On the Road to Large-Scale 3D Monocular Scene Reconstruction using Deep Implicit Functions

Supplementary Material

Paper ID 22

A. Detailed Chamfer-L1 scores across object categories

Table 1 of the main paper provided the main quantitative comparison between our method and the two baseline methods. For each metric we reported two values, corresponding to the quality of reconstruction evaluated across the entire scene mesh, as well as reconstruction quality evaluated at the object level, aggregated across all object categories. In Table A1 we provide additional detail about the latter quantity, by providing a breakdown of the Chamfer-L1 for each object category. We find that we outperform the baseline methods across the majority of object categories in the NuScenes dataset.

Table A1. Chamfer L1 scores (m) across individual object categories on the NuScenes validation set. For this metric, lower scores are better. The 'Mean' column refers to the mean across all object categories, which corresponds to the Chamfer-L1 'Object' column in Table 1 of the main paper.

Method	Car	Truck	Bus	Trailer	Construct. vehicle	Pedestrian	Motorcycle	Bicycle	Traffic cone	Barrier	Other	Mean
Bounding box mesh	0.473	0.665	0.659	0.833	0.768	**0.495**	0.508	0.503	**0.345**	0.453	**0.396**	0.554
PIFu	0.378	0.527	0.524	0.666	0.718	0.571	0.548	0.568	0.448	0.407	0.455	0.528
Ours (GT boxes)	0.333	0.450	0.424	0.538	0.593	0.337	0.243	0.411	0.376	0.356	0.223	0.389
Ours	**0.372**	**0.525**	**0.497**	**0.665**	**0.621**	0.529	**0.455**	**0.465**	0.389	**0.387**	0.409	**0.483**

B. Additional qualitative results

Figure A1 provides additional qualitative examples of the reconstructions produced by our method. For further results, please see the attached demonstration video.

C. Additional comparisons to baselines

We extend the qualitative comparisons between our method and the two baselines: PIFu [30] and a bounding box mesh reconstruction, in Figure A2.

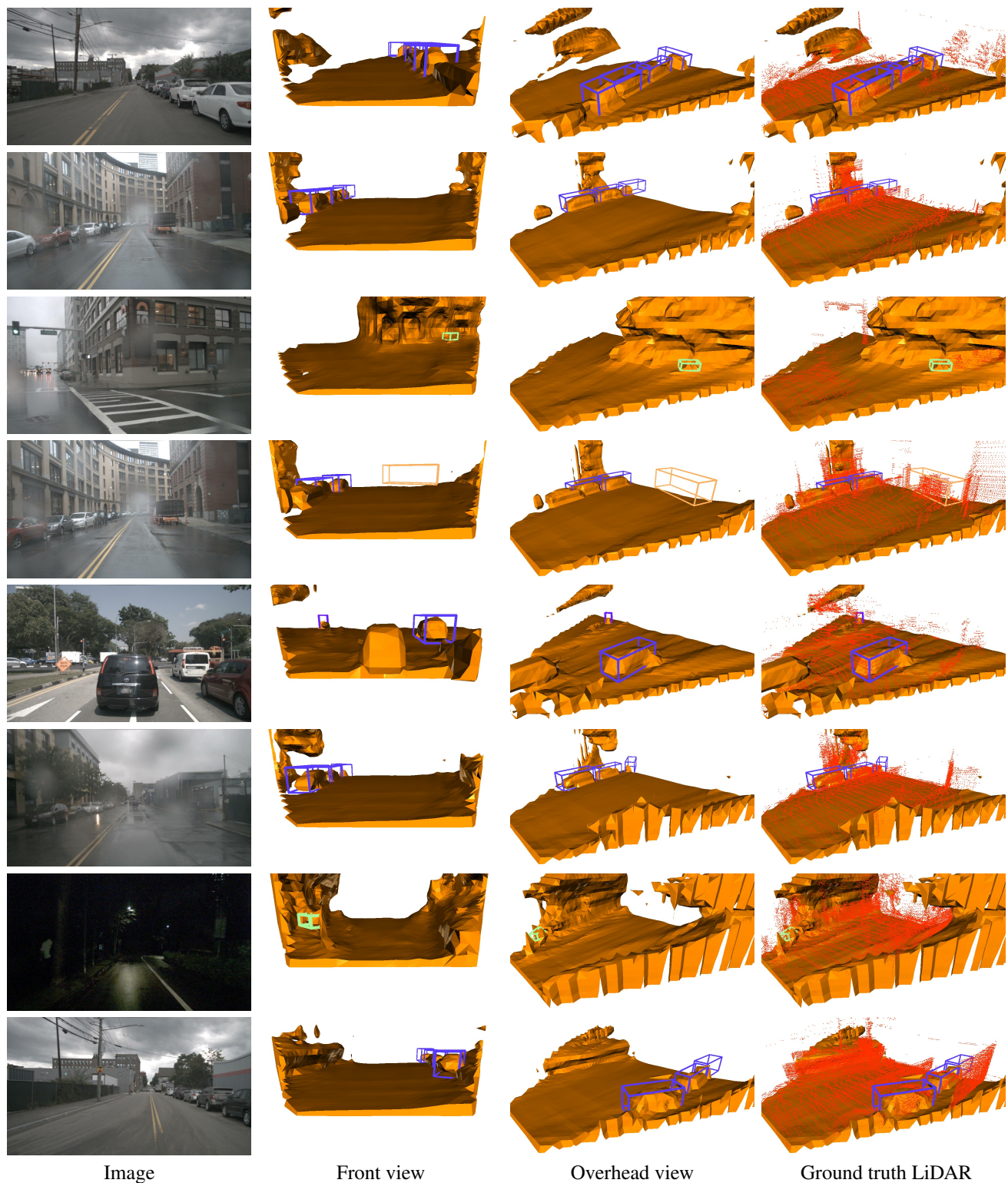| Image | Front view | Overhead view | Ground truth LiDAR |

Figure A1. Additional qualitative examples of our method on the NuScenes validation set, which supplement the results in Figure 4 of the main paper. We show reconstructions produced by our method from the camera perspective and alternative side view. Predicted object bounding boxes, which are generated as part of our method, are shown in blue. We also show the set of densified LiDAR points in red, which provide the ground truth reference points used for evaluation.
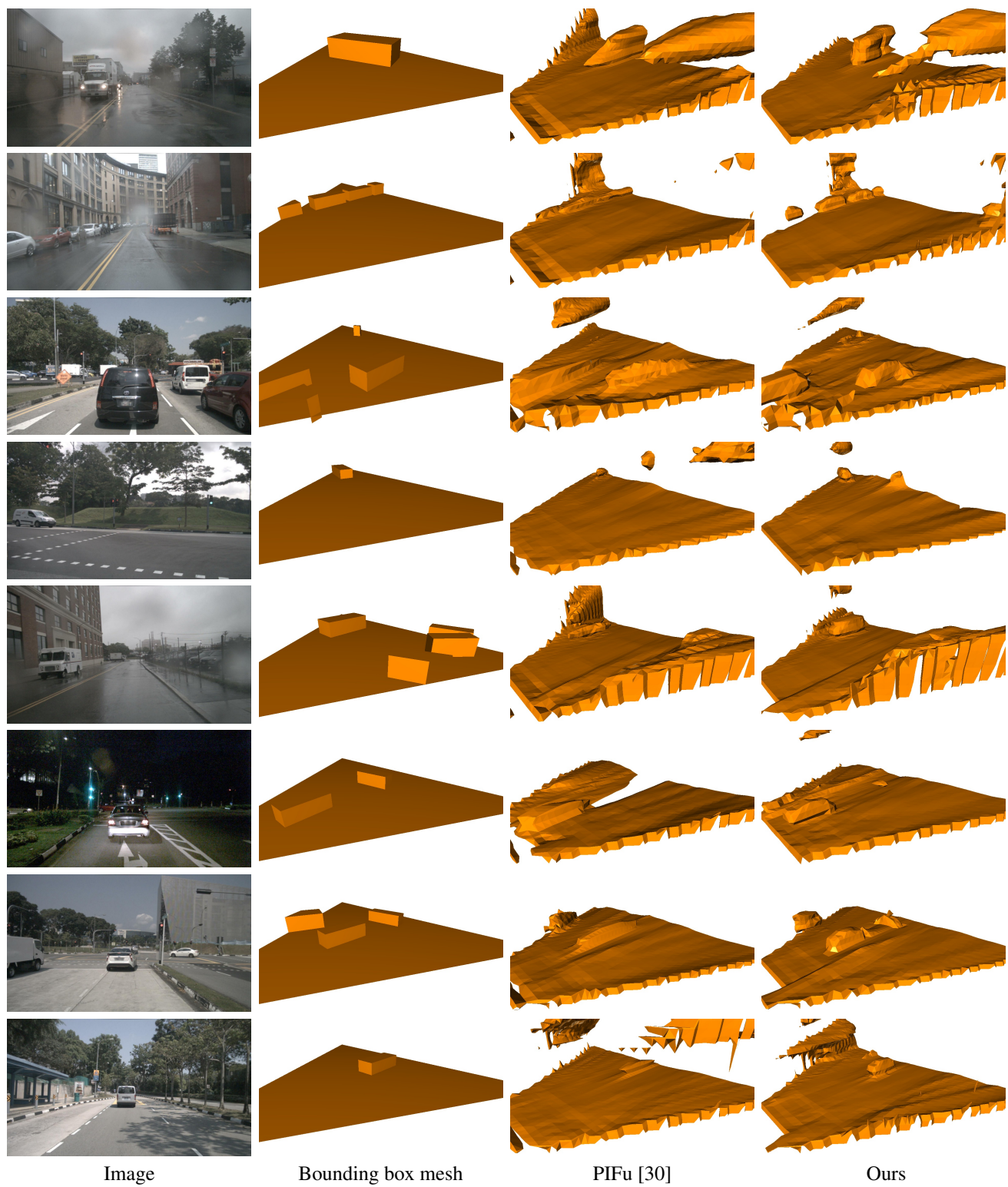
| Image | Bounding box mesh | PIFu [30] | Ours |

Figure A2. Additional qualitative comparisons of ours and baseline approaches on the NuScenes validation set, which extend the results from Figure 5 of the main paper.