

# 가상 표지판 학습 데이터를 이용한 SVM 학습 및 인식

김성주<sup>o</sup>, 박순용

경북대학교 컴퓨터 학부

goddoe2@gmail.com, sypark@vision.knu.ac.kr

## 요 약

일반적인 표지판 인식을 위한 학습 과정에서는 실사 영상을 이용한 기계 학습을 이용한다. 높은 정확도의 인식률을 확보하기 위해서는 각 표지판 마다 100 장 이상의 학습 데이터가 필요하다. 하지만 모든 표지판에 대하여 일정 양 이상의 학습 데이터를 확보하기는 용이하지 않다. 본 논문에서 가상 표지판 데이터를 생성하여 충분한 수의 데이터 세트를 확보 후 HOG(Histogram of Oriented Gradients) 특징을 추출하고 SVM(Support Vector Machine)을 학습시킨다. 가상 표지판 데이터로 학습된 SVM 은 높은 인식률을 보인다.

## 1. 서론

자동차 산업에서 현재 가장 활발히 발전하고 있는 분야 중 하나는 ADAS(Advanced Driver Assistance System)이다. ADAS 란 차량의 외부 혹은 내부의 의미 있는 정보를 운전자에게 전달해줌으로써 운전자의 운전을 도와 주는 시스템이다. ADAS 의 대표적인 기능으로는 차선 검출, 차량 검출, 보행자 검출, 표지판 인식 등이 있다. 특히 주행 중 다양한 정보를 제공하는 도로 표지판의 경우 정확한 인식이 필요하다.

기존에 표지판인식 방법에서는 주로 표지판에 대한 haar like feature 나 HOG 와 같은 특징을 추출한 후 SVM 과 CNN(Convolutional Neural Network) 또는 Adaboost 와 같은 기계학습을 방법을 이용하였다[1].

표지판의 기계 학습을 할 때 높은 인식률을 위해서는 일반적으로 각 표지판 마다 100 장 이상의 학습 영상 데이터를 사용한다. 하지만 수십 종류의 표지판에 대해 다양한 자세와 조명 환경에서 실사 표지판 영상을 획득하기 위해선 많은 시간과 노력이 필요하다. 예를 들면 낙석 도로 주의 표지판, 도로 공사 중 표지판, 터널 표지판 등은 일반적인 곳에는 설치되어 있지 않기 때문에 실사 표지판 영상을 획득하기는 어렵다.

학습 데이터를 획득하기 위한 시간과 노력을 줄이고 다양한 환경의 데이터를 확보하기 위해 가상 학습 영상 데이터를 사용한 연구들이 있었다. 보행자 검출 분야에서는 가상 세계의 보행자들을 이용하여 학습 데이터를 획득한 후 기계 학습한 연구가 있었다[2]. 인간의 3D 모델을 여러 배경 영상들과 합성 후 학습 영상 데이터를 생성한 연구도 있었다

[3]. 하지만 표지판 인식을 위해 가상 학습 영상 데이터를 생성하여 학습한 연구는 찾아보기 힘들다.

제안하는 방법에서는 표지판 표준 영상을 이용해 가상 학습 데이터를 생성한다. 하지만 표지판 표준 영상과 실사 표지판 영상 사이에는 자세와 조명 환경에 따른 차이가 있다. 이런 이유로 표지판 표준 영상을 가상 공간에서 투영한 후 가우시안 블러를 적용한다. 그 결과로 실제 표지판 영상 데이터 학습과 동일한 효과를 낼 수 있는 가상 학습 영상 데이터를 생성 한다.

본 논문은 2 장에서 가상 표지판 데이터에 대한 생성 원리를 기술한다. 3 장에서는 가상 표지판 데이터를 이용한 실험과 결과에 대해 이야기하고 4 장에서는 결론을 내리겠다.

## 2. 가상 표지판 영상 생성 및 학습

표지판 인식을 위해 가상 표지판 영상 데이터를 생성하여 가상 표지판 영상의 HOG 특징을 SVM 학습기로 학습하여 표지판 인식을 수행한다.

### 2.1 가상 표지판 영상 생성

도로 환경에서 실제 표지판은 정면으로 바르게 설치되어 있지 않고 상하 좌우로 약간씩 기울어 진 경우가 많다. 그렇기 때문에 가상 표지판 영상을 생성할 때 투영 변환을 사용하여 표지판 표준 영상을 X 축, Y 축 Z 축으로 회전하여 학습 데이터를 생성 한다.

투영 변환을 수행 하기 위해서는 카메라 기하학을 사용한다. 카메라 좌표계상의 3 차원 물체를 영

상 평면에 투영하여 평면에 투영된 영상을 얻을 수 있다. 카메라 좌표계와 영상 좌표계의 관계는 아래 수식 1 과 같다.

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \cong \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} f \cdot sx & 0 & ox \\ 0 & f \cdot sy & oy \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_1 & r_2 & r_3 & t_x \\ r_4 & r_5 & r_6 & t_y \\ r_7 & r_8 & r_9 & t_z \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

$x_p$                                    $M_i$                                    $M_e$                                    $x_p$

$x_p$  는 영상 좌표계의 점이고  $x_p$  는 카메라 좌표계의 점이다.  $M_i$  는 카메라 내부 파라미터이고  $M_e$  는 카메라 외부 파라미터이다.  $f$  는 초점거리이고  $sx$  는 카메라 좌표계와 영상 좌표계 사이의 스케일 관계이다.  $ox, oy$  는 카메라 좌표계의 원점과 영상 좌표계의 원점을 맞춰준다.  $r_i (i=1\sim 8)$  는 회전 변환 행렬이고  $t_x, t_y, t_z$  는 평행이동 변환 행렬이다. 외부 파라미터  $M_e$  의 회전 변환 행렬과 평행이동 변환 행렬 조정하고 카메라 좌표계의 3 차원 물체를 영상 좌표계에 투영하면 투영 변환된 영상을 얻을 수 있다.

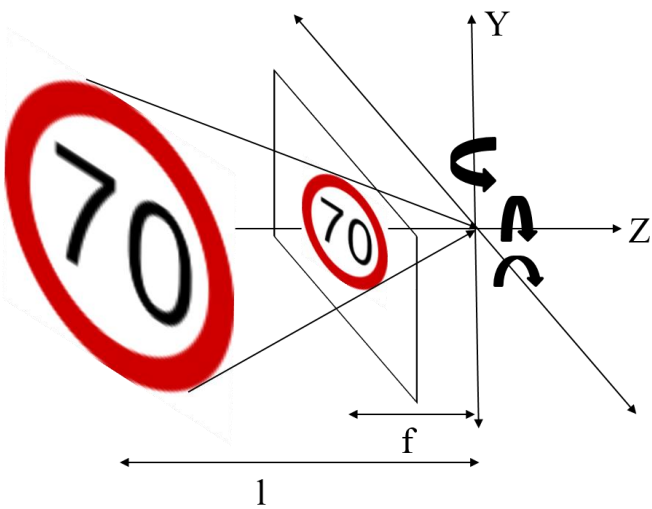


그림 3. 가상 표지판 영상의 투영 변환

가상 표지판 영상을 그림 3 과 같이 좌표계로부터 1 거리에 위치 시키고 회전 변환 행렬로 X 축, Y 축, Z 축을 회전한 후 영상 평면에 투영하여 가상 표지판 영상 데이터를 생성한다.

실제 도로 환경에서는 표지판의 거리가 멀수록 표지판이 흐려져 블러 현상이 생긴다. 가우시안 필터를 사용하면 표지판의 블러 현상을 모사할 수 있다.

### 2.2 가상 표지판 영상 HOG 특징 추출 및 학습

HOG 특징은 조명환경에 강인한 특성을 보이는 특징으로 보행자 인식에 처음 사용되어 현재 여러 검출, 인식분야에 활발하게 사용되고 있다[4].

HOG 특징 검출 과정에서 첫 번째는 인식 객체

의 영상을 가로 세로 일정 픽셀의 셀 영역을 나눈다. 각 셀 내부의 모든 픽셀에 대한 경사(gradient) 방향과 경사 크기를 구하여 경사 방향 별로 경사 크기를 누적하여 해당 셀을 대변하는 히스토그램을 만든다. 그리고 4 개의 셀(Cell)은 하나의 블락(Block)이라고 한다. 블락이 슬라이딩 윈도우와 같이 전체 영상을 한 셀 단위로 슬라이딩 하면서 각 셀을 대변하는 히스토그램들을 한 행으로 이어 붙여 인식 객체의 영상의 HOG 특징을 만든다.

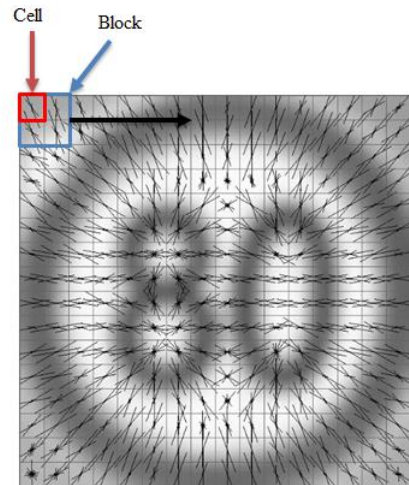


그림 4. HOG 특징 추출

표지판 인식을 위해 SVM 을 사용하여 가상 표지판 표준 영상에 대해 HOG 특징 학습한다. SVM 은 여러 인식 분야에서 가장 많이 사용되는 기계 학습 방법 중 하나이다. SVM 의 핵심 개념은 서로 다른 클래스 사이에 각 클래스를 구분할 수 있는 가장 큰 여백을 가지는 결정 경계를 계산하는 것이다[5]. 이렇게 학습된 SVM 분류기를 이용하여 표지판 인식을 수행 한다.

### 3. 실험 결과 및 분석

가상 표지판 영상 생성 과정 시 입력 가상 표지판 영상의 크기는 256×256(pixel)이다. 가상 표지판 영상의 투영 변환을 수행 하기 위해 가상 환경을 다음과 같이 설정한다. 물체와 카메라 좌표계 사이의 거리  $l = 4000\text{mm}$  로 둔다. 가상 초점 거리  $f = 8\text{mm}$  로 둔다. x, y 축의 회전은 정면에서 -1 도 +1 도 회전하였고 z 축으로는 -5 도 +5 도 회전하였다. 그리고 블러를 위한 가우시안 필터의 크기는 13×13(pixel)로 하였고  $\sigma_x$ 와  $\sigma_y$  는 0~2 까지 증가시키면서 3 단계로 가상 영상 데이터를 블러 시켰다(그림 5). 가우시안 블러의  $\sigma$  값이 클수록 블러 효과가 있다. 한 표지판으로부터 108 가지 가상 표지판 영상을 생성 하였으며 일반적인 도로 환경에서 확보하기 힘들었던 10 가지의 표지판에 대해 가

상 표지판 영상 데이터를 생성하였다. 가상 영상 데이터 생성 수는 1080 개이다. 실험에 사용된 컴퓨터는 Core i5(3.4 GHz)CPU에 4GB 메인 메모리를 탑재하였다. 가상 표지판 영상 데이터 생성 시간은 총 1분 걸렸다.










	Z 축 회전 -5도	Z 축 회전 0도	Z 축 회전 +5도
$\sigma_0$			
$\sigma_1$			
$\sigma_2$			

그림 5. Z 축 회전과 3 단계 블러에 대한 가상 표지판 영상 데이터

표 1. 표지판 인식 결과

표지판	표지판 수 (개)	인식 성공 (개)	인식 실패 (개)	인식 률 (%)
이륜차 통행 금지	7	7	0	100.0
자전거 통행 금지	78	74	4	94.9
경운기 통행 금지	10	10	0	100.0
보행자 보행 금지	2	2	0	100.0
우회전 금지	57	53	4	93.0
주정차 금지	44	36	8	81.8
위험	99	99	0	100.0
우로 굽은 도로	59	59	0	100.0
횡풍	40	38	2	95.0
좌 합류 도로	43	42	15	97.7
계	439	420	33	95.7

가상 표지판 영상의 HOG 특징을 추출 하기 위해 가상 표지판 영상을 256×256 크기로 정규화 후 그레이 스케일 영상으로 변환하여 셀 사이즈 8×8(pixel), 블락 사이즈 16×16(pixel), 블락의 슬라이딩 스텝을 한 셀 사이즈, 즉 가로로 8(pixel), 세로로 8(pixel)씩 슬라이딩 하며 HOG 특징을 계산한다. 픽셀의 경사 크기 누적 은 9 방향으로 나누어 9 개의 빈을 가진 히스토그램에 누적한다.

표지판 인식을 위한 SVM 은 OpenCV 라이브러리에서 제공되는 함수를 사용하였다. 학습에 사용된 SVM 은 Linear SVM 이다.

테스트 데이터는 PointGrey 사의 Grasshopper2 를 차량에 설치해 도로를 주행하며 직접 획득한 표지판 영상 데이터이다. 초점 거리 8mm 렌즈를 사용하였다.

수집하기 힘든 표지판 10 가지 표지판에 대해 평균 95.7%의 높은 인식률을 얻을 수 있었다. 주정차 금지 표지판의 경우 표지판 자체가 나무에 의해 가려져있던 표지판이 많아 인식률이 낮게 나온 것으로 추정된다.

#### 4. 결론

표지판은 내부 모양이 단순하고 서로 다른 표지판간에 표지판 테두리 내의 영상 모양 차이가 크기 때문에 경사 특징 기반인 HOG 특징를 사용한다면 가상 표지판 영상 데이터를 이용한 SVM 학습 방법은 굉장히 짧은 시간 내에 많은 종류의 표지판을 학습 데이터를 생성 하여 학습 할 수 있으며 높은 인식률을 가진다.

#### 감사의 글

본 연구는 미래창조과학부 및 정보통신산업진흥원의 IT 융합 고급인력과정 지원사업의 연구결과로 수행되었음 (NIPA-2014-H0401-14-1004). 본 연구는 산업통상자원부 산업원천기술개발사업으로 지원된 연구임(10040927, 운전 안전성 및 편의성 향상을 위한 운전자 시야 중심 차량용 증강현실 정보 제공 시스템 기술 개발)

#### 참고문헌

- [1] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "The German Traffic Sign Recognition Benchmark: A multi-class classification competition," *IEEE International Joint Conference on Neural Networks*, pp. 1453-1460, 2011.
- [2] J. Marin, D. Vazquez, D. Gerenimo, and A. M. Lopez, "Learning appearance in virtual scenarios for pedestrian detection," in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*. 2010, pp. 137-144.
- [3] L. Pishchulin, A. Jain, C. Wojek, M. Andriluka, T. Thormählen and B. Schiele, "Learning people detection models from few training samples," *IEEE Computer Vision and Pattern Recognition*. pp 1473-1480, 2011.
- [4] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *IEEE Computer Vision and Pattern Recognition*, vol. 1, pp. 886-893, 2005.
- [5] C. Cortes and V. Vapnik, "Support Vector Network," *Machine learning*, vol. 20, no. 3, pp. 273-297, 1995