일반논문 (Regular Paper)
방송공학회논문지 제22권 제1호, 2017년 1월 (JBE Vol. 22, No. 1, January 2017)
https://doi.org/10.5909/JBE.2017.22.1.62
ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

색상지도와 멀티 레이어 HOG-SVM 기반의 실시간 신호등 검출 알고리즘

김 상 기[a)], 한 동 석[b)‡]

Real Time Traffic Light Detection Algorithm Based on Color Map and Multilayer HOG-SVM

Sanggi Kim[a)] and Dong Seog Han[b)‡]

요 약

신호등 검출은 첨단운전자보조시스템에서 매우 중요하며 최근 신호등 검출 알고리즘의 연구가 활발히 진행 중이다. 그러나 기존의 영상처리 기반의 신호등검출 알고리즘은 조명의 변화에 민감하다는 문제점이 있다. 이러한 문제점을 해결하기 위하여 본 논문에서는 다음과 같은 신호등 검출 알고리즘을 제안한다. 먼저 제안하는 컬러맵과 HSV(hue-saturation-value)를 이용하여 신호등의 후보영역을 검출한다. 이후 검출된 신호등 후보영역으로부터 HOG(histogram of oriented gradient) 서술자와 SVM(support vector machine)을 이용하여 신호등을 검출한다. 검출된 신호등 영상을 이용하여 제안하는 Multilayer HOG 서술자를 이용하여 신호등의 방향 정보를 결정한다. 실험결과에서 확인할 수 있듯이 제안하는 알고리즘은 높은 검출성능과 실시간 처리가 가능하다.

Abstract

Accurate detection of traffic lights is very important for the advanced driver assistance system (ADAS). There have been many research developments in this area. However, conventional of image processing methods are usually sensitive to varying illumination conditions. This paper proposes a traffic light detection algorithm to overcome this situation. The proposed algorithm first detects the candidates of traffic light using the proposed color map and hue-saturation-value (HSV) Traffic lights are then detected using the conventional histogram of oriented gradients (HOG) descriptor and support vector machine (SVM). Finally, the proposed Multilayer HOG descriptor is used to determine the direction information indicated by traffic lights. The proposed algorithm shows a high detection rate in real-time.

Keyword : traffic light detection (TLD), support vector machine (SVM), histogram of oriented gradient (HOG), advanced driver assistance system (ADAS)

---

a) 경북대학교 대학원 전자공학부(Kyungpook National University, Graduate School of Electronics Engineering)
b) 경북대학교 IT대학 전자공학부(Kyungpook National University, School of Electronics Engineering, College of IT Engineering)
‡ Corresponding Author : 한동석(Dong Seog Han)
E-mail: dshan@knu.ac.kr
Tel: +82-53-950-8669
ORCID: http://orcid.org/0000-0002-7769-0236
※ 이 논문은 2016년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 지역신산업선도인력양성사업 성과임(No. 한국연구재단에서 부여한 과제번호 : 연구사업 통합시스템에서 확인)(NRF-2016H1D5A1909920).
· Manuscript received September 30, 2016; Revised December 19, 2016; Accepted December 21, 2016.

Ⅰ. Introduction

Recently, advanced driver assistance system (ADAS) is widely used in autonomous vehicle system. Among the ADAS techniques, traffic light detection is one of the important issues of which there have been many research developments to get more accurate detection of traffic lights.

Vision-based systems can be integrated with sensors such as global positioning system (GPS) to discover the potential presence of traffic light regions. Research works [2] and [3] used GPS to acquire traffic light regions robustly and reduce noise such as varying illumination conditions and background noise. However, these systems require additional cost and accurate equipment such as sensors and GPS. To reduce the additional cost, many researchers have focused on purely vision based systems and they used the features on traffic lights such as color, shape and illumination.

De Charette et al [5], [6] used the white-top-hat-filter and morphological filter to obtain candidates for traffic lights. Then an adaptive template matching is applied to the result of candidate regions to detect traffic lights. Masako et al. [7] used the normalized (red-green-blue) RGB color space to detect candidates of traffic lights. Then they detect edges for traffic lights to compare the shape. However, these methods would be affected by varying illumination. A lot of research has been carried out in the hue-saturation-value (HSV) color space for the detection of traffic lights. The HSV color space is less sensitive to varying illumination. Hwang et al. [8] used the HSV color space by adopting an appropriate threshold and the center finding with a Gaussian mask to detect candidates of traffic lights. Then traffic lights are detected by using a suggested existence-weight map. Shen et al. [9] used hue and saturation to obtain traffic lights. Chen et al. [10] used the normalized RGB and HSV which are called a joint color space to obtain candidates of traffic lights. Then the proposed histogram of oriented gradient (HOG) feature is used to detect traffic lights. John et al. [4] used GPS information to discover potential traffic light regions robustly. Then traffic lights are detected based on the convolutional neural network by training various datasets for traffic lights. However, this method has high complexity so that the processing time is too long to be implemented in real-time.

This proposed algorithm uses a color map in the saliency map [1] and the HSV color space. One of the advantages using both color map and HSV is reduced noise compared to other color spaces. The noise in non-traffic light regions is reduced from the result of the color map. Then the conventional HOG feature is used to detect traffic lights and the support vector machine (SVM) classifier finally decides the existence of the traffic lights in the test images. The direction information in the traffic light is detected by using the proposed M-HOG with linear SVM. The proposed vision-based algorithm can detect both traffic lights and direction information. This paper is organized as follows. Section II presents a detailed review of the proposed traffic light detection algorithm. Comparative analysis and performance of the proposed algorithm are given in Section III and conclusions are drawn in Section IV.

Ⅱ. System Model for Proposed Traffic Light Detection Algorithm

The proposed algorithm is a purely vision-based algorithm to detect traffic lights and extract the directional information. Fig.1 shows the proposed algorithm. The operation of the proposed algorithm is as follows: At first, the input image is converted to a color map to detect candidates of traffic lights, then the red or green color is detected. If the color of a candidate is not red or green, the candidate should be removed. The color is decided from the HSV color space. At the second step, the conventional HOG descriptor with linear SVM is used to detect traffic lights from the candidates. At the third step, the M-HOG feature is extracted to determine the direction in

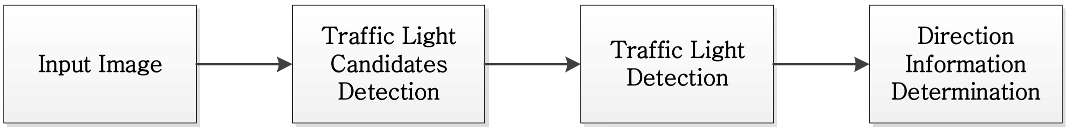


그림 1. 제안하는 알고리듬의 개요
Fig. 1. Block diagram of the proposed algorithm

formation in the detected traffic lights. Finally, the trained-linear SVM classifier is used to decide the direction in the traffic lights. When the direction information is decided the linear SVM classifier utilizes the fact that traffic light lamps are consisted with the shape of a circle or arrow. The detailed process is described in the next section.

1. Color Map with HSV

Traffic lights are defined by fixed colors and shapes which are easily noticeable to human eyes. Candidates of traffic lights are normally detected in the color space. Among the color spaces, RGB and HSV are widely used to obtain such candidates. When the candidates for traffic lights using the color space are detected, the appropriate spectrum threshold value defining red and green colors is applied to the detected candidates. Adopting the threshold value of color space in road environments does not yield a good performance because of color distortion and varying illumination conditions.

In this paper, the color map is used to detect noticeable parts of images and then detect the candidates using HSV to solve the problem. The color map is one of the features on the saliency map proposed by Itti et al [1]. The saliency map is composed of a linear combination of the color map, intensity map, and orientation map of the divided Gaussian

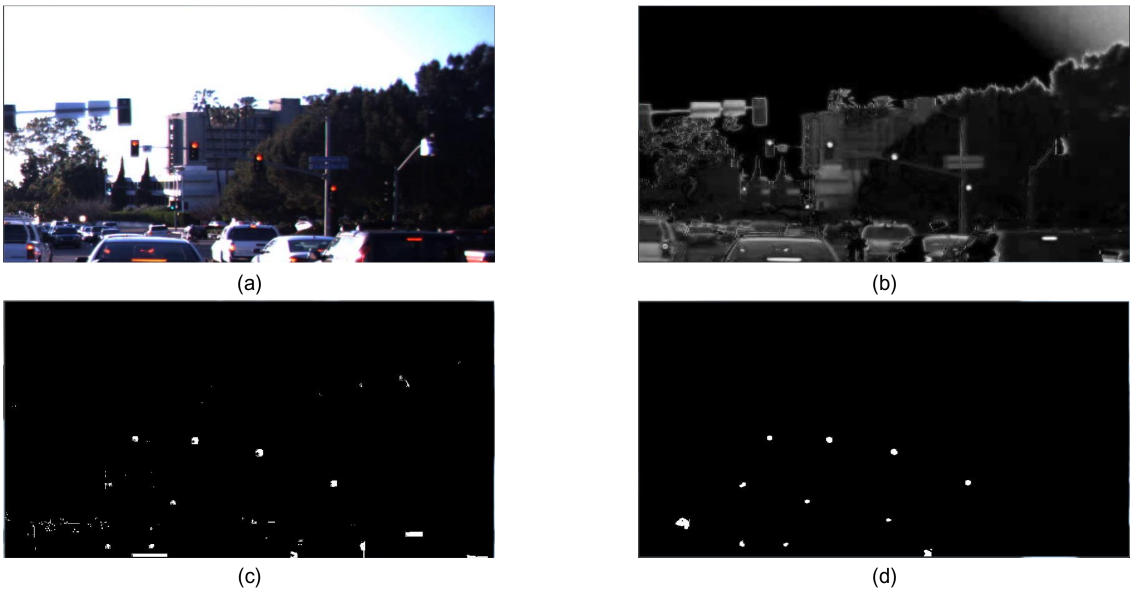


그림 2. 신호등 후보군 선택 절차: (a) 입력 영상, (b) 입력영상으로부터 색상지도 결과영상, (c) 색상지도와 HSV 적용 영상 (d) (c)결과영상으로부터 노이즈필터 적용 결과영상
Fig. 2. The procedure of traffic light candidate selection: (a) An input image, (b) Result of color map from the input image, (c) Result of color map with HSV color space to obtain candidates of traffic light, (d) Result of Noise filter from the previous result in (c)

pyramid to obtain noticeable parts in the image. However, it is hard to apply the saliency map to road environments because of the high complexity. The proposed color map has the same architecture of the color map in the saliency map. However, it does not use a pyramid structure, but just subtracts the blur image from the input image to obtain noticeable parts. The result of the proposed color map is shown in Fig. 2(b). As seen in Fig. 2(b), it is obvious that the proposed color map can detect noticeable parts with traffic lights in the image. However, the colors both red and green are not considered in this part. In order to remove candidates in the color space which define neither red nor green, the image is converted to HSV from the result of the color map as

$$H=\arctan\left(\frac{\sqrt{3}(G-B)}{(R-G)+(R-B)}\right)$$
$$S=1-\frac{\min(R,G,B)}{R+G+B} \quad (1)$$
$$V=\frac{R+G+B}{3}$$

Where H, S and V are hue, saturation and value in the HSV color apace.

From the HSV using (1), H and S are used to detect colors of red and green and the candidates with other colors are deleted from the candidates by adopting threshold values as shown in Fig. 2(c). The threshold value was empirically set using red and green traffic light dataset obtained from experimental environment. The candidate traffic lights are detected if pixel range satisfies the threshold.

2. Noise filter

As shown in the previous result of Fig. 2(c), there is still much noise in the image. This noise occurs due to varying illumination conditions, background noise and taillights. To reduce the noise, we propose a noise filter. Traffic light emitting units have fixed shapes. Their average size can be known from the training datasets. We find the connected component and calculate the area for each candidate to remove noise. The condition to remain as a candidate is as follows:

$$T_1 \leq O \leq T_2 \quad (2)$$

Where O is the size of an object in the image, T_1 and T_2 are minimum and maximum areas, respectively. However, the filter for average size of traffic light cannot remove noise when the size is within the minimum and maximum range defined by T_1 and T_2 respectively. In order to remove these noise, we consider aspect ratio regarding all blobs as follows:

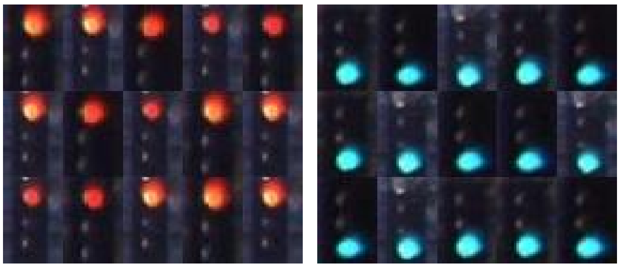
$$AR=\frac{\max(width,height)}{\min(width,height)} \quad (3)$$

Where AR is the aspect ratio, $width$ and $height$ represent the width and height of the bounding box. If AR is larger than the threshold T_{AR}, the object is registered as a candidate for traffic light. The range of the aspect ratio may be change due to dynamic road environments causing the varying illumination condition. Therefore, the threshold for aspect ratio was set at 1.5 empirically in this paper. The results of noise filter is shown in Fig. 2(d).

3. Traffic Light Detection

In the previous section, we detected traffic light candidates and color attributes of red or green. Traffic lights consist of three emitting units with different positions. The traffic light can be detected from the result of emitting unit detection by using traffic light architecture as explained in Section III. Then candidate regions are divided into two images of red and green. Each image in Fig. 3 is used to extract the HOG feature. The process to obtain the HOG feature is as follows: Candidate regions are estimated from the body size of traffic lights using color attributes. The size of candidate regions are normalized as $(width \times height = 32 \times 64)$ Then the detection window size of HOG-descriptor is normalized as (32×64) and block and cell size are (8×8) and (4×4) respectively to extract a conventional HOG feature. Then extract the conventional HOG feature for each color.

The training dataset shown in Fig. 3 is used to extract the HOG feature by applying the above steps. Linear SVM classifier is used to train HOG feature for training dataset. Then trained linear SVM classifier for traffic lights is used to detect traffic lights.



(a) (b)

그림 3. 신호등 검출을 위한 훈련영상: (a) 적색 신호등 훈련영상 (b) 녹색 신호등 훈련영상

Fig. 3. Training samples to detect traffic lights: (a) training samples for the red traffic light (b) training samples for green traffic light

4. Direction Information Determination

In the previous section, we detected traffic light with color. Therefore, we don't consider color when the direction of traffic light is determined. The orientations of the gradient are normally calculated for circle and arrow images to determine the direction information for the traffic light. Another way to identify the direction of the traffic light is to use the conventional HOG with linear SVM classifier. However, those methods cannot achieve a good performance because of changing illumination and color distortion. Thus, we determine the direction of the traffic light by using the multilayer HOG method. The multilayer HOG descriptor can be obtained as follows: First, we divide emitting unit regions into R, G, B and grayscale from the result of traffic light. Then the image size for each layer is normalized as $(width \times height = 24 \times 32)$ The detection window size of mltilayer HOG is normalized as (24×32) The block and cell size are (8×8) and (4×4) respectively for purposes of extracting the multilayer HOG feature. For each cell, we accumulate gradient magnitude weighted votes for gradient orientation into 9 bins. Then, the multilayer HOG descriptor decides the maximum gradient magnitude for each layer. The whole process for multilayer HOG descriptor is shown Fig. 4

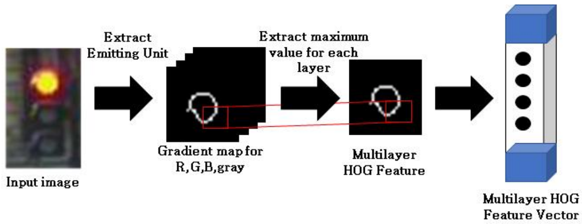
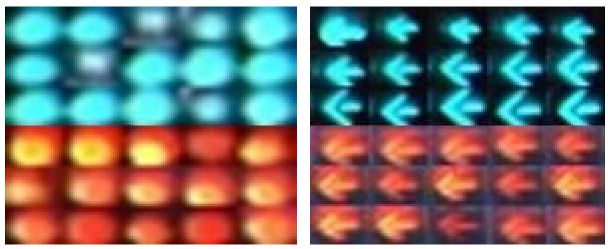


그림 4. multilayer HOG descriptor 처리 절차

Fig. 4. The process of for multilayer HOG descriptor

The training dataset shown in Fig. 5 is used to determine the direction of traffic lights by using the multilayer HOG feature described in (4). Then the linear SVM classifier is trained using the multilayer HOG feature in order to classify the direction information which is used to decide the direction in the traffic lights. The result of traffic light detection and result of directional information is shown in Fig. 6.



(a) (b)

그림 5. 신호등 방향정보 추출을 위한 훈련영상: (a) 원 영상 (b) 화살표 영상

Fig. 5. Training sample to extract direction information: (a) training sample to extract circle (b) training sample for the arrow

Ⅲ. Experiment Result

1. Dataset

In this paper, the USA dataset [14] and french dataset [13] are used to train the proposed algorithm. To train the dataset of the proposed algorithm, 5,000 and 8,000 traffic light images for positive dataset and negative dataset were used for the traffic light detection, respectively. The negative images from the test dataset were collected randomly. For the direction identification, 8,000 circle images and 8,000 arrow images for positive dataset, and 13,000 background images for negative dataset were used. Table 1 shows our training dataset.

표 1. 훈련 영상
Table 1. Training Dataset

Dataset	Traffic Light	Circle	Arrow
RED	5,000	5,123	6,371
GREEN	5,000	2,877	1,629
NEGATIVE	8,000	13,000	13,000

2. Performance Evaluation

In order to verify the performance of the proposed algorithm, we utilize French dataset for traffic light detection and USA dataset to detect direction information. The methods of measurement for the proposed algorithm are presented as

$$\mathrm{Precision} = \frac{\mathrm{TruePositives}}{\mathrm{TruePositives}+\mathrm{FalsePositives}} \qquad \mathrm{Recall} = \frac{\mathrm{TruePositives}}{\mathrm{TruePositives}+\mathrm{FalseNegatives}} \tag{5}$$

The performance evaluation of traffic light detection methods are shown in Table 2. [6],[7] are based on image processing algorithm. As shown in Table 2, image processing methods are relatively of low performance compared to the proposed method because of appearance variations and illumination variations. Image processing methods can't classify similar objects to traffic light such as tail light because of multiple threshold problems. However, The proposed algorithm is relatively robust to illumination variations and the losses of the traffic lights by setting multiple threshold are less than [6],[7].

Table 3 and 4 shows the result of extracting direction information of the traffic lights using Multilayer HOG descriptor and conventional HOG descriptor. The conven-

그림 6. 신호등 및 신호등의 방향 정보 검출 결과영상
Fig. 6. The result of the proposed algorithm for both traffic light detection and extract direction information

tional HOG descriptor detects the edge using difference of the brightness value from the grayscale. However, this method cannot distinguish between an arrow and a circle due to illumination variations. To solve this problem, Multilayer HOG descriptor is applied to four channel such as R, G, B and grayscale to extract robust features. As shown table 3 and 4, the proposed Multilayer HOG features show better performance compared to the conventional HOG descriptor. The advantage of the Multilayer HOG descriptor is that it is less sensitive to change the illumination and represents the shape for the object compared to the conventional HOG descriptor.

표 2. 제안 알고리즘의 성능
Table 2. Performance of Proposed algorithm

Method	Precision	Recall
proposed	92.16%	93.84%
[6]	84.5%	53.5%
[7]	61.22%	93.75%

표 3. Multilayer HOG 서술자를 이용하여 도출된 방향정보 인식성능에 대한 혼동행렬
Table 3. Confusion Matrix for Performance of Direction Information using Multilayer HOG descriptor

Class	Go	Go Left	Stop	Stop Left
Go	97.2%	7.3%	0%	0%
Go Left	2.8%	92.3%	0%	0%
Stop	0%	0%	96.4%	6.4%
Stop Left	0%	0%	3.6%	93.6%

표 4. HOG 서술자를 이용하여 도출된 방향정보 인식성능에 대한 혼동행렬
Table 4. Confusion Matrix for Performance of Direction Information using HOG descriptor

Class	Go	Go Left	Stop	Stop Left
Go	94.6%	10.7%	0%	0%
Go Left	5.4%	89.3%	0%	0%
Stop	0%	0%	93.1%	9.6%
Stop Left	0%	0%	6.9%	90.4%

3. Computation Time

To evaluate the computation time of the proposed algorithm, computer simulations have been conducted in a computer having the Intel core CPU i7-4790 processor with the frequency of 3.60 GHz and with 8GB RAM. The algorithm is implemented in C++ and the proposed algorithm is a non-optimized C++ code using a single thread. The image size of the dataset is 1280×960 However, to reduce computation time, image is resized as 780×400. The computation time for the proposed algorithm is a nearly real time algorithm (≈ 13.7 fps) and the computation time for each part is present in Table 4. As shown in Table 4, the proposed algorithm can perform in nearly real time and it might be applied in practice.

표 4. 제안하는 알고리즘의 처리시간
Table 4. Computation Time For Proposed algorithm

Algorithm	Computation Time
Color Map	30ms
Noise Filter	12ms
Traffic Light Detection	31ms
Total	73ms

Ⅳ. Conclusion

In this paper, we suggested that a purely vision based algorithm to detect traffic lights and extract direction information robustly. The proposed algorithm detects candidates of traffic light and color information using a color map and the HSV color space. Then a noise filter is used to remove noise from the previous result. Then conventional HOG with linear SVM is used for detection of traffic light and then multilayer HOG feature is used to determine direction of traffic lights finally. The proposed algorithm was evaluated by using a public dataset. From the experimental result, the proposed algorithm shows that high accuracy for both detection of traffic light and direction for traffic light is achieved in a minimal computation time. In our future work, we will remove empirical thresholds by learning the optimal threshold automatically. Additionally, we will combine our method with tracking algorithm to get better performance and reduce the computation time.

참 고 문 헌 (References)

[1] L. Itti, C. Koch, and E. Niebur, “A model of saliency-based visual attention for rapid scene analysis,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 11, pp. 1254-1259, Nov. 1998.

[2] J. Levinson, J. Askeland, J. Dolson and S. Thrun, “Traffic light mapping, localization and state detection for autonomous vehicles,” in proc. IEEE International Conference on Robotics and Automation, Shanghai, pp. 5784-5791, May. 2011

[3] N. Fairfield and C. Urmson, “Traffic light mapping and detection,” in proc. IEEE International Conference on Robotics and Automation, Shanghai, pp. 5421-5426, 2011.

[4] V. John, K. Yoneda, B. Qi, Z. Liu, and S. Mita, “Traffic light recognition in varying illumination using deep learning and saliency map,” in proc. IEEE International Conference on Intelligent Transportation Systems, Qingdao, pp. 2286-2291, Oct. 2014

[5] R. de Charette and F. Nashashibi, “Real time visual traffic lights recognition based on spot light detection and adaptive traffic lights templates,” in proc. IEEE Intelligent Vehicles Symposium, Xi'an, pp. 358-363, Jun. 2009.

[6] R. de Charette and F. Nashashibi. “Traffic light recognition using image processing compared to learning processes,” in proc. IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, pp. 333-338, Oct. 2009.

[7] G. Siogkas, E.Skodras and E. Dermatas. “Traffic lights detection in adverse conditions using color symmetry and spatiotemporal information,” in proc. International Conference on Computer Vision Theory and Applications, 2012.

[8] T. Hwang, I. Joo, and S. Cho, “Detection of traffic lights for vision-based car navigation systems,” in proc. Advances in Image and Video Technology, First Pacific Rim Symposium, Taiwan, pp. 682-691, Dec. 2006.

[9] Y. Shen, U. Ozguner, K. Redmill, “A robust video based traffic light detection algorithm for intelligent vehicles,” in proc. IEEE Intelligent Vehicles Symposium, Xi'an, pp. 521-526, Jun. 2009.

[10] Q. Chen, Z. Shi, Z. Zou, “Robust and real-time traffic light recognition based on hierarchical vision architecture,” in proc. IEEE International Congress on Image and Signal Processing, Dalian, pp. 114-119, Oct. 2014.

[11] V. John, K. Yoneda, B. Qi, Z. Liu, and S. Mita, “Saliency map generation by the convolutional neural network for real-time traffic light detection using template matching,” IEEE Transaction on Computational Imaging, vol. 1, no. 3, September. 2015

[12] Robotics Centre of Mines ParisTech. (2015) Traffic lights recognition (TLR) public benchmarks. [Online]. Available: http://www.lara.prd.fr/benchmarks/trafficlightsrecognition

[13] U. Laboratory for Intelligent and Safe Automobiles . (2015) Vision for intelligent vehicles and applications (viva) challenge. [Online]. Available: http://cvrr.ucsd.edu/vivachallenge/

저 자 소 개

김 상 기

- 2015년 2월 : 경운대학교 디지털전자공학과 학사
- 2015년 3월 ~ 현재 : 경북대학교 전자공학부 석사
- ORCID : http://orcid.org/0000-0002-1336-7253
- 주관심분야 : 컴퓨터비전, 머신러닝

한 동 석

- 1987년 : 경북대학교 전자공학과 공학사
- 1989년 : KAIST 전기 및 전자공학과 공학석사
- 1993년 10월 ~ 1996년 8월 : 삼성전자 기술총괄신호처리 연구소 선임연구원
- 1996년 8월 ~ 현재 : 경북대학교 전자공학부 교수
- 2006년 7월 ~ 2008년 7월 : 정보통신연구진흥원 디지털 TV/방송사업단 단장
- 2011년 ~ 현재 : 경북대학교 ICT·자동차융합연구센터장
- ORCID : http://orcid.org/0000-0002-7769-0236
- 주관심분야 : 통신신호처리, 지능형 교통 시스템