DSP 기반 Object Recognizer 설계

신성윤[O], 조승표[*], 조광현[**]
[O]군산대학교 소프트웨어학부,
[*](주)에이치브레인,
[**]군산대학교 수학과
e-mail: s3397220@kunsan.ac.kr[O], spcho@hbrain.co.kr[*], gwanghyun@kunsan.ac.kr[**]

DSP-based Object Recognizer Design

Seong-Yoon Shin[O], Seung-Pyo Cho[*], Gwanghung Jo[**]
[O]School of Computer Sci. and Software, Kunsan National University,
[*]Hbrain Co. Ltd.,
[**]Dept. of Mathematics, Kunsan National University

● 요 약 ●

차량에는 사각지대가 운전자에게 쉽게 보이지 않으며, 안전지대를 보이게 하는 시스템이 많이 있다. 본 논문에서는 사각지대 차량 안전 지원 시스템의 사물인식 개발을 위한 요구사항, 사물인식 네트워크 구조, 훈련 기법에 대해 살펴본다.

키워드: 사각지대(blind spot), 차량(vehicle), 안전지대(safety zone)

I. Introduction

Have you ever been startled by another vehicle or person that suddenly appeared out of sight while driving? This is because of the blind spot, the area that the driver cannot see. Caution should be exercised as it can lead to a serious accident if rearward vision is not secured due to blind spots [1].

A blind spot refers to an angle at which an object cannot be seen with the naked eye or an angle at which a mirror cannot reflect an object. In particular, the reason why the blind spot of a car is dangerous is because it is easy to lead to an accident. In fact, it is said that 8 out of 10 drivers have experienced an accident due to a blind spot. In addition, many novice drivers have difficulty changing lanes or parking due to blind spots [1].

II. Requirements

○ The fusion sensor part consists of a camera part and a TOF sensor part.

○ The basic operating principle of the fusion sensor is to recognize the type of object through camera image information about the object location collected by the TOF sensor by matching the pixel coordinate system of the camera and the distance coordinate system of the TOF sensor in advance.

○ In order to prevent collisions in blind spots around large vehicles such as buses and trucks, the performance and functional requirements of tableware, which is an object, are as follows.

- Possible to recognize objects (pedestrians, bicycles, two-wheeled vehicles) in the blind spot of the vehicle while the vehicle is stopped / driving at low speed
- Fast object recognition performance of 30 frames per second
- Even if part of the object is partially covered by surrounding obstacles, it must be recognizable
- The camera must be able to recognize the object regardless of movement (still, walking) while moving
- Must be able to extract object information by separating it from surrounding obstacles
- Must be able to fuse with the TOF sensor part of the fusion sensor
- Small size, low power consumption, high performance device should be connected to vehicle embedded system
- Possible to input video of at least 4 channels
- Must be able to connect regardless of digital camera SerDes

chip

- Deep learning-based image recognition processing must be completed within the device itself, and external interfaces such as Ethernet, serial, and GPIO must be supported so that recognition results can be delivered to external devices.
- It must be able to be configured to operate independently of the external system, and it must be manufactured in a modular form according to the system design to be able to be configured as an embedded system.

III. Network Architecture

- The existing VGG-16 is inserted at the front of the SSD, and the kernel size of the first convolutional layer is adjusted so that 4 channel inputs can be entered.
- Apply 5 convolution layers to Extra Feature Layers at the back
- Extra Feature Layers are used to predict bounding boxes of various sizes through feature maps in multiple layers.
- Using feature maps from a total of 6 convolution layers
- If k boxes can be predicted in each feature map, in a feature map of size m x n, the network gives a total of c class scores and 4 box coordinates as a result, resulting in a total of (c + 4)kmn bounding box outputs. being

IV. Training Technique

- Use of 5,000 training data
- Fixed input data size to 600x150
- Experimental data uses KITTI's object detection data set
- The initial value of the learning rate is set to 0.0001
- 100,000 iterations
- When the iteration exceeds 600,000, the learning rate is reduced by a factor of 1/10.
- The training method uses stocastic gradient descent
- Cost Function:

$$L(p_i, t_i) = \frac{1}{N}\sum_i (L_{conf}(p_i, p_i^*) + \alpha L_{loc}(t_i, t_i^*))$$

  - i : index of bounding boxes
  - p_i : Probability of predicted label
  - p_i^* : ground-truth label
  - t_i : 4 coordinate values of the predicted box
  - t_i^* : coordinate values of the ground-truth box

• Mini-batch size set to 16

V. Conclusion

In this paper, the object recognition development requirements of the blind spot vehicle safety support system and the object recognition network structure and training techniques were examined.

REFERENCES

[1] How to deal with driving blind spots that are too scary for novice drivers, Bullsone, Available : https://blog.bullsone.com/2534