Fast Hybrid Pipeline for Road and Non-Road Detection

SOURISH CHATTERJEE - EE22MTECH02002

26 April 2023

1 Introduction

Road and non Road Detection has become a really popular research topic in the Navigation Industry especially in the us cases of Autonomous vehicle, Connected Vehicles etc.

Besides obstacle avoidance, road detection can also facilitate path planning and decision making, especially in those situations where lane markings are not visible (for example, because covered by snow or due to poor lightning conditions) or not present (for instance, in certain rural and urban roads).

Camera based approaches are strongly affected by environmental illumination. Therefore, their performance is expected to decrease considerably at night-time or whenever presented with light conditions that deviate from those seen during training.

LIDARs, on the other hand, carry out sensing by using their own emitted light and therefore they are not sensitive to environmental illumination. Hence with lidar, same level of accuracy across the full spectrum of light conditions experienced in daily driving, and for this reason they are particularly suitable for achieving higher levels of driving automation.

2 Literature Survey

2.1 LIDAR-based road and road-edge detection

In [4] road and road edge detection is done using a systematic processing of the elevation data extracted from the LIDAR. The steps used in [4] are as follows

- Identification of local maxima and minima using a Gaussian differential filter
- Feature used for classification of the road is Weighted standard deviation
- Classification is done using the extracted feature and putting a threshold on the parametrized form so that the Number of points also act as parameter
- False alarm mitigation is done using average minimum road width
- Curb detection is done using Hough Transform

The Hough transform is a feature extraction technique used in image analysis, computer vision, and digital image processing. The purpose of the technique is to find imperfect instances of objects within a certain class of shapes by a voting procedure.

2.2 A real-time road detection method based on reorganized lidar data

In [3] they have use a new organisation technique for the LiDAR data that can be easily used to reduce the search time and the complexity. This is done for the top-view LiDAR data. Segmentation is done using a fast hough transform to detect which part of the data belongs to the road part.

2.3 Fast LIDAR-based road detection using fully convolutional neural networks

The problem of road detection in [1] is framed as a pixel-wise semantic segmentation task in point cloud top-view images using an Fully Connected Neural Network. They have done road segmentation in real time on GPU-accelerated hardware. A top-view representation is, in their opinion is more appropriate than a camera perspective representation given that both path planning and vehicle control are executed in this 2D world.

2.4 Progressive LiDAR adaptation for road detection

Despite rapid developments in visual image-based road detection, robustly identifying road areas in visual images remains challenging due to issues like illumination changes and blurry images. To this end, LiDAR sensor data can be incorporated to improve the visual image-based road detection, because LiDAR data is less susceptible to visual noises. However, the main difficulty in introducing LiDAR information into visual image-based road detection is that LiDAR data and its extracted features do not share the same space with the visual data and visual features. Such gaps in spaces may limit the benefits of LiDAR information for road detection. To overcome this issue, we introduce a novel Progressive LiDAR Adaptation-aided Road Detection (PLARD) approach to adapt LiDAR information into visual image-based road detection and improve detection performance. In PLARD, progressive LiDAR adaptation consists of two subsequent modules:

- Data space adaptation, which transforms the LiDAR data to the visual data space to align with the perspective view by applying altitude difference-based transformation
- Feature space adaptation, which adapts LiDAR features to visual features through a cascaded fusion structure.

In particular, PLARD outperforms other state-of-the-art road detection models and is currently top of the publicly accessible benchmark leaderboard[2].

3 Dataset

he road and lane estimation benchmark consists of 289 training and 290 test images. It contains three different categories of road scenes:

- uu - urban unmarked (98/100)
- um - urban marked (95/96)
- umm - urban multiple marked lanes (96/94)
- urban - combination of the three above

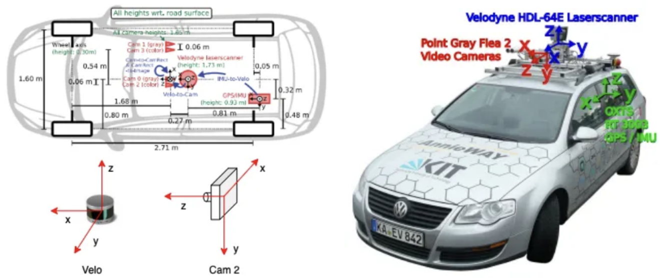

Figure 1: Caption

Ground truth has been generated by manual annotation of the images and is available for two different road terrain types: road - the road area, i.e, the composition of all lanes, and lane - the ego-lane, i.e., the lane the vehicle is currently driving on (only available for category "um"). Ground truth is provided for training images only.

- Data space adaptation, which transforms the LiDAR data to the visual data space to align with the perspective view by applying altitude difference-based transformation.

- Feature space adaptation, which adapts LiDAR features to visual features through a cascaded fusion structure

4 Proposed System

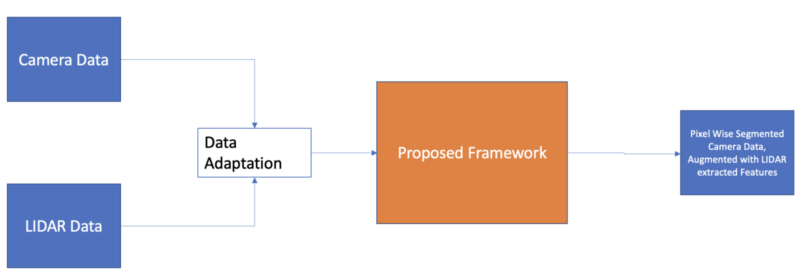

Figure 2: System Diagram

The proposed system as shown in Figure 2, uses the camera data and LiDAR data. The Proposed framework will either fuse the data and create a Altitude Difference Image or use the Z Axis data from the LiDAR data to estimate the part of the data which is road and which is not a road. The output can be of two type as listed below

- Segmented Image highlighting road and not a road
- Z axis points that belong only to the road

5 Implementation

This section will give the details of the implemented framework and details of the each of the processes that will get involved in the pipeline to detect the road and non road using the camera and LiDAR data.

5.1 Implemented Framework Outline

The camera data and LiDAR data are assumed to available at the start of the pipeline and we can use a decision algorithm to choose when we want to choose which route of processing the camera and the LiDAR data. According to the framework defined in Figure 3. we see the 2 pipeline that can be expalined as follows -

- The data apaptation module will map the data from LiDAR space to the camera space and then ADI image will be generated. Further the binary image processing alogorthm will do the segmentation.
- The LiDAR data when used in case of bad lighting condition in the camera is subsampled and processed to detect which part of the Z axis data belong to the road part.

5.2 Computation of Altitude Difference Image

The calibration data helps to augment the lidar data into the space where camera is pointed. This contains a transformation matrix that will translate the 3 dimensional data into the 2 dimensional camera data.

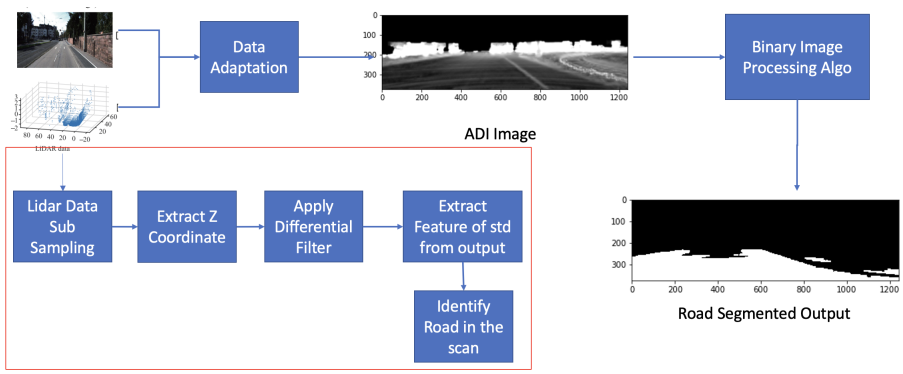

Figure 3: Implemented Framework

This will mantain the size of the image from camera and the size of ADI is same. The mapping is done in space of camera data and interpolation is done since the LiDAR data space is way more sparser than that of input from the camera.

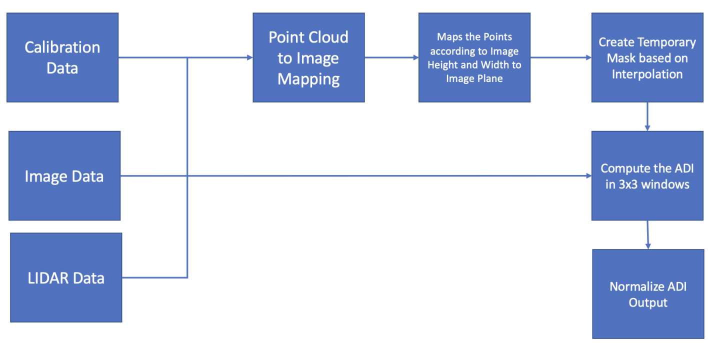

Figure 4: Computation of Altitude Difference Image(ADI)

5.3 Binary Image Processing Algorithm

Figure 5. shows the algorithm used to process the ADI Image to segment the road and non road part of the image. Steps are explained in detail below -

- The binary image is created from the ADI image that has continuous range of pixel intensity in the gray scale domain. This is done by using mean based thresholding. This clears some part of the noise in the image

- We apply erode operation using the kernel of size 9x9 to remove small patches that appear similar to noise in the binary image.
- Since we know that visibility of the camera is fixes hence we choose the empirical value of 200 pixel and blacken the rest of the pixels. This comes from the natural ability of the camera to see up to a limited range.
- Then we remove the blobs from the binary image obtained from the previous stage and further process it to obtain the final segmented output.

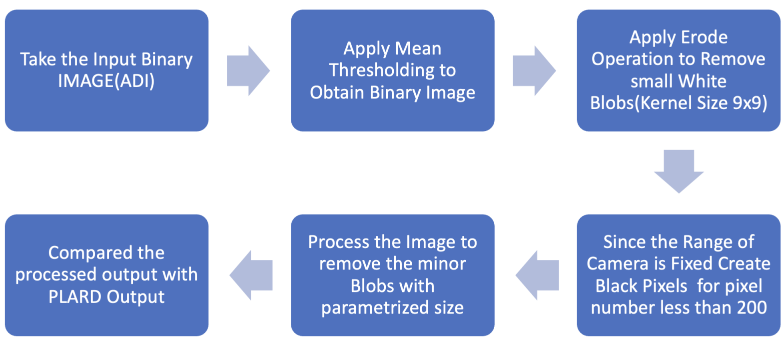

Figure 5: Flow for Binary Image Processing Algorithm

6 Results

Figure 6. shows the results obtained but the algorithm that is implimented in the above mentioned framework. Step 1 does the binary image conversion from the ADI. Step 2 converts the black pixels according the to the visibility. Step 3 will operate the erosion kernel and the minor blob removal. Step 4 is comparison of the same image output from Deep learning based framework of PLARD.

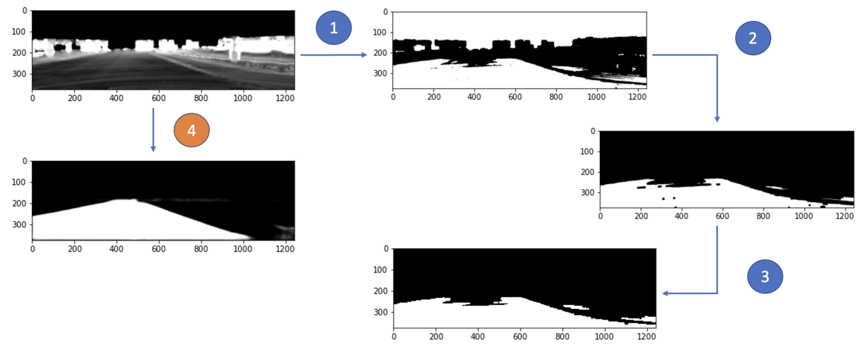

Figure 6: Visualization from Binary Image Processing Algorithm

Figure 7. will help the processing of the LiDAR data. The first plot shows the noisy Z Axis data. On applying the differential filter we obtain the maxima and minima in the data. Further we obtain the Weighted average and standard deviation to remove noise. Noise removal is done by threshold using the standard deviation. Indices of rising maxima and minima are used as the feature of detection for the road part.

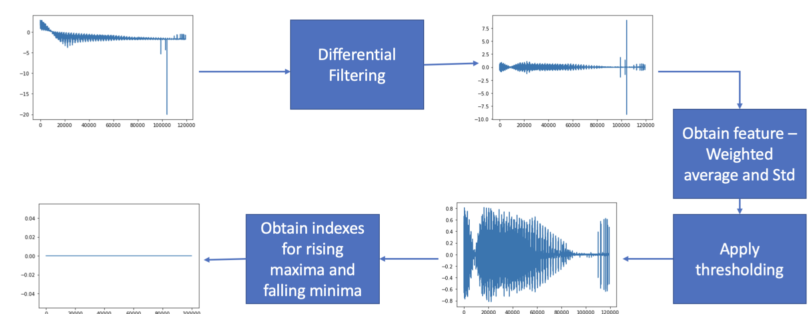

Figure 7: Visualization for Algorithm for LiDAR based road detection

7 Future Work

There are many points that need refinement before we actually implement this algorithm in real-time. Some of which can be resolved by the mentioned methods that are identified as follows -

- Use of motion-based mask to find Hough Transform. As shown in Figure 8. we can use the motion estimation of the vehicle either moving straight, turning left or turning right as a feature to create a mask that can be used to identify the road in those regions only.

- Use adaptive kernel size and thresholding for processing. Current Algorithm feeds on fixed kernel size and thresholds this can be made adaptive for better results.

- Metric for Performance Indication. Currently model has not quantitatively done the comparison of our algorithm with other state of the art available solutions. That also need to implimented.

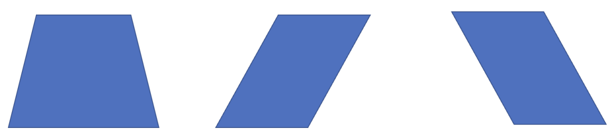

Figure 8: Shapes for Hough Transform Mask

8 Conclusion

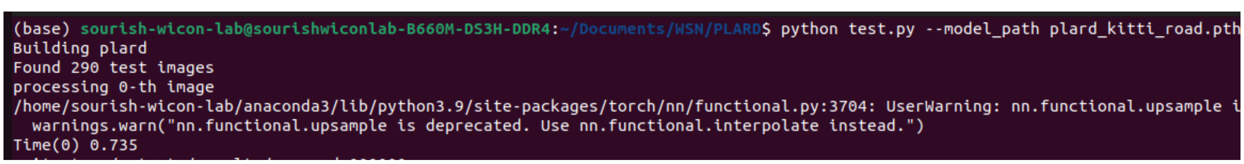

Figure 9: Runtime as per PLARD Model

There exists a Trade-offs that needs to considered for this algorithm. The timing and accuracy of the model can be considered as an element of trade-off between PLARD and Binary Image Processing Algorithm. Also please note that currently Binary Image Processing Algorithm is implemented in Python, once we implement this in C++ we would see faster performing model

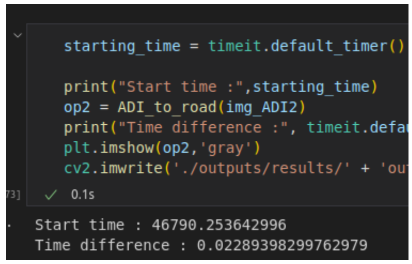

Figure 10: Runtime as per Developed Binary Image Processing Algorithm

References

[1] Luca Caltagirone, Samuel Scheidegger, Lennart Svensson, and Mattias Wahde. Fast lidar-based road detection using fully convolutional neural networks, 2017.

[2] Zhe Chen, Jing Zhang, and Dacheng Tao. Progressive lidar adaptation for road detection. *IEEE/CAA Journal of Automatica Sinica*, 6(3):693–702, 2019.

[3] Fenglei Xu, Longtao Chen, Jing Lou, and Mingwu Ren. A real-time road detection method based on reorganized lidar data. *PLOS ONE*, 14(4):1–17, 04 2019.

[4] Wende Zhang. Lidar-based road and road-edge detection. In *2010 IEEE Intelligent Vehicles Symposium*, pages 845–848, 2010.