MOD : Multi-camera based Local Position Estimation for Moving Objects Detection

JINWOO KIM
Autonomous Driving System Research Group
Electronics and Telecommunications Research Institute (ETRI)
Daejeon, Republic of Korea
jwkim81@etri.re.kr

Yongbon Koo
Autonomous Driving System Research Group
Electronics and Telecommunications Research Institute (ETRI)
Daejeon, Republic of Korea
ybkoo@etri.re.kr

SungHoon Kim
Autonomous Driving System Research Group
Electronics and Telecommunications Research Institute (ETRI)
Daejeon, Republic of Korea
saint@etri.re.kr

***Abstract*— MOD (Moving Object Detection) technology is combined by using recognition, ID tracking, detection and classification by using sensor fusion to get information that local & global position estimation, pose estimation, velocity from around objects in real time over 15 fps. To get more exact information, the object detection and classification should be simultaneously processed. Additionally, this object detection results must have high speed processing time performance in restrict HW platform environment for autonomous vehicle. To solve this problem, we use DARKNET based deep learning method and modified detector to obtain local position estimation [1]. And we use UDACITY self-driving [2] and our additional road environment dataset to learn network. Our main purpose is to get moving object local position information from multi cameras fusion. So, we made the fusion server to synchronize and converse multi objects information from four cameras on our autonomous vehicle. In this paper, we introduce the method to solve the local position estimation from four camera using by detection and our fusion server. Additionally, we implement the method to solve the problem caused by steep slope and curve road environment while driving.**

Keywords*— *Moving Object Detection, deep learning, local position estimation, sensor fusion

INTRODUCTION

As deep learning based technologies develop, the vehicle technology trend also been rapidly changed to high technology with vison, sensor fusion. Especially, detection and classification technology field has been developed exactly and efficiently for autonomous self-driving system. Many filed of recognizing filed including human, motion, tracking objects, need to detect the objects mentally. The reason is that a self-driving environment is different with ADAS system in the side of assistance driving or not. Previous detection algorithm shows the comparatively good accuracy in the low line performance computing like as HOG, Haar-like feature and so on [3].

Object detection is a technology that deals with recognizing classes of objects and their location. Many areas, including face detection, surveillance, and a self-driving car's vision system, need object detection as their core functionalities. Classic object detection systems use the scheme of feature-based methods, such as Haar-like features extraction [3], and histogram of oriented gradients (HOG) and linear support vector machine (SVM) algorithms

Until now, DARKNET [1] based object detection is proper method to detection in real time. Additionally, comparatively DARKNET can be simply modified to adapt specific purpose. It is well known for its fast processing speed and simple architecture. Indeed, YOLO v2 shows the quite good detection results by using five anchors and high speed processing performance to use our autonomous driving system [1]. B.Li. introduces the recent development of our research on transplanting the fully convolutional network technique to the detection tasks on 3D range scan data. By carefully design the bounding box encoding, it is able to predict full 3D bounding boxes even using a 2D convolutional network using by Kitti Data set [5]. We use the UDACITY self-driving car dataset that include annotation data for Car, Truck, pedestrian classes from CrowdAI(1.5GB) and Autti(3.3GB) Car, Truck, Pedestrian [2].

Fig. 1 show the results of four camera (front : fov30 and 120, side:90fov) based moving objects local position estimations for autonomous driving system.

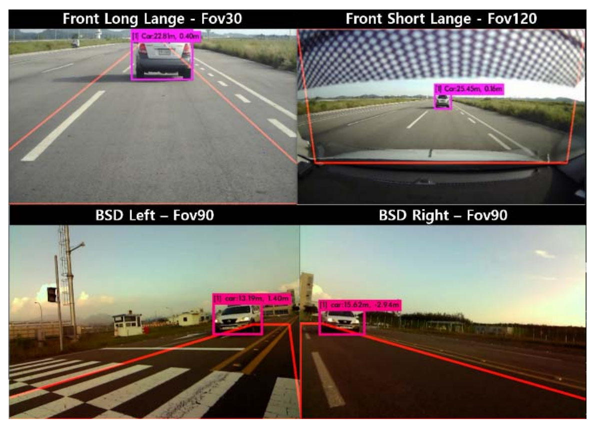

Fig. 1. The results for multi-camera of object detection using YOLOv2

Our main contributions are the following:

- Implementing to get moving object local position information from multi cameras fusion.
- Building up the fusion server to synchronize and converse multi objects information from four cameras on our autonomous vehicle.
- To solve steep slope and curve road constraints, we proposed the method using geometric and steering wheel information from our autonomous vehicle.

Fig. 2 show the system construct for multi cameras and RADAR configuration. Each camera connected DARKNET module conduct the task that detection, classification and local pose estimation. And then our fusion server converse multi camera detector information and load RADAR information.

2375-9356/18/$31.00 ©2018 IEEE
DOI 10.1109/BigComp.2018.00115

IEEE computer society

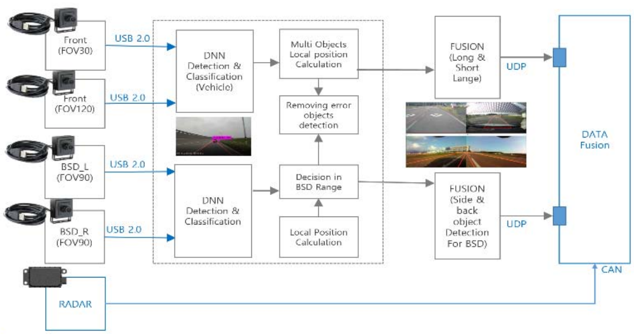

Fig. 2. The system for multi-camera and RADAR based fusion detection

The ‘Auto’ mode is the system priority mode and driving control authority is in the system. And it is possible to the system has control authority, and it is possible to transit to the ‘Auto ready’ mode by the driver's operation.

MAIN PROPOSED METHOD

Fig. 3 show the range of detection for multi cameras configuration. The reason of usage for multi cameras is to cover long and short range of FOV(field of view). To solve long and short range, we use the each 30 and 120 degree of FOV. The cover range is 100 and 30 meter in case of HD resolution (1280X720) using by our trained network with UDACITY and our data set. We also use the two cameras with FOV90 located both side mirror on vehicle to cover the BSD (Blind Spot Detection).

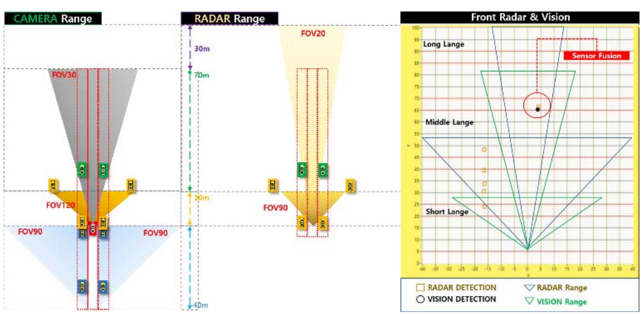

Fig. 3. The distance range of multi-camera based object local position

Actually, we proposed the method to determine more the exact following vehicle position that is on left, back or right lane in steep curve slope environment. To clarify the range of BSD we use the steering wheel angle that distinguish lanes based on current driving lane position instead of lane detection. The button of Fig. 1 shows the results of Blind Spot Detection range using by our proposed method.

To get local position estimation, we actually obtain the physical positon of object on the road. And then, we extract the four position that we previously defined points from image. Finally, we obtain the intrinsic and extrinsic parameter after processing the perspective transform and generator in Fig. 4. And these parameters are used at the detector based on DARKNET and then calculating the local position including the distance of object.

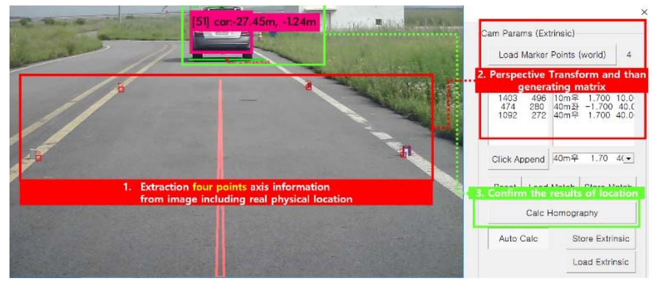

Fig. 4. The method for mono camera based local position estimation

CONCLUSIONS

In this paper, we presented the CUDA-based DARKNET using UDACITY dataset. We also compared its performance with the original DARKNET for NVIDIA GTX TITANX with CUDA 8.0 and cuDNN 5.1 [6].

As future work, we will optimize the DARKNET based detection algorithm and process learning the diverse data from real road environments. Also, we will expand the range of sensors fusion work including RADAR, Lidar. Additionally, our results shows the problem for steep slope that uphill, downhill steep slope environment. This cause the error for distance that is calculated mono camera based local position estimation. We use the yaw, roll, pitching information to reduce the error from diverse road environment up, down, side slope. However, we fully solved the problem that the local position estimation limit of around objects located on the steep slope.

ACKNOWLEDGMENT

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT) (No. 2016-0-00004, Development of Driving Computing System Supporting Real-time Sensor Fusion Processing for Self-Driving Car).

REFERENCES

[1] J. Redmon, DARKNET, https://github.com/pjreddie/DARKNET
[2] https://github.com/UDACITY /self-driving-car/tree/master/annotations
[3] P. Viola, and M. Jones, “Robust real-time object detection,” International Journal of Computer Vision, 2001.
[4] J. Redmon, and A. Farhadi, “YOLO9000: Better, Faster, Stronger,” arXiv preprint arXiv:1612.08242, 2016.
[5] B.Li, T.Zhang and T.Xia, "Vehicle Detection from 3D Lidar Using Fully Convolutional Network," CVPR, 2016.
[6] NVIDIA, CUDA Technology, http://www.nvidia.com/CUDAStefan Lüke, Oliver Fochler, Thomas Schaller, Uwe Regensburger. " Traffic-Jam Assistance and Automation." Springer reference., 1-13, 2015.