Supplementary Material: Polarimetric Imaging for Perception

Michael Baltaxe
michael.baltaxe@gm.com
Tomer Pe'er
tomer.1.peer@gm.com
Dan Levi
dan.levi@gm.com

General Motors
Hertzliya, Israel

1 Dataset Details

The dataset was collected at 6 different localities in Israel. All the locations are small cities with well-defined roads (pavemented, clear sidewalks, standardized parking locations) and with relatively little traffic (to support the self-supervised monodepth framework). The dataset was collected during the month of June, thus, some inclination of the Sun is always present in the data. The 12,627 samples come from the sequential 10 Hz recordings. Additional samples of the dataset are presented in figure 1.

2 Implementation Details

2.1 Free Space Detection

All the Road-Seg networks [2] were trained with an initial learning rate of 0.05. The learning rate was reduced by a ratio of 0.9 (as in the the original SNE-RoadSeg code) every 10 epochs. The training batch size was 8 and the loss used was binary cross entropy. The networks were trained until the evaluation metrics stopped improving on the validation set, which usually occurred after about 100 epochs. The architecture of the two streams of the Road-Seg network was a ResNet-152 with all weights initialized randomly to avoid biases coming from pre-training with large RGB datasets (like ImageNet [1]), compared to the smaller size of the polarimetric data.

2.2 Monocular Depth Estimation

All the monodepth v2 models [3] were trained with the standard parameters provided by the official code. That is, the learning rate was 1e-4 and the batch size was 12. The depth and pose encoders used were ResNet-18 networks. The networks were trained until the evaluation metrics stopped improving on the validation set, usually after about 60 epochs.

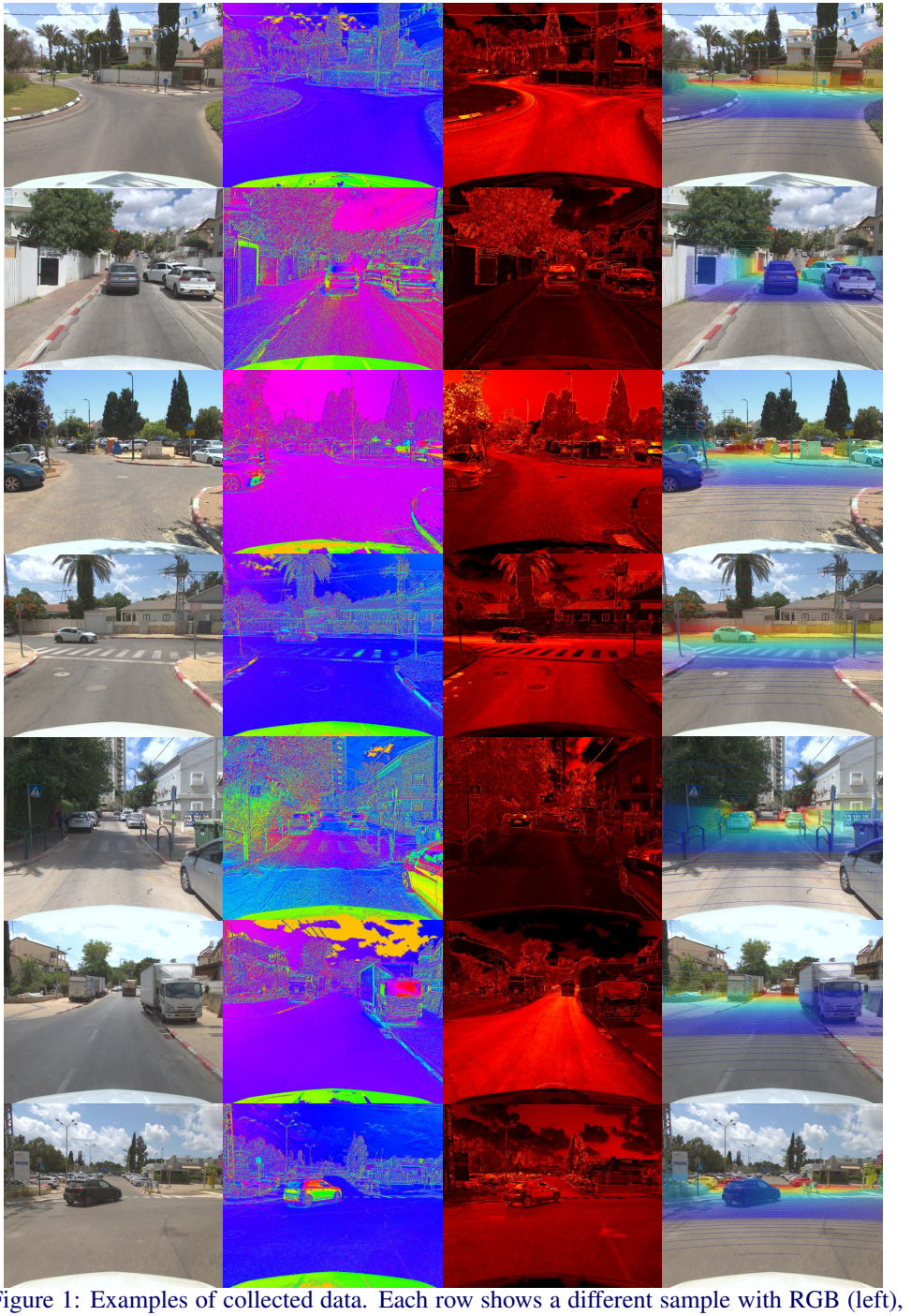

Figure 1: Examples of collected data. Each row shows a different sample with RGB (left), AoLP (middle left), DoLP (middle right) and lidar projected on RGB (right). The cyclic color map in the AoLP images goes from red for 0° to magenta for 179°. In the DoLP images black corresponds to 0 and yellow to 1.

3 Additional Qualitative Results

Figures 2 and 3 present additional qualitative results of the free space detection and monocular depth estimation tasks respectively.

References

[1] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *CVPR*, pages 248–255. Ieee, 2009.

[2] Rui Fan, Hengli Wang, Peide Cai, and Ming Liu. Sne-roadseg: Incorporating surface normal information into semantic segmentation for accurate freespace detection. In *ECCV*, pages 340–356. Springer, 2020.

[3] Clément Godard, Oisin Mac Aodha, Michael Firman, and Gabriel J Brostow. Digging into self-supervised monocular depth estimation. In *ICCV*, pages 3828–3838, 2019.

Figure 2: Qualitative results for free space detection. Yellow, blue and red correspond to true positive, false positive and false negative respectively. Best viewed zoomed in.

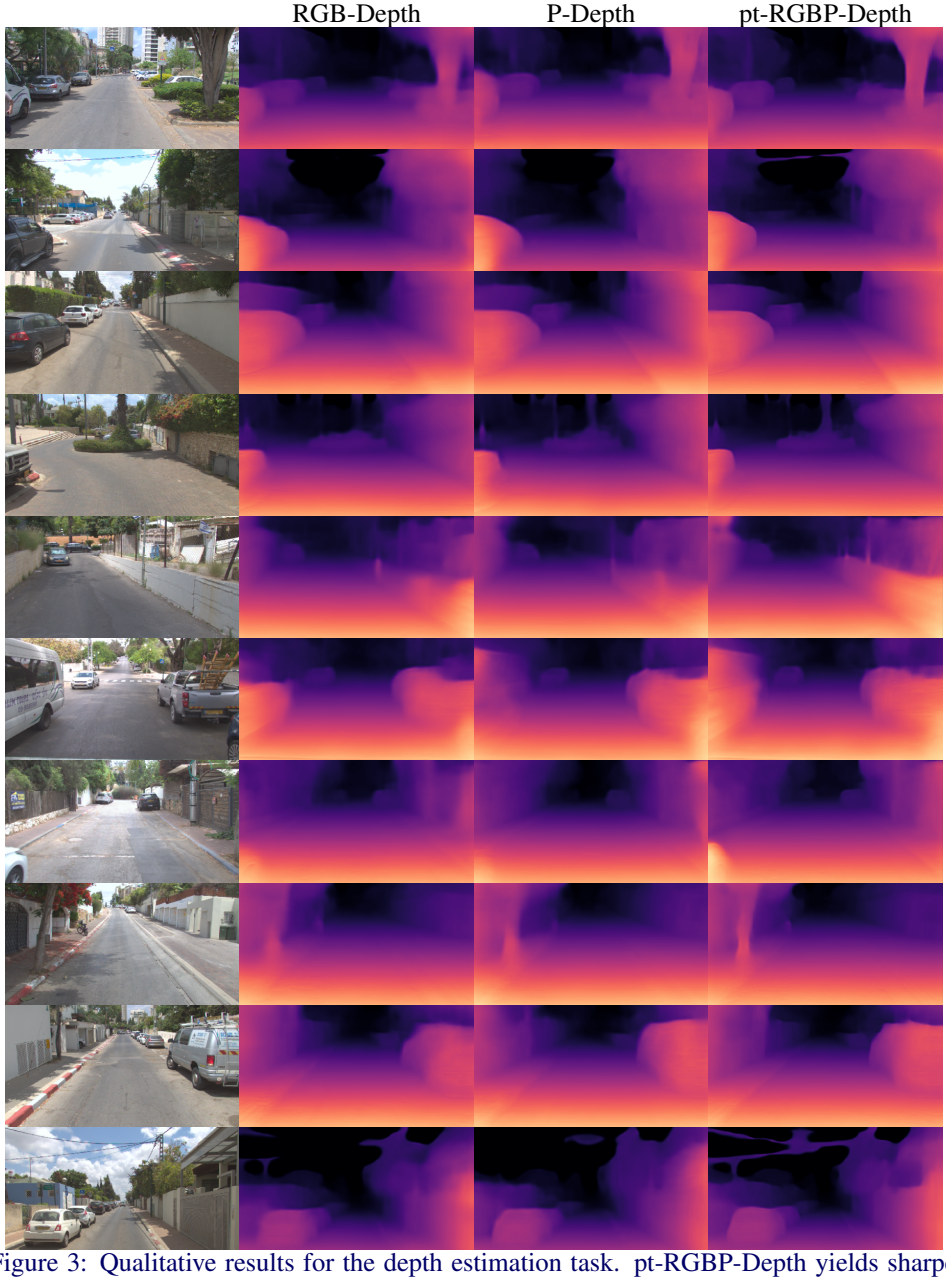

Figure 3: Qualitative results for the depth estimation task. pt-RGBP-Depth yields sharper edges and better recovers all structures. Best viewed zoomed in.