3D Reconstruction with Polarization

GNeRP: Gaussian-guided Neural Reconstruction of Reflective Objects with Noisy Polarization Priors

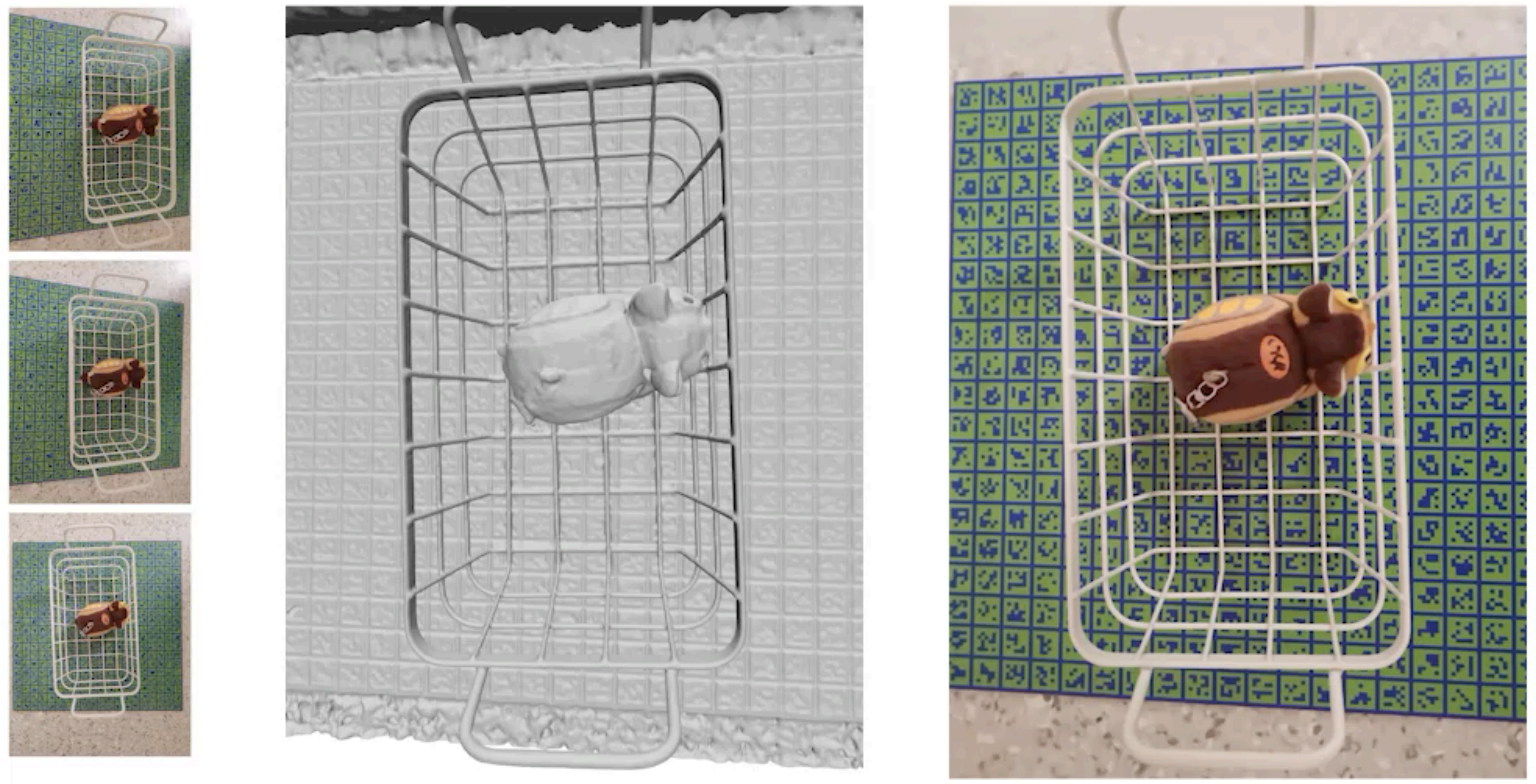
Yang LI Ruizheng WU Jiyong LI Yingcong CHEN

ICLR 2024

Reporter: Yang LI (yli803@cse.ust.hk)

Multi-view 3D Reconstruction

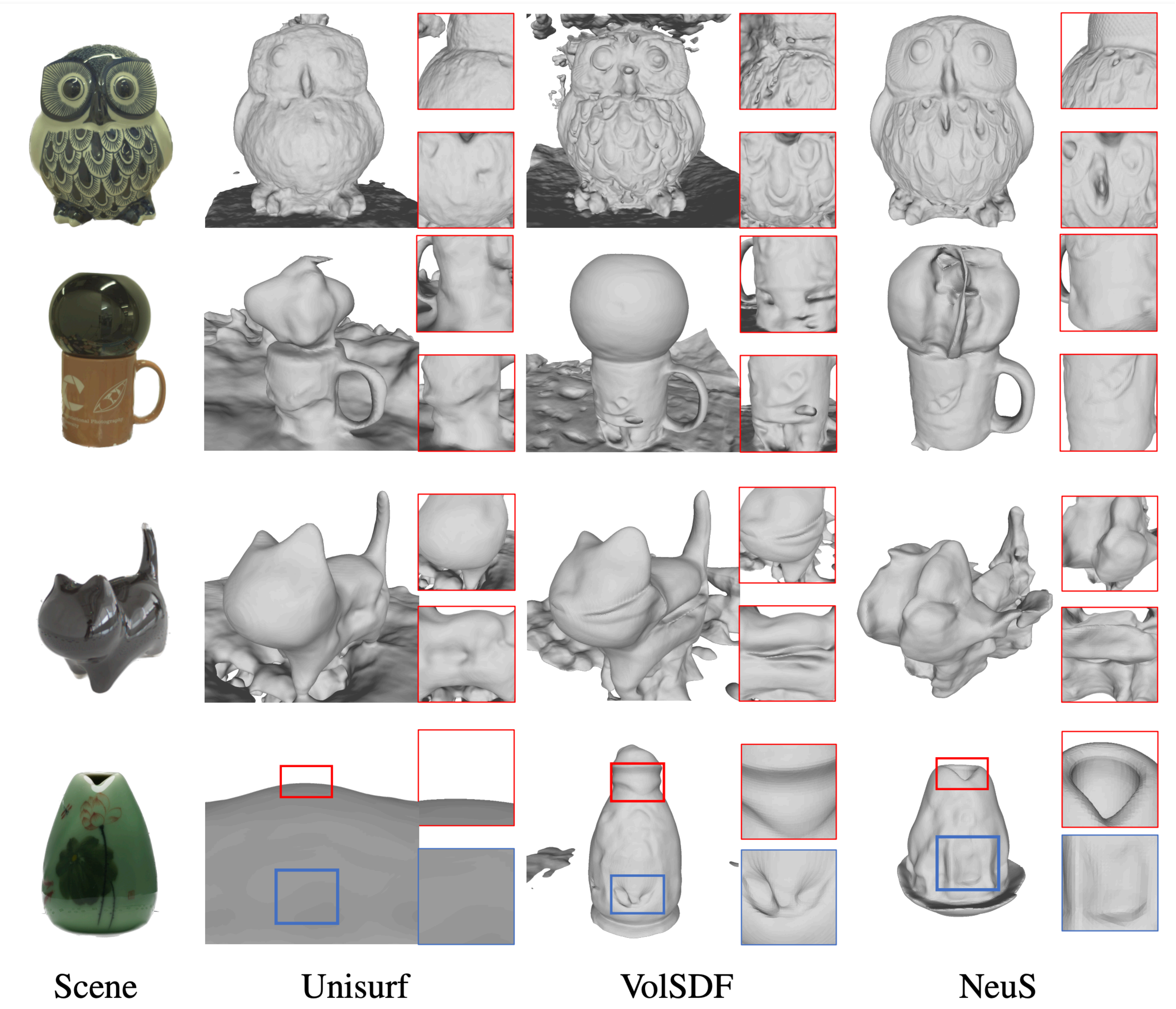
Existing methods have the ability to recover accurate geometry of diffuse objects

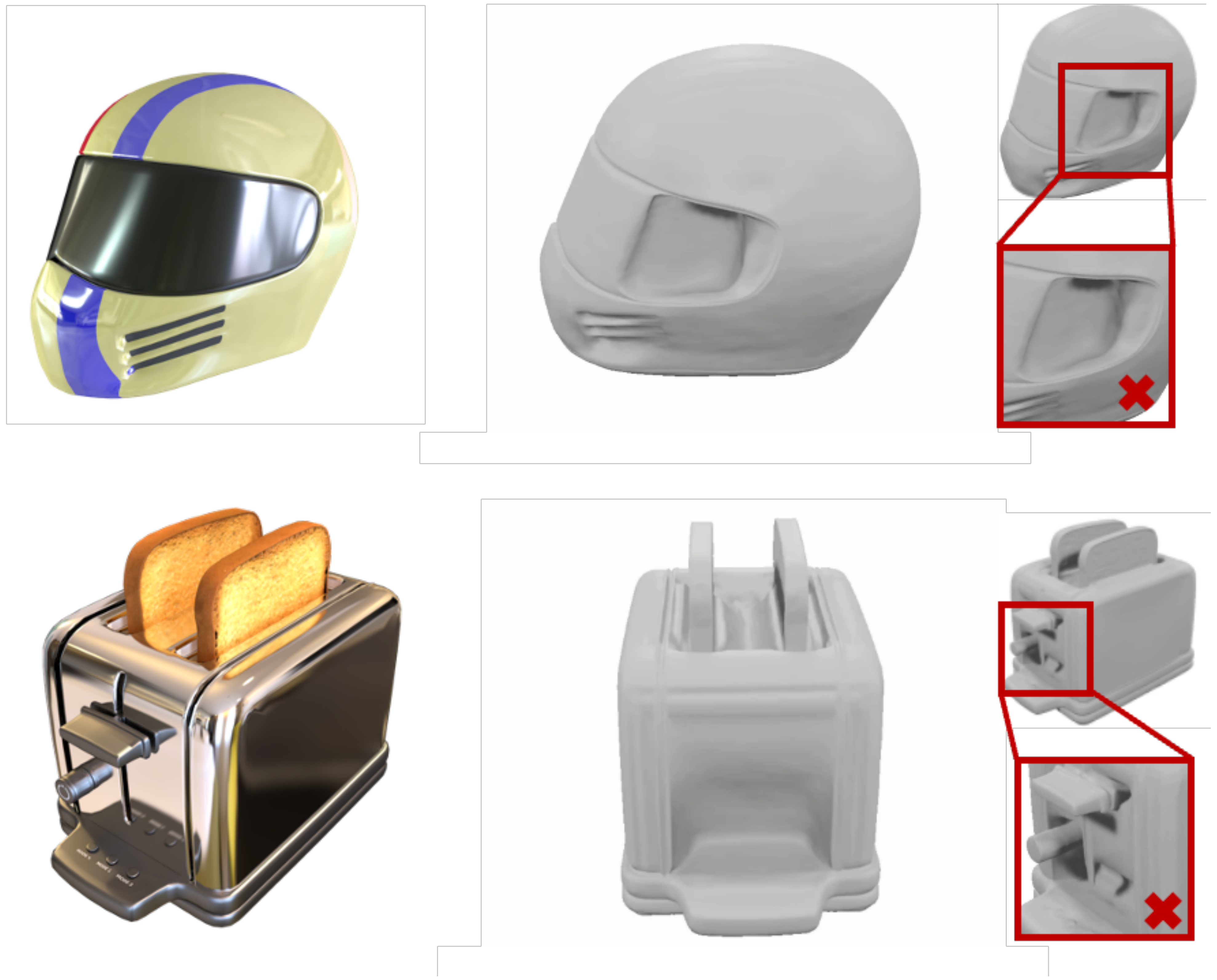


Wang, Peng, et al. "Neus: Learning neural implicit surfaces by volume rendering for multi-view reconstruction." arXiv preprint arXiv:2106.10689 (2021).

Multi-view 3D Reconstruction

However, existing methods are unable to reconstruct reflective objects





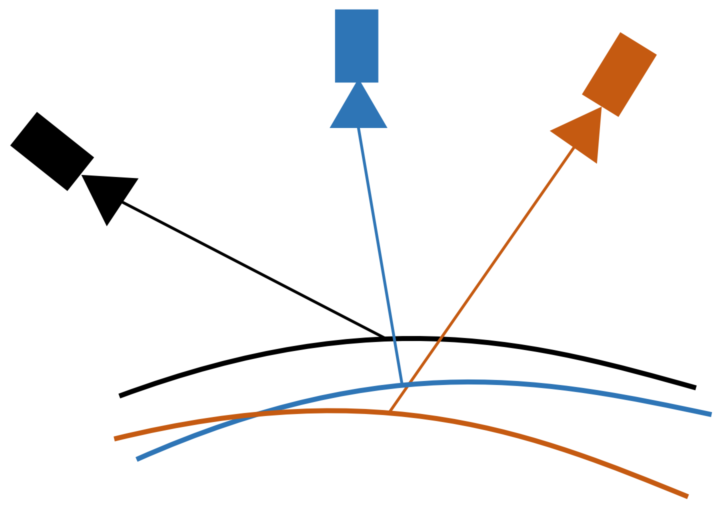
Ge, Wenhang, et al. "Ref-NeuS: Ambiguity-Reduced Neural Implicit Surface Learning for Multi-View Reconstruction with Reflection." ICCV (2023).

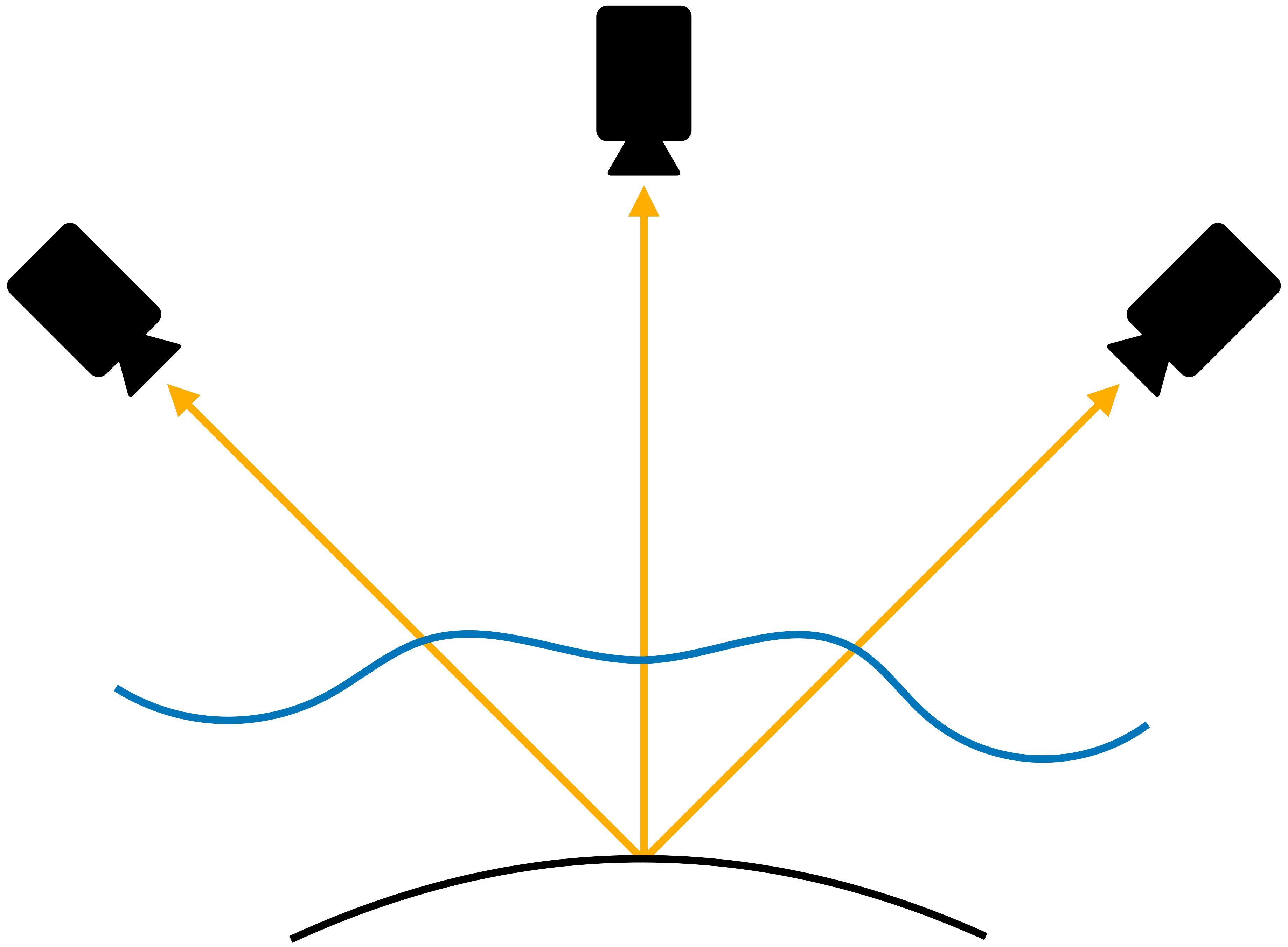
Dilemma in Reflective Objects

Multi-view inconsistency results in high-frequency specular radiance

Multi-view Inconsistency

- Ambiguous surfaces from radiance images

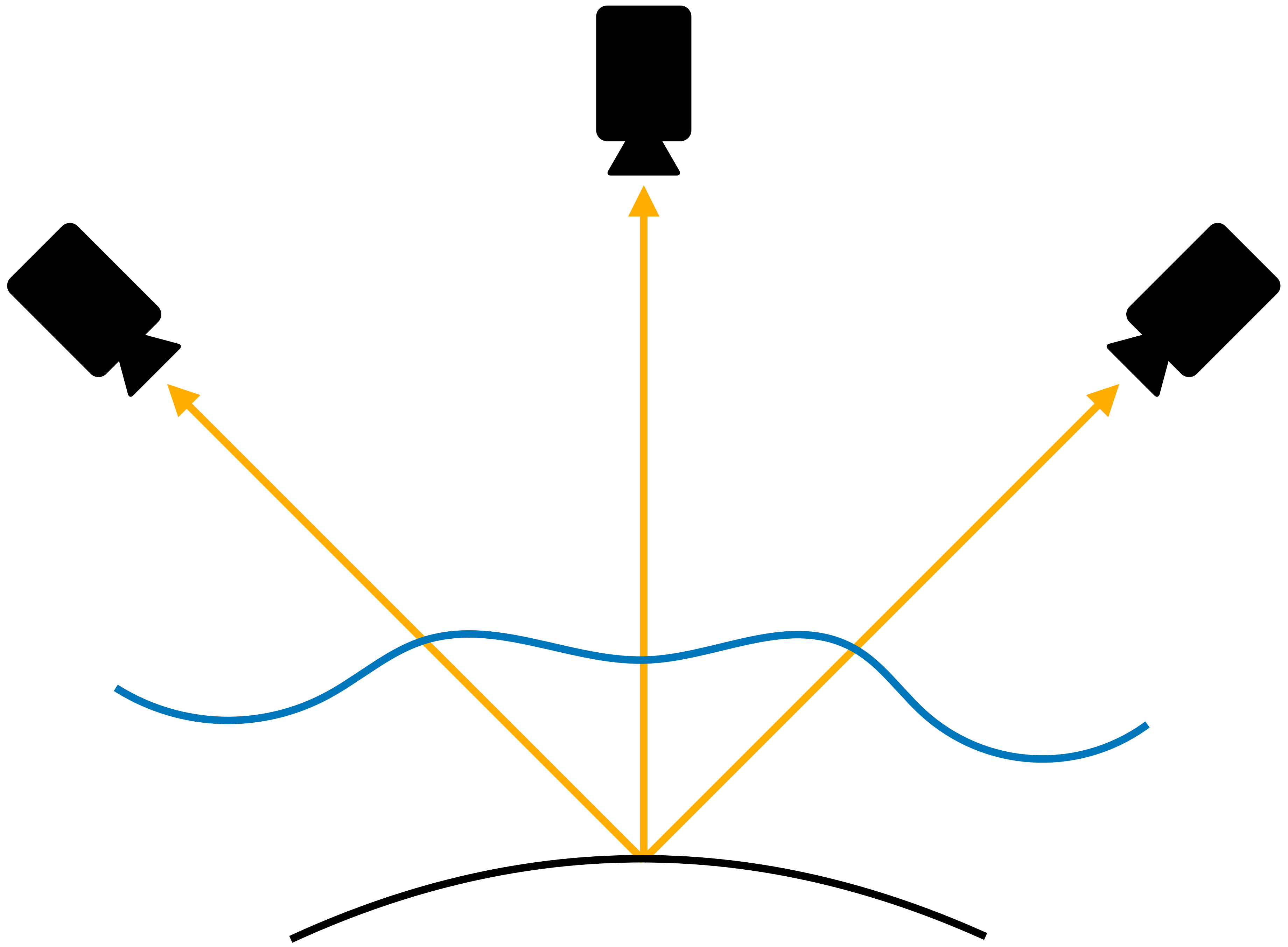




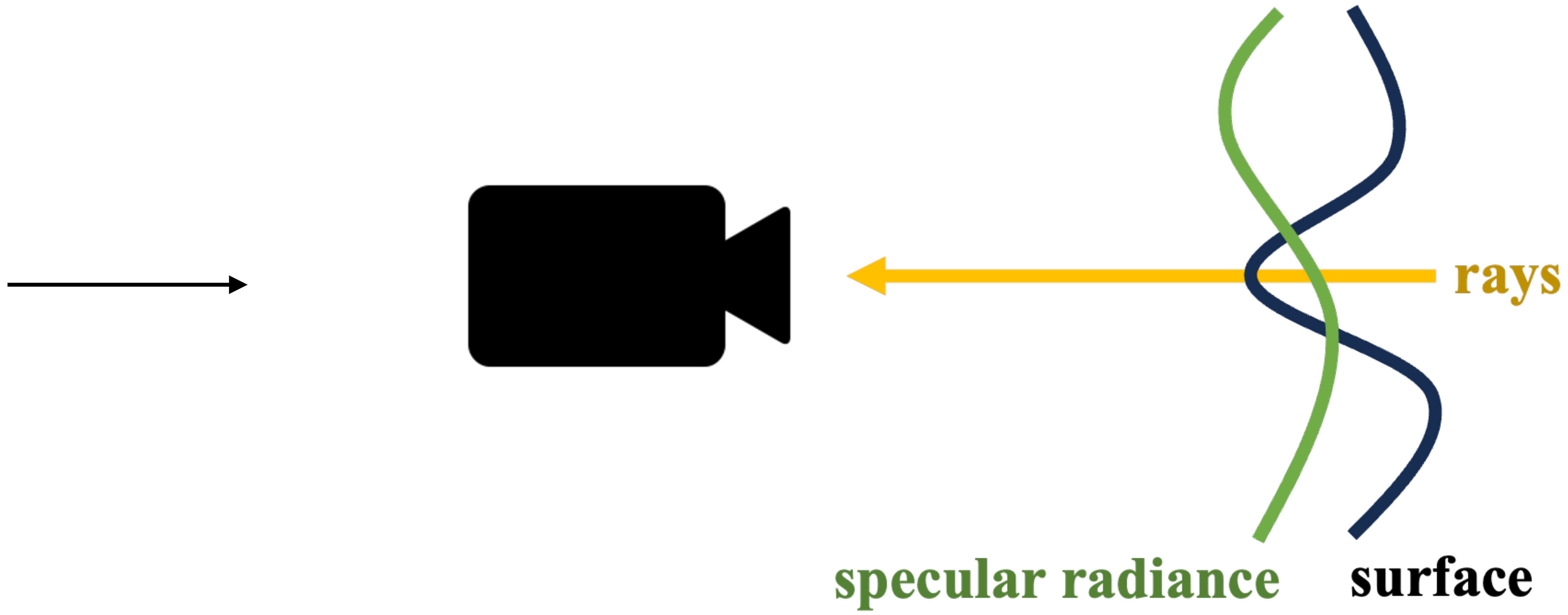
High-frequency Specular Radiance

Dilemma in Reflective Objects

Moreover, disentangle complicated geometry from specular radiance is more intractable

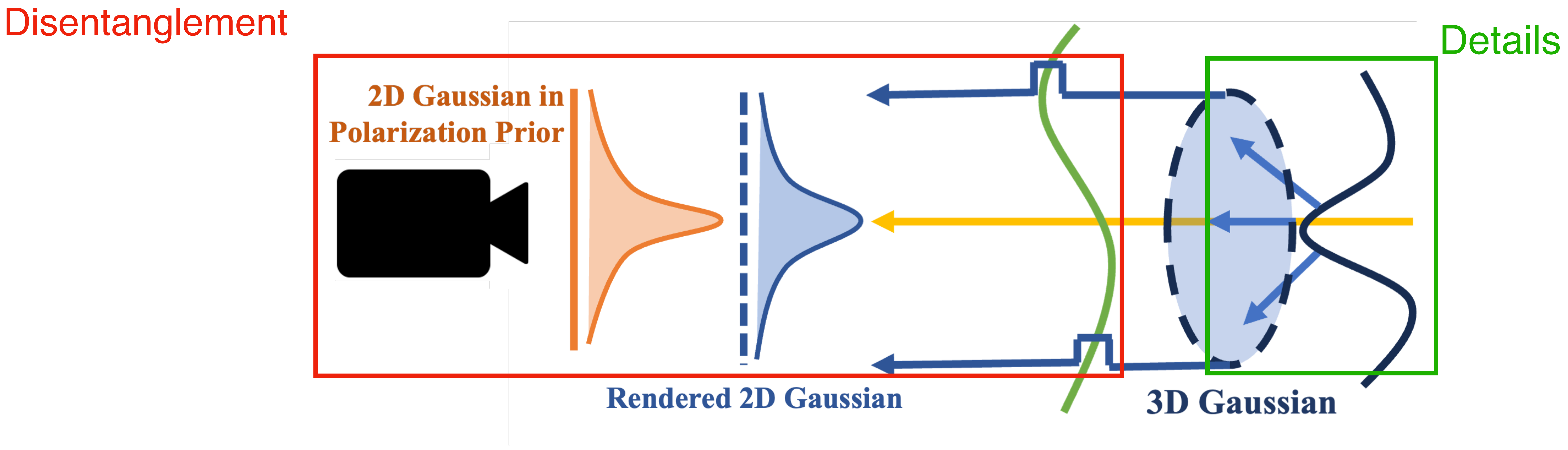


High-frequency Specular Radiance



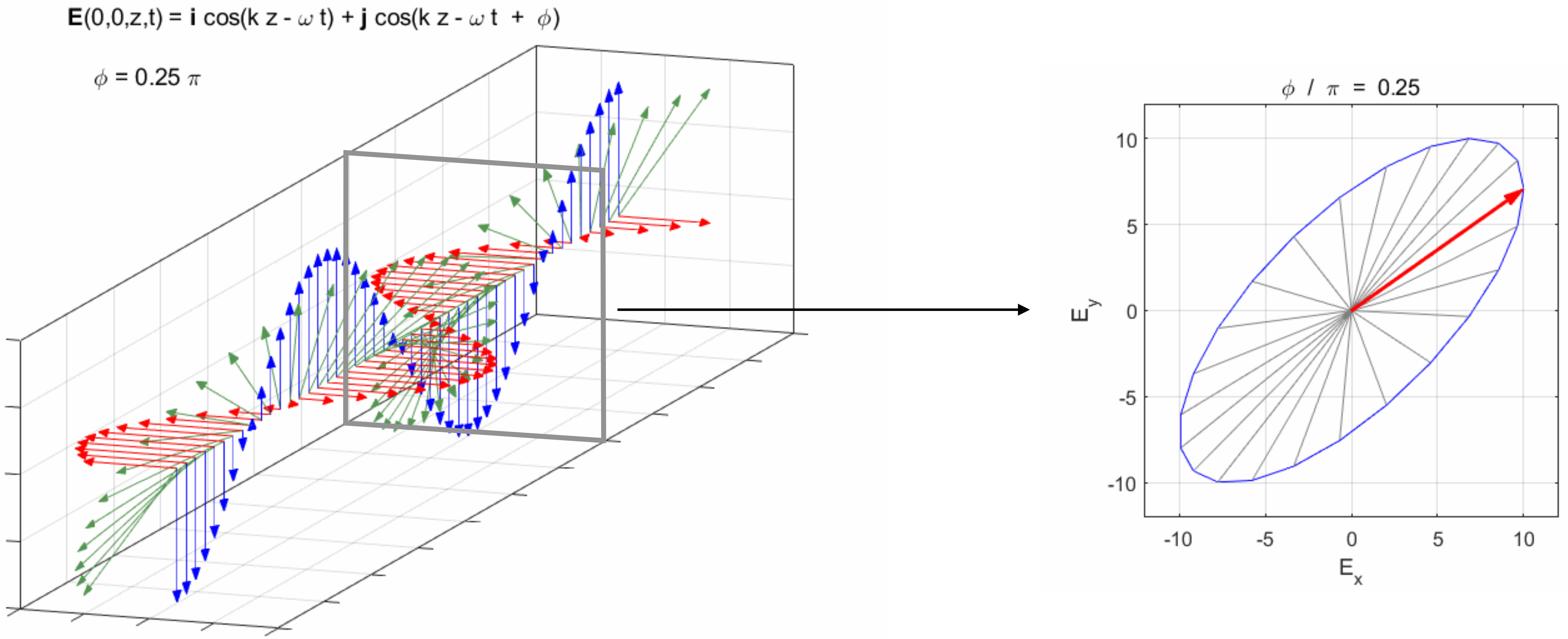
Our Motivation

Gaussian representation of normals is disentanglable and contains details of geometry.



Preliminary of Polarization

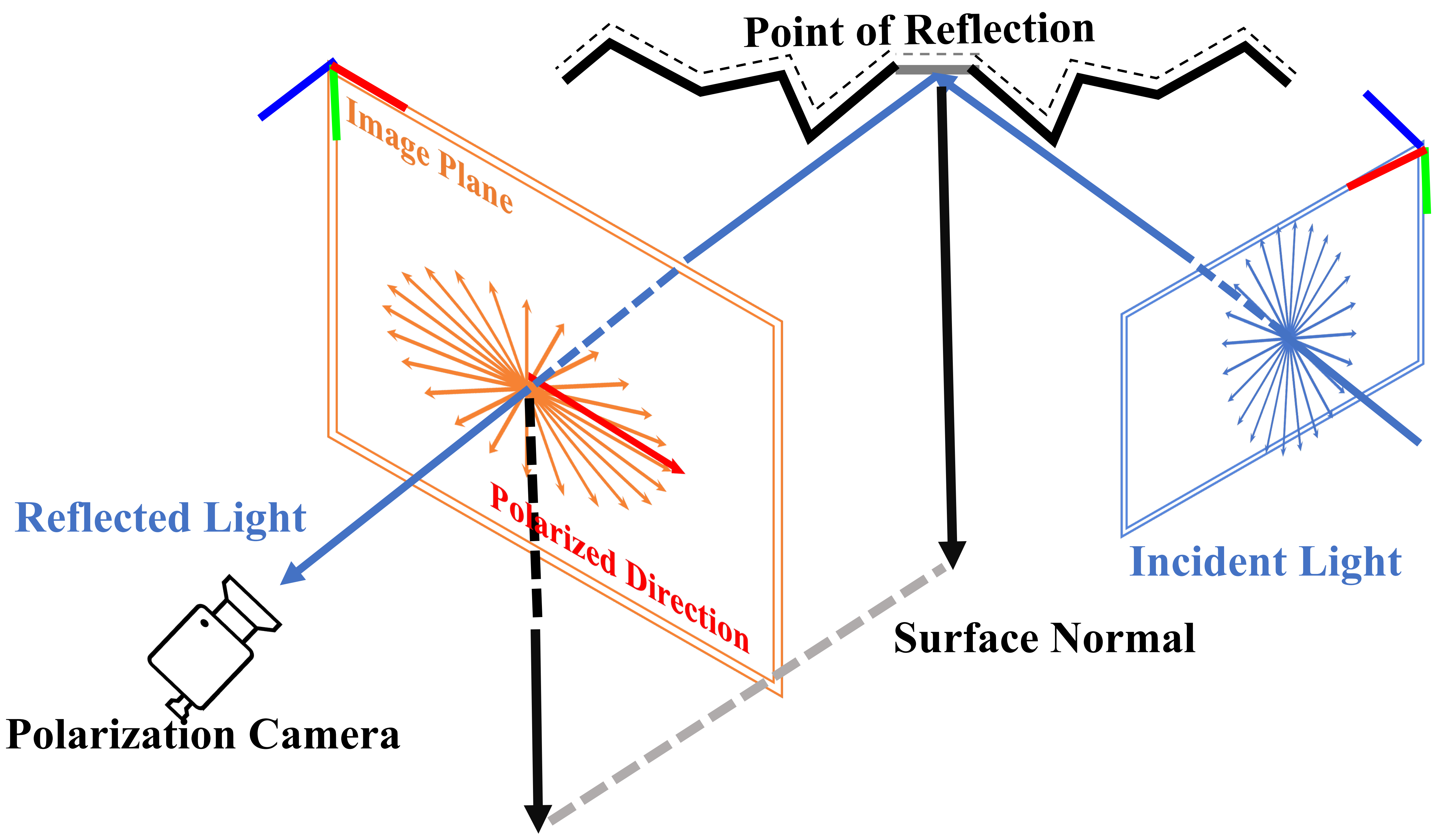
Vibration status of light is a type of electromagnetic wave

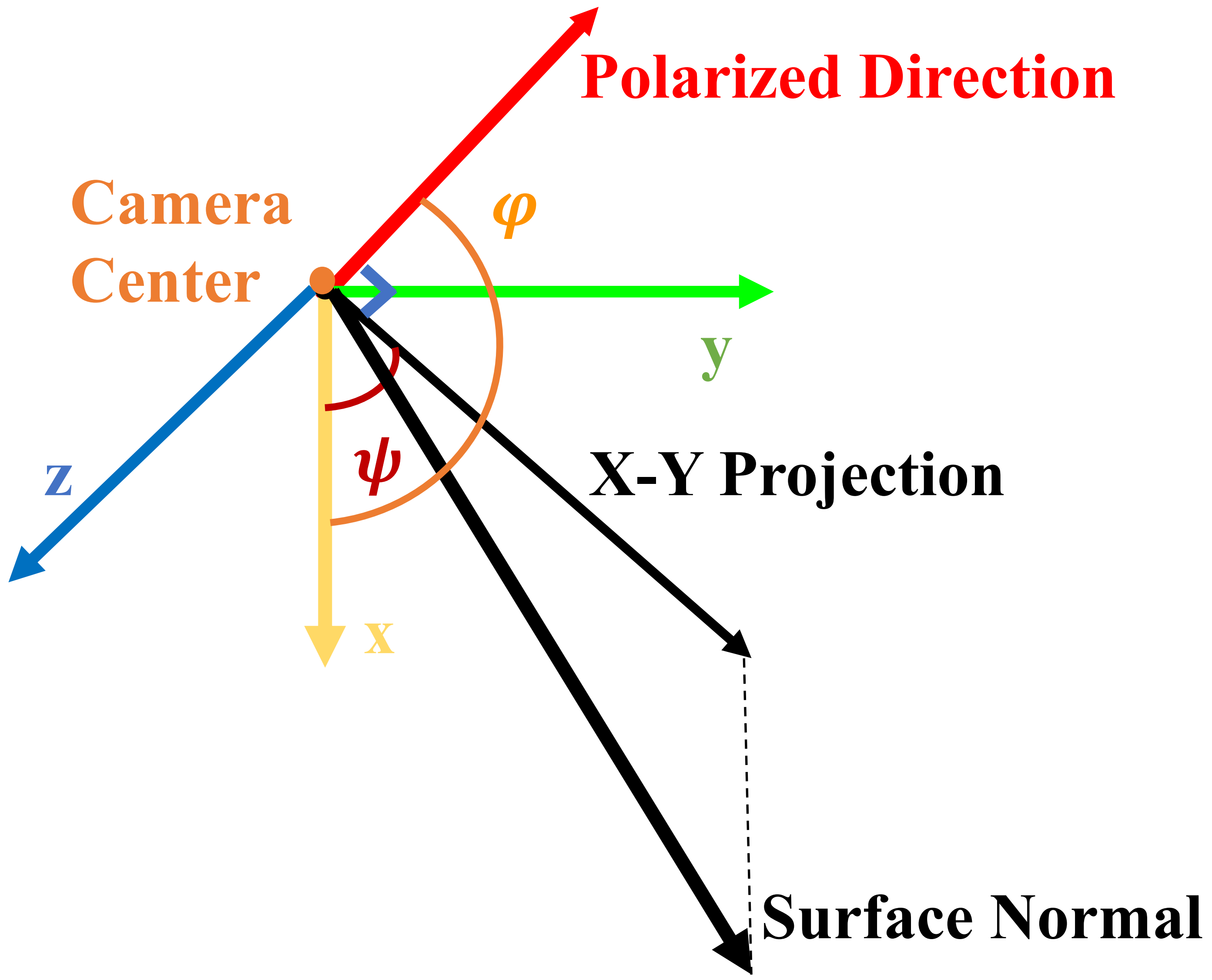


https://d-arora.github.io/Doing-Physics-With-Matlab/mpDocs/op1002.htm

Preliminary of Polarization

Fixed relation to the azimuth angle of surface normals

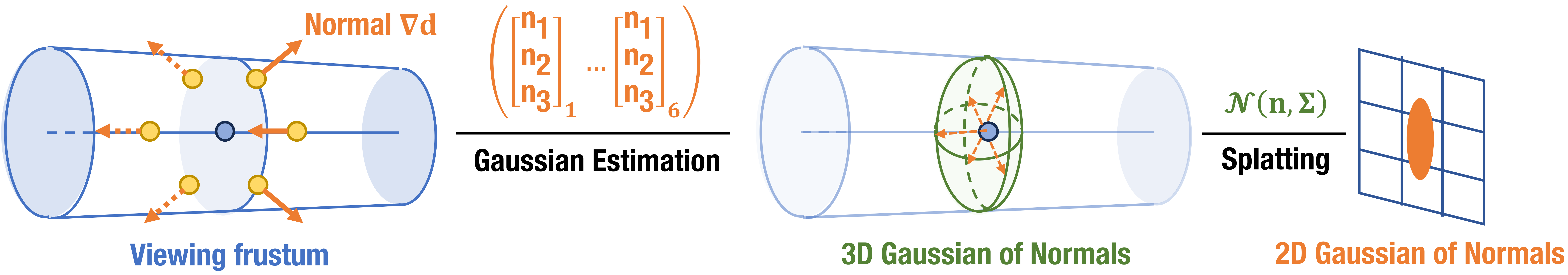




pi/2 difference

Our Method

Gaussian representation of normals



$$\mathcal{G}(\mathbf{x} \mid \mathbf{x}_i) = \mathcal{N}(\mathbf{n}(\mathbf{x}_i), \mathbf{\Sigma}(\mathbf{x}_i)) = \frac{1}{(2\pi)^{\frac{3}{2}} \mid \mathbf{\Sigma}(\mathbf{x}_i) \mid^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{n}(\mathbf{x}_i))^{\mathrm{T}} \mathbf{\Sigma}(\mathbf{x}_i)^{-1} (\mathbf{x} - \mathbf{n}(\mathbf{x}_i))\right)$$

Our Method

Gaussian splatting of normals

Transformation:

$$r''(\mathbf{t}-\mathbf{t}_k) = \mathcal{G}_{V''_k}(\mathbf{t}-\mathbf{t}_k) \text{ (World Space)}$$

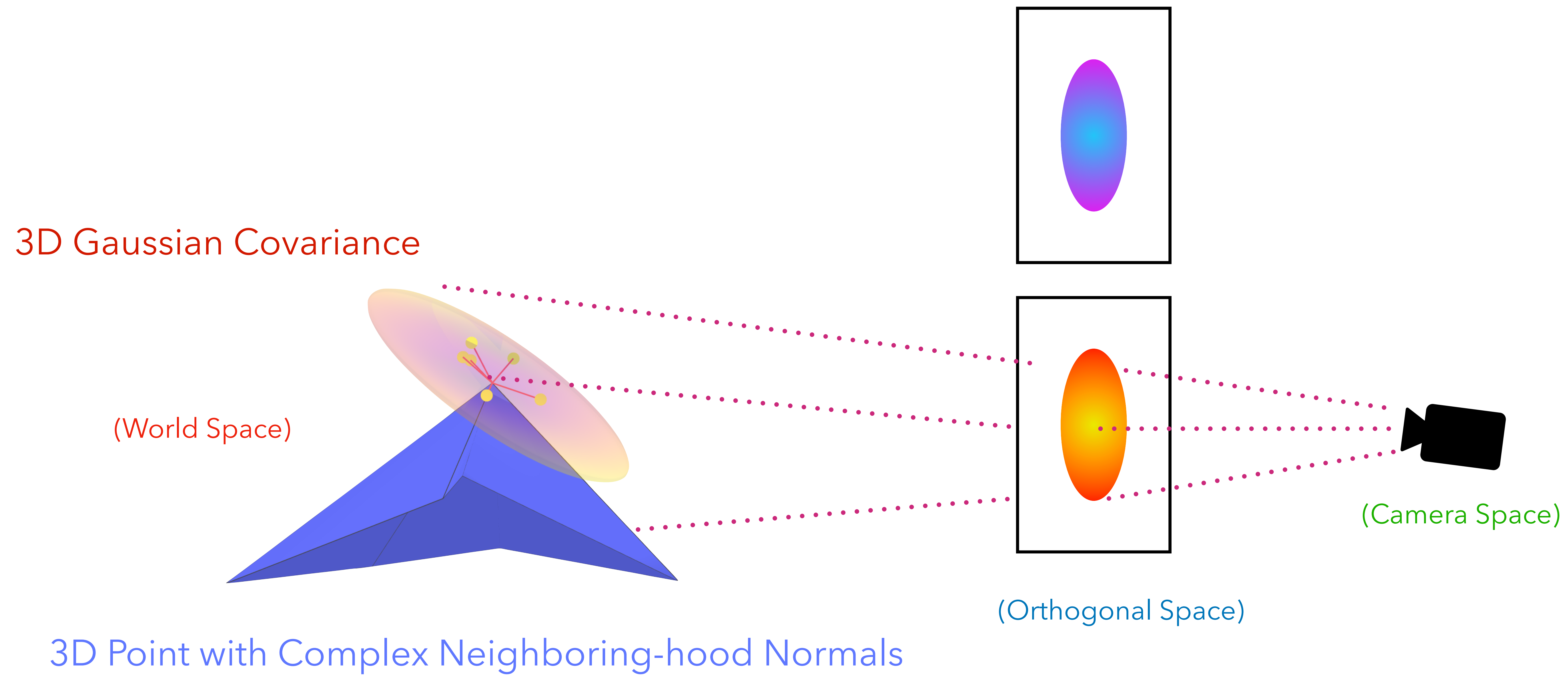
$$\mathbf{u} = \Phi(\mathbf{t}) = \mathbf{W}\mathbf{t} + \mathbf{d}$$

$$r'_k(\mathbf{u}) = \frac{1}{|\mathbf{W}^{-1}|}\mathcal{G}_{\mathbf{W}\mathbf{V''_k}\mathbf{W}^T}(\mathbf{u}-\Phi(\mathbf{t}_k)) \text{ (Camera Space)}$$

$$\mathbf{x} = \mathbf{m_u}(\mathbf{u})$$

$$r_k(\mathbf{u}) = \frac{1}{|\mathbf{W}^{-1}||\mathbf{J}^{-1}|}\mathcal{G}_{\mathbf{J}\mathbf{W}\mathbf{V''_k}\mathbf{W}^T\mathbf{J}^T}(\mathbf{x}-\mathbf{m_{x_k}}(\mathbf{x}_k))$$

(Orthogonal Space)



Zwicker, Matthias, et al. "EWA volume splatting." Proceedings Visualization, 2001. VIS'01.. IEEE, 2001.

Our Method

Gaussian splatting of normals

Transformation:

$$r''(\mathbf{t}-\mathbf{t}_k) = \mathcal{G}_{V''_k}(\mathbf{t}-\mathbf{t}_k) \text{ (World Space)}$$

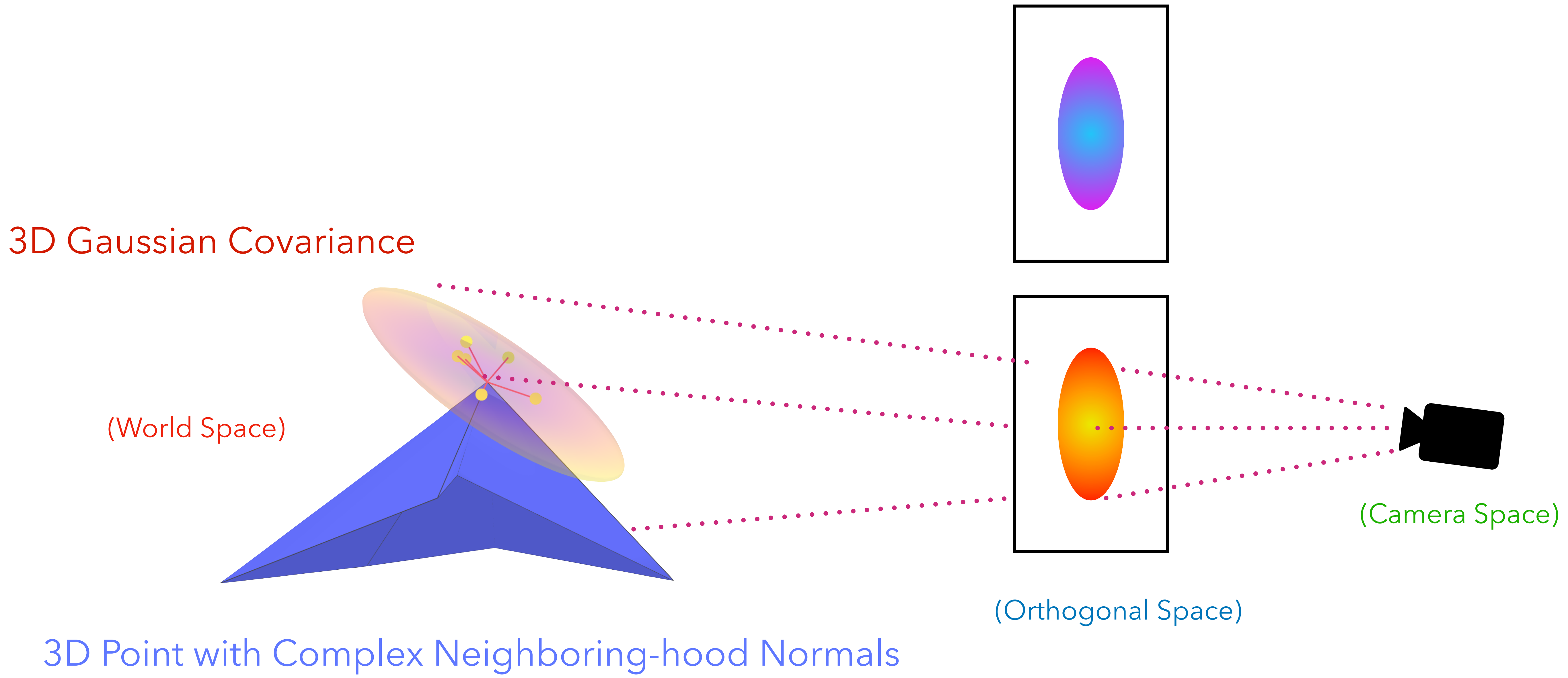
$$\mathbf{u} = \Phi(\mathbf{t}) = \mathbf{W}\mathbf{t} + \mathbf{d}$$

$$r'_k(\mathbf{u}) = \frac{1}{|\mathbf{W}^{-1}|}\mathcal{G}_{\mathbf{W}\mathbf{V}''_k\mathbf{W}^T}(\mathbf{u}-\Phi(\mathbf{t}_k)) \text{ (Camera Space)}$$

$$\boxed{\mathbf{x} = \mathbf{m}_{\mathbf{u}}(\mathbf{u})}$$ Strict Linear

$$r_k(\mathbf{u}) = \frac{1}{|\mathbf{W}^{-1}||\mathbf{J}^{-1}|}\mathcal{G}_{\mathbf{J}\mathbf{W}\mathbf{V}''_k\mathbf{W}^T\mathbf{J}^T}(\mathbf{x}-\mathbf{m}_{\mathbf{x}_k}(\mathbf{x}_k))$$

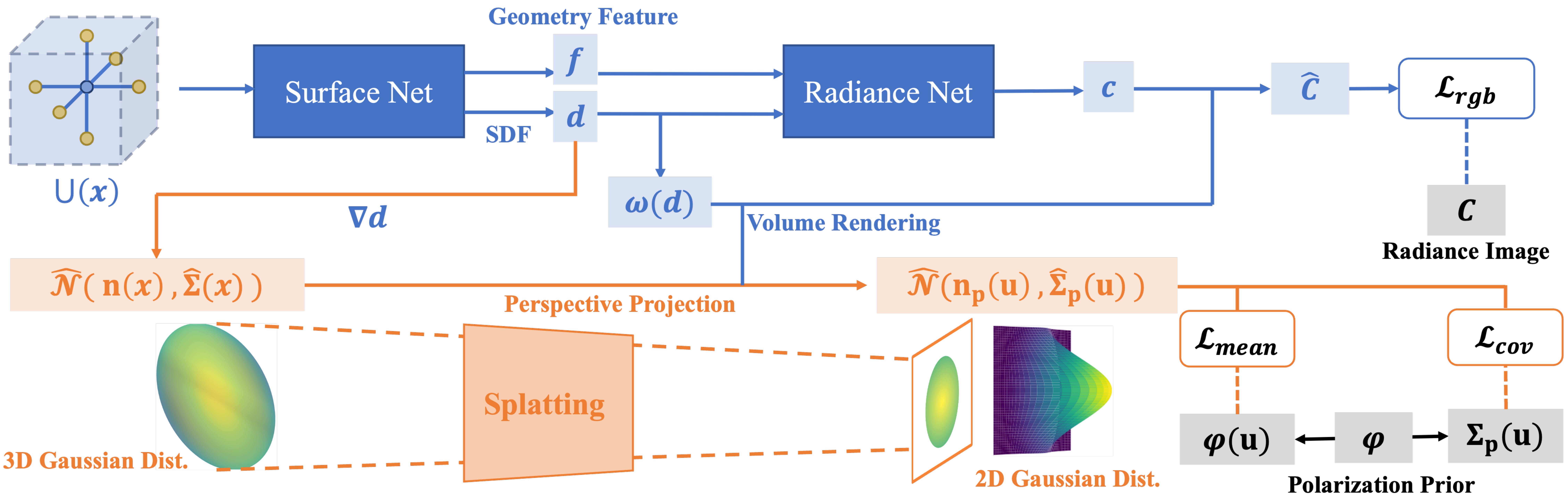
(Orthogonal Space)



Zwicker, Matthias, et al. "EWA volume splatting." Proceedings Visualization, 2001. VIS'01.. IEEE, 2001.

Our Method

Overview



Our Method

Total loss

$$\mathcal{L} = \alpha(1-\boldsymbol{\rho})\mathcal{L}_{\text{color}} + \beta\boldsymbol{\rho}(\mathcal{L}_{\text{mean}} + \mathcal{L}_{\text{cov}}) + \gamma\mathcal{L}_{\text{eik}} + \delta\mathcal{L}_{\text{mask}},$$

$$\mathcal{L}_{\text{color}} = \| \hat{\mathbf{C}}(\mathbf{u}) - \mathbf{C}(\mathbf{u}) \|_2 , \ \mathcal{L}_{\text{mean}} = \| \hat{\boldsymbol{\varphi}}(\mathbf{n}_{\mathbf{p}}(\mathbf{u})) - \boldsymbol{\varphi}(\mathbf{u}) \|_1 ,$$

$$\mathcal{L}_{\text{cov}} = \left(\left\| \frac{\hat{\boldsymbol{\Lambda}}_1}{\hat{\boldsymbol{\Lambda}}_0} - \frac{\widetilde{\boldsymbol{\Lambda}}_1}{\widetilde{\boldsymbol{\Lambda}}_0} \right\|_1 + \beta' < \hat{\mathbf{V}}, \widetilde{\mathbf{V}} > \right)(\mathbf{u}), \ \mathcal{L}_{\text{eik}} = \frac{1}{K}\sum_{i=1}^{K} (\|\nabla_{\mathbf{x}} d(\mathbf{x}_i)\|_2 - 1)^2,$$

In reflective regions, suppress radiance loss and amplify Gaussian loss.

Our Method

Degree-of-polarization reweighing strategy

$$\mathcal{L} = \alpha(1-\boldsymbol{\rho})\mathcal{L}_{\text{color}} + \beta\boldsymbol{\rho}(\mathcal{L}_{\text{mean}} + \mathcal{L}_{\text{cov}}) + \gamma\mathcal{L}_{\text{eik}} + \delta\mathcal{L}_{\text{mask}},$$

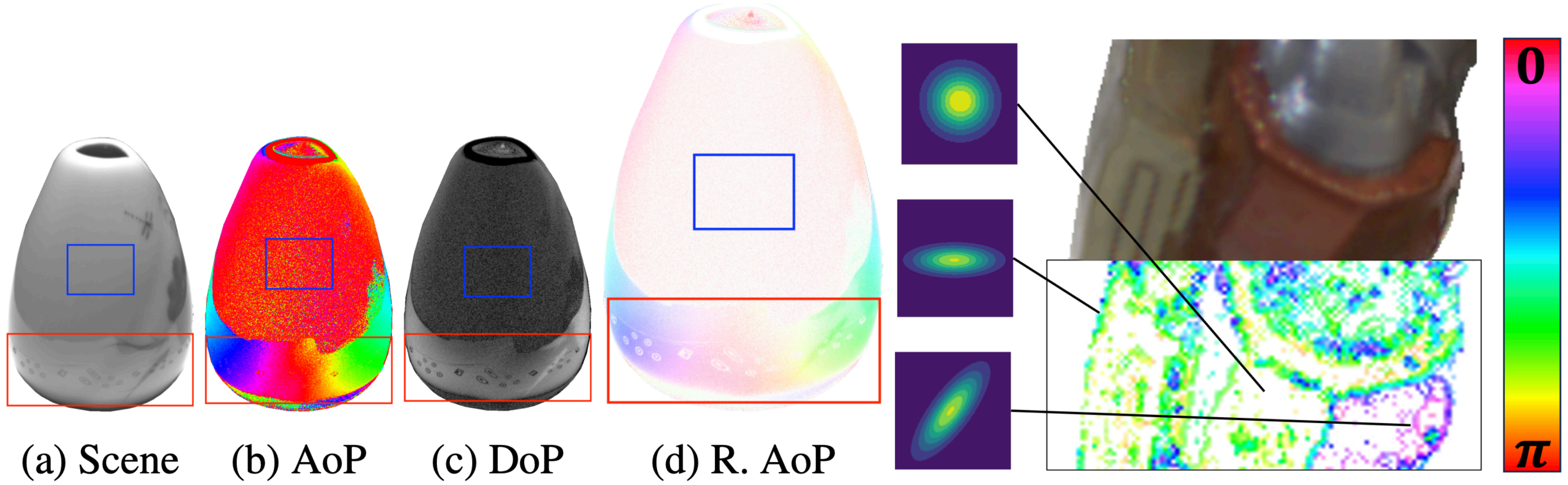
$$\mathcal{L}_{\text{color}} = \| \hat{\mathbf{C}}(\mathbf{u}) - \mathbf{C}(\mathbf{u}) \|_2 \, , \; \mathcal{L}_{\text{mean}} = \| \hat{\boldsymbol{\varphi}}(\mathbf{n}_{\mathbf{p}}(\mathbf{u})) - \boldsymbol{\varphi}(\mathbf{u}) \|_1 \, ,$$

$$\mathcal{L}_{\text{cov}} = \left(\left\| \frac{\hat{\boldsymbol{\Lambda}}_1}{\hat{\boldsymbol{\Lambda}}_0} - \frac{\widetilde{\boldsymbol{\Lambda}}_1}{\widetilde{\boldsymbol{\Lambda}}_0} \right\|_1 + \beta' < \hat{\mathbf{V}}, \widetilde{\mathbf{V}} > \right)(\mathbf{u}), \; \mathcal{L}_{\text{eik}} = \frac{1}{K}\sum_{i=1}^{K}(\|\nabla_{\mathbf{x}} d(\mathbf{x}_i)\|_2 - 1)^2,$$

In **reflective** regions, **suppress** radiance loss and **amplify** Gaussian loss.

Our Method

Degree-of-polarization reweighing strategy



(a) Scene (b) AoP (c) DoP (d) R. AoP

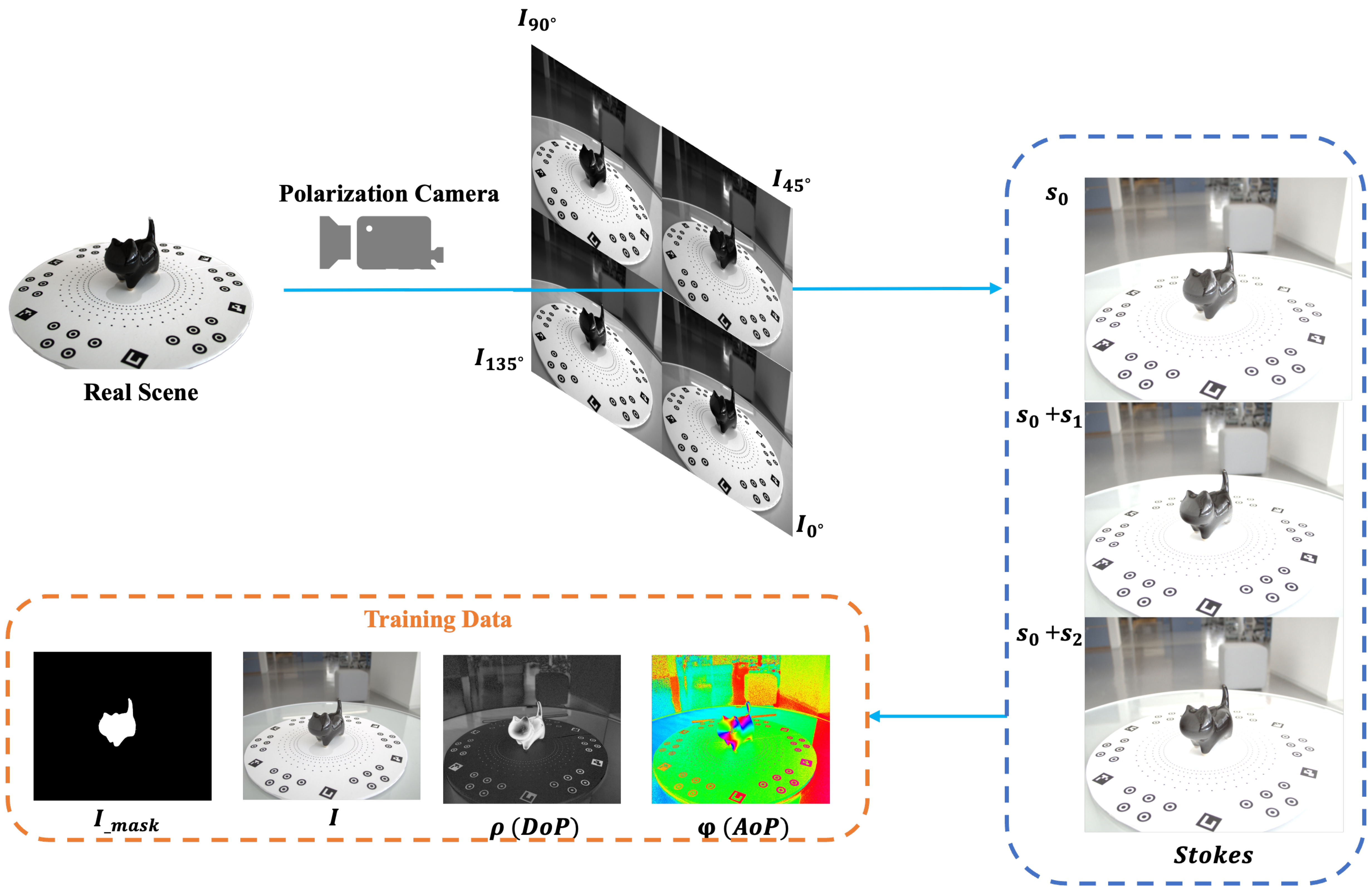
Visualization of DoP Reweighing

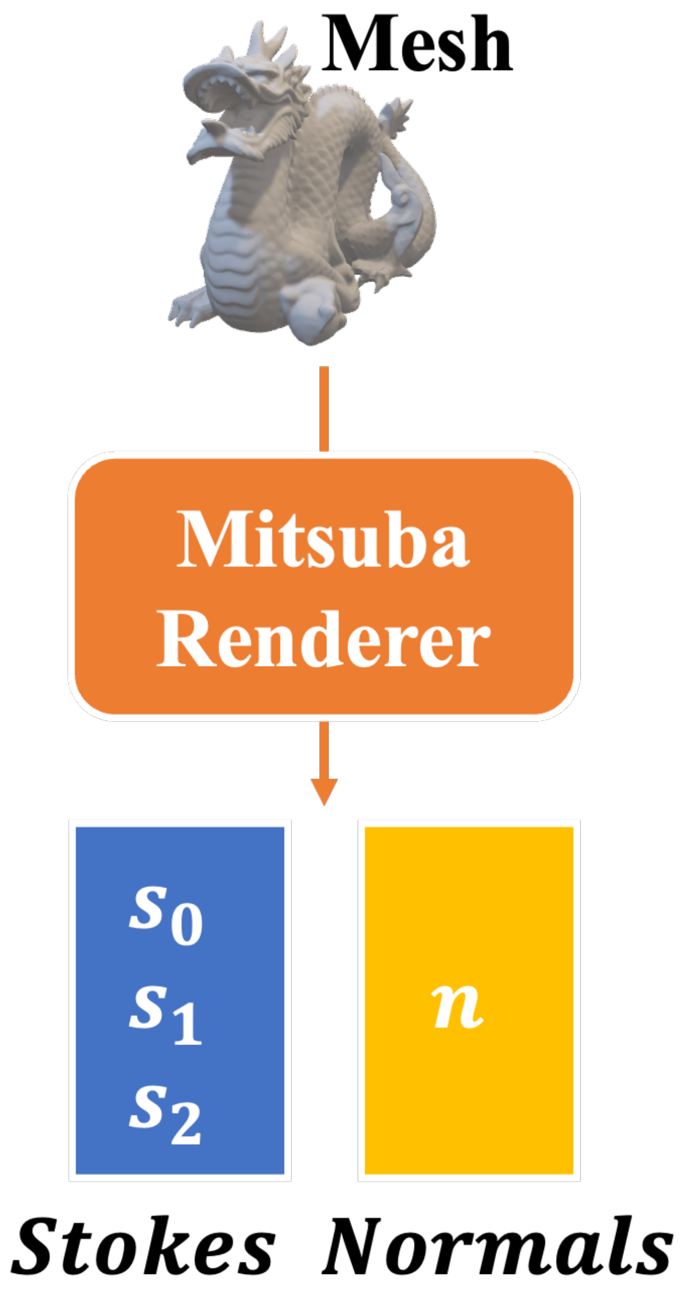
(e) 2D Gaussians in Polarization

Figure 4: Visualization of Reweighted AoP Priors. Red boxes bound specular reflection dominant regions, and the blue boxes bound diffuse ones. (d) is the AoP map reweighted by DoP. Saturation in (e) indicates the degree of anisotropy, and color represents the direction of the singular vector of 2D Gaussians' covariance. A few 2D Gaussians are drawn as ellipses for intuition.

Experiments

Data collection and generation





Experiments

Evaluation protocol

$$\mathrm{CD}\left(P_1, P_2\right) = \frac{1}{2n}\sum_{i=1}^{n}\left|x_i - \mathrm{Nearest}\left(x_i, P_2\right)\right| + \frac{1}{2m}\sum_{j=1}^{n}\left|x_j - \mathrm{Nearest}\left(x_j, P_1\right)\right|$$

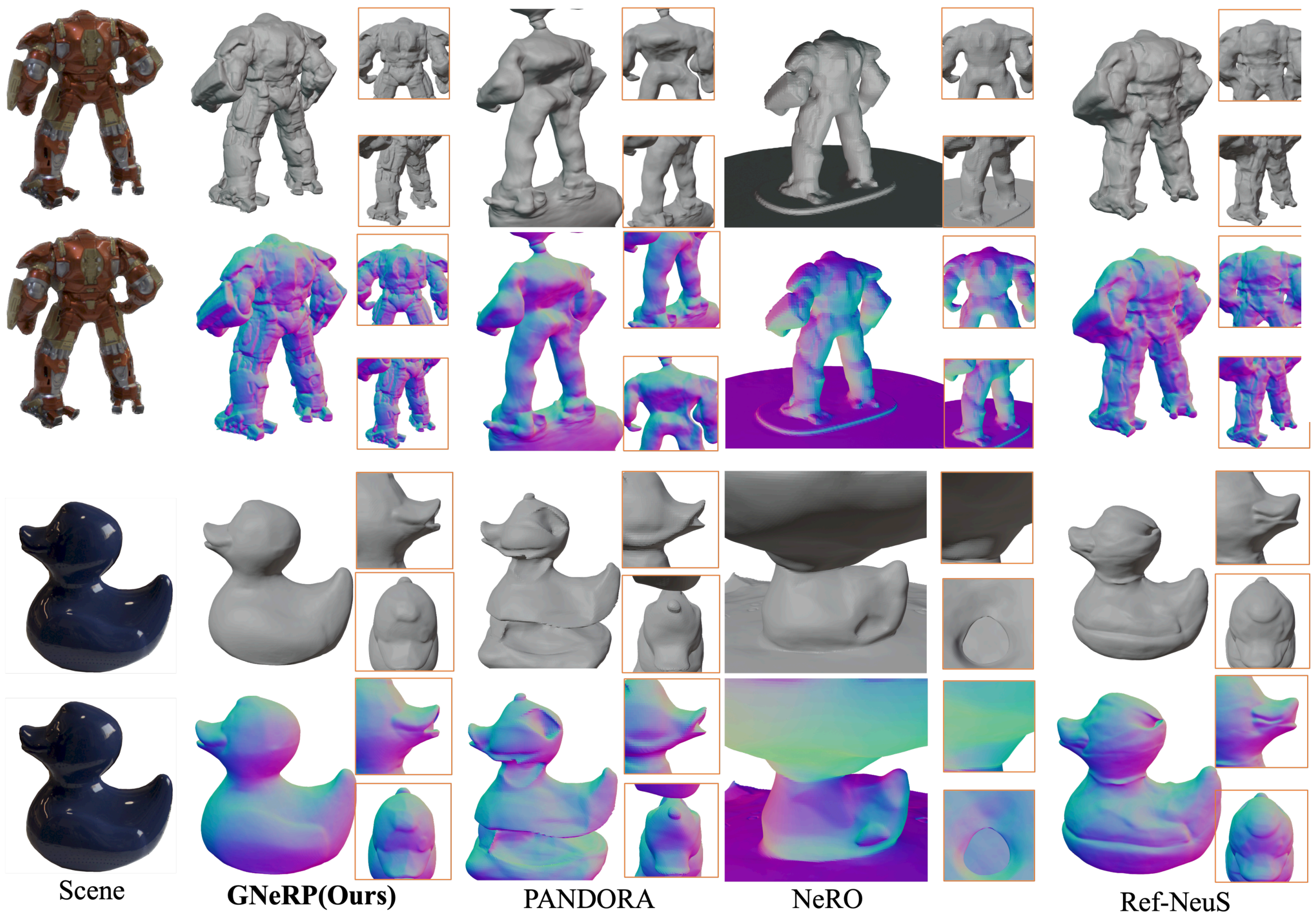
Real & Synthetic Data

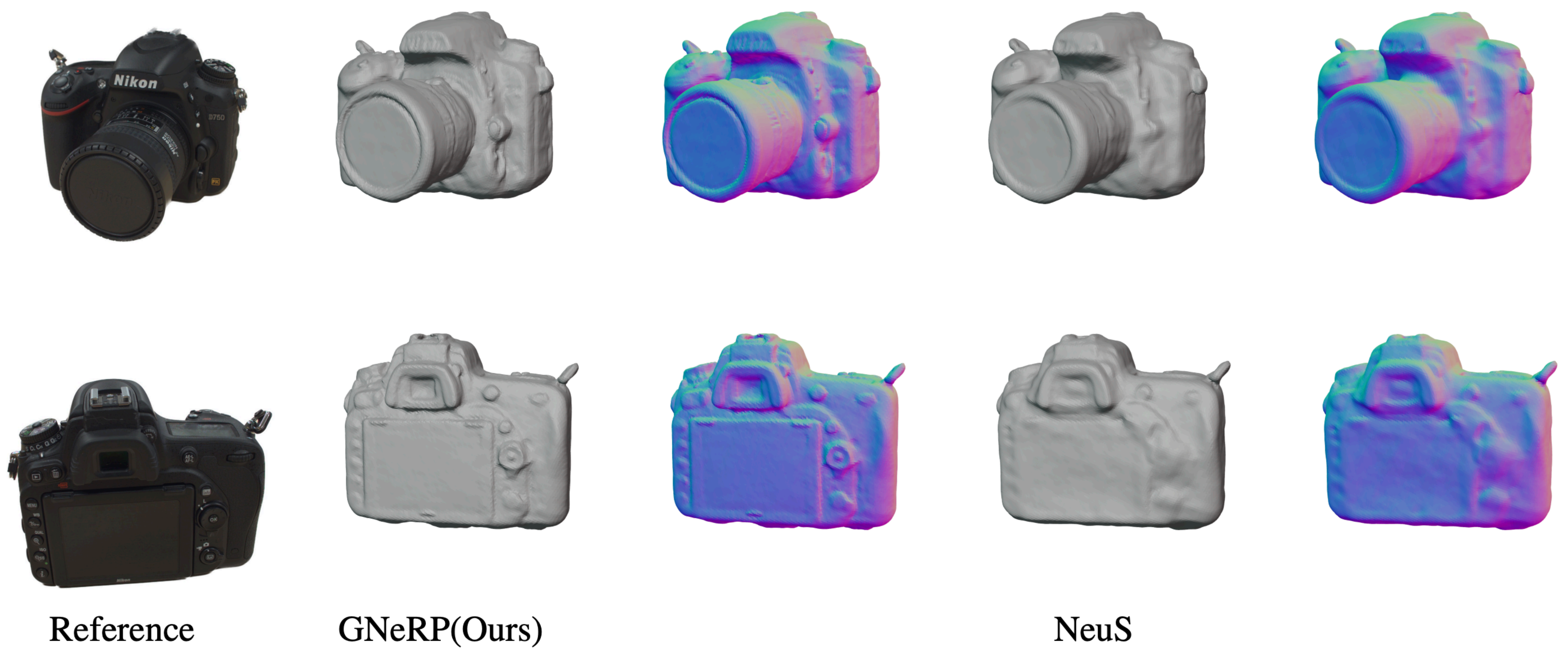
$$\mathrm{MAE}\left(\hat{\mathbf{n}}, \mathbf{n}\right) = \frac{1}{NM}\arccos\left(\frac{\hat{\mathbf{n}}_i^j \cdot \mathbf{n}_i^j)}{\|\hat{\mathbf{n}}_i^j\| \|\mathbf{n}_i^j)\|}\right)$$

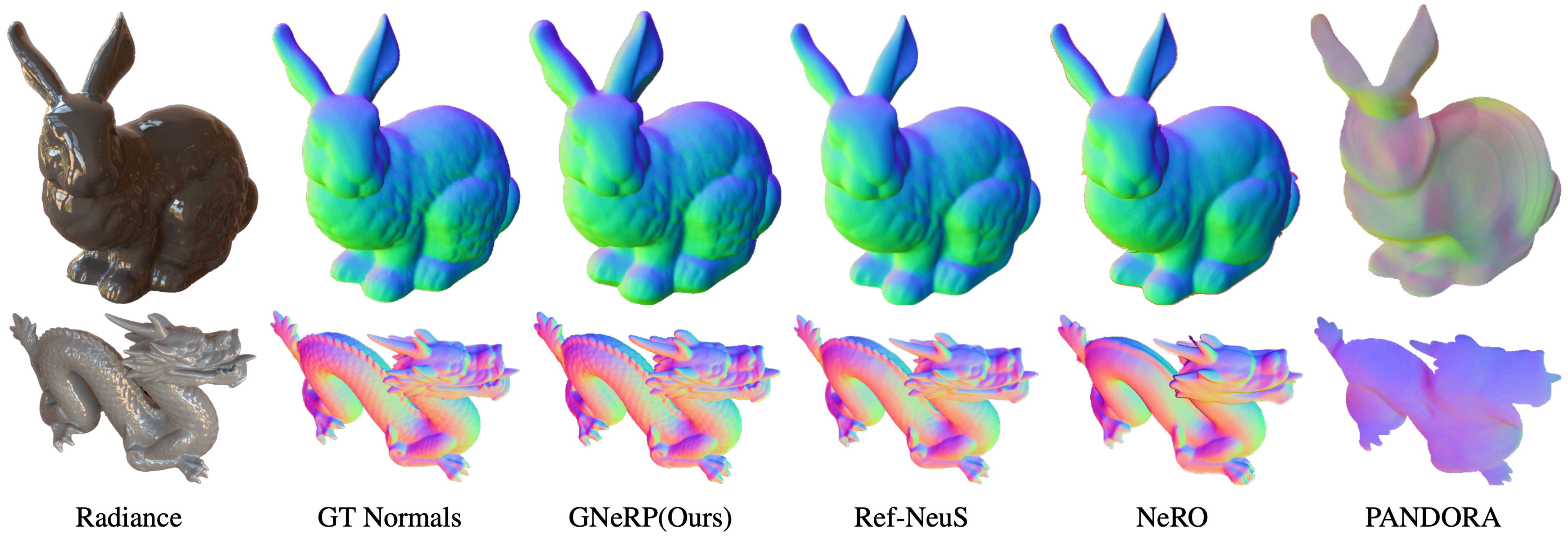
Synthetic Data

Experiments

SOTA results

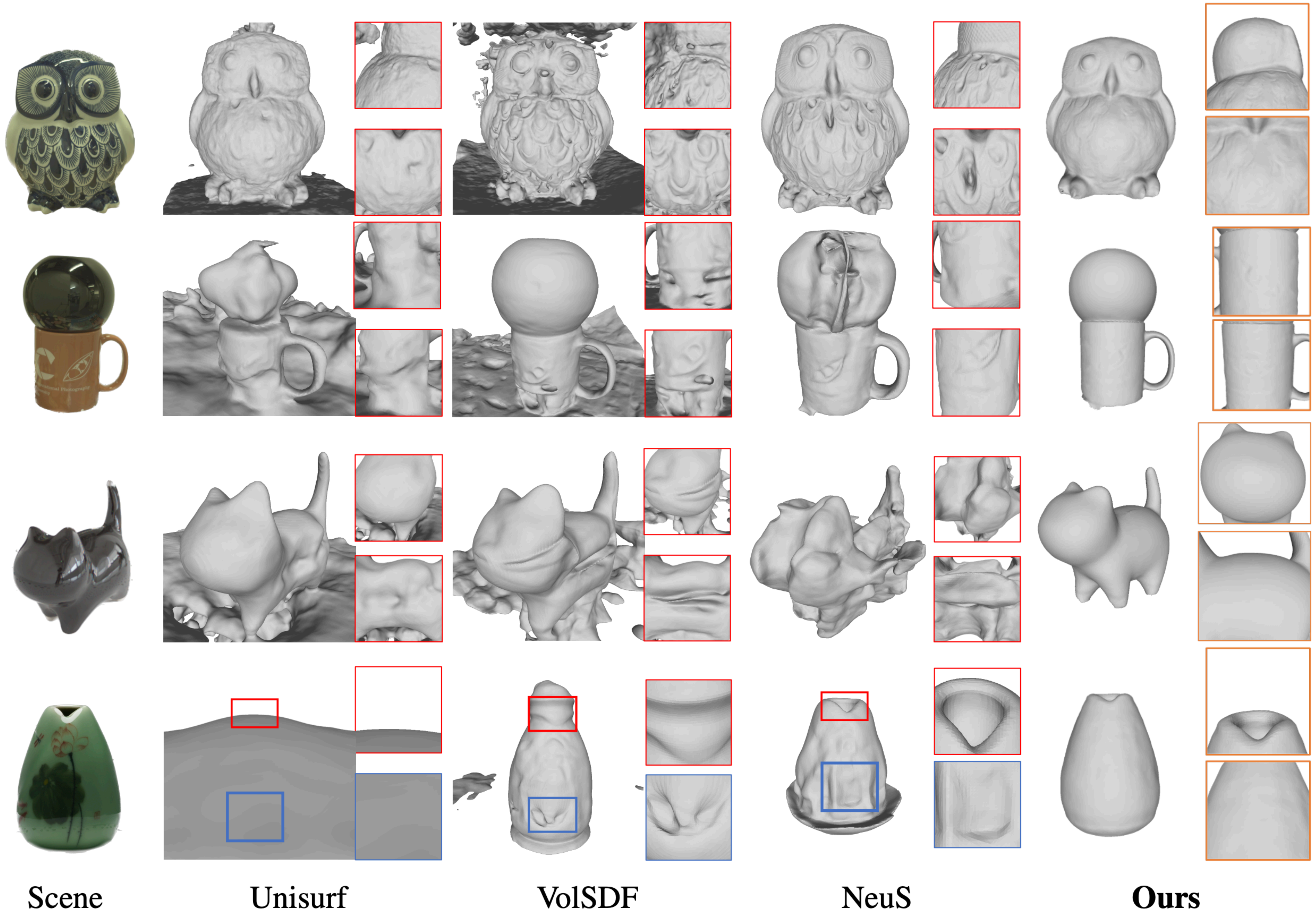






Experiments

SOTA results



Experiments

SOTA results

Methods	**Ironman**	**Duck**	**Cow**	**Snorlax**	**Mean**
Unisurf* (Oechsle et al., 2021)	3.97	10.83	14.43	14.33	10.89
VolSDF (Yariv et al., 2021)	2.72	5.16	5.95	3.20	4.26
NeuS (Wang et al., 2021)	2.28	2.12	3.82	2.11	2.58
Geo-NeuS (Fu et al., 2022)	4.77	10.12	17.48	10.39	10.82
NeuralWarp (Darmon et al., 2022)	12.44	19.78	5.41	20.57	14.55
NeRO* (Liu et al., 2023)	2.29	23.75	2.95	24.30	13.32
Ref-NeuS (Ge et al., 2023)	1.88	1.93	3.66	1.99	2.34
PANDORA † (Dave et al., 2022)	4.61	5.28	7.96	5.73	5.90
GNeRP	**1.34**	**1.63**	**1.39**	**1.05**	**1.35**

Table 1: Quantitative comparison with state-of-the-art methods. The evaluation metric is Chamfer Distance. The lower is better. * indicates the method doesn't use object masks. † refers to the use of polarization priors. The best scores are **bold**, the second best scores are double underlined, and the third best scores are underlined.

Scene	GNeRP (Ours)		Ref-NeuS		NeRO		PANDORA	
	CD ↓	MAE ↓	CD ↓	MAE ↓	CD ↓	MAE ↓	CD ↓	MAE↓
Bunny	**0.72**	**0.78**	1.09	1.03	1.41	2.55	3.77	18.15
Dragon	**0.59**	**1.03**	0.82	1.23	2.15	3.47	5.48	10.98

Experiments

360 video demo