Supplementary material for: P^{2}Net: Patch-match and Plane-regularization for Unsupervised Indoor Depth Estimation

Zehao Yu*[1,2], Lei Jin*[1,2], and Shenghua Gao†[1,3]

[1] ShanghaiTech Univsertiy
[2] DGene Inc
[3] Shanghai Engineering Research Center of Intelligent Vision and Imaging
{yuzh,jinlei,gaoshh}@shanghaitech.edu.cn
https://github.com/svip-lab/Indoor-SfMLearner

1 Surface normal visualization

We provide more visualizations of surface normal prediction on the ScanNet [1] dataset. In our implementation, we directly fit the surface normal from ground truth depth annotation. Black pixels indicate invalid regions where no ground truth depths are provided. Compared to MovingIndoor [3], our surface normal estimation better preserves the boundary of the planar regions, thanks to our superpixel constraint.

2 Point clouds visualization

We further provide some point clouds visualization on NYUv2 [2] and ScanNet [1] dataset in Figure 2.

3 The effect of different patterns.

We compare the effect of different patterns in our Patch-match module. We experiment with different Ns and report the result in Table 1. Setting N to 3 gives best results.

∗ Equal Contribution
† Corresponding author

N	rms ↓	rel ↓	$\delta < 1.25$ ↑	$\delta < 1.25^2$ ↑	$\delta < 1.25^3$ ↑
1	0.629	0.173	0.746	0.939	0.984
2	0.618	0.170	0.748	0.937	0.984
3	**0.612**	**0.166**	**0.758**	**0.945**	**0.985**
4	0.634	0.173	0.741	0.938	0.984

Table 1: Comparison between different patterns in our Patch-match module.

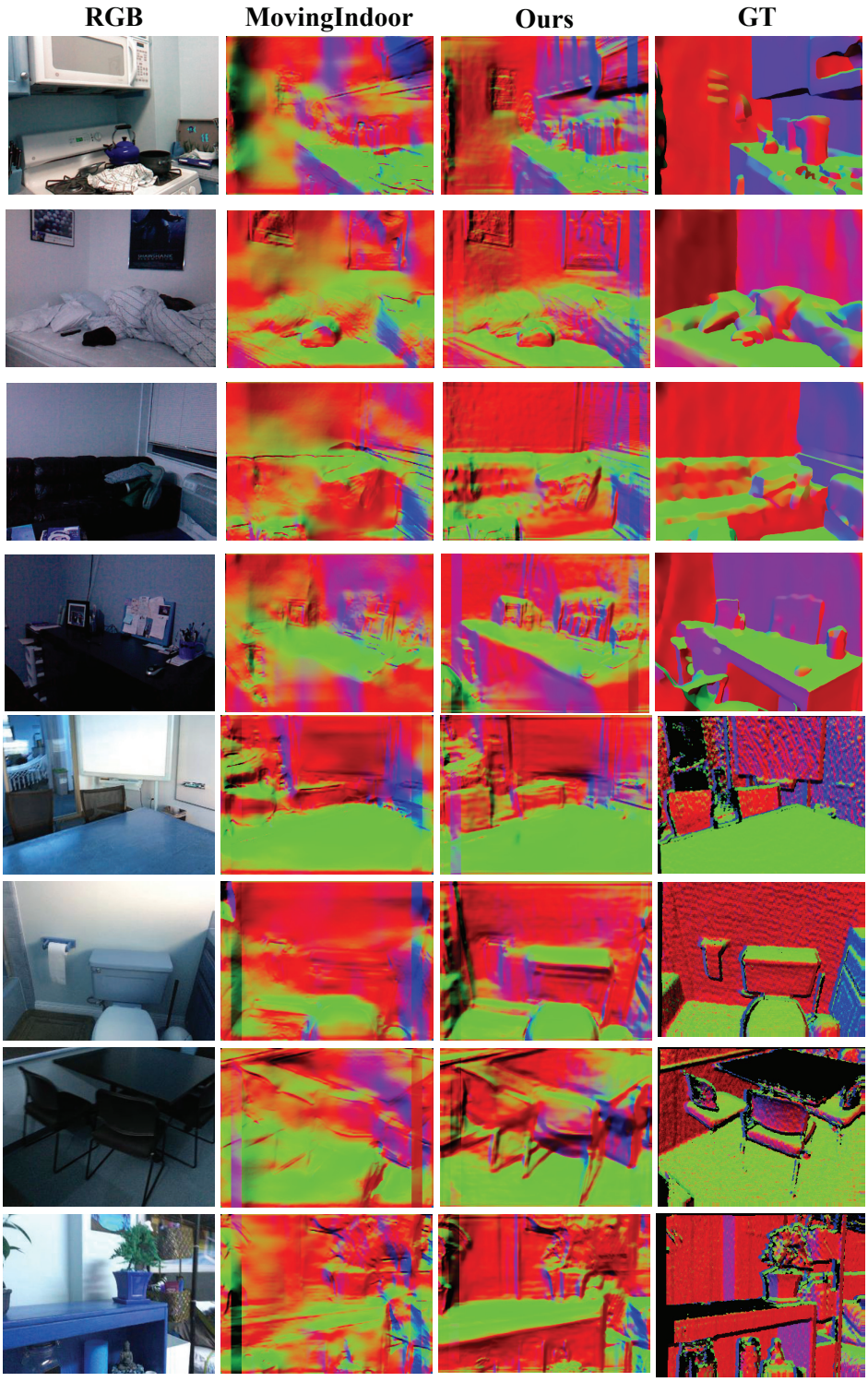

Fig. 1: Visualization of surface normal results on the ScanNet [1] dataset. From left to right: input RGB, MovingIndoor [3], our results and surface normal fitted from ground truth depth. Black pixels in ground truth indicate invalid regions where no depth ground truth are provided.

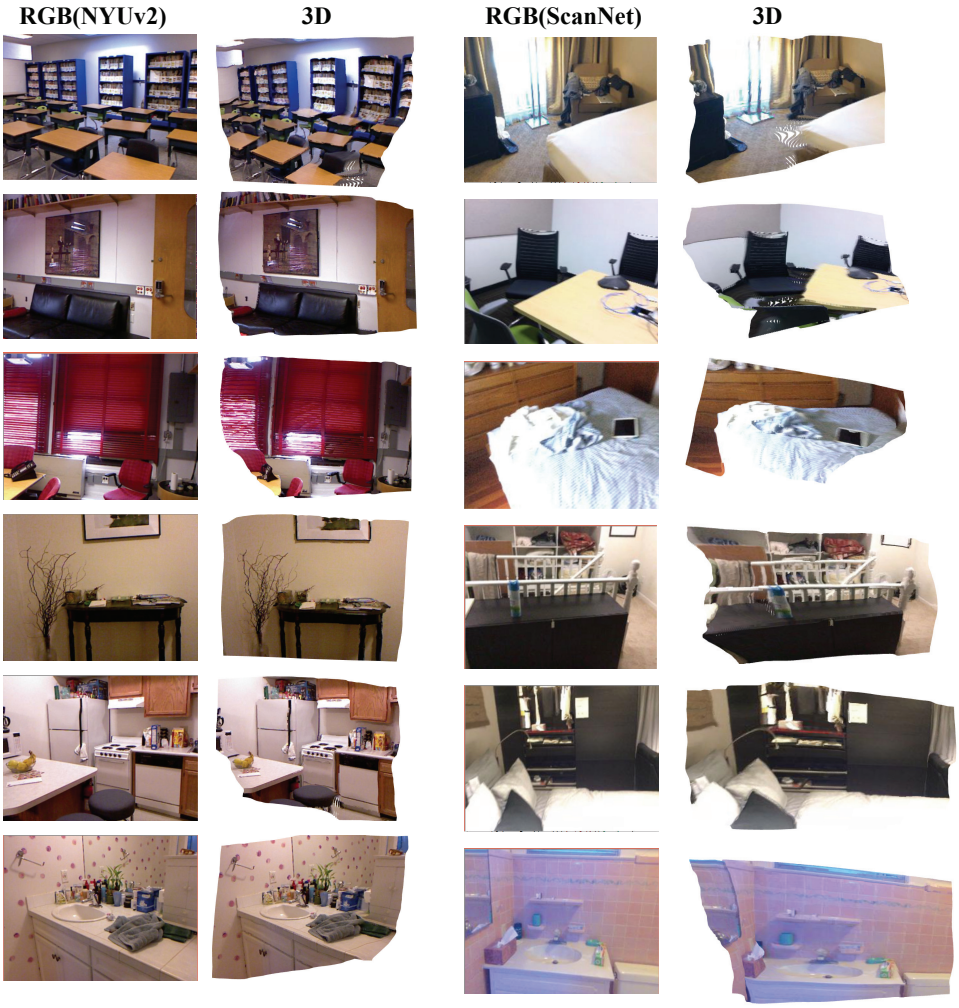

Fig. 2: Point cloud visualization. From left to right: input RGB from NYUv2, point cloud in 3D, RGB from ScanNet, point cloud in 3D.

References

1. Dai, A., Chang, A.X., Savva, M., Halber, M., Funkhouser, T., Nießner, M.: Scannet: Richly-annotated 3d reconstructions of indoor scenes. In: CVPR (2017)
2. Silberman, N., Hoiem, D., Kohli, P., Fergus, R.: Indoor segmentation and support inference from rgbd images. In: ECCV (2012)
3. Zhou, J., Wang, Y., Qin, K., Zeng, W.: Moving indoor: Unsupervised video depth learning in challenging environments. In: ICCV (2019)