Abstract

Detection of various road users, both motorized and non-motorized, plays a vital role in the connected and autonomous vehicle (CAV) environment. Currently, most reported studies present the use of single sensors to collect information of surrounding, including other vehicles. However, due to the complex nature of urban road environment, it becomes difficult to collect efficient and reliable data with a single sensor. Therefore, this thesis presents an attempt to detect the typical users of the road both motorized and non-motorized by integrating the LiDAR and the camera sensors on an instrumented vehicle. The camera provides colour imagery in 2D pixel coordinator system, and hence is efficient in better recognition but, they lack in providing range information. On the other hand, LiDAR provides point clouds in a 3D coordinate system, which is efficient in giving range information, but face challenges in classification and association related issues. Hence, a fusion system to capture the advantages of LiDAR and camera can solve issues related to object detection. In this study, an efficient offline calibration technique for both these sensors is developed by employing a new methodology for determining points of correspondence between the two sensors using a white planar rectangular calibration board. The 3D correspondence points are estimated by fitting a rectangular plane on the laser points lying on the calibration board. Since the original length of the planar board is known, the check for the accuracy of the fitted plane can be performed for better results. The estimated center of the planar board from the LiDAR and those detected using the camera images serve as points of correspondence. Also, the frames from the video camera are extracted such that the camera and LiDAR data are synchronized resulting in a camera frame corresponding to each LiDAR frame.

Once the calibration is over, the 3D LiDAR points are projected to 2D pixel coordinates for a sequence of time frames of interest. On the parallel side, the object detections the camera image frames, are performed based on YOLO V3 (the third version of You Only Look Once), and the location of bounding boxes around the target objects are extracted. The projected points lying inside the bounding boxes are again back-tracked to the original 3D LiDAR points. These 3D points are hence clustered together so that the noise and the unwanted points can be removed and the centroid of the detected object can be efficiently estimated with better accuracy. The centroids are then tracked to find its trajectory of these detected object in each subsequent frames. The key advantages of the fusion technique over other methods is its efficiency to automatically filter unwanted objects present in nearby surrounding and assign vision classification to the point clouds of interest, which provide range information with improved accuracy to reduce false detection rates.