Co3D-10-Category	PSNR↑ / LPIPS↓	FID↓	Noise Ablation	PSNR↑ / LPIPS↓	FID↓	2D Inpainting	PSNR↑ / LPIPS↓	FID↓
pixelNeRF	**17.96** / 0.479	158.50	With noise	17.24 / 0.40	92.23	Determ.	**21.35** / 0.11	9.18
SparseFusion	11.76 / 0.770	257.63	Ours	**18.19 / 0.34**	**56.64**	Ours	20.18 / **0.09**	**4.25**
Ours	17.62 / **0.368**	**66.81**						

Table 1: (left) Evaluation for 10 categories of Co3D. (middle) Ablation with and without noise added to camera poses. (right) 2D inpainting results.

Figure 1: Ablation experiment. Using noisy camera poses at training time leads to lower-quality renderings.
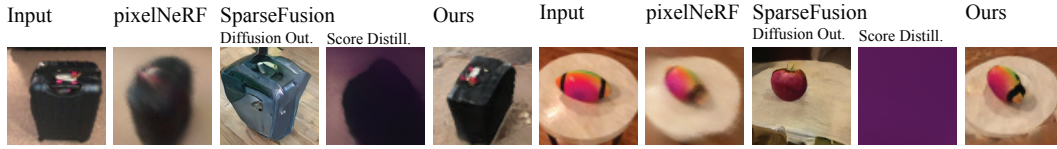

Figure 2: Results for a general model trained on 10 Co3D categories.
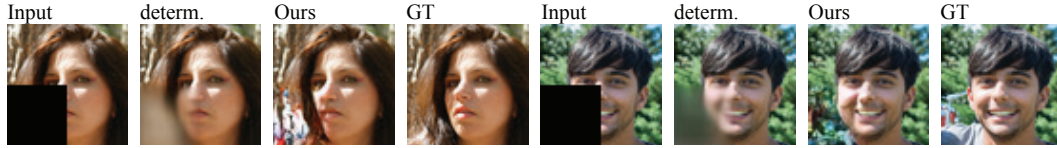

Figure 3: Results on the Objaverse dataset.

Figure 4: Inpainting experiment. Our training dataset has images like the input, with random patches missing. At test time, we can learn to complete the partial observations. Deterministic baseline learns a blurry completion.
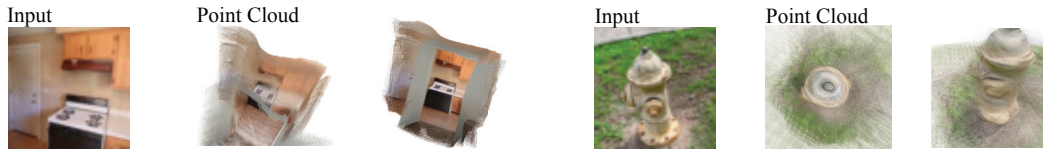

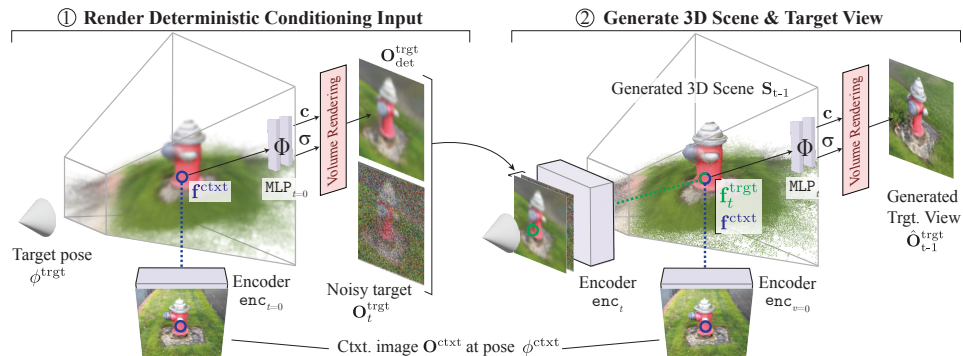Figure 5: Extracted point clouds demonstrate 3D consistent reconstructions, also see point cloud in Fig. 6

Figure 6: Updated Pipeline figure that depicts our inverse graphics pipeline.