RBE 549: HomeWork 2 - Sexy Semantic Mapping

Mandeep Singh
M.S. Robotics Engineering
Worcester Polytechnic Institute (WPI)
Worcester, MA 01609
Email: msingh2@wpi.edu

I. Overview

LiDAR data gives us information about the depth in the scene but we also need to recognise the objects and the semantic information about the scene. For this purpose either we can train a deep learning network where the model will learn label of each point cloud or we can use cameras to obtain a semantic painted point cloud. Here in this homework, we will be doing semantic segmentation on the RGB images obtained form the camera and then using these semantics we will do the semantics for the point cloud as well given that we know the intrinsics and extrinsics for both camera and lidar.

II. Dataset - KITTI 360

The dataset that we are using for point cloud painting is KITTI 360. The dataset has total 9 sequences but we will be using only 1 sequence. Also, the dataset already has semantic segmented RGB images which we will not be using, as we will train out model to get the semantics. We will only use stereo camera images, LiDAR point cloud and calibration data.

III. Point cloud registration - using Point-To-Plane ICP

First of all, we have around 10,000 point clouds in a sequence of data. We need to map all these point clouds in one co-ordinate frame . We will fix our world frame as same as the frame of the first point cloud in the data. Before starting with ICP, convert the .bin files to .pcd file using open3d library in python.
Then we can use either software like Paraview/Cloud Compare or a libraray such as open3d to implement point to plan ICP. The results shown in Fig. 1 are obtained by mapping 10 point clouds using open3d.

IV. Semantic Segmentation for RGB images

To do the semantic segmentation, **DeepLabV3+** neural netork has been used. The network was pre-trained for the Cityscapes dataset with 19 different classes. The pretrained weights worked even for the KITTI 360 dataset as well and an output of the network has been shown in Fig. 3.

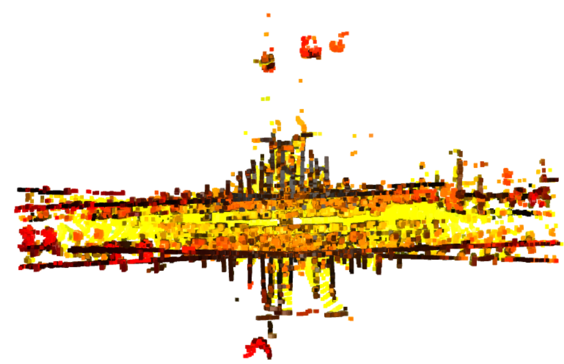

Fig. 1: Raw point cloud map (10 point clouds)

Fig. 2: RGB image

Fig. 3: Segmented RGB image

V. Point cloud Segmentation

To paint the point cloud using RGB semantic segmentation, first we have to find the find the transformation between the LiDAR to camera frame.

Transformation from LiDAR to Camera :

We are given the following data in calibration matrix :

$$T_c^l = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}$$ Transformation from camera to lidar

$$R_{rect} = \text{Rotation rectification for cam0}$$

$$P = \text{Perspective intrinsics of the camera}$$

The transformation from LiDAR to camera will be given by:

$$T_l^c = \begin{bmatrix} R^{-1} & -R^{-1}t \\ 0 & 1 \end{bmatrix}$$

And the Perspective transform from LiDAR to Camera frame (pixel co-ordinates) will be given by:

$$P_l^c = P.R_{rect}.T_l^c$$

Using the above equation we can transform the 3d points in LiDAR frame to camera image plane directly.

3d point painting

Get rid of the point cloud that is outside the given rgb image plane because we have the semantics only of the points that lie in image plane.

The point cloud will be sparse as compared to rgb image. So, find the projection of all points of the LiDAR point cloud in image plane, Then find the corresponding class label of the projected pixel and assign the color corresponding to the pixel to the actual 3d point. We repeat this for all such 3d LiDAR points available.

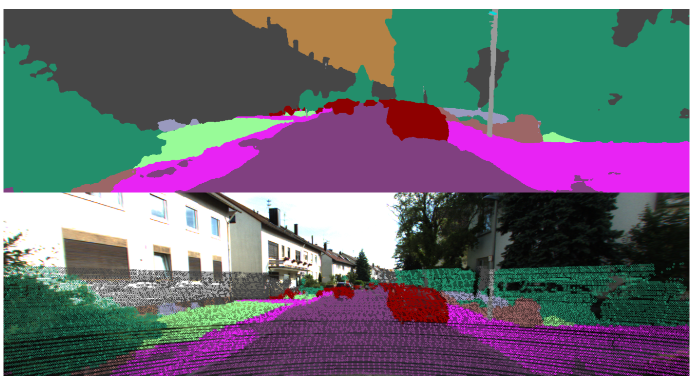

Fig. 4: Segmented RGB image and projected point cloud on image plane

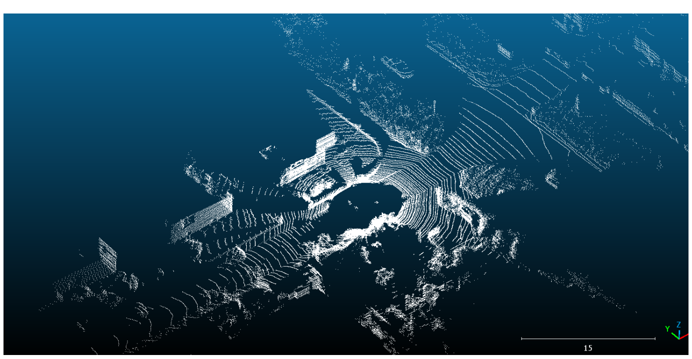

Fig. 5: Raw point cloud

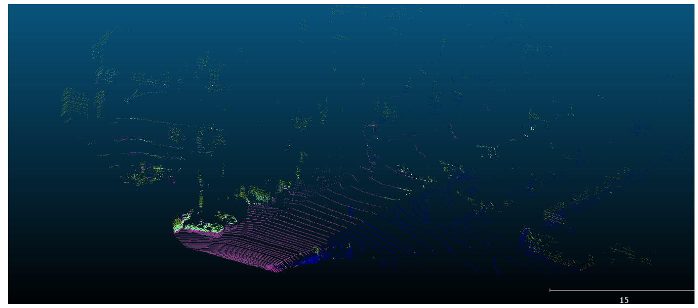

Fig. 6: Painted point cloud

VI. References

- https://arxiv.org/abs/1911.10150
- https://www.cvlibs.net/datasets/kitti-360/documentation.php
- https://github.com/naitri/PointPainting
- https://rbe549.github.io/fall2022/hw/hw2/