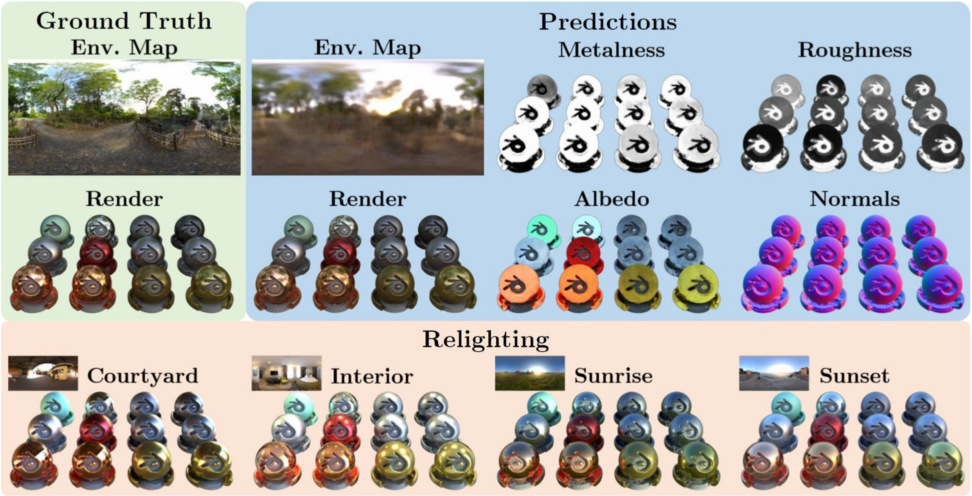
SplitNeRF: Split Sum Approximation Neural Field for Joint Geometry, Illumination, and Material Estimation

Jesus Zarzar, Bernard Ghanem



Introduction

Object Views

Introduction

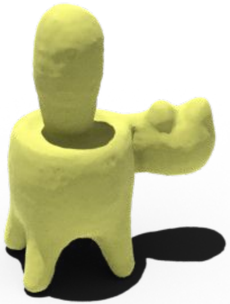
Object Views

Geometry

Introduction

Object Views

Geometry

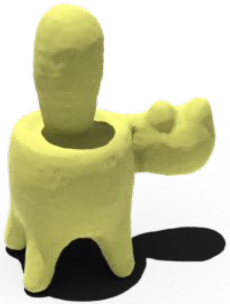


Material Properties

Introduction

Object Views

Geometry



Material Properties

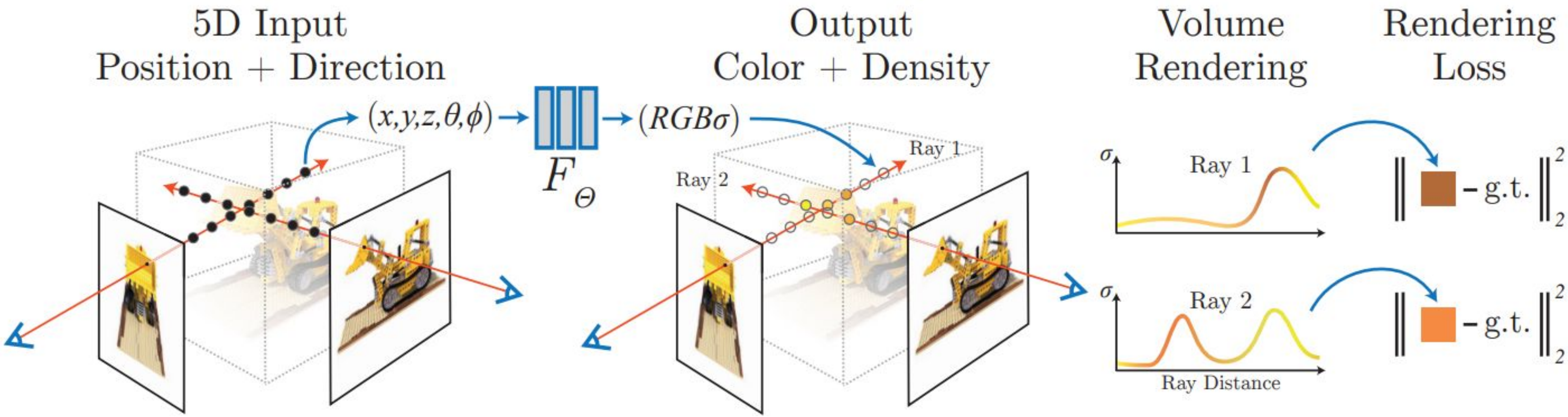
Illumination

Introduction

Object Views

Relighting

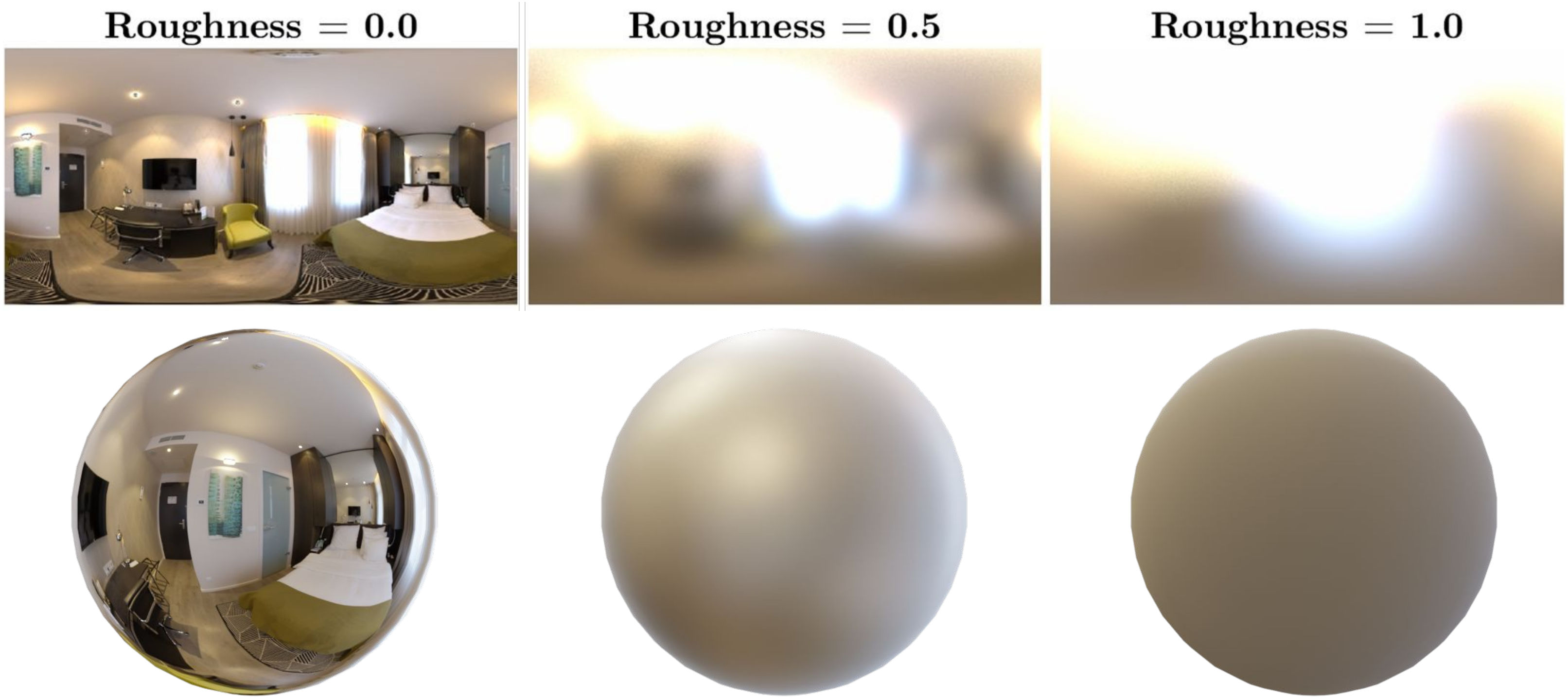
Background: NeRF



$$\hat{C}(\mathbf{r};\theta) = \int_{t_n}^{t_f} T(t)\,\sigma(\mathbf{r}(t))\,\hat{L}_o(\mathbf{r}(t),\mathbf{d})\,\mathrm{d}t, \quad \text{where } T(t) = \exp\left(-\int_{t_n}^{t} \sigma(\mathbf{r}(s))\,\mathrm{d}s\right).$$

Background: PBR Shading

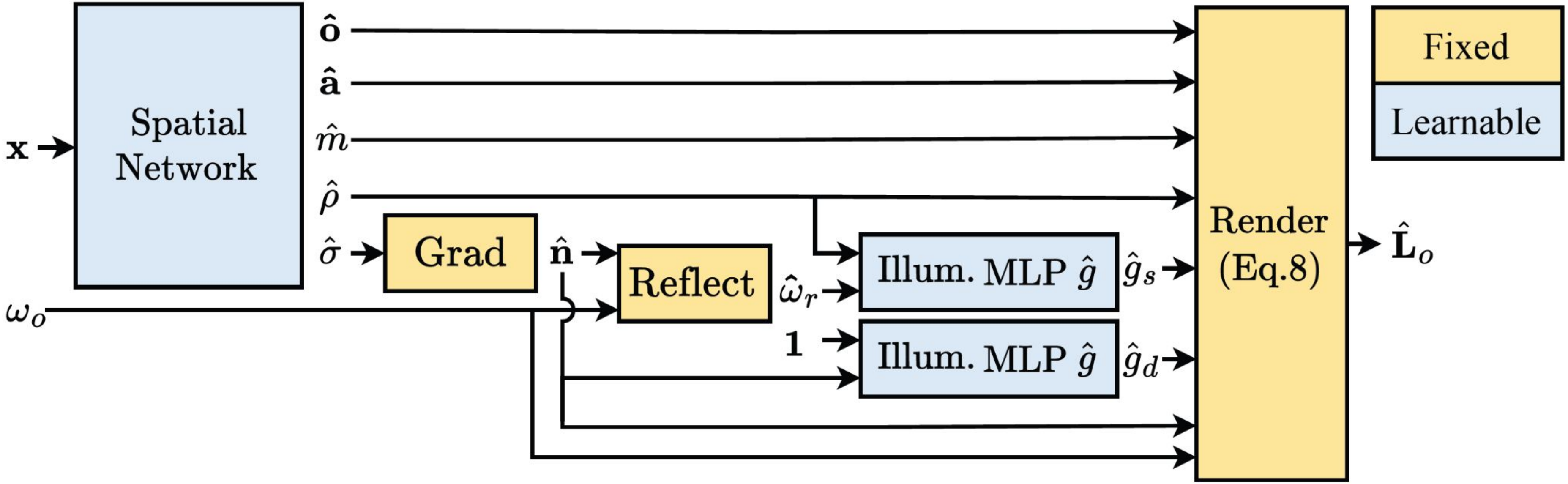
Split-Sum Approximation with Image-Based Lighting



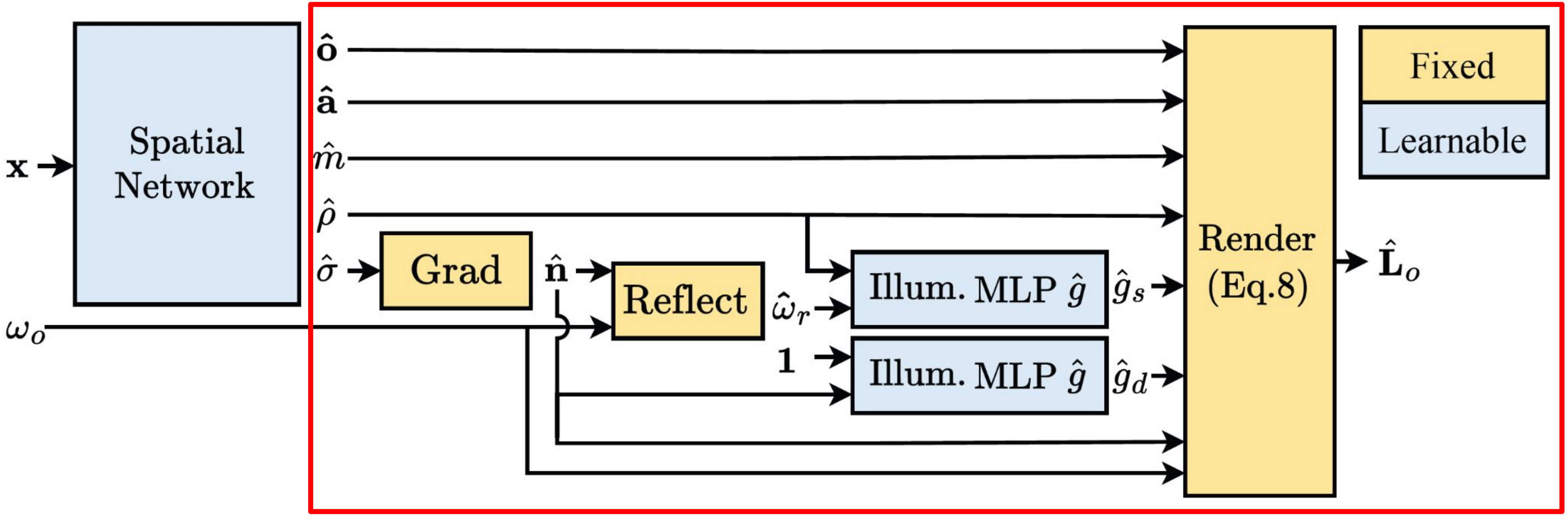
Contributions

1. A novel **MLP representation** for pre-integrated illumination regularized to be **physically accurate**.
2. **Self-occlusion approximation** for pre-integrated lighting with an additional MLP to improve material estimation.
3. Competitive **reconstruction and relighting quality** on both synthetic and **real data** with **~1 GPU-hour** training time.

Method: Pipeline

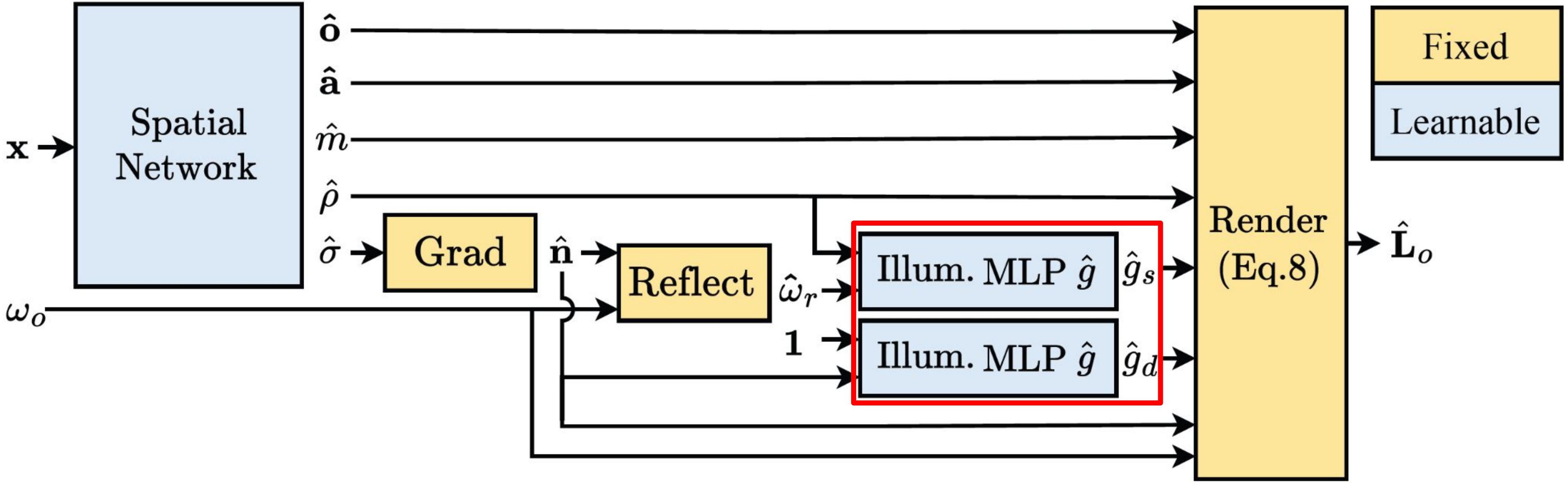


Method: Pipeline



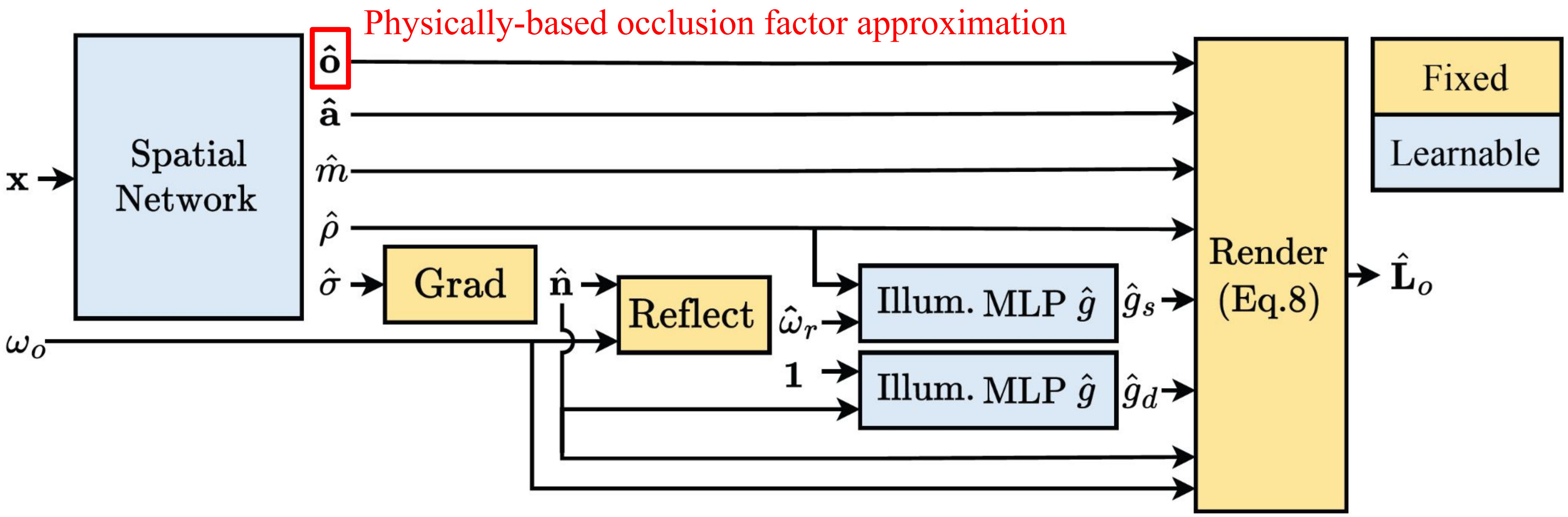
Physically-based radiance prediction

Method: Pipeline

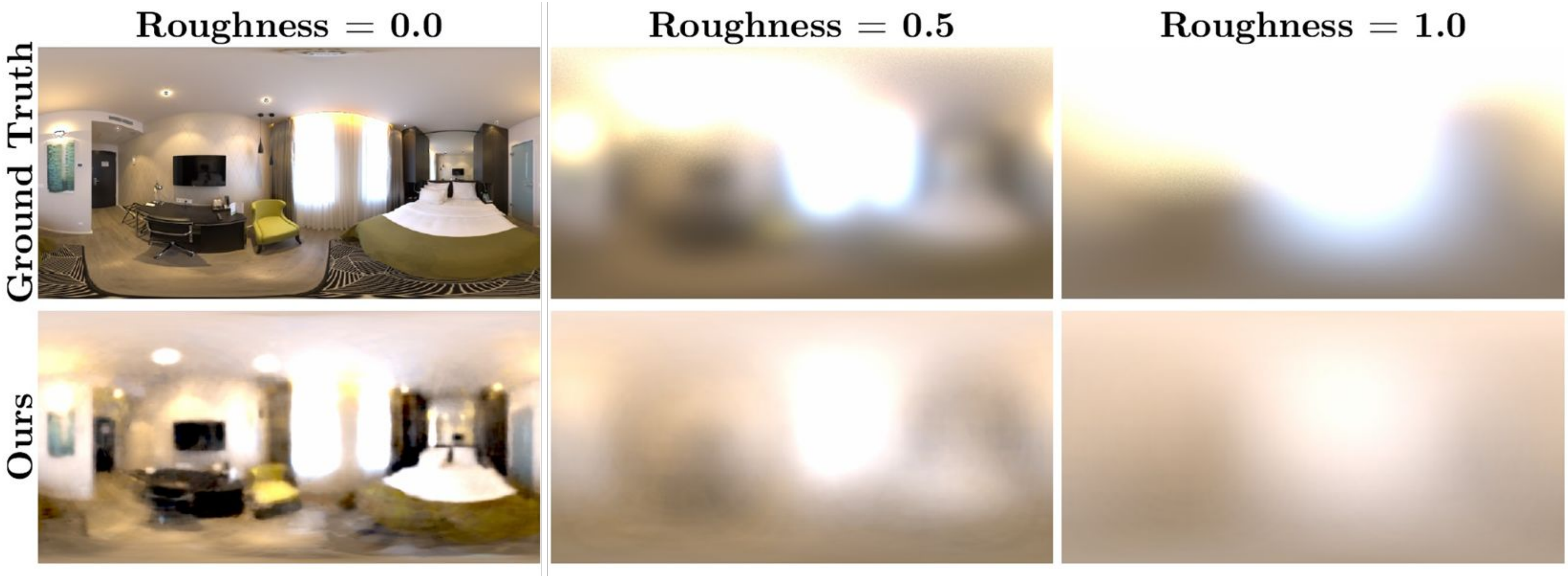


Shared pre-integrated illumination network

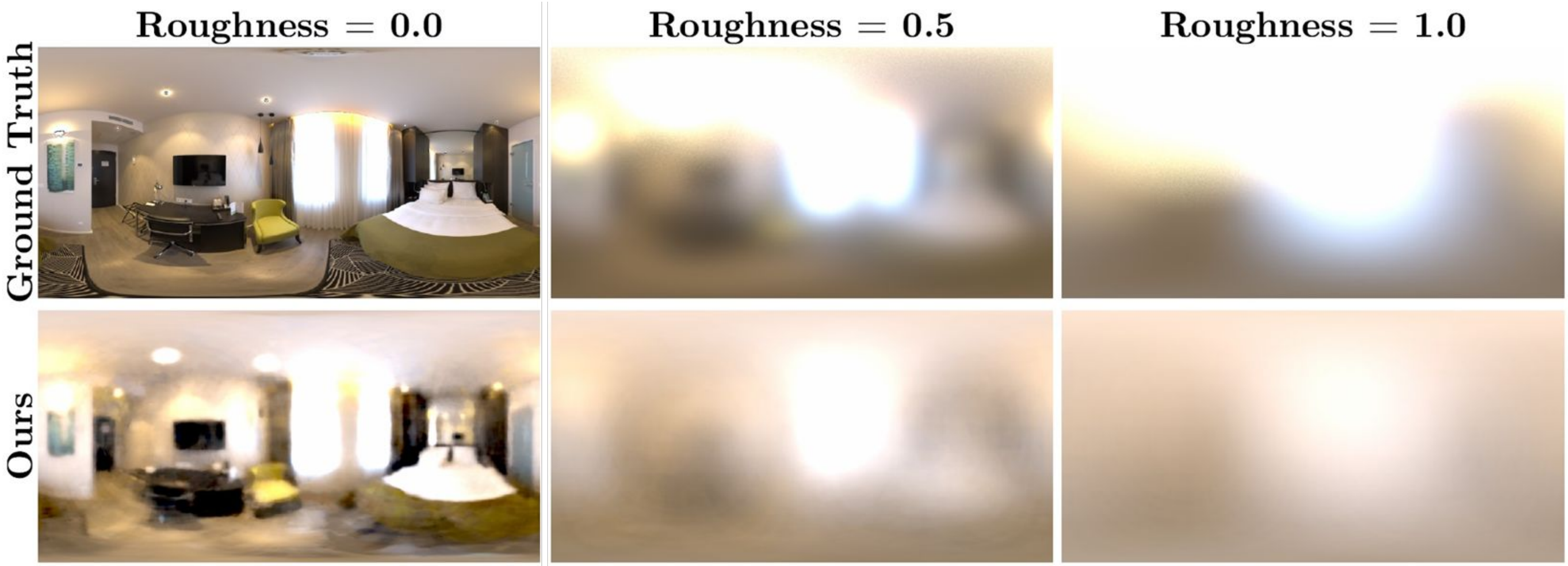
Method: Pipeline



Method: MLP Illumination

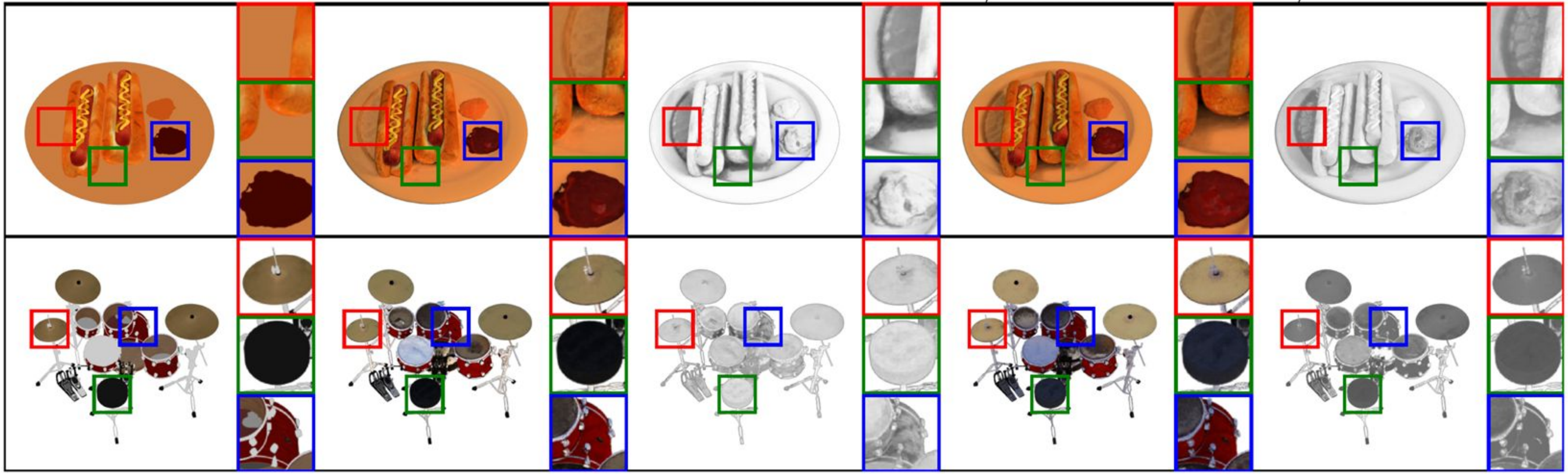


Method: MLP Illumination



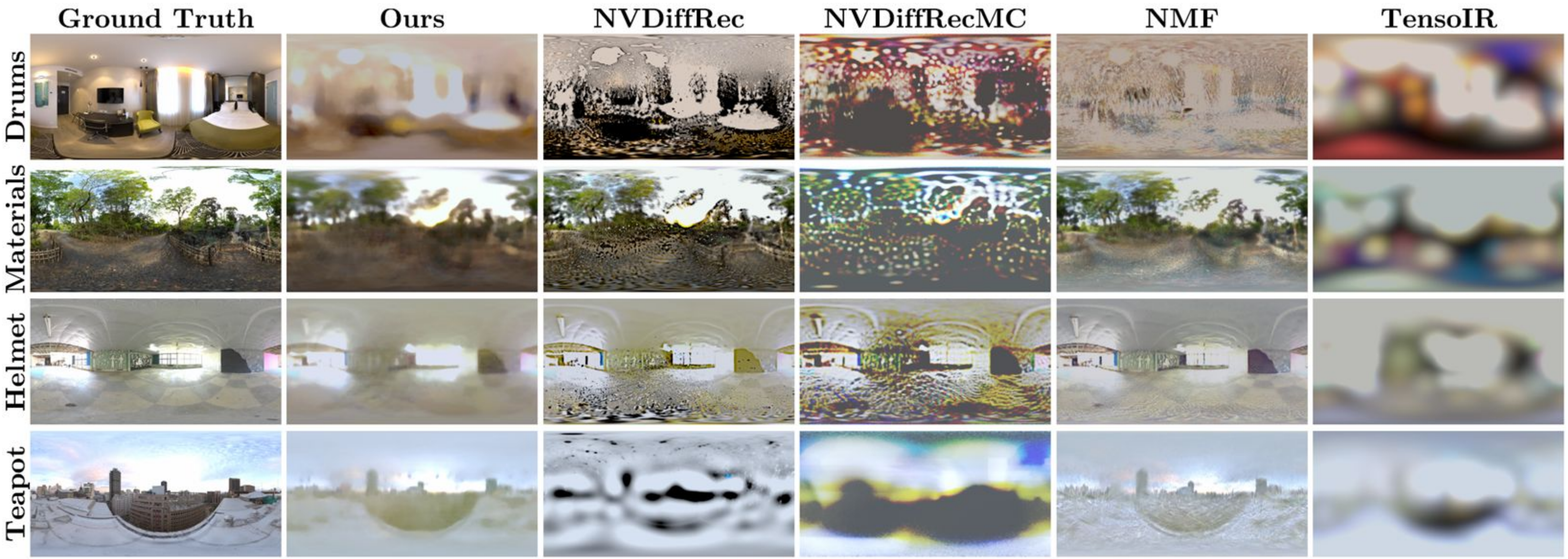
$$\mathcal{L}_D(\theta) = \frac{1}{|\mathcal{S}|}\sum_{s\in\mathcal{S}} |\hat{g}(s) - \bar{g}(s)|_2^2, \quad \bar{g}(s) = \frac{\sum_{\omega_i\in\Omega} D(\omega_i, \omega_s, \rho_s)\hat{g}(\omega_i, 0)\langle\omega_i, \omega_s\rangle}{\sum_{\omega_i\in\Omega} D(\omega_i, \omega_s, \rho_s)\langle\omega_i, \omega_s\rangle}$$

Method: Self-Occlusion Approximation



$$\mathcal{L}_o(\theta) = \frac{1}{|\mathcal{X}|}\sum_{x\in\mathcal{X}} w\,|\hat{o}(x) - \bar{o}(x)|_2^2, \quad \overline{o_d}(x) = \frac{\sum_{\omega_i\in\Omega} L_i\, V_i}{\sum_{\omega_i\in\Omega} L_i}, \quad \bar{o_s}(x) = \frac{\sum_{\omega_i\in\Omega} L_i\, V_i\langle\omega_i, n\rangle}{\sum_{\omega_i\in\Omega} L_i\langle\omega_i, n\rangle}$$

Results: Illumination Estimation



- Predicted illumination for our method and baselines for four different scenes.
- Our representation inherits **smoothness** from the MLP but still captures **high-frequency details** such as indoor objects, trees, and buildings.

Results: Real-Life Object Reconstruction

- Predictions on four scenes from the **real-life** CO3D dataset.
- Our method can successfully recover object geometry, material properties, and illumination for challenging scenes captured in the wild.

Results: Real-Life Object Reconstruction

Method	Normals	Albedo			Relighting			Average
	MAE ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	Runtime
NerFactor	30.49	23.53	0.910	0.109	23.66	0.895	0.120	>20 hr.
NVDiffRec	26.47	23.05	0.901	0.123	21.88	0.880	0.111	0.98 hr.
NVDiffRecMC	25.98	23.84	0.918	0.114	24.06	0.902	0.099	2.95 hr.
NeRO	30.59	22.83	0.897	0.117	23.68	0.907	0.093	18.38 hr.
NMF	24.14	-	-	-	22.23	0.895	0.093	2.91 hr.
TensoIR	22.90	25.21	**0.929**	**0.087**	23.78	0.907	0.100	3.53 hr.
Ours	**17.52**	**25.29**	0.924	0.108	**27.31**	**0.941**	**0.061**	**0.81 hr.**

- Reconstruction and relighting quality of our method vs. baselines on the NeRFactor dataset.
- Our method attains **competitive performances** across all metrics with the **lowest runtime**

Results: Real-Life Object Reconstruction

Method	Normals	Albedo			Relighting			Average
	MAE ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	Runtime
NerFactor	30.49	23.53	0.910	0.109	23.66	0.895	0.120	>20 hr.
NVDiffRec	26.47	23.05	0.901	0.123	21.88	0.880	0.111	0.98 hr.
NVDiffRecMC	25.98	23.84	0.918	0.114	24.06	0.902	0.099	2.95 hr.
NeRO	30.59	22.83	0.897	0.117	23.68	0.907	0.093	18.38 hr.
NMF	24.14	-	-	-	22.23	0.895	0.093	2.91 hr.
TensoIR	22.90	25.21	**0.929**	**0.087**	23.78	0.907	0.100	3.53 hr.
Ours	**17.52**	**25.29**	0.924	0.108	**27.31**	**0.941**	**0.061**	**0.81 hr.**

- Reconstruction and relighting quality of our method vs. baselines on the NeRFactor dataset.
- Our method attains **competitive performances** across all metrics with the **lowest runtime**

Results: Real-Life Object Reconstruction

Method	Normals	Albedo			Relighting			Average
	MAE ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	Runtime
NerFactor	30.49	23.53	0.910	0.109	23.66	0.895	0.120	>20 hr.
NVDiffRec	26.47	23.05	0.901	0.123	21.88	0.880	0.111	0.98 hr.
NVDiffRecMC	25.98	23.84	0.918	0.114	24.06	0.902	0.099	2.95 hr.
NeRO	30.59	22.83	0.897	0.117	23.68	0.907	0.093	18.38 hr.
NMF	24.14	-	-	-	22.23	0.895	0.093	2.91 hr.
TensoIR	22.90	25.21	**0.929**	**0.087**	23.78	0.907	0.100	3.53 hr.
Ours	**17.52**	**25.29**	0.924	0.108	**27.31**	**0.941**	**0.061**	**0.81 hr.**

- Reconstruction and relighting quality of our method vs. baselines on the NeRFactor dataset.
- Our method attains **competitive performances** across all metrics with the **lowest runtime**

Results: Real-Life Object Reconstruction

Method	Normals	Albedo			Relighting			Average
	MAE ↓	PSNR ↑	SSIM ↑	LPIPS ↓	PSNR ↑	SSIM ↑	LPIPS ↓	Runtime
NerFactor	30.49	23.53	0.910	0.109	23.66	0.895	0.120	>20 hr.
NVDiffRec	26.47	23.05	0.901	0.123	21.88	0.880	0.111	0.98 hr.
NVDiffRecMC	25.98	23.84	0.918	0.114	24.06	0.902	0.099	2.95 hr.
NeRO	30.59	22.83	0.897	0.117	23.68	0.907	0.093	18.38 hr.
NMF	24.14	-	-	-	22.23	0.895	0.093	2.91 hr.
TensoIR	22.90	25.21	**0.929**	**0.087**	23.78	0.907	0.100	3.53 hr.
Ours	**17.52**	**25.29**	0.924	0.108	**27.31**	**0.941**	**0.061**	$\boxed{\textbf{0.81 hr.}}$

- Reconstruction and relighting quality of our method vs. baselines on the NeRFactor dataset.
- Our method attains **competitive performances** across all metrics with the **lowest runtime**

SplitNeRF: Split Sum Approximation Neural Field for Joint Geometry, Illumination, and Material Estimation

Jesus Zarzar, Bernard Ghanem