LightSwitch: Multi-view Relighting with Material-guided Diffusion

Supplementary Material

Method	Relighting			
	PSNR-L ↑	PSNR-H ↑	SSIM ↑	LPIPS ↓
LightSwitch	32.02	25.03	0.976	0.027
Neural-PBIR	33.26	26.01	0.979	0.023
IllumiNeRF	32.74	25.56	0.976	0.027
NVDiffrecMC	31.60	24.43	0.972	0.036
RelitLRM	31.52	24.67	0.969	0.032
InvRender	30.83	23.76	0.970	0.046
NeRFactor	30.38	23.54	0.969	0.048
NVDiffrec	29.72	22.91	0.963	0.039

Table 1. 3D Relighting Comparison on Stanford-ORB.

1. Additional Visualizations

We show additional visualizations of LightSwitch's 2D and 3D relighting on BlenderVault 2D data as well as NeRF-Synthetic in Figs. 1-4.

2. Additional Details

Training. LightSwitch was trained in three stages using 8 RTX A6000 GPUs, first by finetuning for single view for 20K iterations on 256×256 data with a batch size of 512. An AdamW 8-bit optimizer was used with a learning rate of $5e-5$. For the second multi-view stage, we train with a batch size of 120, with each batch containing four 256×256 images randomly sampled for the object. This was done for 15K iterations with a learning rate of $2.5e-5$. For the last upsampling stage, a batch size of 28 was used, where each batch contained four 512×512 images. This was done with a learning for 15K iterations with a learning rate of $1e-5$. We repeat the upsampling stage training for Stable-MaterialMV in order to create higher quality material maps.

Stanford-ORB Relighting Evaluation. We report results in Tab. 1, showing LightSwitch is competitive with SOTA while performing significantly faster – relighting a scene in 8 minutes vs. several hours for baselines. Due to lighting changes when moving the capture device, Stanford-ORB has *separate* environment maps for each test view. As our multi-view method relights query views under a common illumination, we disregard the variation and assume the lighting for the first image but this may be suboptimal.

	Mic		Hotdog		Chair		Lego		Materials	
Method	PSNR	LPIPS	PSNR	LPIPS	PSNR	LPIPS	PSNR	LPIPS	PSNR	LPIPS
LightSwitch	30.24	0.025	25.91	0.090	26.65	0.062	23.60	0.081	22.08	0.080
MaterialFusion	30.46	0.045	23.09	0.153	25.40	0.084	21.87	0.145	20.47	0.158
NVDiffrecMC	29.81	0.052	22.88	0.159	25.39	0.083	22.04	0.141	20.50	0.157
TensoIR	30.92	0.024	21.12	0.179	24.82	0.082	21.37	0.100	22.01	0.107
R3DGS	28.87	0.033	20.89	0.179	23.08	0.084	20.38	0.129	20.48	0.101

Table 2. **Relighting on the NeRF-Synthetic Dataset.** We report the performance of all other baselines when their images are encoded/decoded using the Stable Diffusion encoder and decoder before comparison. The VAE causes significant drops in relighting quality for all objects, especially those with reflections sharp fine reflections such as `materials`, which explains our method's struggle on the object.

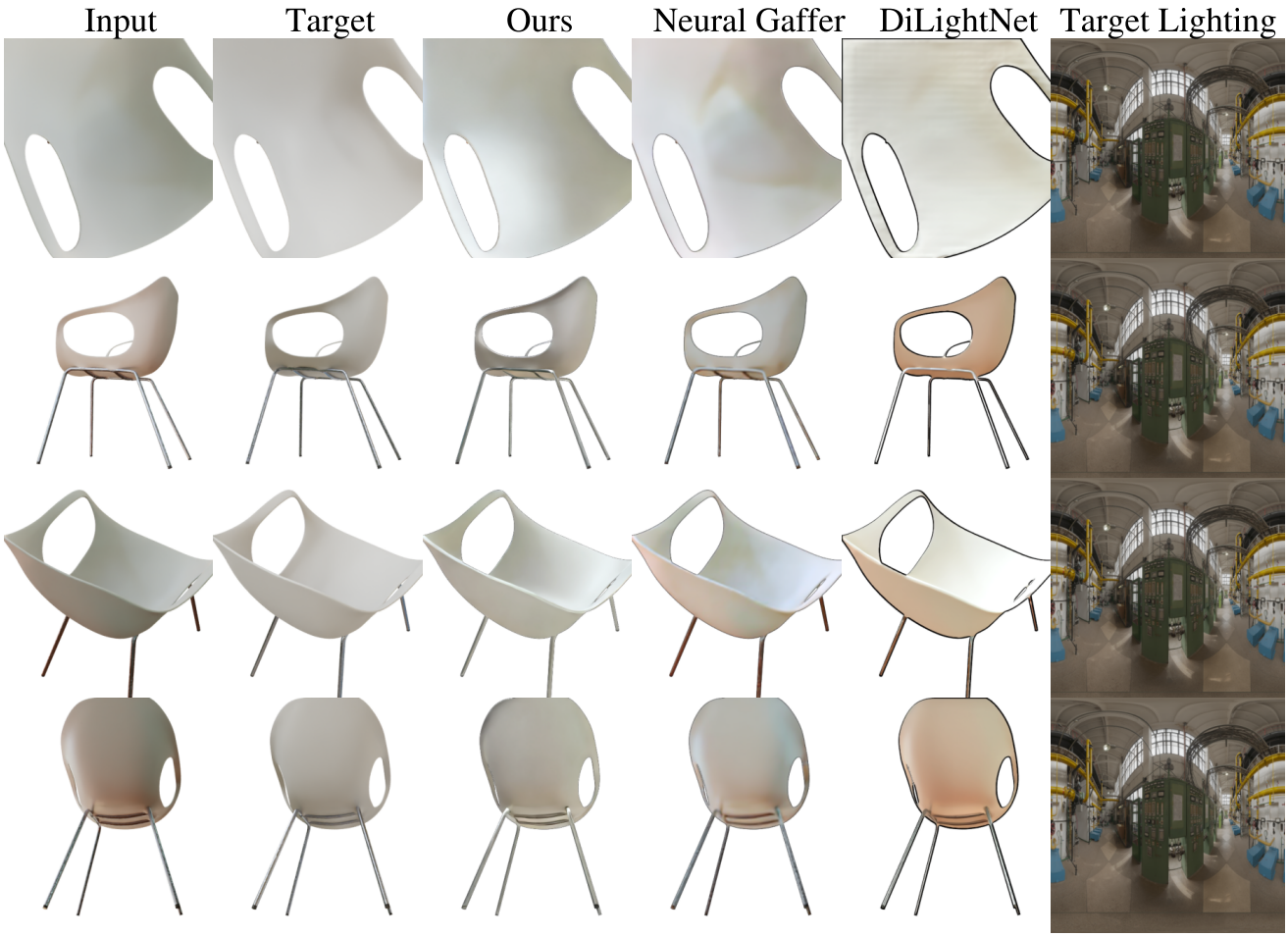

Figure 1. **Additional Visualizations of LightSwitch Relighting on Synthetic Objects.**

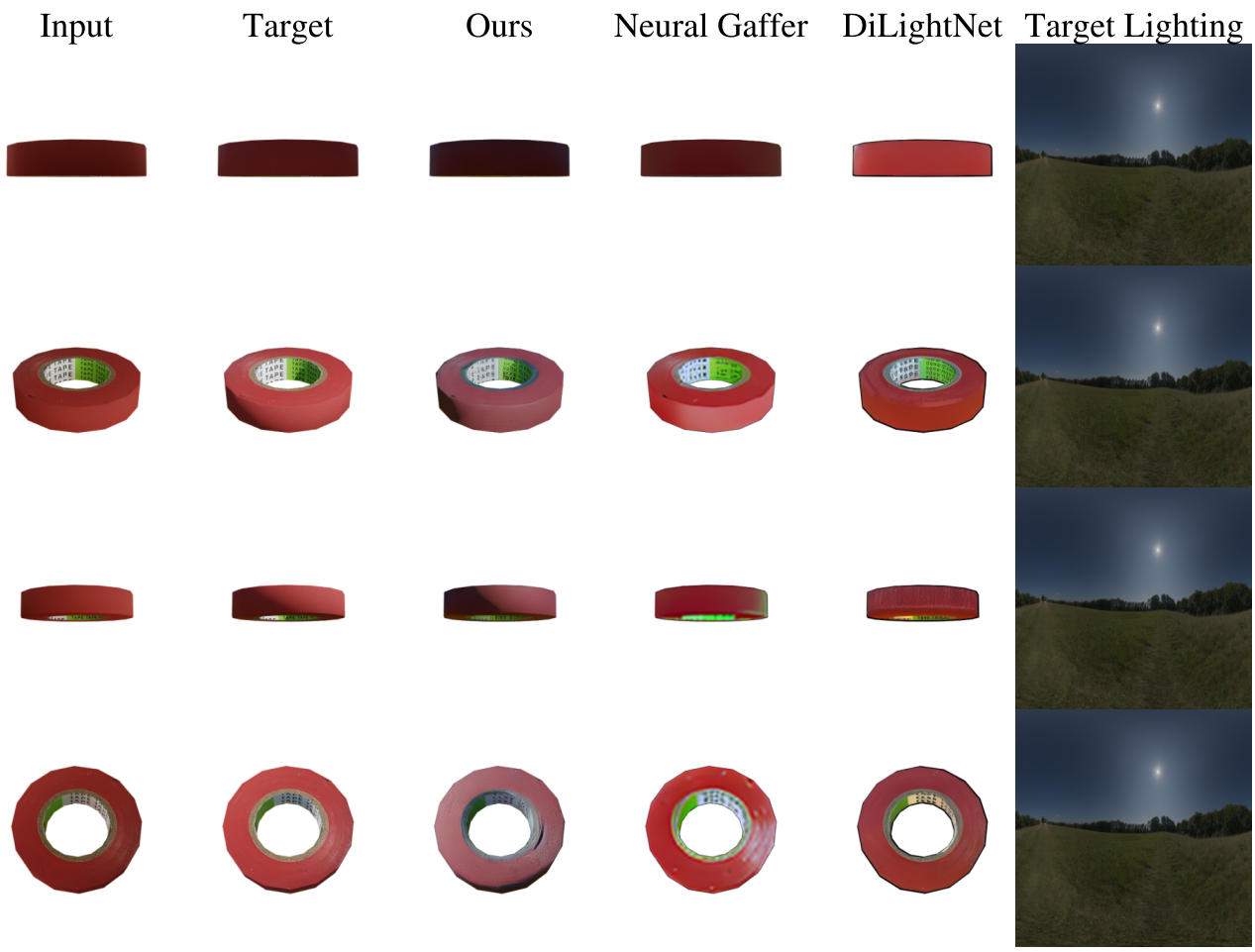

Figure 2. **Additional Visualizations of LightSwitch Relighting on Synthetic Objects.**

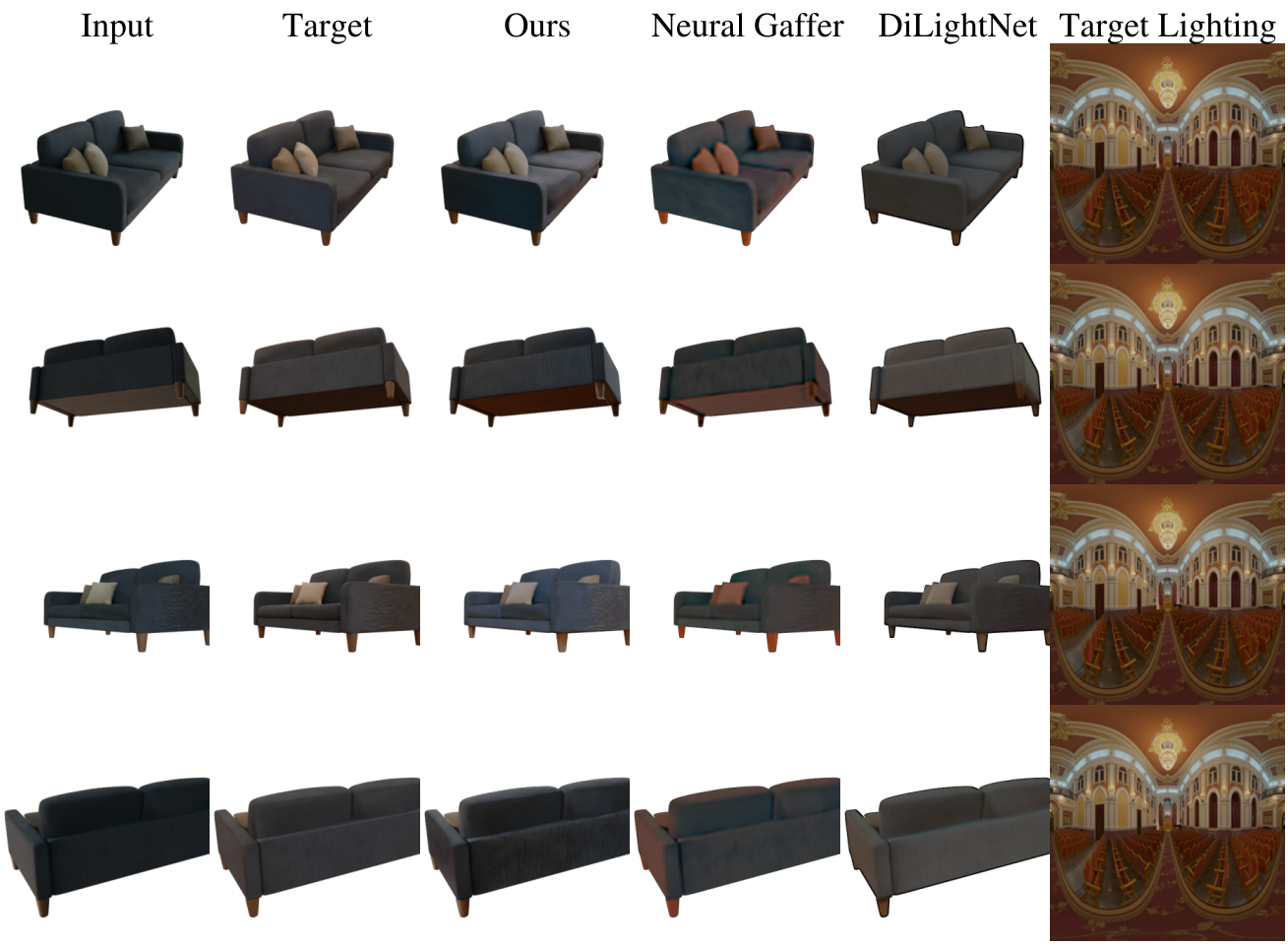

Figure 3. **Additional Visualizations of LightSwitch Relighting on Synthetic Objects.**

Figure 4. **Additional Visualizations of LightSwitch 3D Relightings on NeRF-Synthetic.**