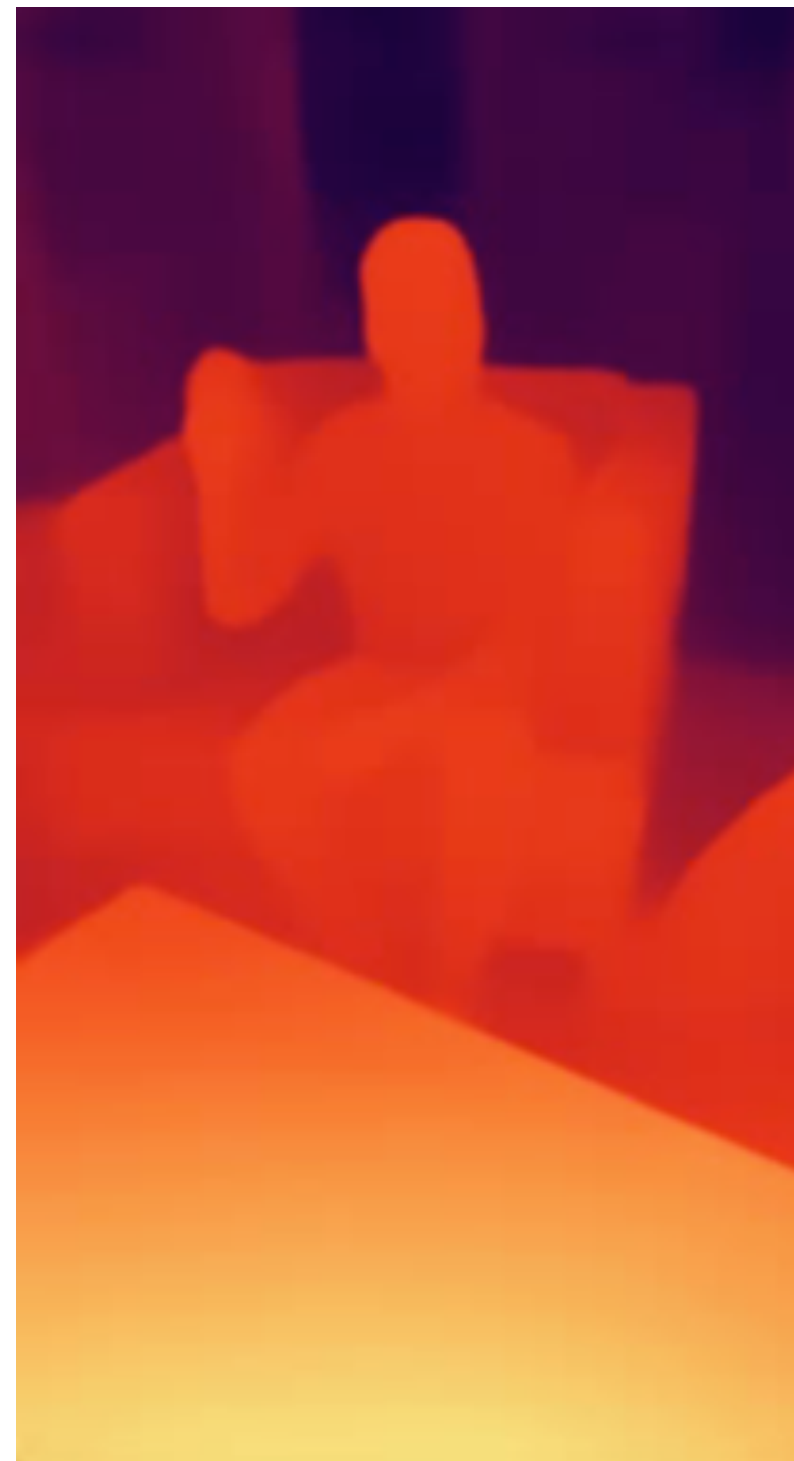*Broxton et al. Immersive Light Field Video with a Layered Mesh Representation, SIGGRAPH 2020.*

Free-Viewpoint Video
Capture scene dynamics and consistent 3D structure.

Input

CVD →

Depth prediction

Warp →

Mesh from novel views

Luo et al. Consistent Video Depth Estimation, SIGGRAPH 2020.
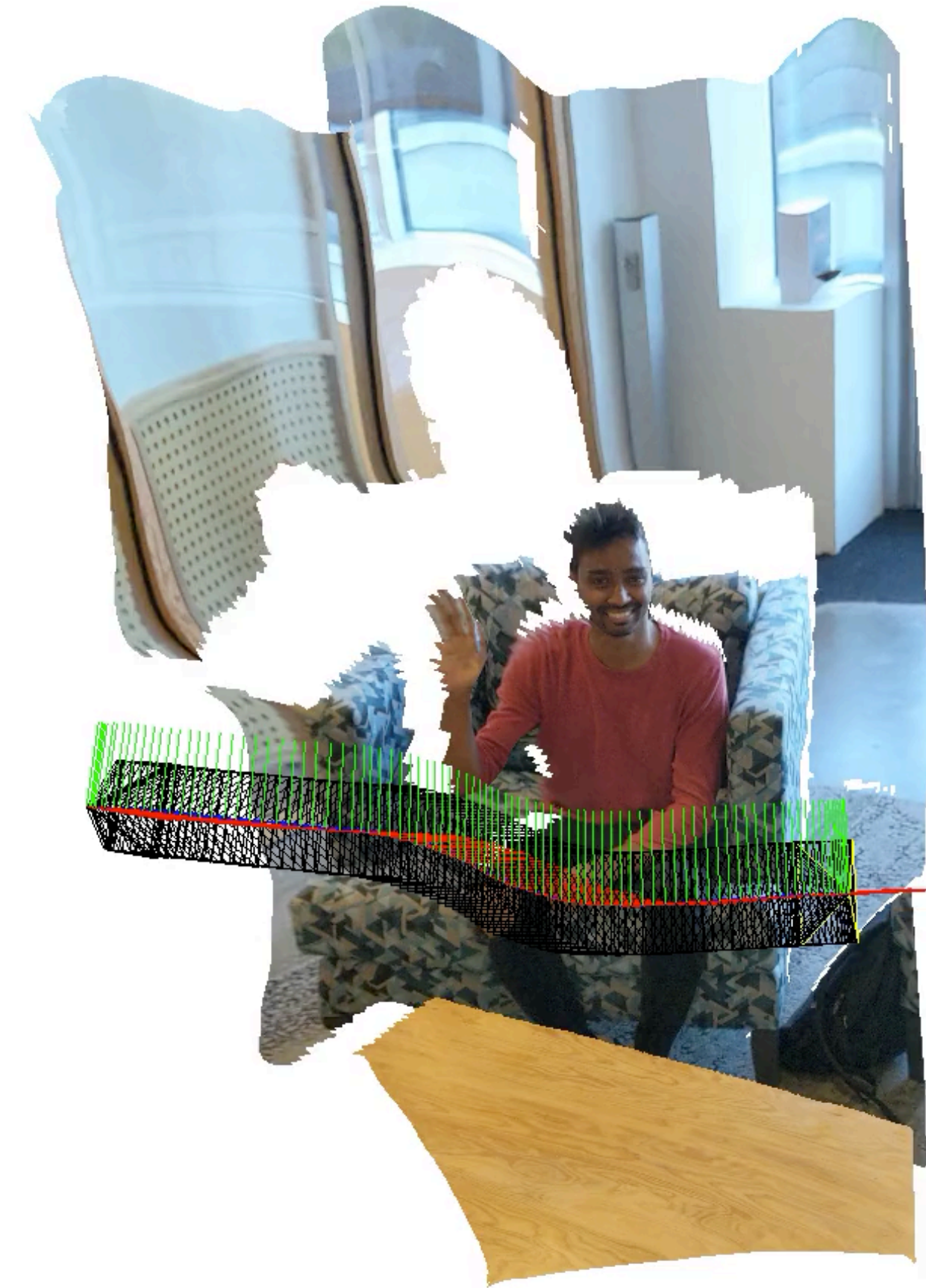
Free-Viewpoint Video

Input Video   **Mesh**   **Inpainted Mesh**

Inpaint →

Gao et al. Flow-edge Guided Video Completion, ECCV 2020.

Implicit representation
NeRF (Neural Radiance Fields)

Mildenhall et al. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, ECCV 2020.
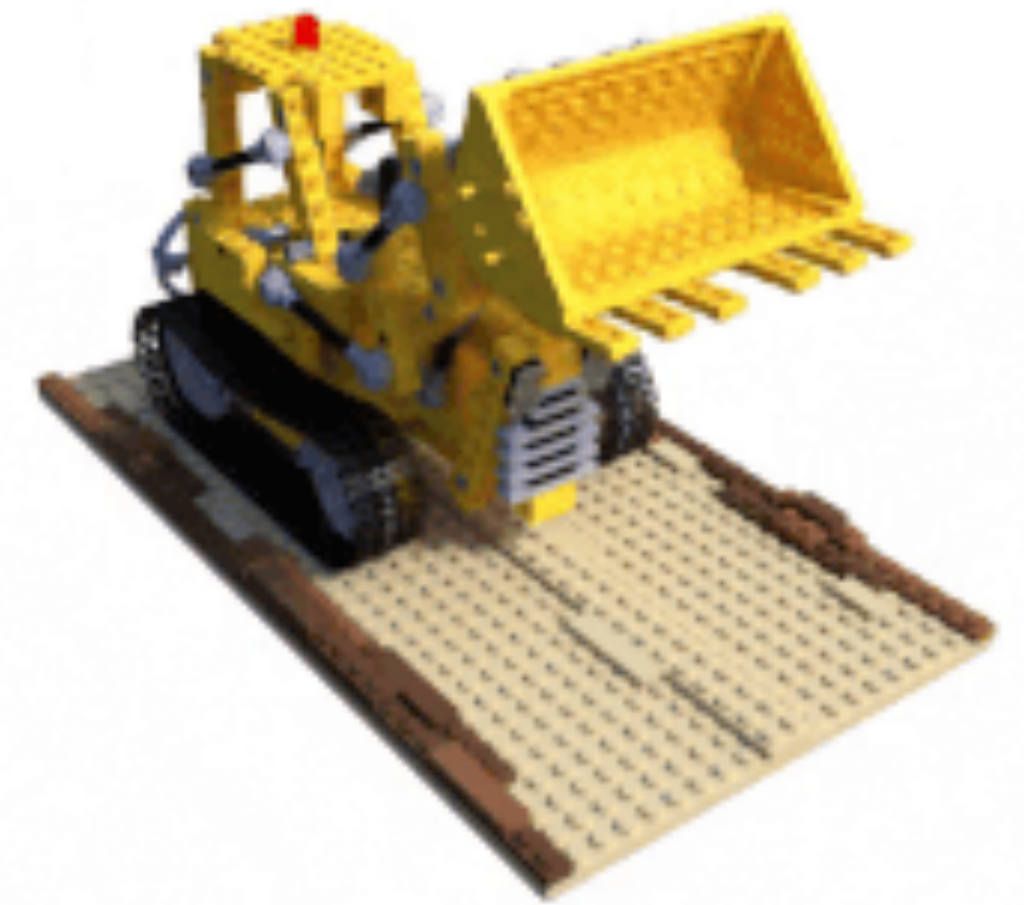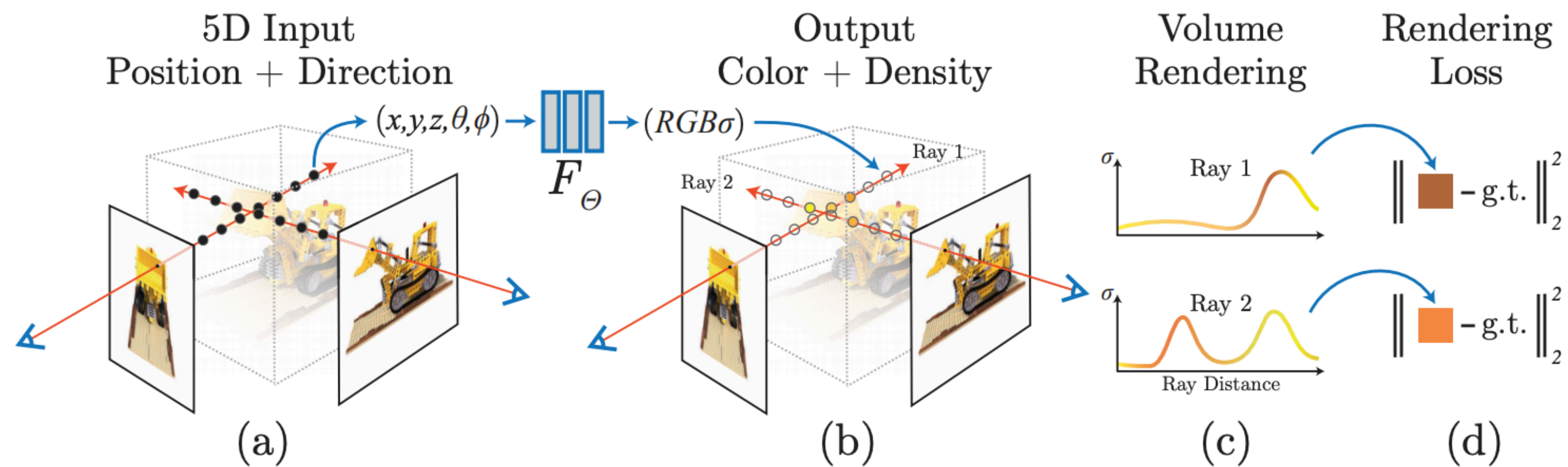
Extension to 3D video
NeRF

5D Input:
Position + Direction

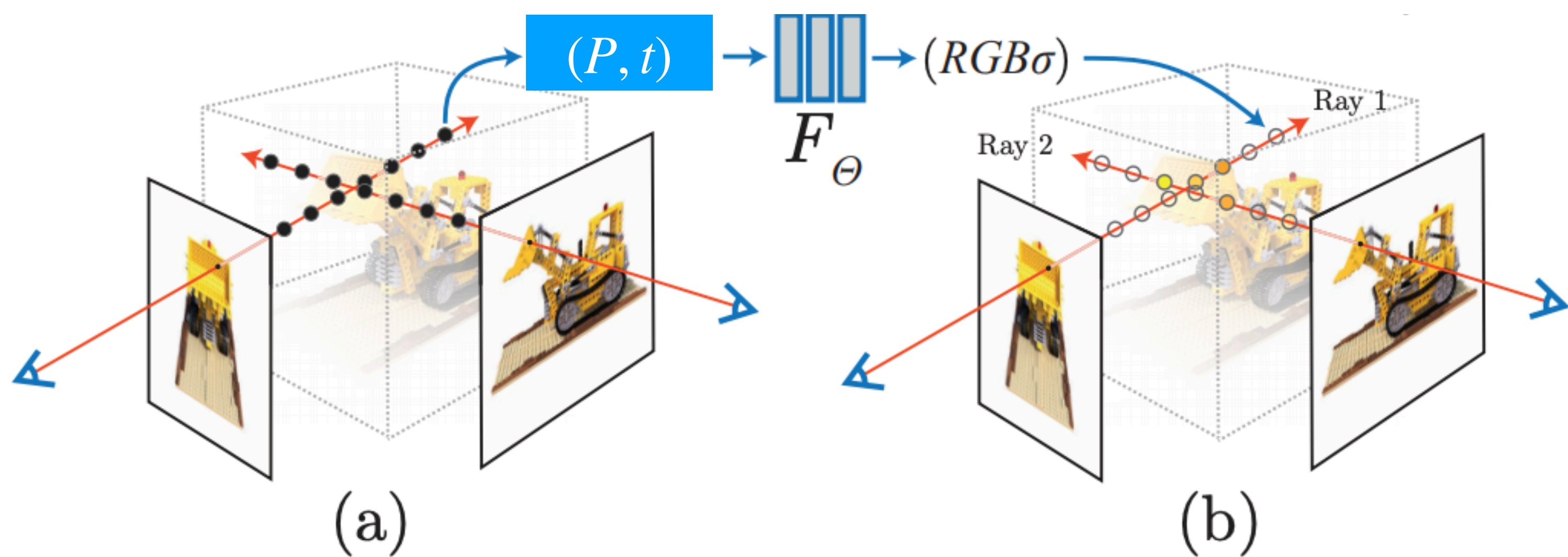$$P = (x, y, z) \qquad v = (\theta, \phi)$$

Output:
Color + Density

(a)

$(P, v) \rightarrow F_\Theta \rightarrow (RGB\sigma)$

Ray 1

Ray 2

(b)
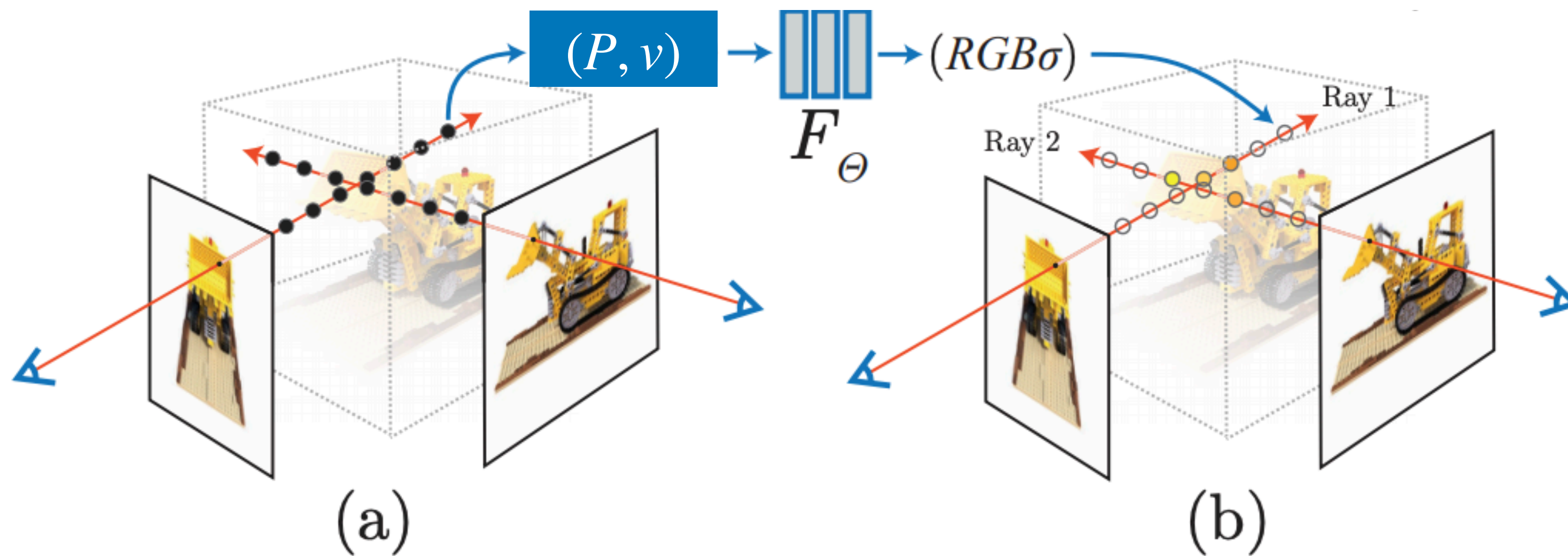
Novel view render

Extension to 3D video
NeRF + time dimension

4D Input:
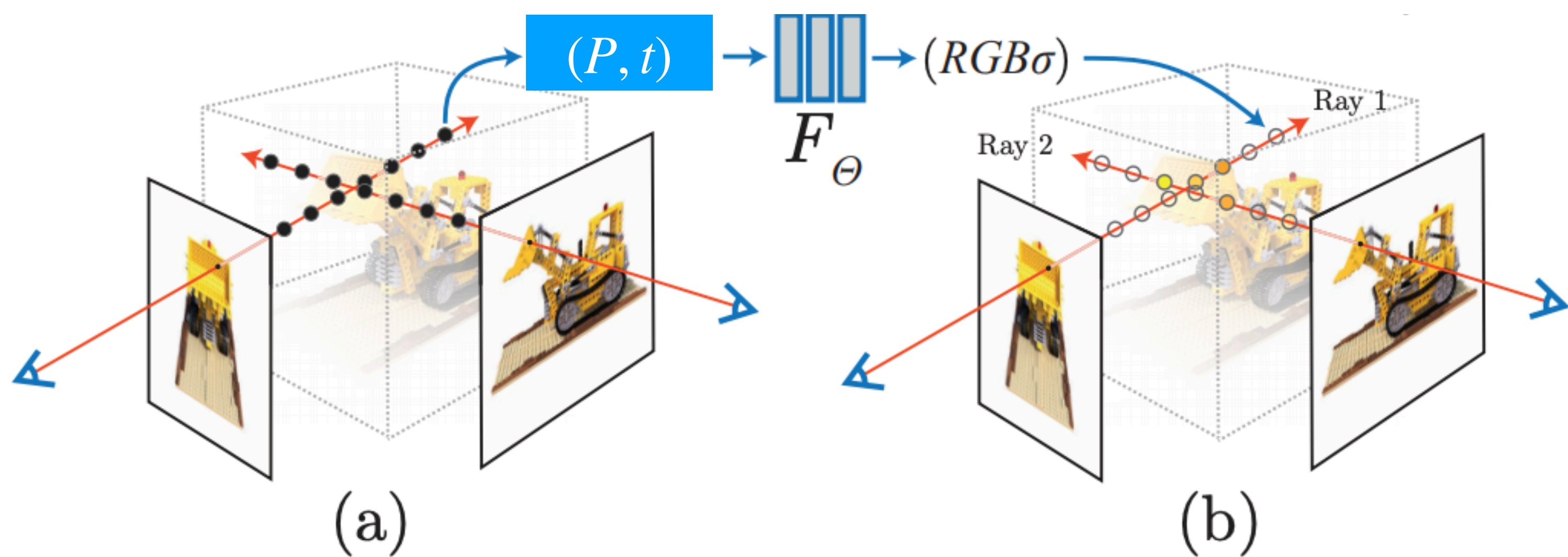Position + Time

Output:
Color + Density

Novel view render

$(P, t) \rightarrow F_\Theta \rightarrow (RGB\sigma)$

Ray 1

Ray 2

(a)

(b)

Extension to 3D video

NeRF + time dimension + geometry constraints

Novel view render
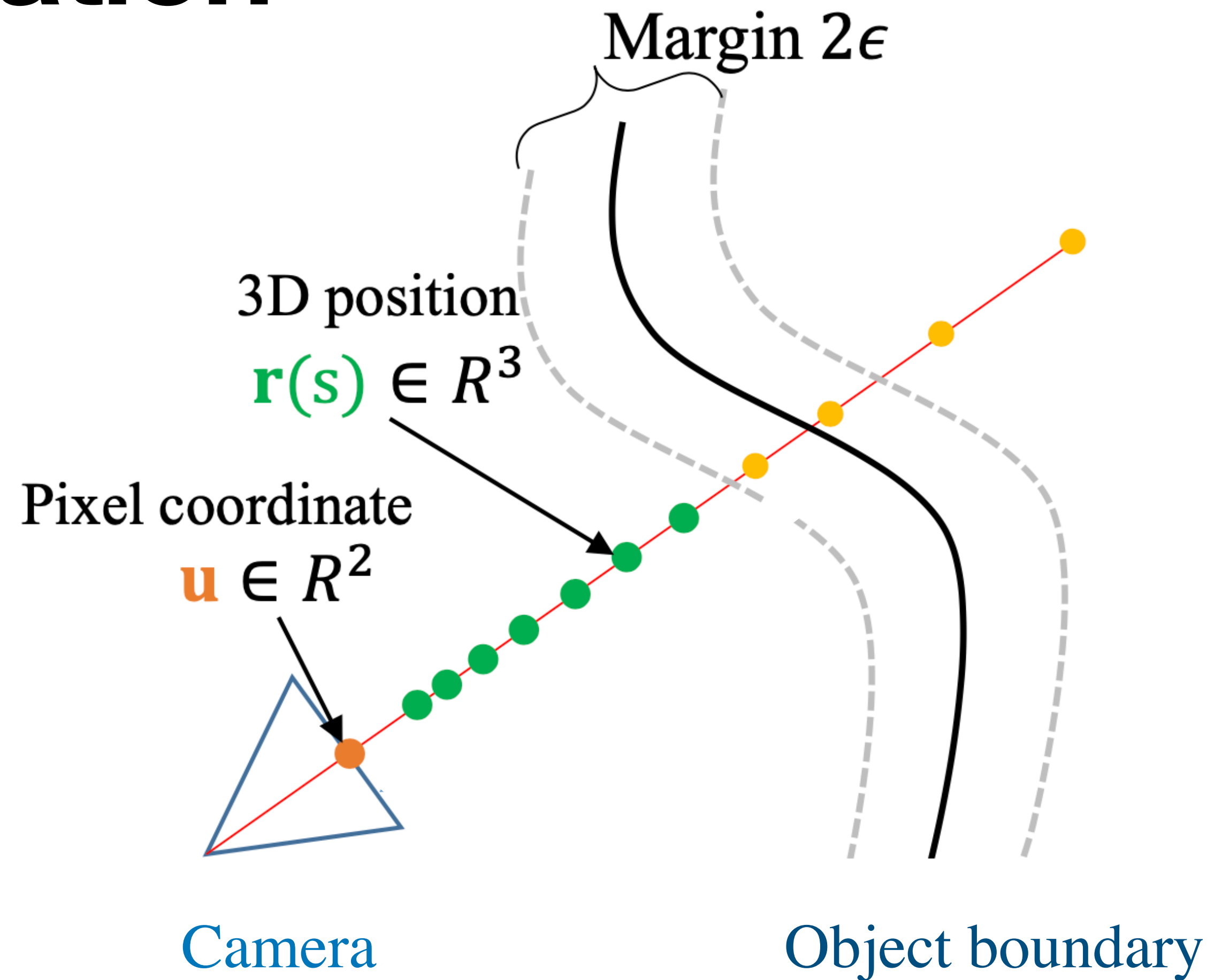
4D Input:
Position + Time

Output:
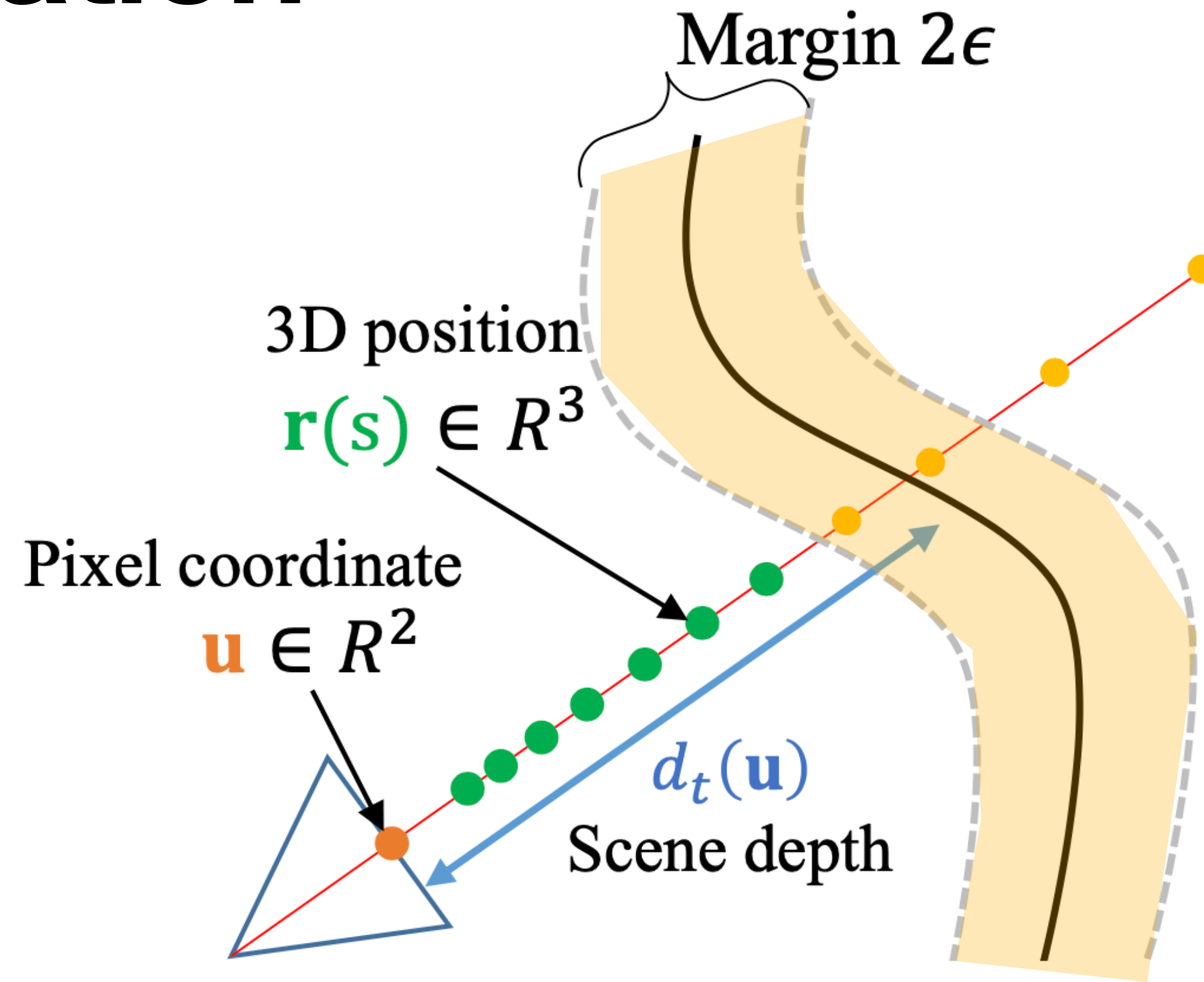Color + Density

Proposed regularization

Depth Loss

Disambiguate time-varying
geometry and appearance

Proposed regularization

Depth Loss

Disambiguate time-varying geometry and appearance

Margin 2ϵ

3D position
$\mathbf{r}(s) \in R^3$

Pixel coordinate
$\mathbf{u} \in R^2$

$d_t(\mathbf{u})$
Scene depth

Ablation

Depth Loss

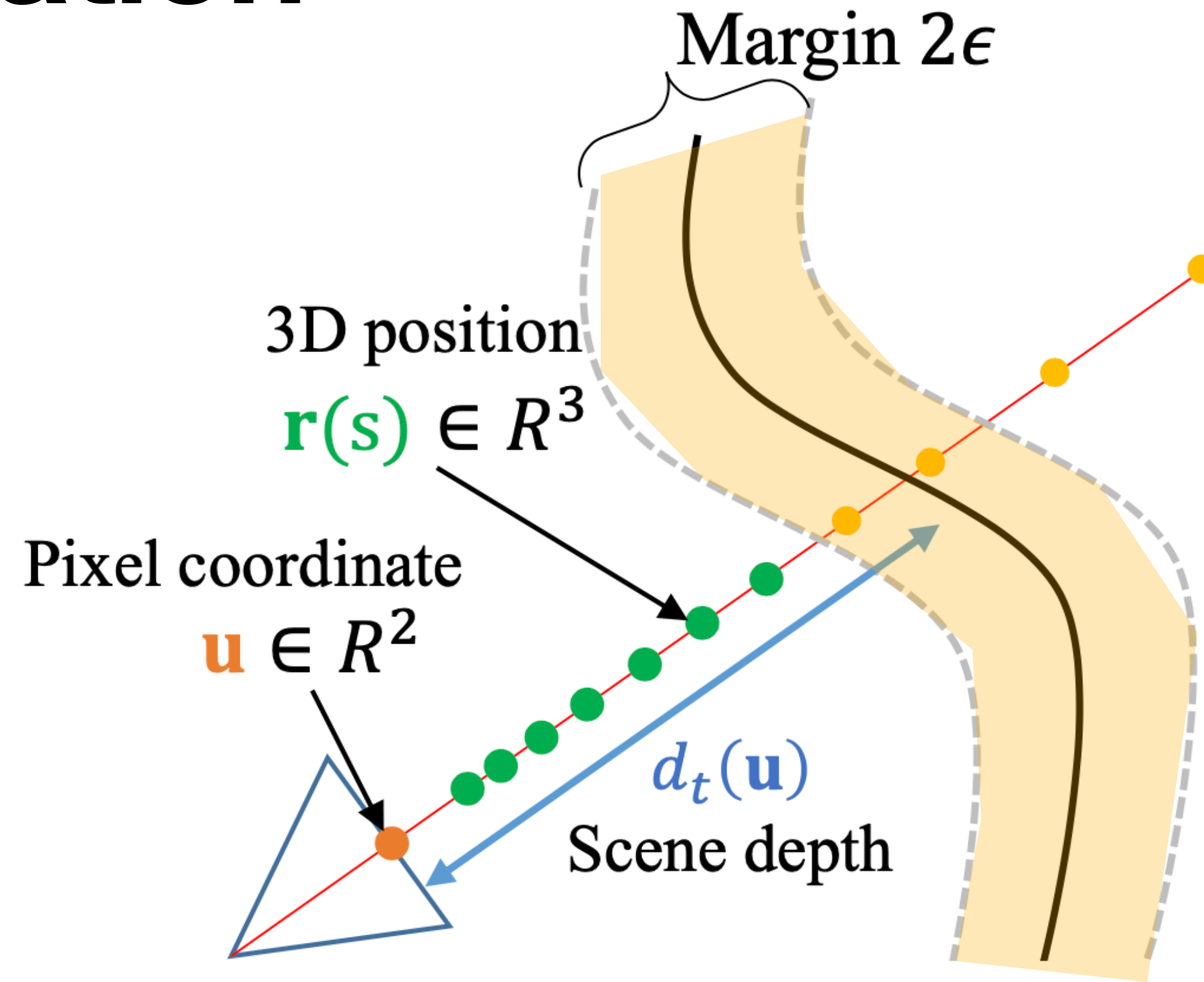Disambiguate time-varying geometry and appearance

Input Novel view Novel view

(a) w/o depth loss (b) w/ depth loss

Proposed regularization

Depth Loss

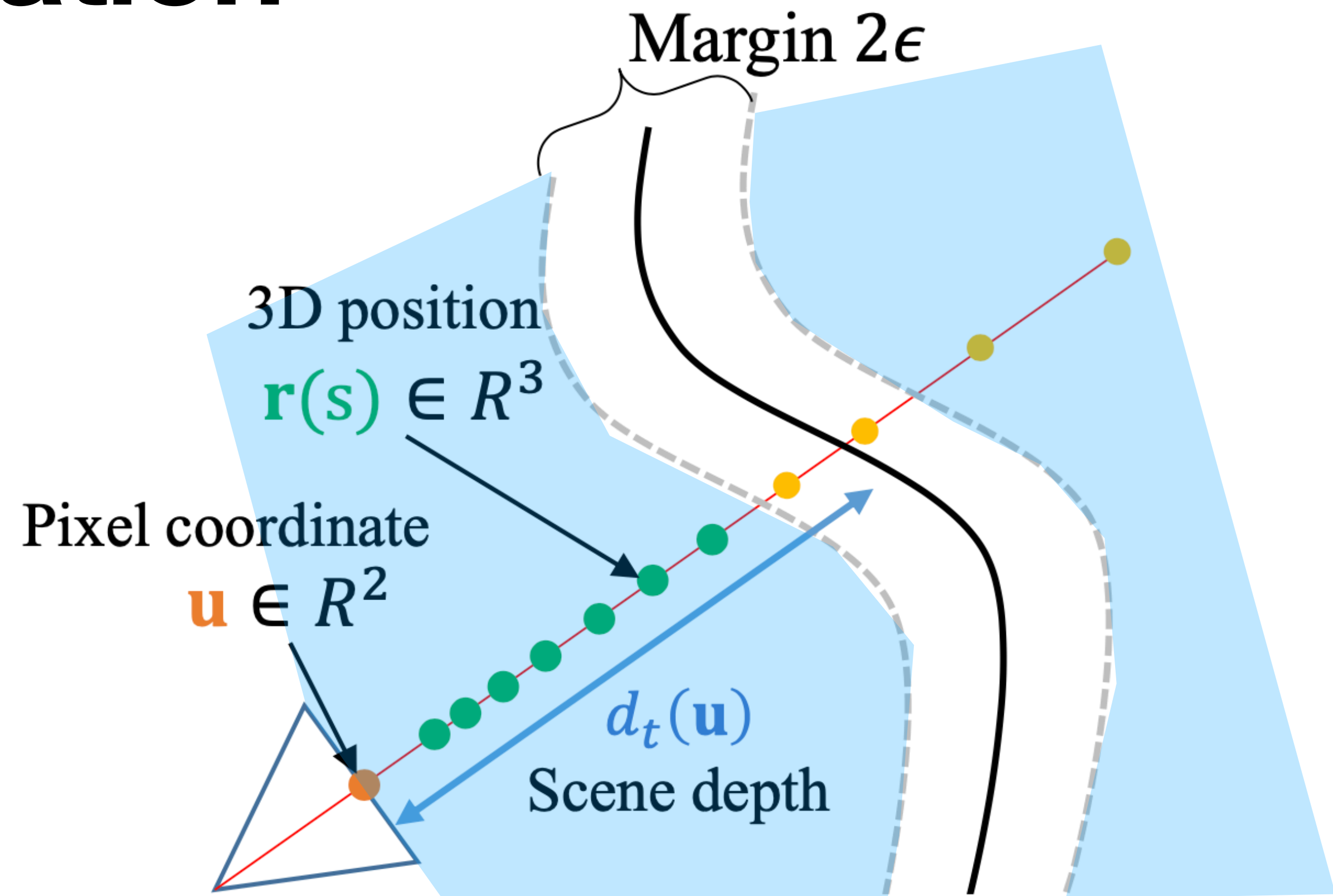Disambiguate time-varying
geometry and appearance

Margin 2ϵ

3D position
$\mathbf{r}(s) \in R^3$

Pixel coordinate
$\mathbf{u} \in R^2$

$d_t(\mathbf{u})$
Scene depth

Proposed regularization

Depth Loss

Disambiguate time-varying geometry and appearance

Static Loss
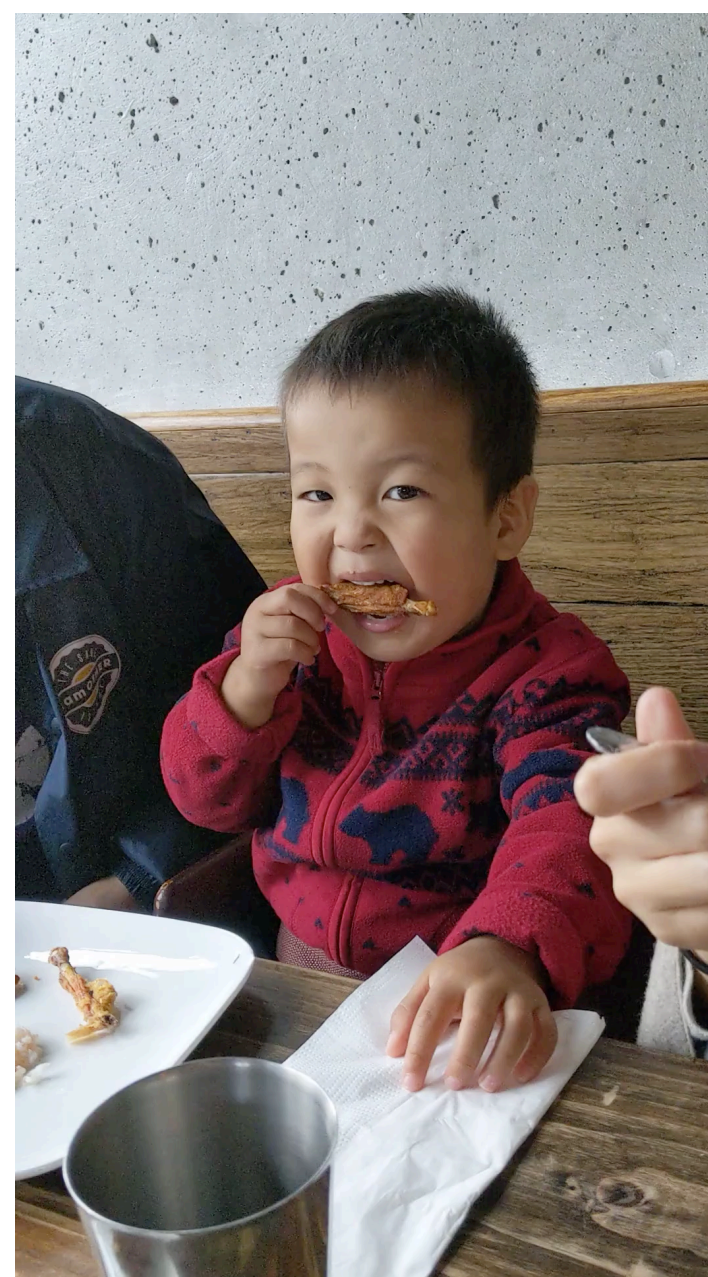
Propagate contents across time at invisible 3D location

Ablation

<u>Static Loss</u>

Propagate contents across time at invisible 3D location

input   (a) w/o static loss   (b) w/ static loss
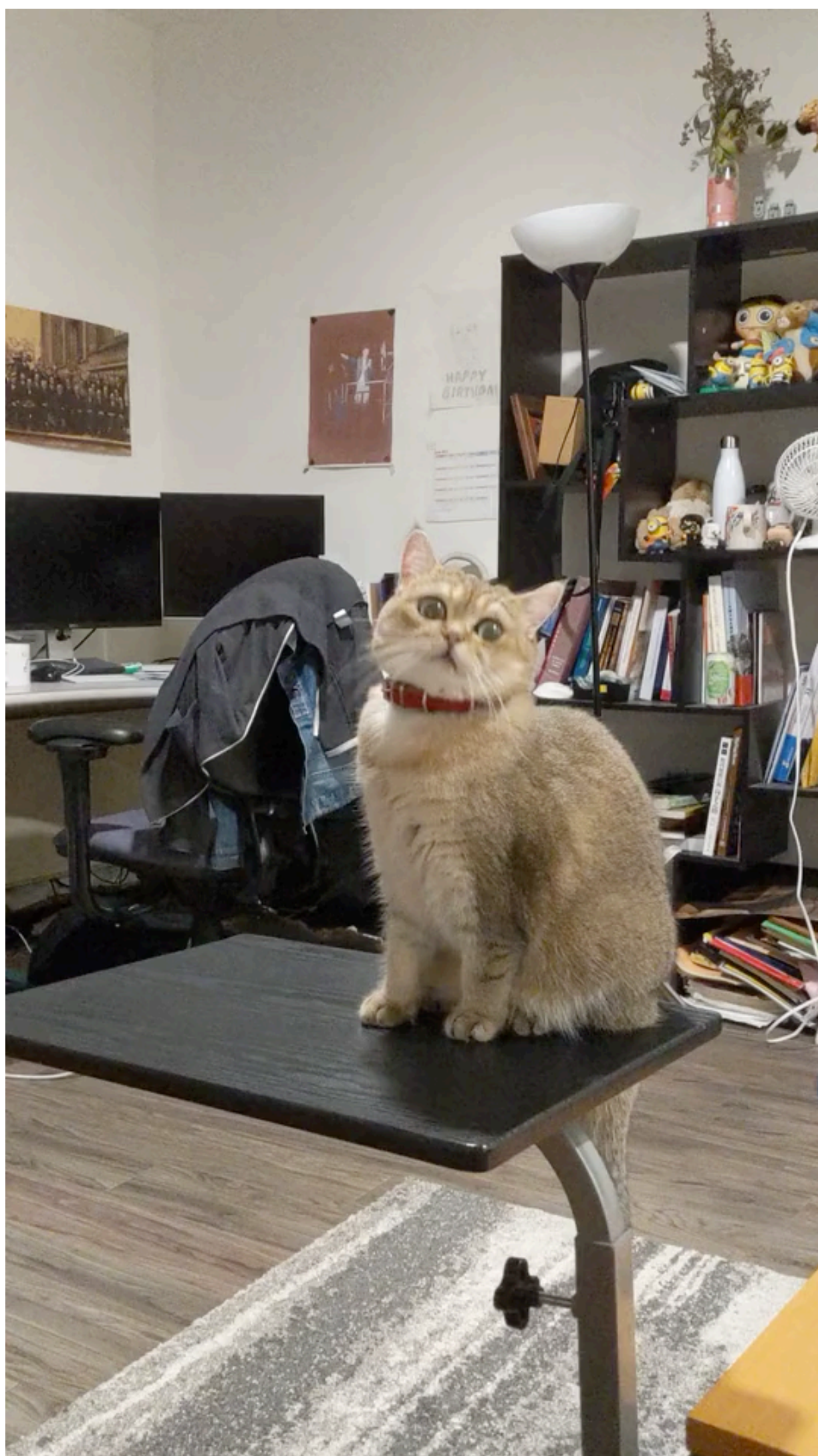
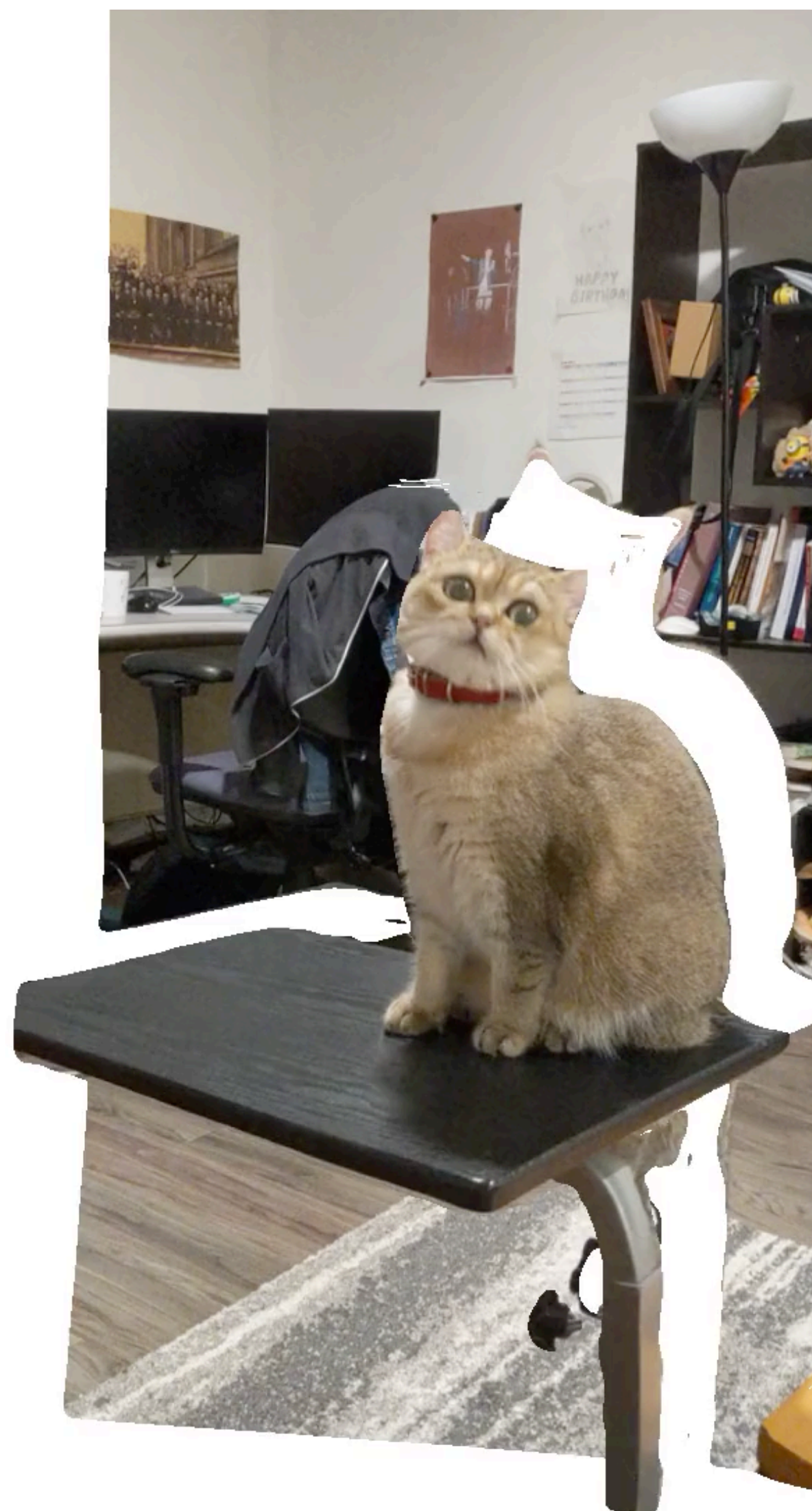**Input Video** | **Mesh** | **Inpainted Mesh**
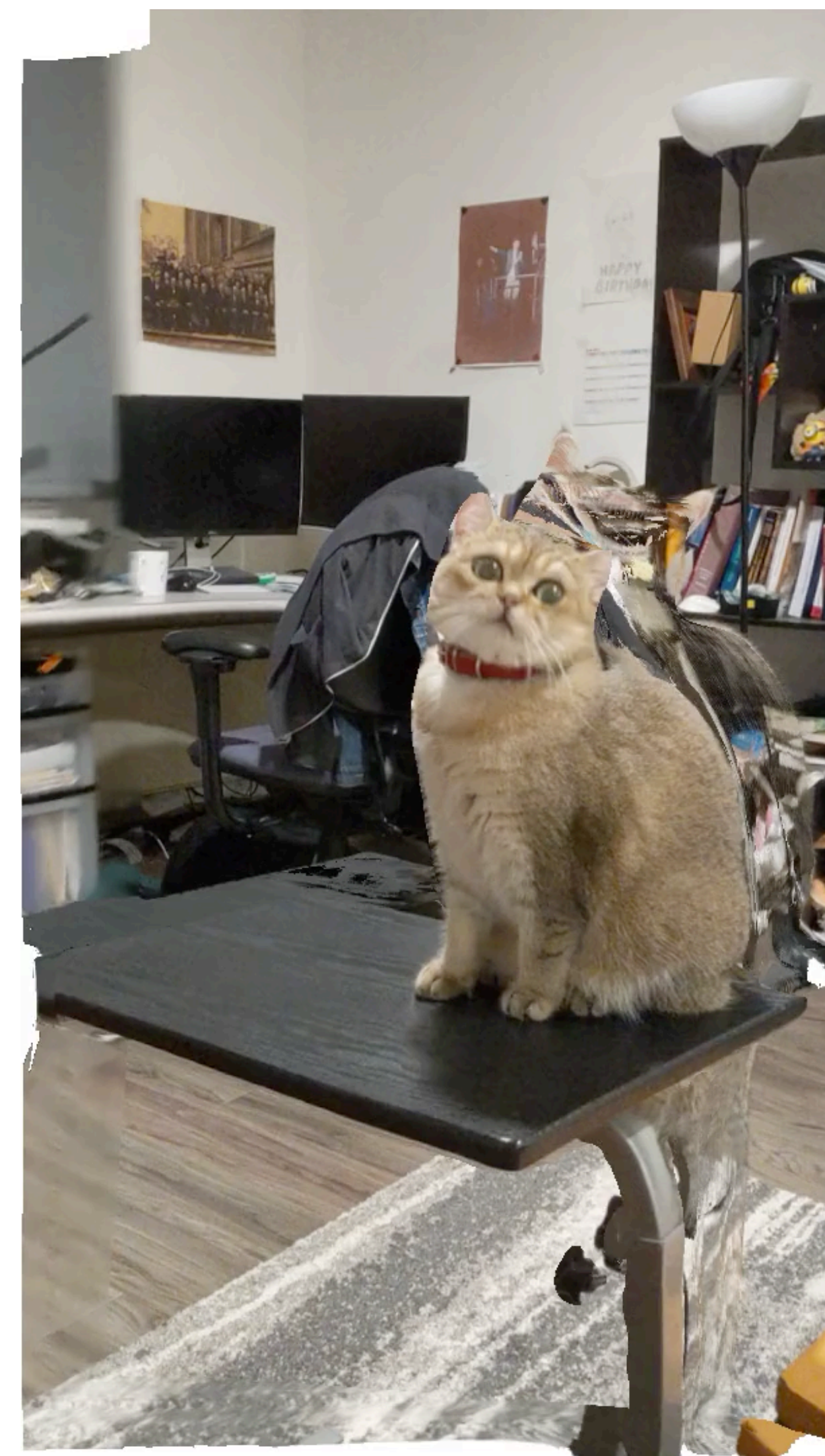
Input Video　　　**NeRF + time**　　　**Ours**

Input Video | **Mesh** | **Inpainted Mesh**

Input Video **NeRF + time** **Ours**

More results

https://video-nerf.github.io
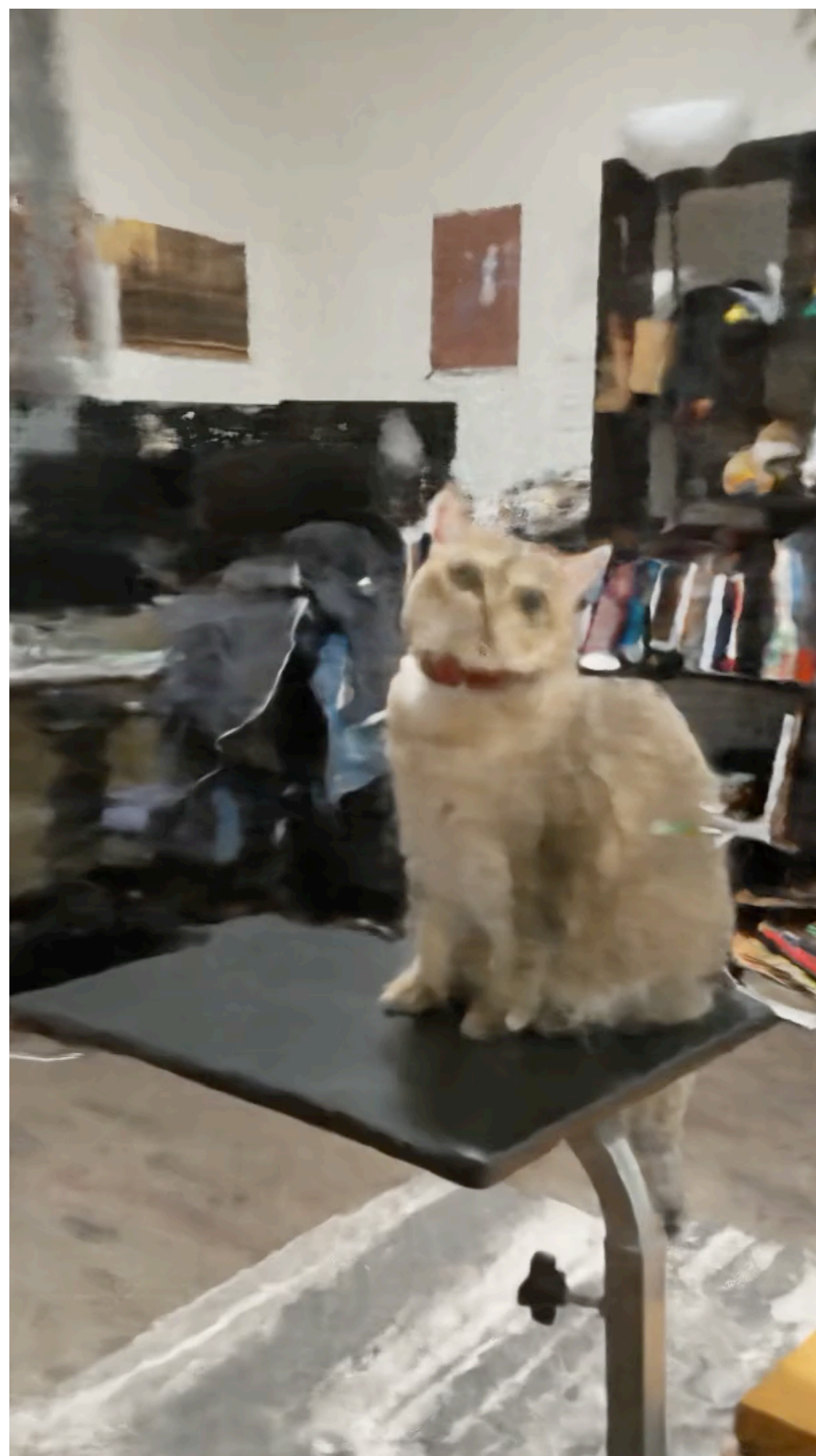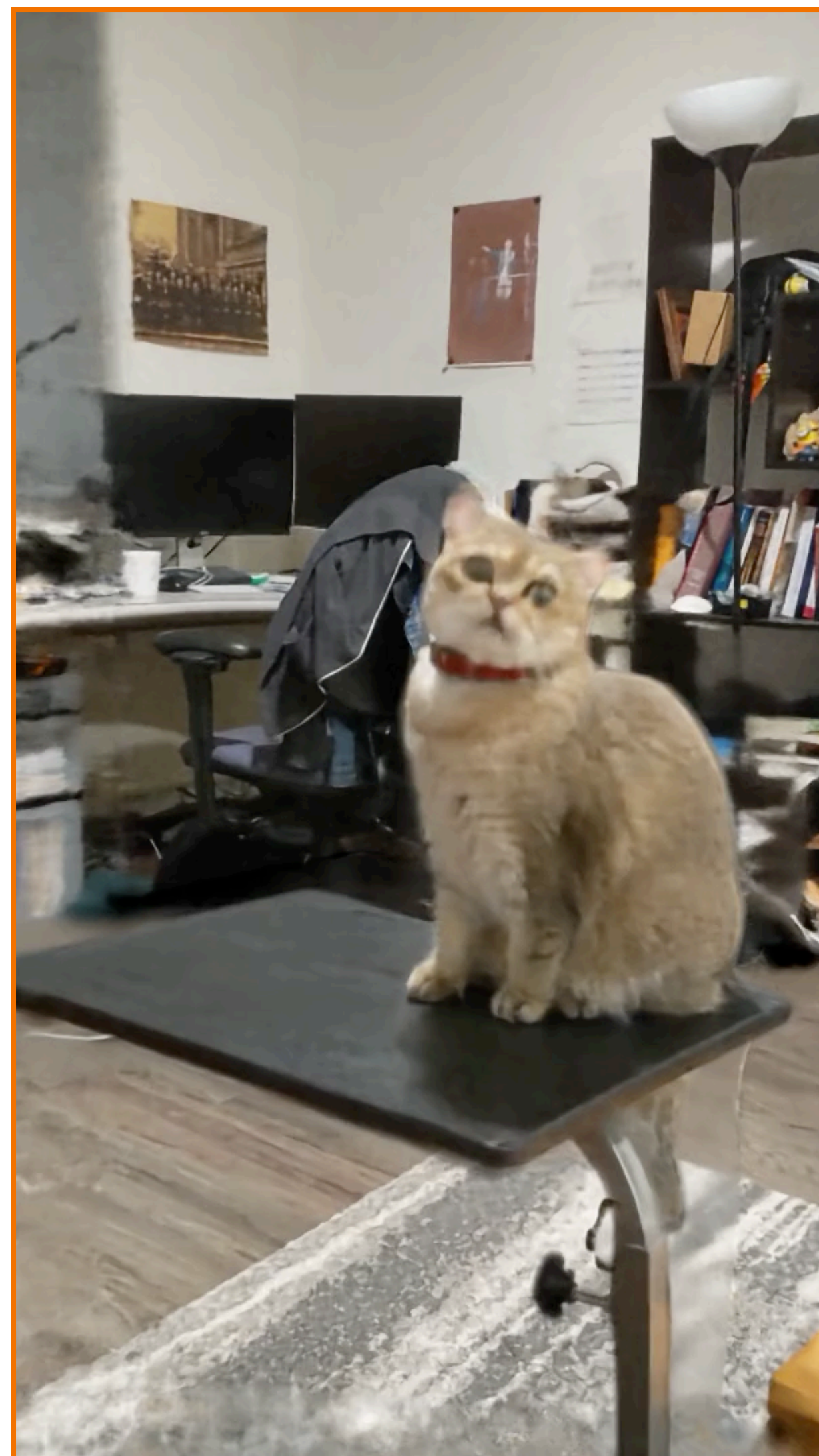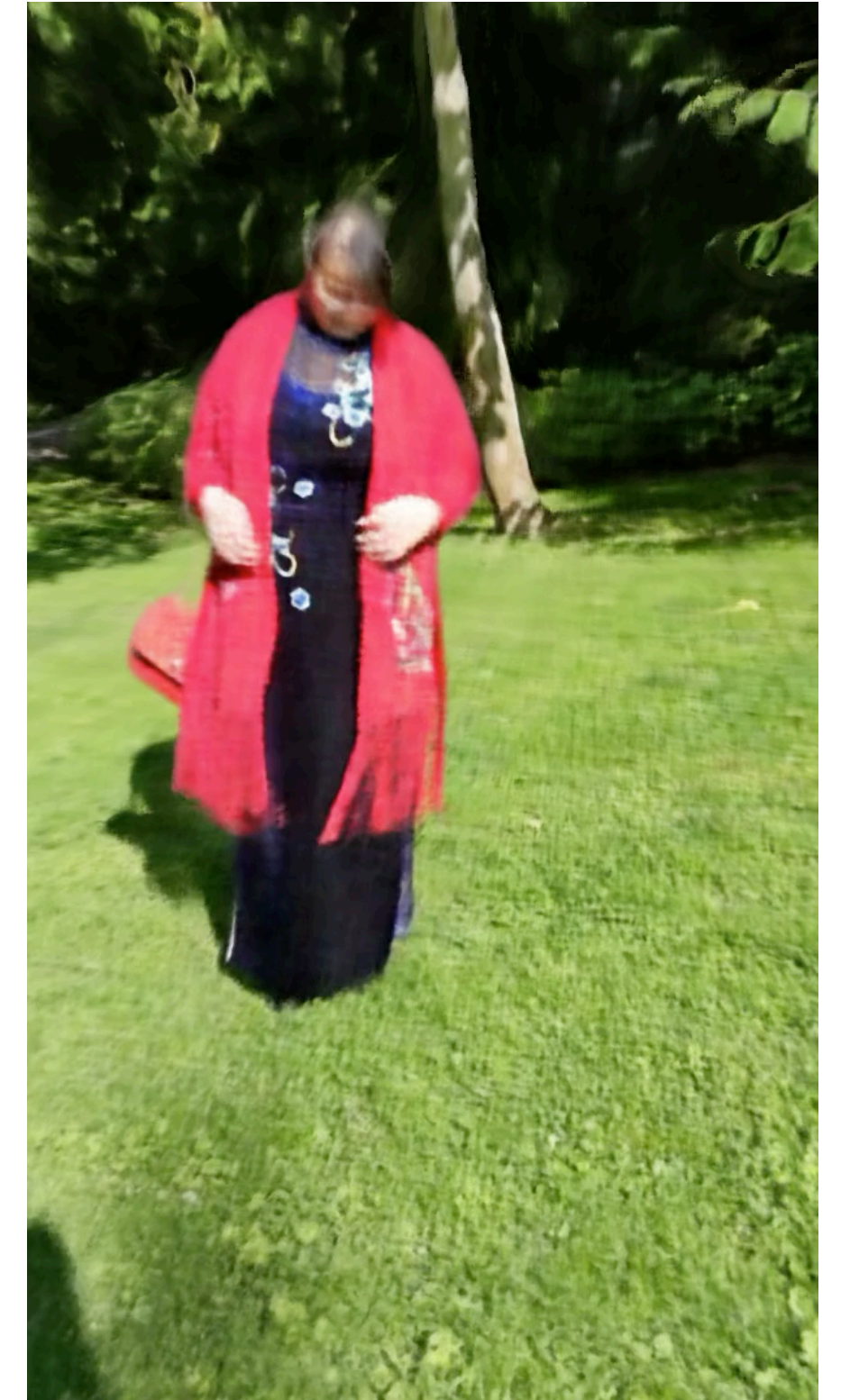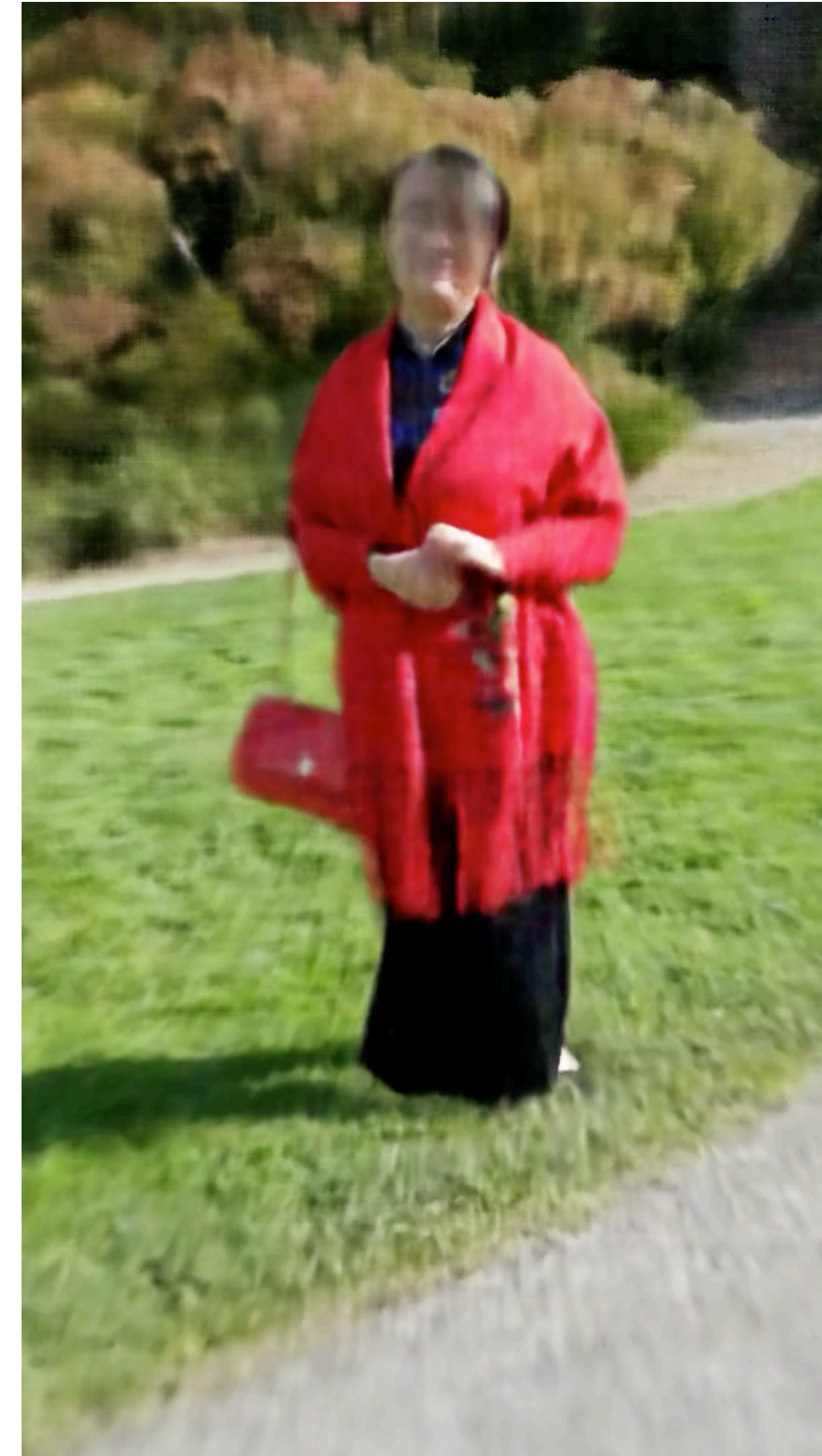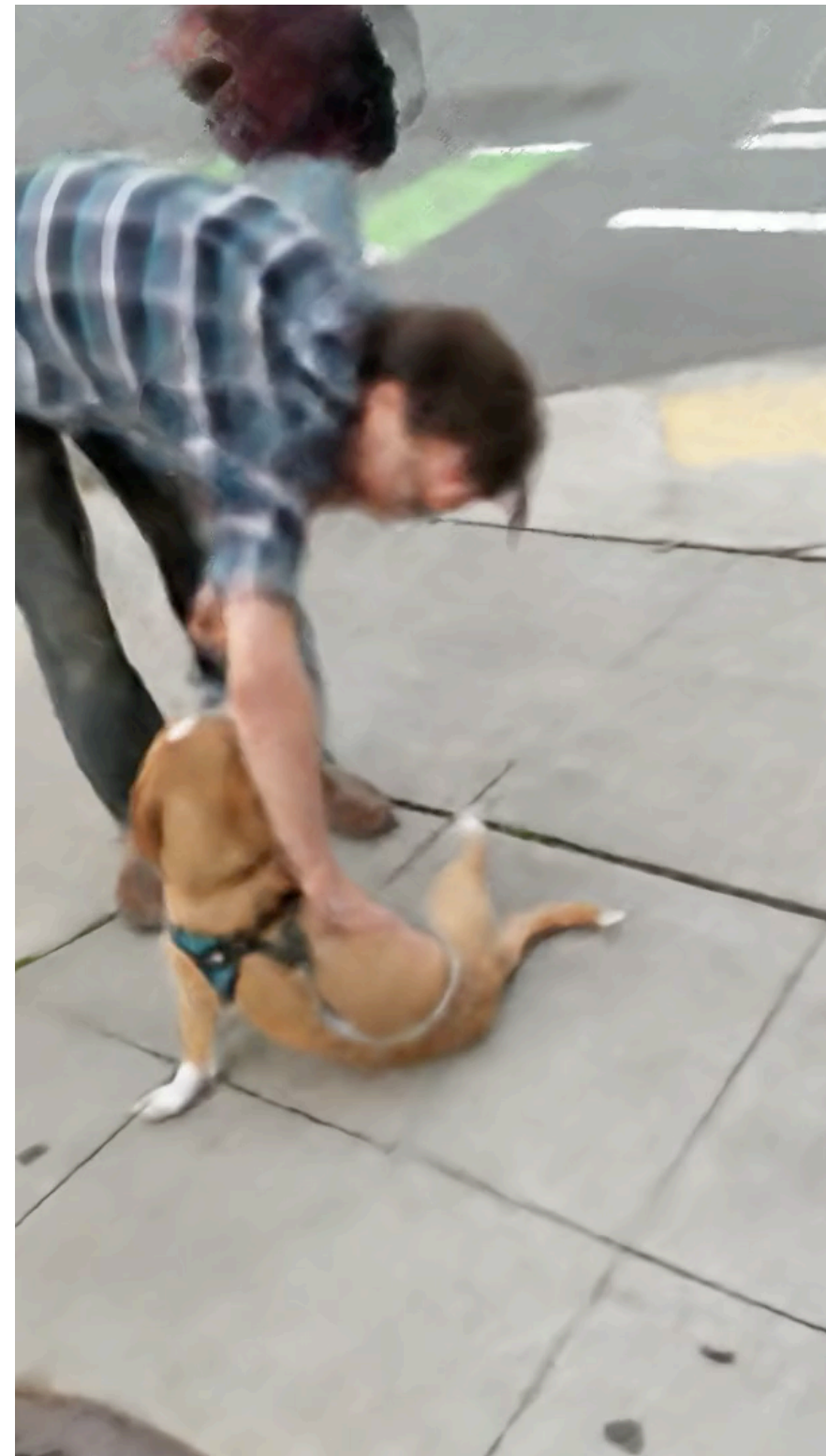
Concurrent works 2020

	Nerfie[1]	DNeRF[2]	NR-NeRF[4]	NSFF[3]	VideoNeRF[6]	N3VS[5]	STaR[7]	NeRFlow[8]
time encoding	latent	t-index	latent	t-index	t-index	latent	N/A	t-index
Input	1 vpf	1 vpf	1 vpf	1 vpf, opt flow, depth	1 vpf, depth	N vpf	N vpf	1 vpf, opt flow
loss	color (fg, bg)	color	color	color, flow, depth	color, depth, static	color	color	color, flow
deformation representation	local SE(3)	offset	offset	scene flow	N/A	N/A	global SE(3) (rigid)	flow (velocity)
FG/BG Separation	No	No	Yes	Yes	No	No	Yes	No
regularizer	smoothness, bg reg	N/A	def sparsity, smoothness	cycle flow, occlusion	N/A	N/A	sigma sparsity, no overlap	flow sparsity, smoothness
quality(1-5)	5	5	5	4	4	4	3	2
scene	(close)	(close)	open	open	open	open	open	open
real/synth	real	synth	real	real	real	real	real/synth	synth

[1] Deformable Neural Radiance Fields, Park et al., Arxiv 2020
[2] D-NeRF: Neural Radiance Fields for Dynamic Scenes, Pumarola et al., Arxiv 2020
[3] Neural Scene Flow Fields for Space-Time View Synthesis of Dynamic Scenes, Li et al., Arxiv 2020
[4] Non-Rigid Neural Radiance Fields: Reconstruction and Novel View Synthesis of a Deforming Scene from Monocular Video, Tretschk et al., Arxiv 2020

[5] Neural 3D Video Synthesis, Li et al. SIGGRAPH submission
[6] Space-time Neural Irradiance Fields for Free-Viewpoint Video, Xian et al., Arxiv 2020
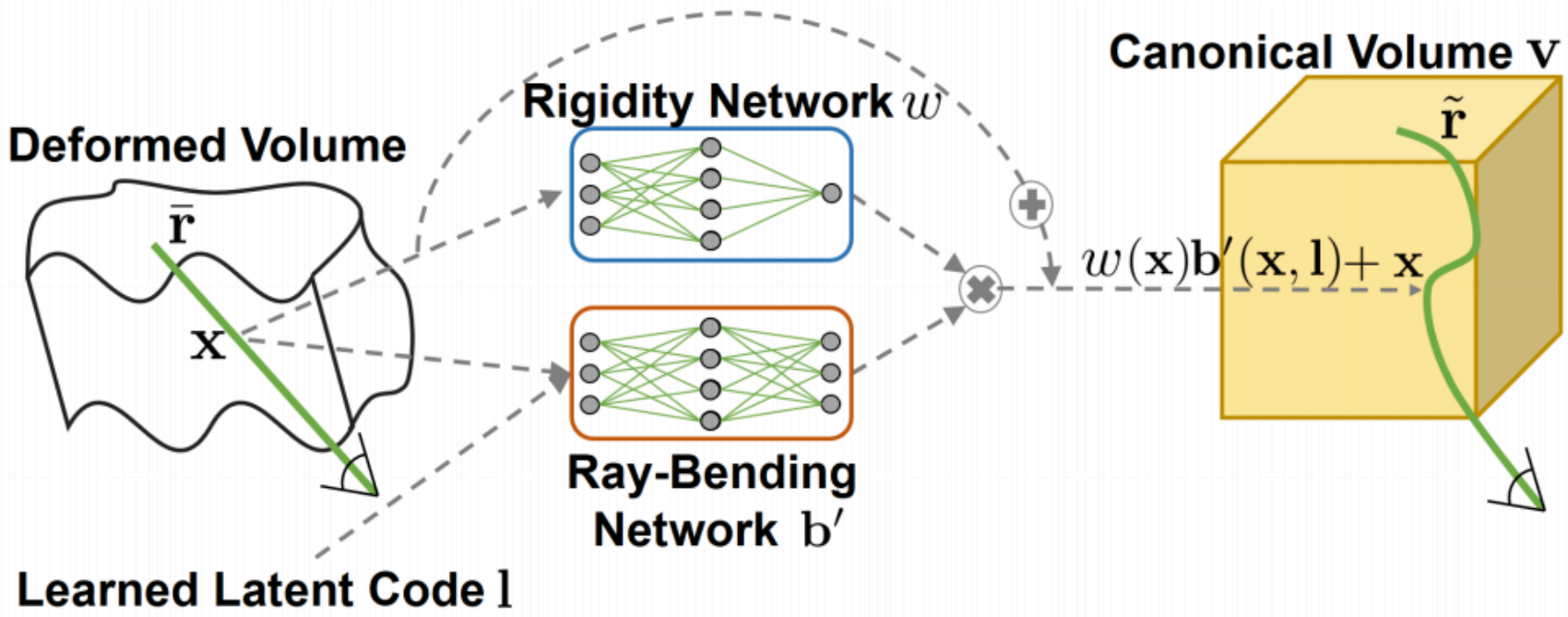[7] Self-supervised Tracking and Reconstruction of Rigid Objects in Motion with Neural Rendering, Yuan et al., Arxiv 2020
[8] Neural Radiance Flow for 4D View Synthesis and Video Processing, Du et al., Arxiv 2020
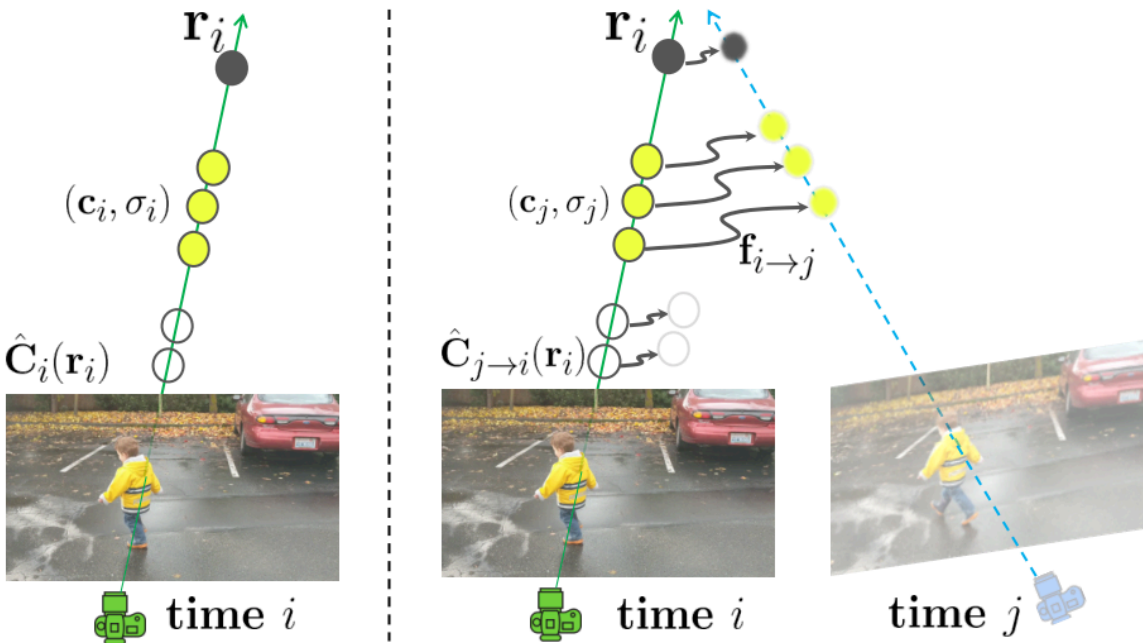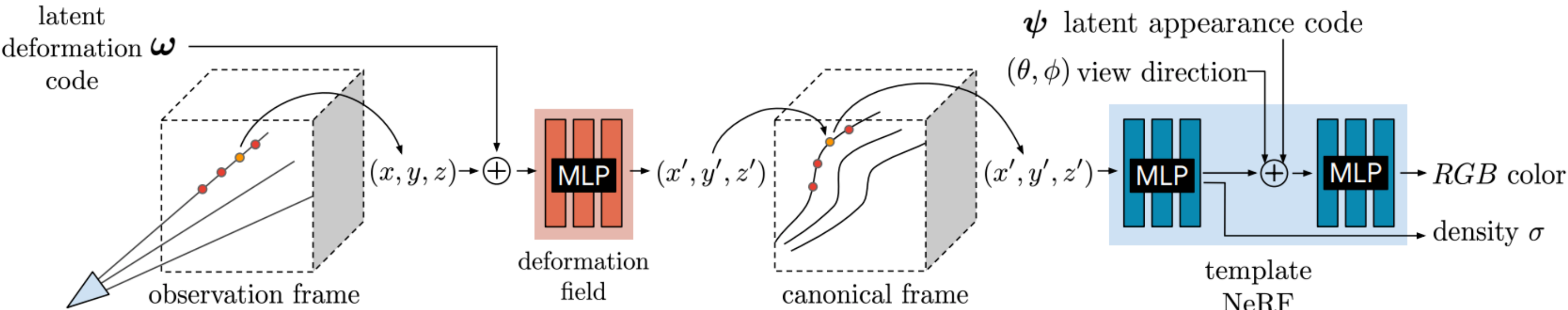
Concurrent works 2020
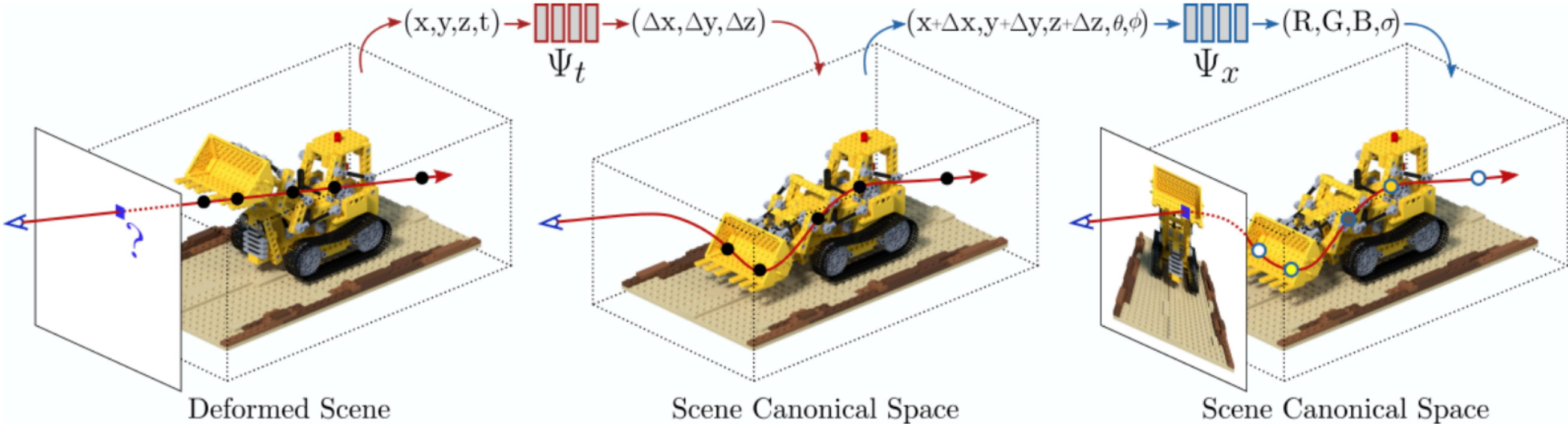
DNeRF

NR-NeRF

NSFF

Nerfie

Thank you!