

## Vanishing point-based 3D object detection method for improving traffic object recognition accuracy

Jeong-In Park\*

\*Ph.D, UFM Systems Co., LTD, Anyang-si, Korea

### [Abstract]

In this paper, we propose a method of creating a 3D bounding box for an object using a vanishing point to increase the accuracy of object recognition in an image when recognizing a traffic object using a video camera. Recently, when vehicles captured by a traffic video camera is to be detected using artificial intelligence, this 3D bounding box generation algorithm is applied. The vertical vanishing point (VP1) and horizontal vanishing point (VP2) are derived by analyzing the camera installation angle and the direction of the image captured by the camera, and based on this, the moving object in the video subject to analysis is specified. If this algorithm is applied, it is easy to detect object information such as the location, type, and size of the detected object, and when applied to a moving type such as a car, it is tracked to determine the location, coordinates, movement speed, and direction of each object by tracking it. Able to know. As a result of application to actual roads, tracking improved by 10%, in particular, the recognition rate and tracking of shaded areas (extremely small vehicle parts hidden by large cars) improved by 100%, and traffic data analysis accuracy was improved.

▶ **Key words:** Vanishing Point, 3-Dimension, Bounding Box, Object Detection, Vehicle Tracking

### [요 약]

이 논문은 영상 카메라를 이용하여 교통 객체를 인식하고자 하는 경우, 영상 내 객체 인식 정확도를 높이기 위해 소실점을 이용하여 객체에 대한 3D 바운딩 박스를 생성하는 방법이다. 최근 인공지능을 이용하여 교통 영상 카메라로 촬영된 차량을 검출하고자 하는 경우 이 3D 바운딩 박스 생성 알고리즘을 적용하고자 한다. 카메라 설치 각도와 카메라가 촬영한 영상의 방향성을 분석하여 종 방향 소실점(VP1)과 횡 방향 소실점(VP2)을 도출하고 이를 기반으로 분석 대상 동영상에서 이동하는 객체를 특정하게 된다. 이 알고리즘을 적용하면 감지된 객체의 위치, 종류, 크기 등 객체 정보 검출이 용이하고, 이를 자동차와 같은 이동류에 적용하는 경우 이를 트래킹하여 각 객체가 이동한 위치와 좌표, 이동속도 및 방향 등을 알 수 있다. 실제 도로에 적용한 결과 트래킹이 10% 향상되었으며 특히 음영지역(큰 차에 가려진 극히 적은 차량 부위)의 인식율과 트래킹이 100% 개선되는 등 교통 데이터 분석 정확성을 향상시킬 수 있었다.

▶ **주제어:** 소실점, 3차원, 바운딩 박스, 객체인식, 차량 트래킹

- 
- First Author: Jeong-In Park, Corresponding Author: Jeong-In Park
  - \*Jeong-In Park (jipark@ufmsystems.co.kr), UFM Systems Co., LTD
  - Received: 2022. 12. 27, Revised: 2023. 01. 16, Accepted: 2023. 01. 17.

## I. Introduction

루프식 검지 장치나 레이저와 같은 기계적인 검출이 아닌 영상 카메라에 의한 동영상 분석함으로써 이를 기반으로 한 프로그램적인 제어에 대한 연구가 다양한 분야에서 진행되고 있다. 교통분야에서도 예외는 아니며, 카메라로부터 획득된 동영상에서 차량 흐름을 분석하고 이를 신호체계에 반영하는 기술은 최근에는 AI를 접목한 단계로 발전하고 있다.

차량의 원활한 흐름을 위해서는 프로그램에 의해 분석된 지능적인 신호체계가 필수적이며, 특히 이를 위해 이동류의 검출과 이동류 트래킹 검출에 대한 정확도가 보장되어야 한다. 이는 객체의 정확한 검출이 보장되어야만 그 도로 상황에 맞는 정확한 신호체계를 제어할 수 있기 때문이다.

동영상에서의 이동류 검출을 위한 알고리즘은 많이 소개되었으나 객체 인식의 정확도 측면에서는 중첩 객체, 큰 객체에 가려진 객체의 인식 오류 등 아직 개선해야 할 부분이 많이 있다.

따라서 이 논문에서 우리는 영상 내 객체 인식 정확도를 높이기 위해 소실점을 이용하여 객체에 대한 3D 바운딩 박스를 생성하는 알고리즘을 소개한다. 또한 이러한 3D 바운딩 박스 생성 알고리즘을 인공지능의 객체 검출 수단에 적용함으로써 검출된 객체 정보를 기반으로 교통 신호 제어 체계에 응용할 수 있는 개선된 객체 감응 알고리즘을 소개한다.

3D 공간에서 원근을 나타내는 라인들이 영상 카메라로 얻어진 2D 공간으로 투영될 때 영상평면에서 한 지점으로 만나게 된다. 이 점을 소실점이라고 하며[1], 3D 영상에서의 소실점 연구는 다양한 분야에서 진행되고 있다. 카메라 렌즈의 중심에서 영상으로 뻗어있는 벡터는 평행한 선 벡터를 의미한다. 따라서 소실점 검출과 영상 내의 적용은 3D 장면을 분석함에 있어서 컴퓨터비전 분야에서 중요한 요소로 활용된다. 소실점은 3D 영상에서 많은 중요한 정보를 제공하며[2, 3], 카메라와 오브젝트 간의 거리나 각종 파라미터 계산에도 유용하게 사용된다[4].

영상에서 소실점을 검출하는 방법은 직선들이 교차하는 지점을 찾는 것이다. 실제 물체의 위치와 영상에서의 위치는 거리 특징을 가지고 있다. 즉, 물체가 소실점에 가까이 있을 수록 영상에서 나타나는 물체의 크기는 더 작아진다. 일반적으로는 소실점이 영상의 중앙 근처에 형성된다는 가정 하에 중앙에 가까워질수록 물체의 크기는 더 작게 보인다. 이 점을 이용하여 3D 정보 추출에 이를 적용할 수

있다. 그러나 이러한 영상 기반 접근방법은 3D 정보 검출 시 다소 정확도가 떨어지는 문제점을 가지고 있다. 이러한 문제점 들을 해결하기 위해 영상 내 특징에 기반한 검지 기법들[5, 6]이 연구되기도 하였다. 이 기법들은 확률 분포 모델이나 기하학적 매핑 모델 등을 기반으로 하는 통계적, 수학적 모델링 분석을 이용하였다.

이처럼 이동류 객체 인식 정확도를 높이기 위해 소실점 기반 3D 객체 탐지기법을 적용하되 좀 더 우수한 검지 성능을 보장하기 위해 검지의 결과물을 학습하고 이를 분석하는 딥러닝(Deep Learning) 알고리즘 기술을 함께 적용하고자 한다.

특히, 딥러닝 기술을 결합하는 이유는 최근 들어 딥러닝 기법을 이용하여 객체 감지, 종류, 트래킹(위치추적) 등을 수행하기 위한 영상 분석 기술에 관한 연구가 활발하게 이루어지고 있고, 이러한 딥러닝을 활용한 영상 분석 기술은 학습을 통해 자체적으로 오차율을 줄이기 때문에 학습량과 활성화 함수에 따라 검지의 정확성 및 정밀도를 높일 수 있는 장점이 있으며, 이에 따라 그 적용 분야가 기하급수적으로 증가하고 있기 때문이다.

이 알고리즘을 적용하고 테스트하기 위한 시스템은 도로상에 설치된 영상 카메라와 이 영상 카메라로부터 획득한 영상에서 오브젝트 추출을 위해 소실점 기반 3D 바운딩 박스를 생성하고 그다음 딥러닝을 적용하여 오류를 줄이고 객체 인식의 정확도를 높일 수 있는 시스템으로 구성되어 있다. 이 때, 특정 차선의 차량 대기가 길어지거나 비상 상황이 발생하는 경우 신호제어기로 제어를 전달하여 연장, 단축 등의 신호체계를 변경할 수 있는 장비도 또한 구성되어 있다.

이 논문에서 소개하는 알고리즘인 소실점을 이용한 3D 바운딩 박스 생성 방법에 따르면 객체 인식률을 개선 시켜 분석의 정확성 및 신뢰도를 획기적으로 높일 수 있다. 또한, 이 시스템이 제공하는 교통신호 제어를 위한 객체 감응형 차량 흐름 제어 방법에 따르면 카메라들에 의해 촬영된 영상을 분석하여 객체 정보 및 교통정보를 생성한 후, 생성된 교통정보에 따라 최적 현시 체계를 생성한 후, 생성된 최적 현시 체계에 따라 교통신호가 제어되도록 구성됨으로써 차량 정체 및 대기시간을 현저히 절감시킬 수 있을 뿐만 아니라 이에 따른 연비 소모 및 오염 가스 배출량을 효과적으로 줄여 사회적 비용 소모를 절감시킬 수 있게 된다.

## II. Related Works

### 1. Vanishing Point Based Object Recognition

소실점 검출 관련해서 본 연구에서 연구 대상 분야는 영상 내 객체의 움직임에 관한 분야로써 도로 모니터링 동영상을 주된 대상으로 한다. 이러한 목적에 맞도록 이 논문에서 소실점을 검출하는 방법은 다음과 같다.

먼저, 영상 내의 엣지 세그먼트 영역을 구한다. 특히, 소실점과 관련이 있는 엣지는 한 면과 다른 면이 연결된 교차점, 카메라 렌즈의 중심을 지나는 벡터와 평행한 에지 성분 등이 핵심 키가 된다. 이러한 에지 성분을 일반적으로 소실라인이라고 하며 소실점을 생성한다. 이 소실라인은 한 개 이상이 될 수 있으며 특정한 지점에서 만나는 것이 정확하지만, 실제로는 어떤 한 점을 형성하기보다는 일정한 범위를 가지며 그 점이 교차하고 있다.

원근법에 따라 물체는 눈에 가까울수록 크게 투영되며, 멀어질수록 점점 작게 투영된다. 이를 선으로 표시하면 시점에서 멀어질수록 선과 선이 만나는 지점 결국 소실점을 형성하게 된다. 이를 위해 하프(Hough) 변환[7] 알고리즘을 적용하고자 한다.

소실점은 영상 내에서 한 개 이상이 될 수 있으며 이는 하프변환을 기반으로 검지한다. 이 후 하프변환으로 추출된 직선을 이용하여 소실점을 추정한다. 이 경우 하프변환으로 얻어진 직선을 바탕으로 소실점을 획득하는 데에는 몇 가지 문제점들이 존재한다.

먼저, 카메라 시점과 동일한 방향으로 평행인 직선들은 소실점을 추정하기에 매우 용이하지만, 만약 카메라 시점 방향과 이 직선이 직각인 경우는 소실점 추정이 어렵다. 두 번째, 소실점은 한 개의 점으로 형성되지 않는 경우가 많은데 이번 논문도 마찬가지인 경우다. 이는 소실점이 영상 밖에서 형성되는 경우이며 혹은 얻어진 직선 성분이 부정확한 경우 등이 원인이다.

이 접근법은 소실점을  $VP(x_0, y_0)$ 라 할 때, 하프 공간에서 식 (1)의 함수인  $J(x_k, y_k)$ 가 최소가 되는 값을 갖는  $(x_0, y_0)$ 를 얻는 것이다.

$$J(x_k, y_k) = \sum_{i=0}^n W_i (\rho_i - x_k \cos \theta_i - y_k \sin \theta_i)^2 \quad (1)$$

$$W_i = \frac{\nu_i}{V} \quad (2)$$

여기서,  $W_i$ 는 가중치로서,  $\nu_i$ 는  $(\rho_i, \theta_i)$  쌍의 수,  $V$ 는 전체 좌표 수 이다. 식(1)은 여러 개의 직선성분에 대응하는 여러 쌍의 정현성분신호를 갖는  $(\rho, \theta)$  공간에서 임의의 점  $(x_k, y_k)$ 과 이 정현성분신호들 간의 차가 최소가 되는 점을 찾는 식이다.

단, 위의 기법[8, 9]은 수직성분 등이 너무 많이 존재하여 소실점 수렴이 잘 되지 않는 경우 본 알고리즘은 소실점을 적절히 추정하지 못한다. 따라서 이 논문에서는 이러한 문제들을 해결하기 위하여 카메라 초점의 방향과 직각 방향의 직선 에지 성분을 제거함으로써 이러한 문제점을 해결하였다.

### 2. Deep Learning Algorithm for Car Object Recognition

최근 대부분의 영상식 카메라는 딥러닝 기반 객체 인식 알고리즘을 이용하여 출시되고 있다. 차량이 카메라 화각에 진입하면 영상식 카메라가 차량을 찍고 이와 동시에 소실점 기반 오브젝트 인식을 수행한다. 뒤이어 딥러닝 알고리즘이 작동하여 인식된 오브젝트에 대해 보정을 하는 등 보다 정확하게 차량 객체를 인식하고 추출한다.

딥러닝은 인간의 정보전달 방식과 유사한 방식으로 노드 상호관계를 정의하고 수많은 필터를 통해 특징들을 추출해서 분류한 다음 추론한 오차를 근거로 가중치를 변경하며 최적의 가중치를 가지는 모델을 생성하는 학습 방법 중 하나이다. 기존 영상에서 특징적인 물체의 노드 간 상호관계와 가중치를 딥러닝 모델로 저장하여 새로운 영상 분석에 활용한다.

CNN (Convolutional Neural Network)은 영상 인식에서 사용되는 대표적인 방법으로 촬영된 이미지를 매트릭스로 수치화 하여 데이터를 분석한다. Fig. 1에서 C1 ~ C5까지는 데이터 처리를 위해 분할된 이미지 영역의 3차원 매트릭스를 분석하고, FC6, FC7에서는 이를 2차원으로 요약/정리하여 활용하며 이렇게 정리된 매트릭스는 딥러닝이나 영상분석에 사용된다[10].

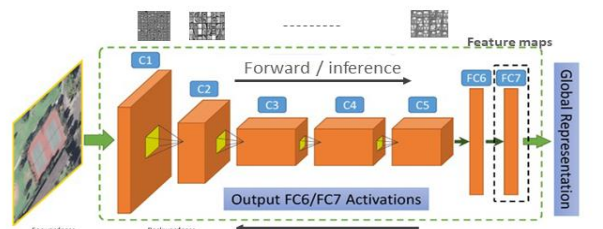


Fig. 1. The structure of CNN

인공지능 모델에서 차량 인식 추출의 오류를 줄이기 위해서 특수 목적을 가진 데이터 셋이 필요하며 이를 이용한 인공지능 모델 개선을 추가 학습하게 되는데 이를 미세조정 (Fine Tuning)이라 한다. 특수목적성 데이터셋을 사용한 예로 미국 알바니 대학교(University of Albany)에서 차량의 촬영 환경 (날씨, 시간 등)에 따라 미세 학습 시 다양한 환경에서의 차량 인식 정확도를 높인 사례를 찾아볼 수 있으며 이를 응용하여 본 논문에서는 YOLO(You Only Look Once)를 도입하여 구현하였다. Fig. 2에서 YOLO는 각 이미지를  $S \times S$ 개의 그리드로 분할하고 그리드의 신뢰도를 계산하여 경계상자의 위치를 조정함으로써 객체 인식이 뛰어나기 뿐만 아니라 실시간 처리가 가능한 장점을 갖는다[11, 12].

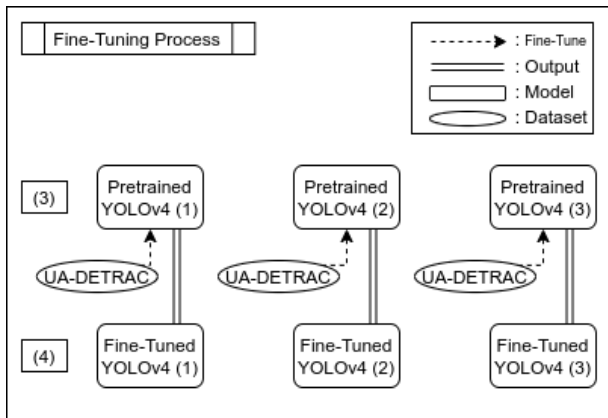


Fig. 2. Introduction of YOLO v4 for fine learning

### 3. Car Queue Detection Development Technology

우리 나라의 대부분의 교차로는 일정한 신호주기 및 현 시 순서에 의해 운영되거나 신호제어기에 내장된 시간대별 신호 시간(TOD), 요일별 신호시간(DOW) 및 특별일 신호 시간 테이블(Holiday Plan Table)에 의해 운영되는 고정주기식 신호체계방식인 일반신호기로 운영되고 있다.

이러한 고정주기식 신호체계방식은 계절별, 요일별, 시간대별로 교통량의 변동이 큰 교통 특성을 잘 반영하지 못하기 때문에 신호시간의 낭비로 도로상에 불필요한 지체를 야기하고 있으므로 감응신호제어 신호운영방식은 고정 시간 시간 제어 운영에 비하여 수요변동이 심한 교차로에서의 지체를 크게 줄일 수 있다.

특히, 좌회전 감응신호제어는 좌회전 교통류 상태를 파악하여 좌회전 신호를 효율적으로 사용함으로써 다른 현 시의 이동류가 남은 신호 시간을 이용하게 하여 연동제어와 동시에 신호 효율을 극대화할 수 있다.

이상의 관련 연구에서 살펴본 바와 같이 개발하고자 하

는 오브젝트 검지 시스템은 Table 1에서와 같이 소실점 기반 3D 객체 검지를 기반으로 차량을 검지하고, 검지한 후 즉시 딥러닝 기술을 이용하여 차량 대수, 대기열 길이를 추출함으로써 검지율 100%를 구현한다. 대기열 차량에 대해 정확한 검지를 기반으로 도로정보에 대한 신뢰성을 보장하고 신호 운영 측면에서 차량이 뜬 외곽 도로의 불필요한 정기적 직진 또는 좌회전 신호 낭비, 불법유턴, 불법 직진 등의 신호체계를 개선함으로써 정확한 통행 차량 현황 정보를 제공하고자 한다.

Table 1. The excellence of the proposal system

	Detection	Recognition	Accuracy
Vanishing Point 3D Detection + YOLO, CNN	★★★★★	★★★★★	99.9% (100%)
Deep Learning Detection (Only)	★★★★☆	★★★★☆	98% ~ 95%
Other Sensor	★★★★☆	★★★★☆	about 95%

## III. System Development

### 1. Vanishing Point-based Moving Object

#### Detection

현실의 3차원 공간은 카메라의 2차원 평면영상에 투영될 때 원근감이 있는 거리정보를 정확히 나타내지는 못한다. 즉, 가까이 있는 물체는 크게 투영되며 카메라로부터 멀리 떨어져 있을수록 점점 작은 크기로 표현된다.

객체를 검출할 때 이러한 3차원 공간상의 깊이 차로 인하여 몇 가지 문제점이 발생 된다. 첫째, 객체 검출을 위해 엷지를 적용할 때, 객체의 크기를 무시하고 동일한 엷지를 적용하는 경우 카메라에서 먼 거리에 있는 객체는 잘 인식하지 못하는 상황이 발생 된다. 둘째, 엷지를 통해 얻어진 객체에 대해 모폴로지 연산과 같은 형태학적 필터링을 적용하고자 하는 경우 모든 객체에 대해서 구성요소의 크기를 동일하게 적용할 때 원하는 처리를 기대할 수 없다. 따라서 이 논문에서는 소실점 기반의 객체 검출 접근 방법을 통해 이 두 가지 문제점을 개선하고자 한다.

#### 1.1 Object Unit Distance Calculation

Fig. 3은 도로의 차로를 위에서 바라본 평면도의 개념도이다. 여기서  $z_1$ 과  $z_2$ 에서 차로의 너비가 각각  $x_1, x_2$  이며  $x_1 = x_2$  라고 가정한다. 이 조건에서, 카메라로 투영

되는 영상은 소실점으로 부터 거리가 각각  $z_1, z_2$ 인 거리에 있는 물체는 원근 투영에 의해

$$\frac{x_1}{z_1} = \frac{x_2}{z_2} \quad (3)$$

와 같이 식(3) 으로 표현가능하다.

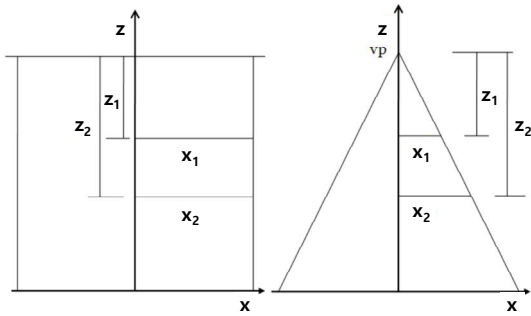


Fig. 3. Real Road and Image Plane

Fig. 3 과 식 (3)에 따라 실제로 동일한 크기의 물체가 영상 면에 나타날 때, 그 물체의 크기는 소실점과 물체 간의 거리에 비례한다. 이러한 특성을 이용하여 도로상의 자동차 주행 영상에 적용하고자 한다.

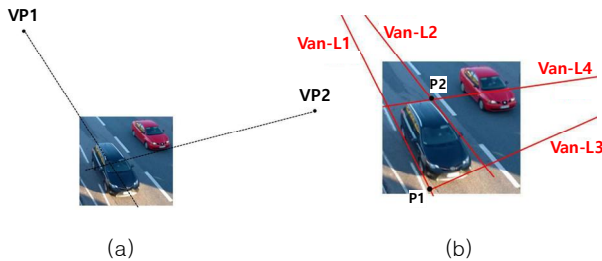


Fig. 4. Two Vanishing Point of Moving Object

Fig 4에 나타난 영상에서 소실점은 VP1, VP2로 표시된 점이다. 자동차를 촬영하기 위한 카메라는 정해진 각도를 부여하는 것이 아니라 자동차가 쉽게 식별될 수 있는 위치에 카메라를 설치한다. 따라서 자동차의 영상은 반듯한 사각형이 아닌 사다리꼴 또는 평행사변형에 가깝다. 실제로 카메라와의 거리가 거의 비슷한데도 불구하고 (b)영상에서의 P1과 P2 는 상당히 차이가 나는 것처럼 투영되어 있다.

이 논문에서는 이 점을 고려하여 소실점과 소실점이 중첩되는 지점을 대각선으로 마주 보게 배치하여 하나의 사각형을 도출 함으로써 전체적인 바운딩 박스를 완성하게 된다.

## 1.2 Creating a Bounding Box Considering The Distance from The Vanishing Point

본 절에서는 앞절에서 제시한 차량 객체 거리 검출 방법을 기반으로 차량에 바운딩 박스를 그리는 방법과 이를 이용하여 객체를 검출하는 방법을 설명한다. Fig. 5는 객체 검출 단계를 나타낸 흐름도이다.

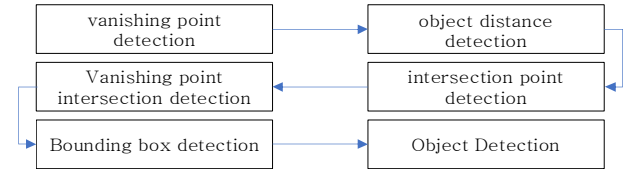


Fig. 5. Vehicle Object Detection Process

### 1.2.1 Change Area and Object Distance Detection

만약  $x_i$ 를 입력 영상이라고 하고,  $x_b$ 를 배경 영상이라고 할 때, 변화영역은

$$y = \begin{cases} 1, & \text{if } (|x_i - x_b| > 7) \\ 0, & \text{else} \end{cases} \quad (4)$$

를 통해 얻는다.

먼저 객체별로 크기를 계산한다. 만일 영상  $R$ 이 유한 개의 영역  $R_i$ 개로 구성되어 있을 때, 영역  $R_i$ 중에서 화소의 개수가 지정된 오티지  $T_r$ 보다 클 경우 움직이는 객체로 간주 되고 그렇지 않으면 잡음으로 처리된다. 즉,

$$\alpha(R_i) > T_r \quad (5)$$

이면 영역  $R_i$ 는 움직이는 객체로 간주할 수 있다. 이때, 소실점으로부터의 거리를 고려하여

$$\alpha(R_i) > kT_r \text{ 및 } k = \beta \cdot (b_i - p) \quad (6)$$

(6)과 같이 변형된 형태를 적용하고자 한다. 이 때,  $k$ 는 소실점으로부터의 거리에 비례하는 값을 의미하며,  $p$ 는 영상에서 얻어진 소실점의 위치를 나타낸다.  $b_i$ 는  $R_i$ 의 소실선과 접하는 위치를 나타내고 있다.

### 1.2.2 Bounding Box Detection

카메라 설치 각도와 영상 카메라가 촬영한 영상의 방향성을 분석하여 종방향 소실점 VP1과 횡방향 소실점 VP2를 도출한다.

VP1에서 자동차 객체의 외연에 접하도록 이어진 좌단 하부 소실선 Van-L1과 우단 상부 소실선 Van-L2를 생성하고 VP2에서 자동차 객체의 외연에 접하도록 이어진 전단 하부 소실선 Van-L3와 후단 상부 소실선 Van-L4를 생성하여, Van-L1과 Van-L3의 교차점 P1과, Van-L2와 Van-L4의 교차점으로 P2를 생성한다. (Fig. 4 (a)(b))

자동차 객체 최좌단이 접하는 좌측 후방 수직선 Ver-L1과, 객체 최우단이 접하는 우측 전방 수직선 Ver-L2를 생성하여, Ver-L1과 Van-L1의 교차점 P3과 Ver-L2와 Van-L3의 교차점 P4를 생성한다.

그리고 이어서 Ver-L1과 Van-L4의 교차점 P5, Ver-L2와 Van-L2의 교차점 P6가 생성된다. VP2에서 P3를 지나는 소실선 Van-L5와 상기 VP2에서 P6를 지나는 소실선 Van-L6를 생성한다.

P1을 지나는 좌측 전방 수직선 Ver-L3와, P2를 지나는 우측 후방 수직선 Ver-L4를 생성하여, Ver-L3와 Van-L6의 교차점 P7, Ver-L4와 Van-L5의 교차점 P8이 생성된다.

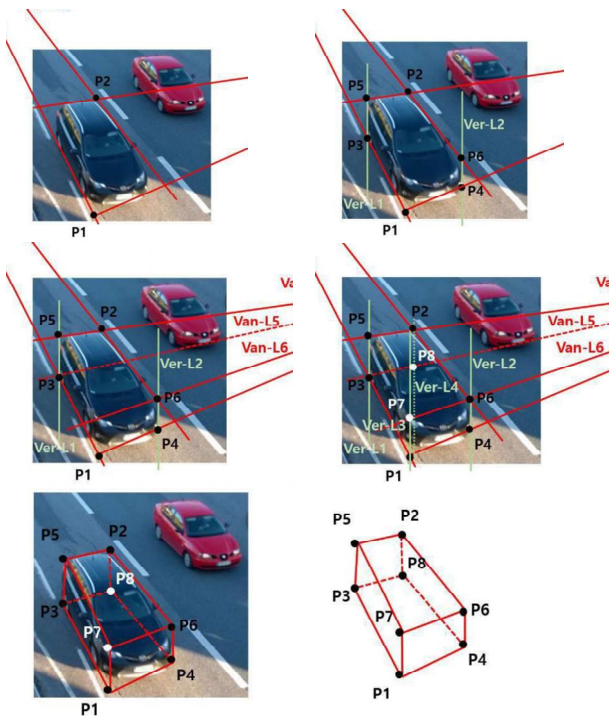


Fig. 6. Bounding Box and Object Detection

결과적으로 P2, P5, P7, P6가 상단 꼭지점을 이루고, P1, P4, P8, P3가 하단 꼭지점을 이루는 6면체 바운딩 박스가 생성된다.

### 1.2.3 Applying Morphology

바운딩 박스를 검출함에 있어서 모폴로지 연산을 적용한 이유는 형태학적 요인으로써 영상 필터링인 팽창과 침식 연산 등을 통해 객체의 외형 부분을 매끄러운 형태로 만들고자 하는 데 있다. 본 논문에서는 얻어진 영역에 대해 닫힘연산(메움 등)을 통해 객체를 추출한다.

## 2. System Detailed Implementation

Fig. 7은 객체 감응형 차량 흐름 제어 방법이 적용되는 차량 흐름 제어 시스템을 나타내는 구성도이다.

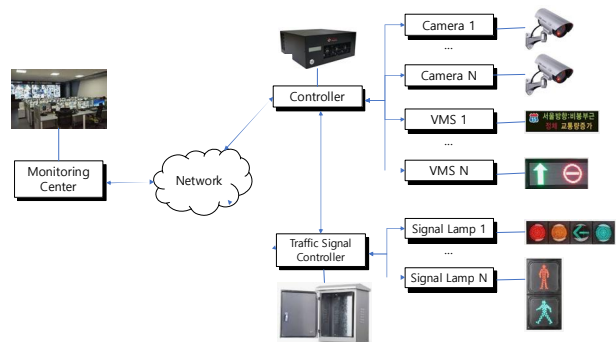


Fig. 7. Vehicle Flow Control System

이를 간단히 살펴보면 도로의 각 차로를 촬영하는 카메라들과 딥러닝 기반의 객체분석 알고리즘을 이용하여 각 카메라의 촬영에 의해 획득된 영상들을 분석하여 객체 분석정보를 검출한 후, 검출된 객체분석정보에 대응하는 최적 신호체계를 생성한다.

생성된 최적 신호체계에 따른 감응신호를 교통신호 제어기로 전송하는 컨트롤러와, 컨트롤러로부터 전송받은 최적 신호체계에 따른 감응신호에 따라 신호등들의 동작을 제어하는 교통신호 제어기로 구성되어 있다.

편의 기능으로 갓길 또는 신호등 지주에 설치되어 컨트롤러의 제어에 따라 기 설정된 경고문구를 디스플레이 하는 VMS와 같은 전시수단들과, 컨트롤러로부터 전송받은 객체 분석 정보, 교통정보, 위반정보, 위반 영상을 저장 및 모니터링 하는 관제센터서버와, 컨트롤러 및 관제센터서버 사이의 데이터 이동 경로를 제공하는 통신망으로 이루어진다.

Fig. 8은 이 논문의 주요 알고리즘에 대한 구성도로써 객체 감응형 차량 흐름 제어 방법은 입력 영상으로부터 객체를 감지하기 위한 딥러닝 알고리즘의 모델을 도로 환경에 적합하도록 최적화시키고 동시에 카메라 촬영에 의해 영상 획득 시 딥러닝을 이용하여 영상을 분석하여 객체(차량 및 보행자)를 검출한다.

이 검출된 객체정보를 기반으로 신호체계를 수립함으로써 영상분석을 통한 객체 감지 및 트래킹의 정확성을 현저히 높임에 따라 신호 제어의 신뢰도를 높일 수 있을 뿐만 아니라 교차로의 차량 대기시간이 현저히 절감되며, 현장에서 실시간 영상처리가 가능하게 된다.

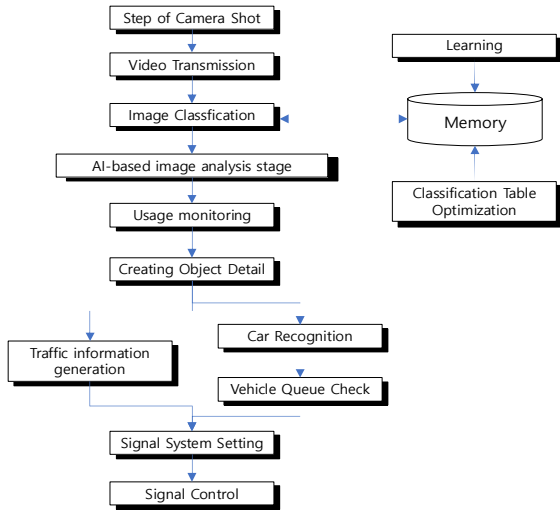


Fig. 8. Object Sensitive Vehicle Flow Control Method

학습단계는 미리 설정된 주기마다 진행되며, 학습단계는 특정 주기 동안 수집된 영상과 기 설정된 객체종류를 활용하여, 영상 및 객체종류 간의 상관관계를 학습할 수 있는 학습데이터를 생성하고 생성된 학습데이터를 활용하여 객체영상 및 객체종류 간의 상관관계에 대한 파라미터 값들의 집합인 추출모델을 도출한다.

또한 이 시스템의 객체 감응형 차량 흐름 제어 방법은 딥러닝 알고리즘으로써 잘 알려진 합성곱 신경망(CNN, Convolution Neural Network) 기반의 YOLO 모델을 적용시킴으로써 객체 인식률을 더욱 높일 수 있게 된다.

딥러닝 알고리즘의 학습 시, 인식대상을 종래에 80종류에서, 도로에서 자주 볼 수 있는 5종류로 제한하여 학습이 이루어지도록 구성됨으로써 객체 인식률을 더욱 개선시킬 수 있다. 이 경우 복수개의 GPU(Graphic Processing Unit)들을 이용하여 영상을 분석함으로써 카메라들의 촬영에 의해 획득된 고용량의 영상들의 실시간 처리 및 분석이 가능하다.

## IV. System Testing and Performance Evaluation

이 시스템을 종합적으로 시험 및 성능을 평가하기 위하여 우선 시험환경을 설정해야 한다. 시험 환경은 현장 Test-Bed의 카메라 영상 장비 및 제어기 등을 구성으로 하는 장비 인프라 환경과 이를 구동하는 프로그램 환경으로 나눌 수가 있다. 장비 인프라 환경은 이 논문을 위해서 구축한 것은 아니며 본 연구과제(사사 참조)를 수행하기 위해 경기도 안양시의 협조를 얻어서 선정된 교차로에 장비를 설치하여 구축한 환경이다.

### 1. Performance Evaluation

시스템 성능 평가는 사실상 결과가 100%로 도출되었기 때문에 성능은 만족할 만한 수준으로 도출되었다. 성능 측정 방법은 차량 인식에 있어서 소실점 기반 3D 객체 탐지 알고리즘을 적용한 경우와 아닌 경우의 검지율을 측정하였으며, 이 알고리즘을 적용하지 않은 경우는 단지 영상 카메라로만 검지하여 이 알고리즘을 적용하기 전후의 성능 개선 변화를 평가하였다.

Table 2. The Comparison of Performance Evaluation

	Test conditions	Result
Detection Rate	With 3D object detection algorithm based on vanishing point	100%
	Without this algorithm, only deep learning is used -YOLO, CNN	95%

검지율에 있어서 이 알고리즘의 적용 유무에 따라 다소 검지율의 차이가 발생함을 알 수 있다. 딥러닝 만으로는 조금 차이가 발생한 검지율을 이 소실점 기반 3D 객체 탐지알고리즘을 적용하면 상호 보완적으로 작동함을 알 수 있었다.

또한, Fig. 9와 같이 5종의 차 종류에 따라 바운딩 박스가 정확하게 그려져 있는 것을 볼 수 있다. 이 또한 차량이 중첩되거나 일부 숨겨져 있는 경우 그 부분만을 바운딩 박스를 인식함으로써 누락 없는 차량의 검지가 가능함을 알 수 있다.

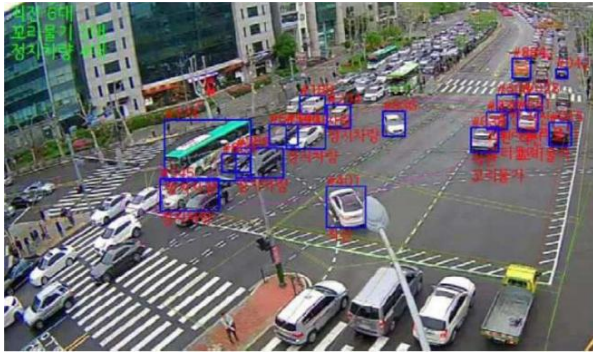


Fig. 9. Vehicle Recognition for Signage Changes

Fig. 10은 Table 2.에 따라 성능을 비교한 결과이다.

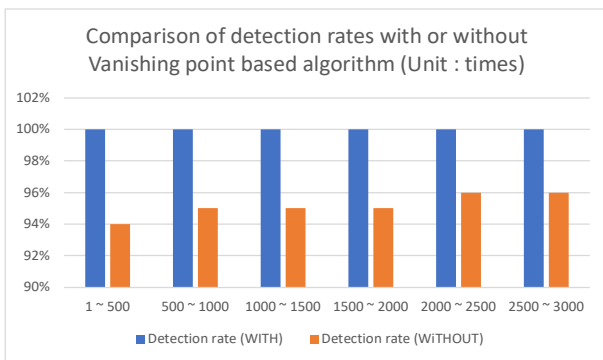


Fig. 10. The result of performance evaluation comparison

### V. Conclusions

이 논문에서 우리는 영상 내 객체에 대한 인식 정확도를 높이기 위해 소실점 기반 3D 객체 탐지 방법을 소개하였다. 그리고 이 알고리즘을 적용한 시스템은 카메라들에 의해 촬영된 영상을 분석하여 객체정보 및 교통정보를 생성한 후, 생성된 교통정보에 따라 최적 현시체계를 생성한 후, 생성된 최적 현시체계에 따라 교통신호가 제어되도록 구성함으로써 차량 정체를 및 대기시간을 현저히 절감시킬 수 있을 뿐만 아니라 이에 따른 연비소모 및 오염가스 배출량을 효과적으로 줄여 사회적 비용 소모를 절감시킬 수 있는 장점이 있다.

이 알고리즘으로 테스트 해본 결과 답러닝 만으로는 약간의 오차가 발생할 수 있는 부분을 이 3D 기반 객체 탐지 알고리즘이 상호보완함으로써 100%에 이르는 인식율을 보장할 수 있게 되었고, 차량의 종류 5종을 지정하여 바운딩 박스를 적용하였기 때문에 잘 보이지 않는 차, 중첩된 차량에 대해서도 완전히 검지할 수 있었다.

본 시스템의 제약사항으로는 카메라의 각도에 따라 소실점 검색에 다소 차이가 발생하지만 강풍이나 충격에 의해 비정상적으로 틀어진 각도의 경우를 제외하고는 대체로 소실점이 정상적으로 탐색되었다.

향 후 연구 방향으로서는 차량과 같은 도로상의 이동류 뿐만 아니라 3차원 공간상의 이동류에 대해 확대 적용하여 연구를 하고자 한다.

### ACKNOWLEDGEMENT

This thesis was conducted with the support from the Ministry of SMEs and Startups' 2020 SME Network Technology Development Project (R&BD). (S3029645)

### REFERENCES

- [1] Y. Horry, K. I. Anjyo, and K. Arai, "Tour into the picture : using a spidery mesh interface to make animation from a single image." Proceedings of SIGGRAPH '97, pp. 225~232, Aug. 1997 <https://doi.org/10.1145/258734.258854>
- [2] F. A. van den Heuvel, "vanishing point detection for architectural photogrammetry." International Archives of Photogrammetry and Remote Sensing, vol. XXXII part5, pp. 652 ~ 659, 1998, [https://www.isprs.org/PROCEEDINGS/XXXII/part5/652\\_XXXII-part5.pdf](https://www.isprs.org/PROCEEDINGS/XXXII/part5/652_XXXII-part5.pdf)
- [3] H. W. Kang, S. H. Pyo, K. Anjyo, S. Y. Shin, "Tour Into the Picture using a Vanishing Line and its Extension to Panoramic Images" Eurographics 2001, Vol. 20, No. 3, Sep. 2001 <https://doi.org/10.1111/1467-8659.00506>
- [4] Joseph Redmon and Ali Farhadi, "YOLOv3: An Incremental Improvement", arXiv preprint arXiv:1804.02767, Apr. 2018. <https://doi.org/10.48550/arXiv.1804.02767>
- [5] Yuan Wang, Tianyue Shi and Peng Yun, et al., "PointSeg: Real-Time Semantic Segmentation Based on 3D Lidar Point Cloud", arXiv preprint arXiv:1807.06288, Sep. 2018. <https://doi.org/10.48550/arXiv.1807.06288>
- [6] Ross Girshick, "Fast R-CNN", Proceedings of the IEEE international conference on computer vision, Santiago, Chile, pp.1440-1448, Apr. 2015. <https://doi.org/10.48550/arXiv.1504.08083>
- [7] Shaoqing Ren, Kaiming He, Ross Girshick, et al., "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", Advances in neural information processing system,



- Montreal, Canada, pp.91-99, Jun. 2015. <https://doi.org/10.48550/arXiv.1506.01497>
- [8] Liu Mingjie, Cheng Bin Jin and Xuenan Cui, et al., "Online multiple object tracking using confidence score-based appearance model learning and hierarchical data association", *IET Computer Vision*, Vol.13, No.3, pp.312-318, Apr. 2019. <https://doi.org/10.1049/iet-cvi.2018.5499>
- [9] Morgan Quigley, Brian Gerkey and Ken Conley, et al., "ROS: an open-source Robot Operating System", *ICRA workshop on open source software*, Vol.3, No.3.2, p.5, Jan. 2009. <http://robotics.stanford.edu/~ang/papers/icraoss09-ROS.pdf>
- [10] B. Wu, X. Zhou, S.Zhao, X. Yue and K.kutzer, "SqueezeSegV2: Improved Model Structure and Unsupervised Domain Adaptation for Road-Object Segmentation from a Lidar Point Cloud", *ICRA*, Vol. 6, No. 3, pp. 1721-1737, Jun, 2019. <https://doi.org/10.48550/arXiv.1809.08495>
- [11] Xiaozhi Chen, Huimin Ma, Ji Wan, et al., "Multi-View 3D Object Detection Network for Autonomous Driving", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, Hawaii, pp.1907-1915, Jun. 2017. <https://doi.org/10.48550/arXiv.1611.07759>
- [12] Pan Wei, Lucas Cagle and Tasmia Reza, et al., "Lidar and Camera Detection Fusion in a Real-Time Industrial Multi-Sensor Collision Avoidance System", *Electronics*, Vol.7, No.84, Jul. 2018. <https://doi.org/10.48550/arXiv.1807.10573>

## Authors



Jeong-In Park received Ph.D. degree in Computation Engineering from In-Cheon University in 2015. He is currently the head of research at UFM Systems Co.LTD's Lab. His primary research interests include big

data, image recognition, and AI.