

Ground Segmentation by CNNs in Velodyne Data

Exploration and semantic mapping of the environment surrounding a mobile robot plays a key role in robot's perception systems. Commonly used ground detection systems are often based on primitive features with low discriminative capabilities. On the contrary, state of the art technique for ground detection using Loopy belief propagation achieves the robustness by compromising its time efficiency (approximately 145 seconds per single Velodyne HDL-64E frame).

In our recent work, we deployed convolutional neural networks for the task of ground segmentation in sparse Velodyne point cloud data. We have developed a semi-automatic tool for ground annotation (see red points in the figure below) and the transformation of sparse LiDAR data into a dense 2D matrix which can be further processed by CNN. We have designed and trained multiple networks with shallow topologies (3-5 convolutional layers) using the Caffe framework.

Our experiments show, that we have significantly reduced the time requirements of ground segmentation process to the fraction of the time required by the state-of-the-art while preserving the same precision (mAP 99% for annotated frames of KITTI dataset). Processing of a single Velodyne frame by our network using only CPU takes 140 ms on average. Using the GPU acceleration (NVIDIA GeForce GTX 770), the processing time is further reduced to 7ms per frame.

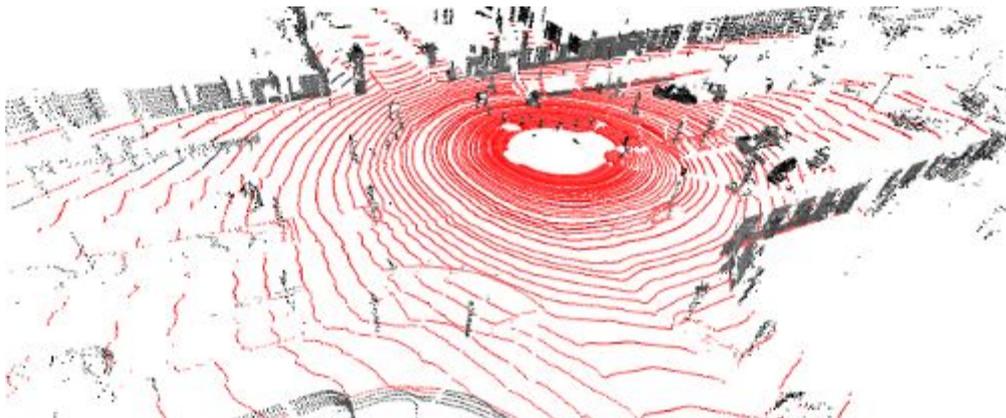


Figure: Ground segmentation in the single Velodyne LiDAR frame.

Video reference:

1. [CNNs for Ground Segmentation in Velodyne data](#)