Obstacle Classification using 3D point cloud and Image data in self driving car application

Indian Institute of Technology, Kanpur

CS365 Course Project

Siddharth Tanwar (13699)
Akash Panda (13063)

Mentor - Prof. Vinay Namboodiri

Abstract

Following is the documentation of obstacle detection and classification using 3D point cloud and Image data applicable to driving cars in urban settings. Camera information coupled with Point Cloud data is used to identify cars in the camera frame. Cascade Classifiers[3] are trained on various features extracted from images to classify cars. The point cloud data is partitioned to identify obstacles using Velodyne Height Map algorithm which is further clustered using nearest-neighbour type algorithm. This data is used to generate ROIs in images. The implementation is tested on the KITTI Vision Benchmark Suite - Object Detection Evaluation 2012 dataset[2].

Contents

1 Introduction

The coming age of Artificial Intelligence has set new standards for our future. One of the major applications of AI in this vision of future is self-driving cars. Cars that are able to navigate, avoid obstacles and are able to adapt to changes in their environment. Obstacle detection and recognition are going to be vital to the realization of our dream of self-driving cars.

2 Motivation

There has already been a lot of progress in this aspect of computer vision. Several programs have already been used to detect and recognise obstacles. Most of these programmes recognize obstacles from images taken from a camera mounted on the car. Though there are image based programmes which do a pretty good job at recognising obstacles, there are bound to be cases where only the data from the images are not going to be enough to recognize the objects correctly. Recently, some developments were made in using Lidar sensors to generate 3D point cloud data of the environment in order to detect obstacles. We drew our inspiration from these recent developments to improve the performance of present image based object detectors with the assistance of 3D point cloud data taken from Lidar sensors.

3 Problem Formulation

The present obstacle detectors work in the following way, a camera fixed on the car provides us with images at a certain rate, we process the images to detect and recognize obstacles based on trained classifiers, we segment the image and then use the camera calibration matrix in order to extract information about the 3D location of the obstacle from the 2D image. This information is then used to change the speed, steer, or hit the brakes in order to avoid the obstacle. A Lidar sensor mounted on the car could provide us with the 3D data of the obstacles directly. So, the objective is to find a way to process the 3D point cloud data generated by the Lidar sensors in order to detect and segment the obstacles. Then we could try to reconcile this data with the 3D data extracted from the images in order to get a one to one correspondence between the image and the point cloud data and in the process remove all the false positives present in the image data.

For this project we use Velodyne HDL-64E Laserscanner data and Point Grey Flea 2 (FL2-14S3C-C) Color Camera data, both of which have been fitted on a car driving in an urban setting.

4 Fundamentals

4.1 Histogram of Oriented Gradients

The essential thought behind the histogram of oriented gradients descriptor[1] is that local object appearance and shape within an image can be described by the distribution of intensity gradients or edge directions.This is a feature descriptor used in image processing and computer vision for the purpose of object detection and recognition. This technique counts the occurrences of gradient orientation in localized portions of image.

Figure 1: The Setup

src : http://www.cvlibs.net/datasets/kitti/setup.php

4.2 Cascade Classifiers

Cascading is a particular case of ensemble learning based on the concatenation of several classifiers where the information collected from a given classifier is used as additional information by the next classifier.

The classifier used in this project was cascade of weak learners (decision stumps). Boosting technique was used to get accurate classifier from a cascade of weak learners.

4.3 Weak Learners

A weak learner is defined to be a classifier which is only slightly correlated with the true classification. It's labeling of examples is only slightly better than random guessing. The weak learners used here are Decision stumps. A decision stump is a machine learning model that consists of a one level decision tree. It has one root node and two leaf nodes. Decision stumps make decisions based on the value of a single input feature.

4.4 Boosting

It belongs to a family of machine learning algorithms that convert weak learners into strong ones. The training set used for each members is chosen based on the performance of the earlier stages. Incorrectly predicted examples are chosen more often thanthose that were correctly predicted. Thus, boosting produces new classifiers that are good at predicting examples that the previous classifiers were poor at.

4.5 Clustering

Clustering refers to the task of grouping objects in such a way that the objects in the same group are more similar to each other than they are to the members of other groups.Cluster analysis provides an abstraction from individual data objects to the clusters in which those data objects reside.

5 Methodology

We were provided labelled images (labels included cars, cyclist, pedestrian etc.) adjusted for distortion. Along with that corresponding Point Cloud velodyne data with transformation matrix from Global frame to camera frame. A brief outline of the various sections is :

- **Training the Classifier**- The Cascade classifier consisting of decision stump stages was trained using the images taken from the KITTI data set in order to detect and classify traffic obstacles like cars.
- **Identifying obstacle Cloud**- Use Velodyne height map algorithm to separate from the point cloud the obstacles of interest. PCL libraries[4] are used to handle Point Cloud data.
- **Partitioning obstacle cloud into different obstacles**- Use graph search techniques to further partition the obstacle cloud into individual clusters of obstacles.
- **Identifying ROIs in images from cloud data**- Use extrinsic camera transformation matrix to superimpose the 3D point cloud in 2D image frame and form a bounding box ROI for each cluster.
- **Detection of obstacle in image frame**- Detects cars in ROIs in the test images using Viola-Jones algorithm[5]

5.1 Training the Classifier

As discussed above the Classifier used in this project is a cascade of decision stumps. Decision stumps are weak learners by themselves. Their classification is only slightly better than guessing. But by the use of boosting techniques a weak learner can be converted into a strong learner and that is what our classifier does.

The images that were used for the training of the classifier were taken from the data set generated by KITTI's on-board camera. The HOG features were extracted from these images after applying proper bounding boxes to the positive samples.

This classifier was further used as a detector.Each stage of the classifier labels the region defined by the current location of the sliding window as either positive or negative. Positive indicates that an object was found and negative indicates no objects were found. If the label is negative, the classification of this region is complete, and the detector slides the window to the next location. If the label is positive, the classifier passes the region to the next stage. The detector reports an object found at the current window location when the final stage classifies the region as positive.

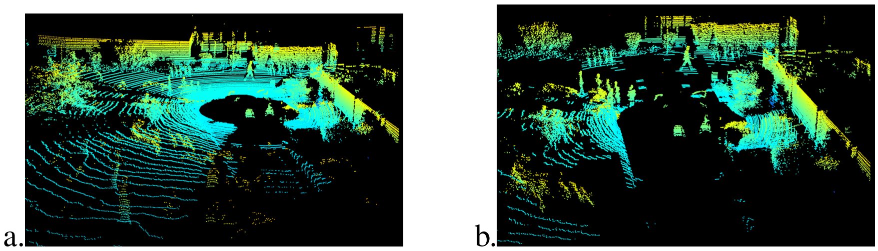

Figure 2: a) Entire Point Cloud b) Obstacle Point Cloud[2]

5.2 Identifying Obstacle Cloud

The algorithm adopted from Velodyne height map ROS package is as follows :

- Divide the x-y Plane into a square grid of equal sized cells.
- Of all the points in a cell if the difference between the z coordinate of any 2 point is greater than a threshold then add all points in the cell to the obstacle point cloud

5.3 Partitioning obstacle cloud into different obstacles

The algorithm takes the projected square grid used in previous section and marks a cell positive if an obstacle is present else the cell is marked zero. Now using basic DFS algorithm clusters are identified in the 2D grid and all points above the grid cells in a particular cluster fall in the object cluster. Using this method number of possible obstacles are identified in a frame.

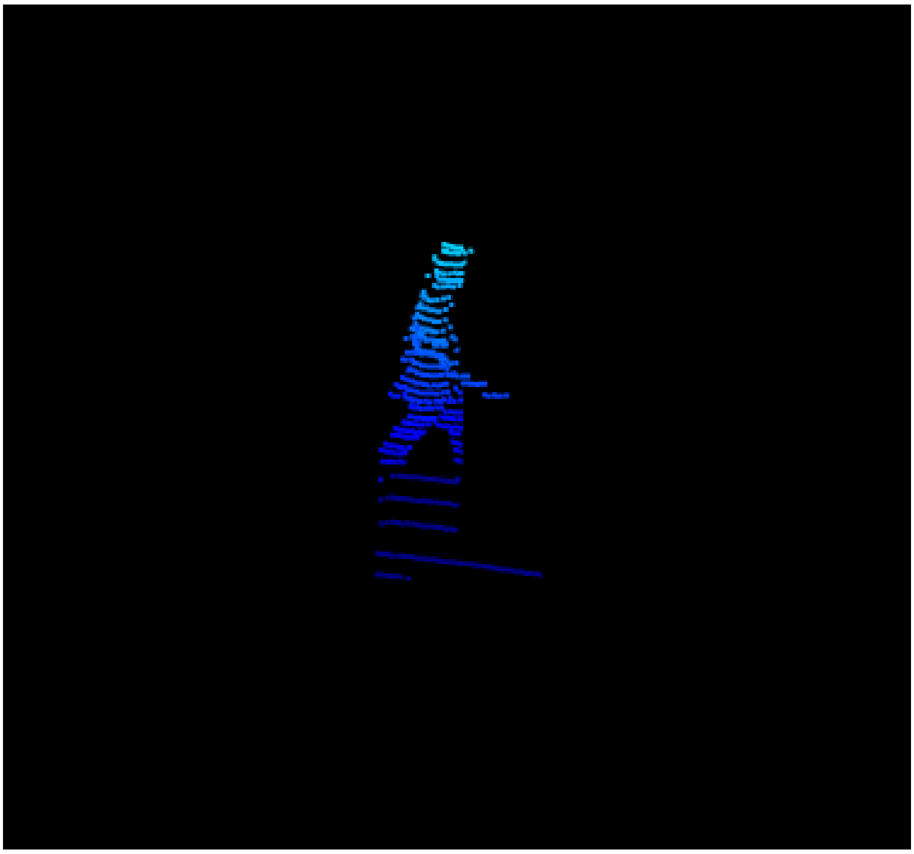

Figure 3: A pedestrian segmented from Obstacle cloud

5.4 Identifying ROIs in images from cloud data

Now the segmented point clouds are converted to homogeneous 3-D coordinates and pre-multiplying by the extrinsic camera matrix and the transformation matrix (R—t) (3x4 matrix) from velodyne sensor to camera results in conversion to homogeneous 2D coordinates in image plane. Among these points the ones lying in the camera frame are taken and used to generate bounding boxes.

$$x = P * Tr * y \tag{1}$$

where x is 2D homogeneous image coordinate, P is extrinsic camera matrix ,Tr is transformation matrix and y is homogeneous 3D point.

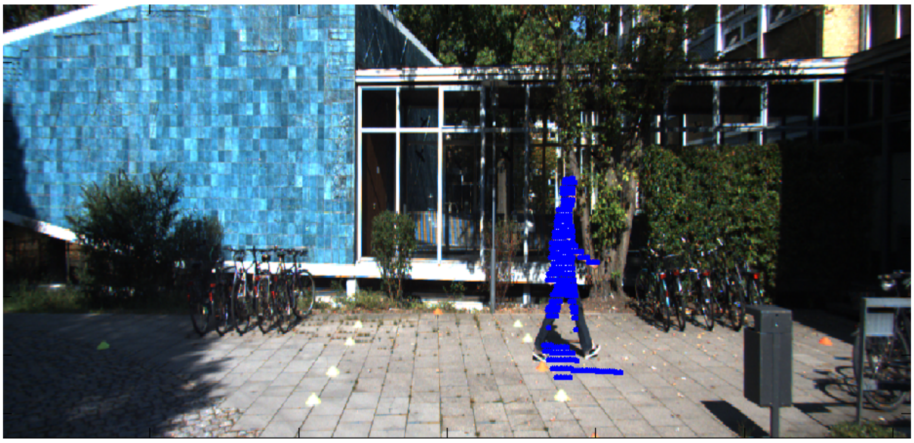

Figure 4: The pedestrian segmented earlier from Obstacle cloud superimposed on image

6 Experiments and Results

6.1 Dataset

We are training our Classifier on images obtained from KITTI vision benchmark suite's object detection evaluation 2012 dataset[2] which is a collection of colored images. It also consists of corresponding point cloud datasets. All of these are available for various environments. For our project we used the data for the urban environment. These datasets were collected using the set up in shown in figure1. The recording platform is a Volkswagen Passat B6. The laser scanner spins at 10 frames per second, capturing approximately 100k points per cycle. The vertical resolution of the laser scanner is 64. The cameras are mounted approximately level with the ground plane.

6.2 Results

Our results are as follows:

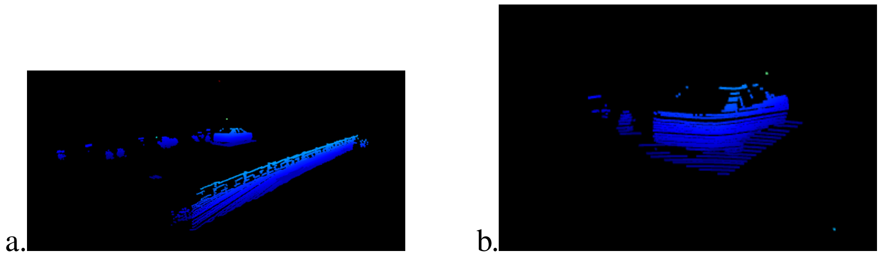

Figure 6: a) Obstacle Point Cloud b) Car seperated from Point Cloud[2]

Figure 5: Car detected in a frame using cascade classifiers alone

Figure 7: Imposition on image from point cloud

.

Figure 8: a) Identification using Cascade classifier leading to false positives b) Removed the false positives after using point cloud information[2]

7 Discussion

Due to lack of availability of labels in point cloud data along with images, implementing learning techniques on point cloud appeared difficult thus we stuck to simply using it as an assisting measure. Accuracy improves because the ROI is reduced. This also reduces computation time. Due to the fair simplicity of the Point Cloud algorithms real time application is easily possible.

8 Conclusions and Future work

- **Using 3D point cloud learning**:Implementing learning techniques on point cloud by identifying features efficiently is a hard task and may be done to reduce dependence on images.
- The analysis on classs of cars done in this project may be extended to other classes and a combined algorithm may be developed that segments all kind of obstacles.

9 Acknowledgement

We thank Prof. Vinay Namboodiri, Dept. of CSE for his valuable support throughout the project guiding us from time to time and looking into the project when it was needed. We thank the instructors of the course CS365 for their quality teaching. We thank PCL and Matlab community for their user friendly libraries.

References

[1] DALAL, N., AND TRIGGS, B. Histograms of oriented gradients for human detection proc. ieee computer society conference on computer vision and pattern recognition. volume 1,.

[2] GEIGER, A., LENZ, P., AND URTASUN, R. Are we ready for autonomous driving? the kitti vision benchmark suite. In *Conference on Computer Vision and Pattern Recognition (CVPR)* (2012).

[3] R., L., A., K., AND PISAREVSKY, V. Empirical analysis of detection cascades of boosted classifiers for rapid object detection. In *25th DAGM Symposium on Pattern Recognition. Magdeburg, Germany*.

[4] RUSU, R. B., AND COUSINS, S. 3D is here: Point Cloud Library (PCL). In *IEEE International Conference on Robotics and Automation (ICRA)* (Shanghai, China, May 9-13 2011).

[5] VIOLA, PAUL, AND JONES, M. J. Rapid object detection using a boosted cascade of simple features proc. 2001 ieee computer society conference on computer vision and pattern recognition, 2001. volume: 1.