

Supplementary Material: Deep View Synthesis From Colored 3D Point Clouds

Zhenbo Song¹, Wayne Chen², Dylan Campbell^{2,3}, and Hongdong Li^{2,3}

¹ Nanjing University of Science and Technology, Nanjing, Jiangsu, China

² Australian National University, Canberra ACT 0200, Australia

³ Australian Centre for Robotic Vision

1 Video Link

A video rendering a smooth camera trajectory for the ICL-NUIM dataset and a discretized camera trajectory for the NYU dataset, as well as additional qualitative results, is available at the following URL: <https://youtu.be/Ja9jCPxR744>.

2 Ablation Study

To demonstrate the importance of the coarse network, we trained RefineNet without the coarse network input (RGBD input only) and show in Table 1 that it performs significantly worse than the full system on the NYU V2 dataset. While the performance of the coarse network is quantitatively similar to *invsgm* [?], it is qualitatively very different and provides orthogonal information. The coarse network recovers the main structure of the scene, that is, the spatial layout, while RefineNet fills in the details for realism.

Table 1. Ablation study for the coarse network.

	Max Points	MAE ↓	PSNR ↑	SSIM ↑
<i>invsgm</i>	>20000	0.15	14.5	0.54
Ours-RefineNet-Only	4096	0.16	13.8	0.38
Ours-Coarse	4096	0.15	14.3	0.41
Ours-Refined	8192	0.12	16.4	0.57