2016 12th World Congress on Intelligent Control and Automation (WCICA)
June 12-15, 2016, Guilin, China

Cooperative Fusion for Road obstacles detection using Laser Scanner and Camera

Shashibushan Yenkanchi and Q.M. Jonathan Wu

Abstract— In order to account for robustness of Automotive safety applications, fusion of data from multiple sensors is of remarkable importance to know the position of road obstacles. Challenges arise in Multi Sensor Data Fusion (MSDF) due to sensor uncertainty, multiple occluding targets and clutter by changing weather conditions. The proposed architecture address the problem by fusing information cooperatively from Laser scanner and monocular camera for robust detection of scene objects in the vehicle environment. The Fusion steps in the proposed method involve the application of the M-estimator SAmple Consensus (MSAC) algorithm for ground plane removal and density based clustering of laser data. Then the filtered laser objects are projected on the image plane and the corresponding region of interest (ROI) is extracted to localize the potential targets. Experimental results on challenging scene sequences of benchmark data sets prove the robustness of proposed fusion architecture for detecting vehicles on the road.

I. INTRODUCTION

One of the challenging tasks, researchers and auto manufacturing industries are facing is to increase road traffic safety and reduce the number of fatalities. Today's modern world applications in automotive industry known as Advanced Driver Assistance Systems (ADAS) will rely on numerous sensors that are attached on vehicle to accurately perceive the surrounding environment. As it can be noted that there is no accurate, error free sensor manufactured till date, most research is put forth in designing the complex signal processing strategies to increase the reliability of the overall system leveraging capabilities of individual sensors to process information synergeticly. This increases the ability of the vehicle to detect the potential threats on the road either actively or passively. Current trends in the ADAS architectures and the future advancements can be found in [1].

Some of the related popular projects to use MSDF for automotive safety include ProFusion2 which is a part of EU co-funded integrated project PReVENT [2]. Similar work has been considered by EUCLIDE and PAROTO research projects [3], [4] which rely on infrared and radar sensors with focus on vision enhancement for forward collision warning applications. Information from Radar, Video and Laser were used by CARSENSE [5]. Many sophisticated techniques have been developed over the recent years, which resemble the above projects aiming at a specific task of increasing reliability and push the technology towards autonomous driving making the driver mere spectator in the near future.

In order to achieve the aforementioned goals many algorithms in literature have been proposed for efficient data fusion. Detailed review and techniques involved in design of fusion architectures can be found in [6] which, further classifies the process into sub categories as in Data association, State estimation and Decision fusion. Thus, this work further explores techniques for efficient MSDF.

Although there are many algorithms to detect the vehicles just by considering single sensor, the proposed algorithm takes advantage of a laser scanner in the image plane which rely on the projected filtered 3D points. The aim of this method is to localize the targets accurately so the above mentioned data fusion steps such as Data association and State estimation can be performed. There are various methods which use model fitting, for example Gaussian mixture model to estimate the parameters of the pixel intensity for localization. However, we need the prior knowledge of the density means to initialize the process. Unlike the methods that need prior initialization, the proposed algorithm employs a window based technique to form the clusters along with density based clustering algorithm which uses the neighboring information to gather the Laser data points. Based on these clusters, ROIs are obtained to from the binary mask which has potential advantages, such as histogram based cues from an image can be added to increase the confidence of the detected objects.

The rest of the paper is organized as follows: Section II explains the proposed architecture, this gives the overview of the algorithms introduced. Section III provides Analysis of the experimental results obtained. Finally, Section IV concludes the paper with the scope of Future work.

II. PROPOSED ARCHITECTURE

The approach proposed in this work for cooperative fusion architecture is shown in fig. 1 in terms of block diagram. It clearly explains the steps undertaken in the proposed fusion methodology. As a preprocessing step, both the laser scanner and camera sensors need to be timely registered in a common vehicle coordinate system.

Shashibushan Yenkanchi and Q. M. Jonathan Wu are both with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON, N9B 3P4, CANADA. (e-mail: yenkanc, jwu@uwindsor.ca).

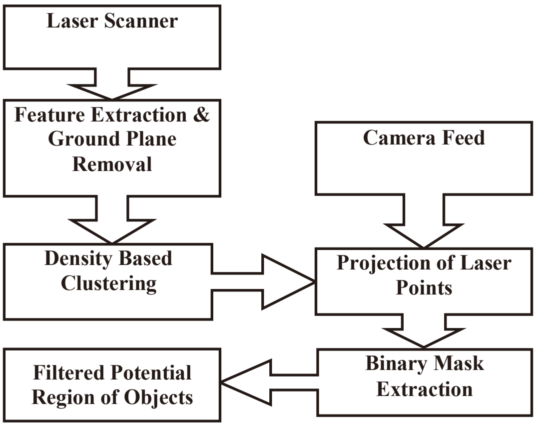


Fig. 1. Block Diagram of the Proposed Method.

A. Feature Extraction

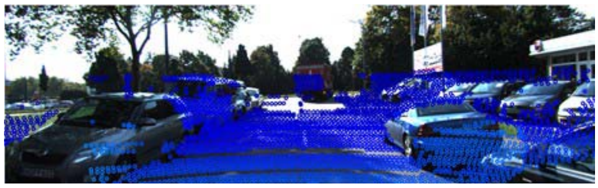
Since real world objects are made of primitive geometric shapes in nature, the point cloud data obtained from laser scanner contain information about local saliencies such as surface normal and curvature at each of the respective 3D point locations. One of the first preprocessing steps involved is to fit a planar model by estimating the above mentioned saliencies and be able to accurately identify and remove the points belonging to the ground surface.

In the computer vision literature some of the classic algorithms used to perform this task are Least squares, PCA, and RANSAC. Some of the recent work and analysis of above mentioned algorithms can be found in [7], [8]. These algorithms use re-sampling techniques which iteratively generate the candidate solutions into inliers and outliers by using the minimum number of observations required to estimate the model parameters that fit a plane. In this proposed work MSAC algorithm which is a variant of RANSAC algorithm is used. This algorithm was first proposed by Torr and Zisserman [9] for the model fitting on to the 3D point cloud. Later, Vosselman and Klein [10] showed some of the Advantages of using MSAC algorithm in comparison to RANSAC.

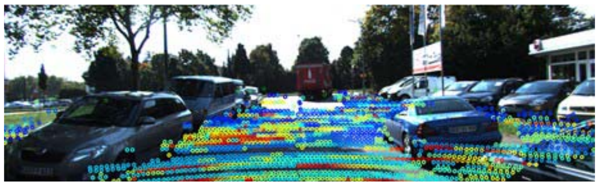
An example of ground plane removal is shown in fig. 2. Where, (a) shows the scene with projected raw laser points, (b) shows the laser points belong to the ground plane, and (c) shows the laser points belong to vertical surface.

B. Density Based clustering

After extracting the ground plane the remaining point cloud data in the scene consists of the points associated with the potential obstacles on the road which may be coming from vehicles, pedestrians, walls, road traffic signs, etc. The task of the clustering algorithm is to group the points coming from the same object into one. 3D points coming from the laser may vary in each frame depending on the arbitrary shapes. In order to cluster the data of this type most promising methods include graph-theoretic clustering methods which clusters the data based on the neighbors [11]. This proposed work implements one of the above classes of algorithm called Dbscan proposed by Ester et al. [12].



(a). The concerned scene with projected raw laser points.



(b). The laser points belong to ground plane.

(c). The laser points belong to vertical surface.

Fig. 2. An Example of Ground Plane Removal.

The algorithm takes two parameters as an input specified by the user. Parameter's include the search radius € and the minimum number of neighboring points θmin to be considered while clustering. The advantage of this algorithm is it does not require the prior initialization of the cluster's to be formed. This operates by starting with arbitrary point P and looks for the € points in θmin neighborhood. The new cluster is started if this condition is satisfied by marking P as cluster center and all the neighbor points belonging to this core. In the next step θmin neighborhood of all these points is checked and the cluster is grown arbitrarily. The points which do not satisfy this condition are marked as outliers. The example can be seen in fig. 3 with $\theta min = 4$ points to form a cluster. Euclidean distance metric is used to do the computation. The complexity of this algorithm is $O(n\log(n))$ where n is the number of points. Worst case complexity is given by $O(n^2)$.

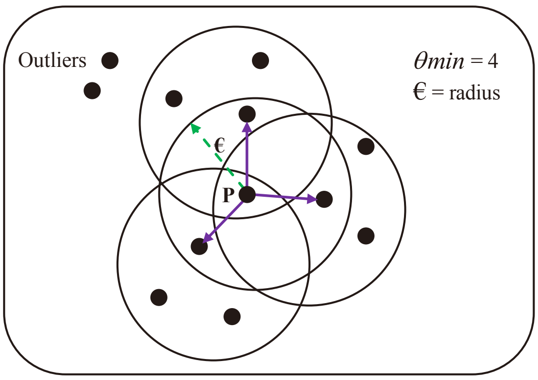


Fig. 3. An Example of Dbscan Algorithm.

C. Binary mask Extraction

The clustered laser scene objects are projected onto the image plane by using camera calibration matrix. In order to localize the potential targets in the scene a binary mask is created by searching for the projected laser points in the image region. A $N \times N$ size window is run scanning for the projected points. Window looks for the min number of points to be considered in order to confirm the object projection. The Algorithm below explains the procedure to get the mask considering the specified window size as the input.

The Algorithm

```
BinMask= False
For   i = 1 : N : end of X coordinate
   For j = 1: N : end of Y coordinate
      For k = 1: Size(Projected  laser Points)
        If  (in the Image Plane)
           Extract Projected Points
        End
      End
        If  Size(Extract Projected Points) > Threshold
           BinMask = True
        End
   End
End
```

III. EXPERIMENTS AND RESULTS

This section demonstrates the performance of proposed architecture which was applied on KITTI benchmark datasets [13]. Different scene sequences were considered in the analysis with varying number of target types. Table 1 provides a short description of the data sequences used in the experiments. The experiments were carried out in MATLAB environment on 2.4 GHz Pentium 4 processor running a windows 7 operating system.

TABLE I. DESCRPTION OF DATASEQENCES

Data seq.	Description
DS#1	This scene is taken from KITTI dataset with sequence ID 2011_09_26_drive_0059. Consisting of 52 cars, 3 vans and 5 pedestrians.
DS#2	This scene is taken from KITTI dataset with sequence ID 2011_09_26_drive_0018. Consisting of 11 cars, 2 vans and 2 Trucks.

A. Visual results of the experiment

This section depicts the visual experimental results obtained by the proposed methodology. As it can be seen in fig. 4, the first sequence (a) DS#1 has challenging scene situation due to varying lighting conditions and it is tedious to detect the parked vehicles in the shade by using a single camera. Therefore, with the help of Laser scanner, the proposed method is able to localize both the moving vehicle and parked vehicles on the street in the form of ROI in a binary image plane as shown in (b). The fused image of laser reflections from the obstacles that are projected onto the camera image can be visualized in (c). The same is true for the second sequence DS#2 in fig. 5 of traffic junctions. The ROI includes the area of occluded targets as well.

The size and shape of the obtained ROI are influenced by input parameters that are fed into clustering and mask extraction blocks down the signal processing chain in the proposed methodology. Fine tuning of these parameters is vital in obtaining well defined ROI shapes that localize potential road obstacles in a binary image. However, the parameters essential for calibration and registration of the sensors are not considered in the scope of this study. Instead, for this experiment registered camera and laser scanner images are used along with calibration matrices provided in the dataset for projection of laser points onto the image plane.

For density based clustering of laser points the input parameter θmin was chosen to be 5, which specifies the number of nearest neighbors considered to group objects into different classes of clusters. Choosing the neighbors value too low or high may eventually result in irregular shapes of ROI due to the addition of clutter points which may not actually belong to the same object in 3D space. The value of 5 neighbors was heuristically chosen for this experiment and it shows consistent behavior in localizing the targets for different scene scenarios. As the number of targets present in the scene varies in accordance with the background it is tedious to group individual objects. Hence, the scope for improvement in the optimization of tuning parameters is more and different graph-theoretic clustering algorithms could be applied in order to verify the performance of classification under various scene circumstances.

(a). A Sample Scene in DS#1.

(b). Corresponding Binary Mask of the Scene in (a).

(c). Fitered Detected Targets for the Scene in (a).

Fig. 4. Results of the detection on scene sequence 1.

(a). A Sample Scene in DS#2.

(b). Corresponding Binary Mask of the Scene in (a).

(c). Filtered Detected targets for the scene in (a).

Fig. 5. Results of the detection on scene sequence 2.

The contribution of this paper emphasizes on the binary mask extraction which defines the area enclosed by the potential road obstacles. The parameters which are crucial for this block are the optimal window size and the threshold that checks for the minimum number of laser points in the window to confirm the presence of a target. Based on the heuristics, a 50×50 window size was selected and the threshold value was set to consider a minimum of 5 laser points. The ROI shapes can further be optimized by experimenting with different combinations of these parameters for varying scene conditions.

Classification is based on the length of the clusters formed. The offset point was chosen based on carefully examining the size of clusters formed in number of frames to filter out the walls and outliers such as poles, benches or a dustbin from that of vehicles present in the laser field of view. The cluster density corresponding to the walls is very high, while that belonging to the clutter objects is too low.

IV. CONCLUSION AND FUTURE WORK

In this paper, a road obstacle detection method was presented using a cooperative fusion of a laser scanner and monocular camera. The objective was to obtain well defined ROI of the obstacles in a binary image that ensures complete knowledge of the surrounding environment of the vehicle on which the sensors are attached. Experimental results show that the proposed algorithm is able to detect the road obstacles in challenging scene sequences. The Density based clustering is proved to be efficiently working in combination with the window based technique. There is scope for improving the detection results, by changing the window size and the threshold to collect a minimum number of Laser points for the binary mask creation. The Future work involves extracting color histograms from the camera image which overlap with the window of laser points. This can further be used in the application of data association and state estimation for Robust detection and tracking.

REFERENCES

[1] K. Bengler, K. C. Dietmayer, B. Farber, M. Maurer, C. Stiller and H. Winner, "Three decades of driver assistance systems: review and future perspectives," *IEEE Intelligent Transportation Systems Magazine*, vol. 6, no. 4, pp. 6–22, 2014.

[2] S.B. Park, F. Tango, O. Aycard, A. Polychronopoulos, U. Scheunert and T. Tatschke, "ProFusion-2 – sensor data fusion for multiple active safety applications," *Proc. of 13th Intelligent Transportation Systems*, London, 2006.

[3] A. Polychronopoulos, U. Scheunert and F. Tango, "Centralized data fusion for obstacle and road borders tracking in a collision warning system," *Proc. of the ISIF 7th International Conference on Information Fusion*, Stockholm, Sweden, 28/06-01/07, pp. 760–767, 2004.

[4] C. Blanc, L. Trassoudaine, Y. Le Guilloux and R. Moreira, "Track to track fusion method applied to road obstacle detection," *Proc. of the ISIF 7th International Conference on Information Fusion*, Stockholm, Sweden, 28/06-01/07, pp. 775–782, 2004.

[5] C. Coue, T. Fraichard, P. Bessiere and E. Mazer, "Using bayesian programming for multi-sensor data fusion in automotive applications," *Proc. of IEEE Intelligent Vehicle Symposium*, Versailles, France, June 18–20, 2002.

[6] F. Castonedo, "A review of data fusion techniques," *The scientific world Journal*, 2013.

[7] A. Nurunnabi, D. Belton and G. West, "Robust segmentation in laser scanning 3D point cloud data," *International Conference on Digital Image Computing and Applications*, Fremantle, Australia, 3–5 December, pp. 1–8, 2012.

[8] A. Nurunnabi, D. Belton and G. West, "Robust statistical approaches for local planar surface fitting in 3D laser scanning data," *ISPRS journal of photogrammetry and Remote Sensing*, pp. 106–122, 2014.

[9] P.H.S. Torr and A. Zisserman, "MLESAC: a new robust estimator with application to estimating image geometry," *Journal of Computer Vision and Image Understanding*, vol. 78, no. 1, pp. 138–156, 2000.

[10] G. Vosselman, and R. Klein, "Visualization and structuring of point clouds," *Airborne and Terrestrial Laser Scanning*, vol. 1, pp. 43–79, Whittles Publishing, Scotland, UK, 2010.

[11] K. Klasing, D. Wollherr, and M. Buss, "A clustering method for efficient segmentation of 3D laser data," *IEEE International Conference on Robotics and Automation*, pp. 4043–4048, 2008.

[12] M. Ester, H. P. Kriegel, J. Sander and X. Xu, "A density based algorithm for discovering clusters in large spatial databases with noise," *2nd International Conference on Knowledge Discovery and Data Mining*, pp.226–231, 1996.

[13] "Object Detection and Orientation Estimation Benchmark," 2012. [Online]. Available: http:// www.cvlibs.net/datasets/kitti/eval-object.php. [Accessed 2015].