Supplementary Material

1. SemanticKITTI Stationary Scenes

Moving object annotations exist for the SemanticKITTI dataset [4, 1, 3]. To evaluate our approach against the state-of-the-art, we considered extracting all stationary frames from the SemanticKITTI validation sequence (sequence 08). The sequences in the SemanticKITTI dataset correspond to the sequences of the KITTI Odometry dataset [4]. All of these sequences have a Raw counterpart in the Raw KITTI dataset, which, among others, contains velocity measurements for each sequence from a high precision GPS/IMU inertial navigation system. In particular, the "forward velocity" measurements from the IMU unit provide the crucial information about the vehicle's velocity in m/s. These measurements allow automatically determining the number of moving or stationary frames in SemanticKITTI.

The correspondences between Raw KITTI and the Odometery sequences used in SemanticKITTI have been outlined in, *e.g.*, the Supplementary material of [2] in Table 2. However, we can not consider stationary frames from other sequences except the validation sequence. That is because they are part of the training dataset, and we evaluate against models trained on this data. The sequences from the test set can also not be considered for evaluation because the annotations are not publicly available.

Figure 1 shows the distribution of the velocity measurements from the IMU unit in the SemanticKITTI validation sequence. We find that less than 25 stationary frames exist in the validation sequence of SemanticKITTI. This amount of frames is not enough for a meaningful evaluation against the state-of-the-art.

For the sake of completion, we also provide the distribution over the velocity measurements from the training sequences in Figure 2. We can see that less than 200 frames appear in the full SemanticKITTI training dataset.

2. Raw KITTI Annotation Details

We annotated all moving objects with the point cloud labeler tool from the SemanticKITTI website (http://www.semantic-kitti.org/resources.html#labeling). In the point cloud labeler tool, we remove the ground with a setting of 0.3 and annotated all moving objects with their respective SemanticKITTI label.

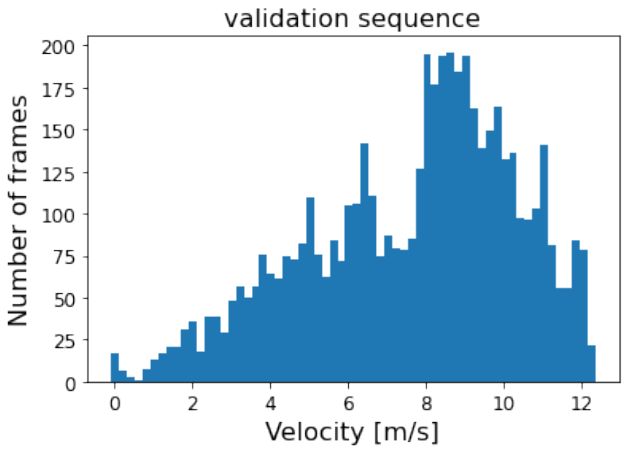

Figure 1. Overview of velocity measurements for the SemanticKITTI validation sequence

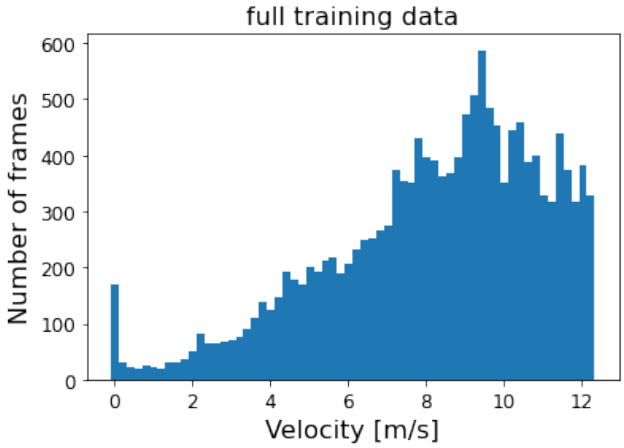

Figure 2. Overview of velocity measurements for the full SemanticKITTI training dataset

References

[1] J. Behley, M. Garbade, A. Milioto, J. Quenzel, S. Behnke, C. Stachniss, and J. Gall. SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences. In *Proc. of the IEEE/CVF International Conf. on Computer Vision (ICCV)*, 2019.

[2] Changhao Chen, Stefano Rosa, Yishu Miao, Chris Xiaoxuan Lu, Wei Wu, Andrew Markham, and Niki Trigoni. Selective

sensor fusion for neural visual-inertial odometry. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10542–10551, 2019.

[3] Xieyuanli Chen, Shijie Li, Benedikt Mersch, Louis Wiesmann, Jürgen Gall, Jens Behley, and Cyrill Stachniss. Moving object segmentation in 3d lidar data: A learning-based approach exploiting sequential data. *IEEE Robotics and Automation Letters*, 6(4):6529–6536, 2021.

[4] A. Geiger, P. Lenz, and R. Urtasun. Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 3354–3361, 2012.