



## Traffic light recognition exploiting map and localization at every stage



Chulhoon Jang<sup>a</sup>, Sungjin Cho<sup>a</sup>, Sangwoo Jeong<sup>a</sup>, Jae Kyu Suhr<sup>c</sup>, Ho Gi Jung<sup>b</sup>,  
Myoungho Sunwoo<sup>a,\*</sup>

<sup>a</sup> Department of Automotive Engineering, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Republic of Korea

<sup>b</sup> Major of Information & Communications Engineering, Korea National University of Transportation, 50 Daehak-ro, Chungju-si, Chungbuk 27469, Republic of Korea

<sup>c</sup> School of Intelligent Mechatronics Engineering, Sejong University, 209 Neungdong-ro, Gwangjin-gu, Seoul 05006, Republic of Korea

### ARTICLE INFO

#### Article history:

Received 5 February 2017

Revised 30 June 2017

Accepted 5 July 2017

Available online 12 July 2017

#### Keywords:

Traffic light recognition

Localization and digital map

Intelligent vehicles

Intelligent transportation systems

Computer vision

Object detection

### ABSTRACT

Traffic light recognition is being intensively researched for the purpose of reducing traffic accidents at intersections and realizing autonomous driving. However, conventional vision-based approaches have several limitations due to full image scanning, always-on operation, various different types of traffic lights, and complex driving environments. In particular, it might be impossible to recognize a relevant traffic light among multiple traffic lights at multiple intersections. To overcome such limitations, we propose an effective architecture that integrates a vision system with an accurate positioning system and an extended digital map. The recognition process is divided into four stages and we suggest an extended methodology for each stage. These stages are: ROI generation, detection, classification, and tracking. The 3D positions of traffic lights and slope information obtained from an extended digital map enable ROIs to be generated accurately, even on slanted roads, while independent design and implementation of individual recognition modules for detection and classification allow for selection according to the type of traffic light face. Such a modular architecture gives the system simplicity, flexibility, and maintainable algorithms. In addition, adaptive tracking that exploits the distance to traffic lights allows for seamless state estimation through smooth data association when measurements change from long to short ranges. Evaluation of the proposed system occurred at six test sites and utilized two different types of traffic lights, seven states, sloped roads, and various environmental complexities. The experimental results show that the proposed system can recognize traffic lights with 98.68% precision, 92.73% recall, and 95.52% accuracy in the 10.02–81.21 m range.

© 2017 Elsevier Ltd. All rights reserved.

### 1. Introduction

Recently, the traffic light recognition (TLR) problem has become an important research focus for driver assistant systems (DAS) and autonomous driving. Perception of traffic lights (TL) at intersections and crosswalks is a necessary function for compliance with traffic regulations and to prevent fatal road accidents (De Charette & Nashashibi, 2009a; Fairfield & Urmson, 2011; Gomez, Alencar, Prado, Osorio, & Wolf, 2014; Jensen, Philipsen, Møgelmose, Moeslund, & Trivedi, 2016; Kim, Shin, Kuk, Park, & Jung, 2013; Levinson, Askeland, Dolson, & Thrun, 2011; Yu, Huang, & Lang, 2010). There are two approaches to TLR: a communication-based method and

a vision-based method. The communication-based method utilizes wireless communications between the vehicle and surrounding infrastructure, in addition to providing TL states to surrounding vehicles (Bazzi, Zanella, & Masini, 2016; Ferreira & d'Orey, 2012; Florin & Olariu, 2015; Younes & Boukerche, 2016). This method may be the most faithful way to deliver traffic information, but it requires substantial investment for the installation of wireless devices in all traffic lights and vehicles (Fairfield & Urmson, 2011; Jensen et al., 2016; Levinson et al., 2011; Yu et al., 2010). On the other hand, the vision-based method utilizes a camera installed on a windshield for TLR and has considerable potential to increase functionality for customers. However, this method can also be easily affected by disturbances in the surrounding area such as objects with similar shapes and colors.

Conventional vision-based approaches have been researched in various ways (Diaz, Cerri, Pirlo, Ferrer, & Impedovo, 2015; Jensen et al., 2016). These approaches have several limitations in DAS or autonomous driving as they attempt to meet stringent requirements. First, the camera requires a wide field of view (FOV)

**Abbreviations:** TLR, traffic light recognition; TL, traffic light; FOV, field of view; DAS, driver assistant system; ROI, region of interest; LIDAR, light detection and ranging; HOG, histogram of oriented gradients; SVM, support vector machine.

\* Corresponding author.

E-mail addresses: chulhoonjang@gmail.com (C. Jang), sungjincho215@gmail.com (S. Cho), sangwoojeong0318@gmail.com (S. Jeong), jksuhr@gmail.com (J.K. Suhr), hogijung@ut.ac.kr (H.G. Jung), msunwoo@hanyang.ac.kr (M. Sunwoo).

and high-resolution to cover short and long ranges for TLR. The high-resolution image increases the amount of information and requires large amounts of computing power. Second, many false positives may be generated since a TL is a very tiny object with few distinctive features; in particular, it is hard to distinguish between TLs and the brake lights of preceding vehicles on slanted roads, or to identify TLs against complex backgrounds with lights and colored signs (Gomez et al., 2014; Jensen et al., 2016; Omachi & Omachi, 2010). Third, the various compositions of TLs lead to sophisticated problems of classification and decision making: red, green, yellow, and arrows can be combined in many different ways (Jensen et al., 2016). And last, recognition of TLs in multiple lanes and multiple TL conditions is a much more challenging problem than simple recognition of TL states (Jensen et al., 2016).

Various methods to utilize localization and map information for TLR have been proposed as this approach uses prior knowledge contained within a map to make online perception simpler and more efficient (Barnes, Maddern, & Posner, 2015; Fairfield & Urmson, 2011; John, Yoneda, Qi, Liu, & Mita, 2014; Levinson et al., 2011; Lindner, Kressel, & Kaelberer, 2004; Tae-Hyun, In-Hak, & Seong-Ik, 2006; Ziegler et al., 2014). In other words, prior knowledge leads to improvements in the accuracy of recognition and reduces algorithm complexity. For instance, the recognition algorithm does not need to operate continuously as perception begins only when the distance to the facing TL is within a certain threshold. As a result, resources can be saved and efficiency enhanced.

This paper introduces a method to extend localization and map information to all stages of TLR and improve recognition performance. The TLR system is divided into four stages: ROI generation, detection, classification, and tracking. ROI generation based on the 3D positions of TLs is a useful way to reduce scan areas within images, but conventional approaches do not take into account slope conditions and so may fail to recognize TLs located on slanted roads. The proposed ROI generation method not only exploits the 3D position of TLs, but also exploits slope information to create precise candidate regions even in uphill or downhill conditions. Furthermore, it is difficult to design a generalized recognition module due to different conventions between countries and non-standardized TLs; thus, this paper introduces a modular architecture to selectively apply specific detectors and classifiers according to the TL type of facing TLs. Lastly, nearest neighbor filters are commonly used for tracking, but track loss or unnecessary track generation might occur with fixed association parameters. As such, this paper proposes an adaptive approach using distance to TL to enhance seamless tracking from long to short range.

This paper provides the following contributions

- (1) It presents a method that reliably compensates for road slopes to precisely generate ROIs by applying linear interpolation between constant slope regions and flat regions.
- (2) It suggests a modular architecture for TLR. Individual recognition modules for all TL types are trained separately offline and selected according to the online retrieved TL type.
- (3) It introduces adaptive data association to compensate for perspective deformation by using distance to TL.

This paper evaluates the proposed methods using on-road experimental data, and the results of a comparison with a standalone vision method show that extending localization and map information to an advanced perception system can secure computational efficiency, reduce false positives, and improve the recognition rate.

## 2. Related works

Much research has been conducted on TLR, along with the recent introduction of survey papers (Diaz et al., 2015; Jensen et al., 2016). The conventional approach is a pure vision based approach

(Almagambetov, Velipasalar, & Baitassova, 2015; Angin, Bhargava, & Helal, 2010; Anh, Ramanandan, Anning, Farrell, & Barth, 2012; Borrman et al., 2014; Cai, Gu, & Li, 2012; Chen, Shi, & Huang, 2015; Chin-Lun, Shu-Wen, & Jyh, 2009; Chiu, Chen, & Hsieh, 2014; Chung, Wang, & Chen, 2002; De Charette & Nashashibi, 2009a, 2009b; Diaz-Cabrera & Cerri, 2013; Diaz-Cabrera, Cerri, & Medici, 2015; Diaz-Cabrera, Cerri, & Sanchez-Medina, 2012; Fan, Lin, & Yang, 2012; Gomez et al., 2014; Jang, Kim, Kim, Lee, & Sunwoo, 2014; Jensen et al., 2015; Jianhua, 2015; Jianwei et al., 2010; Jongwon, Byung Tae, & In So, 2013; Kim et al., 2013; Kim, Kim, & Yang, 2007; Li, Cai, Gu, & Yan, 2011; Michael & Schlipzing, 2015; Nienhuser, Drescher, & Zollner, 2010; Omachi & Omachi, 2009, 2010; Philipsen, Jensen, Mogelmose, Moeslund, & Trivedi, 2015; Philipsen, Jensen, Trivedi, Mogelmose, & Moeslund, 2015; Shi, Zou, & Zhang, 2015; Sooksatra & Kondo, 2014; Trehard, Pollard, Bradai, & Nashashibi, 2014; Xuemei, Yanmin, Minglu, & Qian, 2012; Yehu, Ozguner, Redmill, & Jilin, 2009; Ying, Chen, Gao, & Xiong, 2013; Yu et al., 2010; Zhang, Fu, Yang, & Wang, 2014; Zixing, Yi, & Mingqin, 2012; Zong & Chen, 2014), but there has been less work on utilizing localization and map information for TLR (Barnes et al., 2015; Fairfield & Urmson, 2011; John et al., 2014; Levinson et al., 2011; Lindner et al., 2004; Tae-Hyun et al., 2006; Ziegler et al., 2014). Lindner et al. (2004) first introduced the concept of localization and map information based TLR. The system in that paper consisted of three parts: detector, tracker, and classifier. That paper proposed three different methods for detection, including color-based, shape-based, and cascade classifier-based methods. Among these methods, the paper found that a fusion method integrating a color-based detector, differential GPS, and map information was superior. In particular, the paper showed that when the system has 1 m of positioning accuracy and 1° of heading uncertainty, false positives can be reduced by five times. Tae-Hyun et al. (2006) introduced a guidance module for the TLR system. The guidance module provides the position of the ego vehicle, intersection information, and TL location with several points of attribute data. The map information is utilized for three purposes: as a task trigger for algorithm operation, to limit the search area in an image, and to estimate the size of a TL. Fairfield and Urmson (2011) introduced in detail several advantages when using extended digital map information for TLR: restricted scan areas, the composition of a robust classifier, and enhanced tracking performance. That paper highlighted a method for TL mapping by using image-to-image association and least squares triangulation. As a result, a huge TL map containing over 4000 sites was added to Google maps, and TLR was successfully conducted with a TL map and LIDAR (Llght Detection And Ranging) based precise localization during both day and night. Levinson et al. (2011) proposed another method for TL mapping that sequentially applies tracking, back-projection, and triangulation. The distinguishable contribution here is probabilistically estimating the TL state using various factors such as sensor data and uncertainty along with the relationship of multiple TLs at intersections. John et al. (2014) proposed a method to limit scan area using GPS and map information. In particular, normal and low illumination conditions are categorized into two specific environments and TLR is conducted based on a saliency map and convolutional neural network. In Ziegler et al. (2014), there are two modes for TLR: an offline mode that registers the distinctive visual features around TLs in a database, and an online mode that compares and matches registered features with extracted features in real-time to determine the location of the TL. This method successfully recognized 155 TLs on the Bertha Benz Historical Route. Barnes et al. (2015) introduced a probabilistic framework based on integrating a TL map and localization uncertainty for TLR to enhance online detection performance. This approach generates scale-space candidate regions, and evaluates detection scores based on classification results and the prior distribution of 3D occurrences.

# of Bulbs	2		3		4	
Types and Status	Layout	Status	Layout	Status	Layout	Status
Vertical	Layout	Status	Layout	Status	Layout	Status

Fig. 1. Types and states of TLs in South Korea.

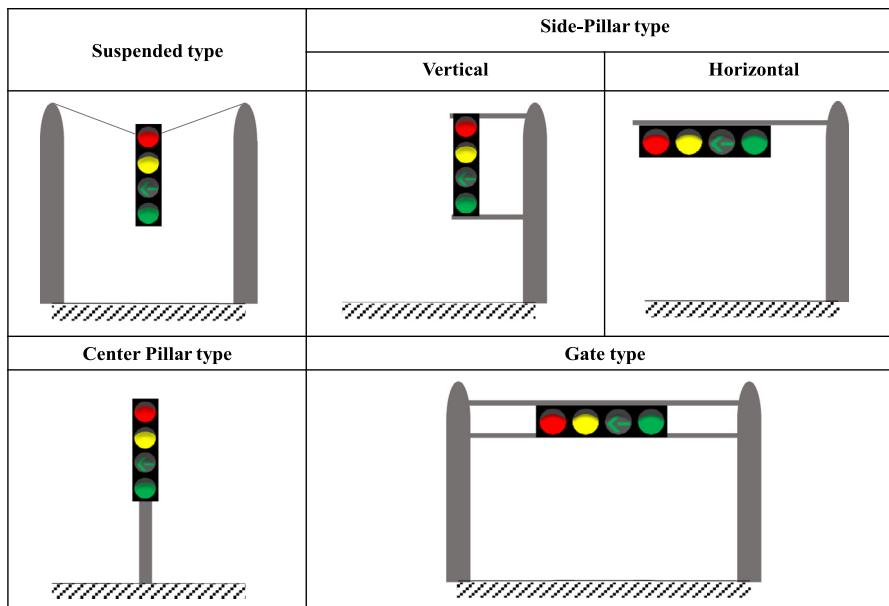


Fig. 2. Various installation methods for TLs.

### 3. System overview

#### 3.1. Traffic light convention

TLs in South Korea follow the Vienna Convention on Road Signs and Signals like most countries in Europe ([United Nations Economic Commission for Europe, 2006](#)), whereas the Federal Highway Administration in the USA regulates TLs through the Manual on Uniform Traffic Control Devices ([Federal Highway Administration, 2015](#)). Fig. 1 lists the various types and states of TLs in South Korea. One signal is composed of a 355 mm × 355 mm black box with colored bulbs inside (diameter 300 mm). Depending on the number of signals, three types of TLs exist: two bulbs, three bulbs, or four bulbs. All of them can stand horizontally or vertically depending on the installation environment at an intersection. For each type, two bulbs have STOP / GO classes, three bulbs have STOP / WARNING / GO classes, and four bulbs have STOP / WARNING / LEFT TURN / LEFT TURN and GO / GO classes. In South Korea, 98% of TLs are categorized into horizontal three bulbs or horizontal four bulbs; thus, this paper uses a recognition method to perceive these two types of TLs.

Fig. 2 describes various installation methods for TLs. There are five different types; however, most TLs in Korea are installed as side-pillar horizontal types.

#### 3.2. Traffic light map database

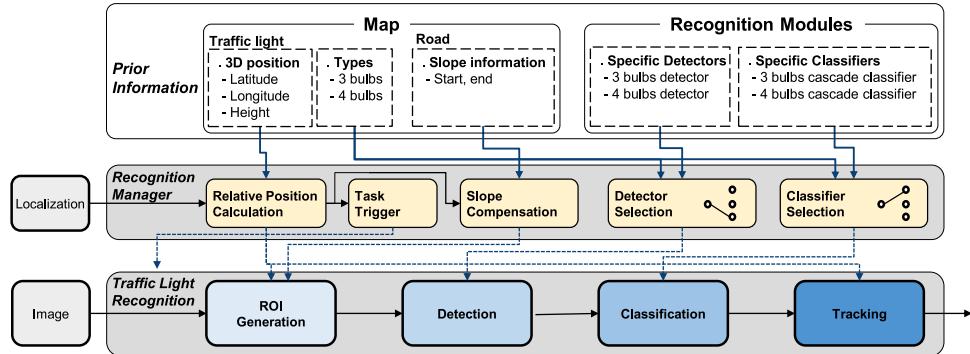
To build a TL map, a high-precision (under 10 cm Circular Error Probability) localization system and a 3D laser scanner are utilized. A probe vehicle records all data, including point clouds and the position of the vehicle with time stamps, when driving through intersections. In offline mode, TL positions are manually selected on an integrated 3D point cloud map. As shown in Table 1, a database stores the latitude, longitude, and height of TLs. Indexes for specific detectors and classifiers are assigned according to type of TL. Furthermore, slope information for the road is included for height compensation during ROI generation. The database is created and managed by the OpenStreetMap standard interface.

#### 3.3. Test vehicle configuration

Several hardware requirements should be considered during the design process of the TLR system. First, focal length must be



**Fig. 3.** (a) System configuration, (b) sample image of color camera.



**Fig. 4.** System architecture.

**Table 1**  
Database specification for traffic light maps.

Items	Unit/Description
Latitude	Degree
Longitude	Degree
Height	Meter
Height at constant slope	Meter, $Z_{ts}$
Type of detectors	Index
Type of classifiers	Index
Slope ID	Index
Start of constant slope	Meter, $d_s$
End of constant slope	Meter, $d_e$
Start of flat road	Meter, $d_f$

carefully configured. A long focal length is more suitable for long range detection; however, this leads to a narrow FOV which can disable detection at short range. Therefore, a focal length that maximizes both detection range and FOV should be selected. Second, a high-resolution image contains more information to describe an object, whereas a low-resolution image has less information at the same focal length. This means that resolution can affect maximum detection range. However, significant processing data causes increases in computation time; therefore, a resolution that guarantees real-time performance for an algorithm should be selected. Third, the camera needs to be tilted slightly to see the upper part of the road, because TLs are generally installed over a specific height. The tilt angle enables the recognition of TLs even when a TL is near the ego vehicle. In particular, installation positions are entirely different according to the intersection; therefore, these environmental factors should be determined through experimental data during the design process.

For the experiments, this paper used a color camera (Basler, acA2000-165uc) with an 8 mm focal length. The image provides a resolution of  $2040 \times 1086$  pixels with a wide FOV ( $70 \times 40^\circ$ ). For instance, if a TL with a 300 diameter lies far away at 100 m, the bulb is projected onto the image as 4 pixels. The camera is installed

with a tilt angle of  $10^\circ$  below the windshield. Also, a localization system (RT3002 of Oxford Technical Solutions) was adopted to provide ego vehicle position with high accuracy. The computing unit had an i5-6300U@2.4GHz with 8GB of RAM, and the algorithm was run in Visual Studio, Windows. Fig. 3 depicts the hardware configuration and a sample image.

## 4. Methodology

### 4.1. System architecture

The proposed TLR system consists of three layers as shown in Fig. 4. The prior information layer contains prior knowledge regarding TLs such as positions, types, and slope information; moreover, the prior information layer provides well-trained individual modules for recognition. The recognition manager layer fuses prior information and real-time positioning data, and then configures the TLR layer sub-modules. The last layer performs vision-based object recognition in four steps. First, ROIs are created in the image through 3D position projection of TLs while considering marginal areas. The HD map provides TL global positions, and the relative 3D position between the ego-vehicle and TLs can be calculated by subtracting the position of ego-vehicle. Second, detection locates TLs in the image. Third, classification identifies the states of the TLs. Last, tracking estimates the positions of the TLs and compensates for the negative effects of false positives or missing TLs. The proposed architecture offers several advantages.

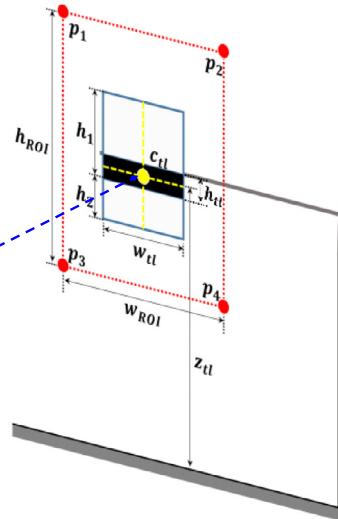
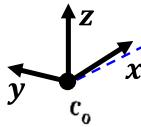
First, the 3D positions of TLs from the map are applied to determine if the algorithm runs or not, in addition to being utilized for ROI reduction on images. This approach can result in a drastic reduction of false positives when compared with vision-standalone approaches. For instance, it is possible to design the algorithm to be robust against similar objects, such as the backlights of vehicles or spot noise generated by external lighting sources, and concentrated searches on ROI can improve the missing rate. Furthermore, it is easy to ensure real-time performance, even at high

$c_o$ : origin of ego vehicle coordinate  
 $c_{tl}$ : center of traffic light ( $x_{tl}, y_{tl}, z_{tl}$ )  
 $h_{tl}$ : height of traffic light box  
 $w_{tl}$ : width of traffic light box

$h_1$ : margin for positive pitch motion  
 $h_2$ : margin for negative pitch motion

$h_{ROI}$ : height of ROI  
 $w_{ROI}$ : width of ROI

$p_1$ : top left corner of ROI  
 $p_2$ : top right corner of ROI  
 $p_3$ : bottom left corner of ROI  
 $p_4$ : bottom right corner of ROI



**Fig. 5.** ROI generation with precise position information.

resolutions, since this approach consumes less computational resources than the full scan method.

Second, ROI generation can be generated properly even in slanted road conditions by utilizing slope information from the map. In practice, there are many cases where TLs are installed on slopes; this situation frequently occurs in South Korea. In such cases, conventional projection from world coordinates to image coordinates is not valid because the calibration parameters of the cameras are derived assuming flat road conditions. Moreover, measuring the slope through sensors is not easy or accurate. Thus, this proposed slope compensation enables the appropriate creation of ROIs on slopes.

Third, it is possible to design and implement individual recognition modules and apply them strategically. TL types vary according to each country's regulations; for example, there are nine types in California, USA, eight types in China, and five types in South Korea. The diversity of target objects can cause design complexity when creating a generalized structure for perception. On the other hand, the type of the facing TL is informed by the recognition manager layer in a localization and map-aided system. This means that it is possible to apply specific recognition modules for perception. This modular architecture might be extended to specify several modules based not only on types of TL, but also on time of day, weather, and other factors.

Lastly, distance to TL information can allow for seamless tracking. Unlike traffic signs, the state of a TL changes recursively; therefore, steady tracking of the TL from long to short range is needed. The proposed adaptive tracking algorithm can overcome loss of tracks and missing problems.

#### 4.2. Traffic light recognition based on precise localization with an HD map

TLR operates only in specific zones that are created based on relative position calculation using localization and map information. This feature functions as a task trigger to manage the activation or deactivation of the TLR system, conserving resources and preventing false detections caused by full activation.

##### 4.2.1. ROI generation and slope compensation

ROI generation is one of the most efficient ways to increase computational efficiency and decrease false positive rates. Precise localization and a TL map can provide more accurate areas on im-

ages. In general, the method of ROI generation is performed as follows:

- i Positioning the ego-vehicle in global coordinates using the precise localization system
- ii Extracting a list in global coordinates of neighboring TLs placed in detection zones on the TL map
- iii Calculating the relative position between the ego-vehicle position and neighboring TLs in local coordinates
- iv Image projection of ROIs that consist of four corners to consider marginal spaces instead of actual TL sizes

To generate ROIs, safety margins should be contained. There are several reasons for doing so. First, the image can be affected by the dynamic motions of the ego-vehicle; in particular, a rapid pitch motion causes fatal failures in ROI generation because image projection assumes a flat road environment. Second, precise localization and the TL map have data uncertainties caused by sensor inaccuracy and human error. Last, projection error gradually increases according to the extension of distance due to camera characteristics; therefore, the safety margin should compensate for these perspective effects.

Fig. 5 describes ROI generation based on precise localization and map information. All calculations are conducted in local coordinates and the unit is meters. An origin that is the ego-vehicle is  $c_o$ , and the center of a TL is  $c_{tl}$ . ROI is defined by  $h_{ROI}$  and  $w_{ROI}$ . The maximum height of the TL box at the maximum positive pitch of the ego-vehicle is  $h_1$ , whereas the minimum height of the TL box at the maximum negative pitch of the ego-vehicle is  $h_2$  as in (1). The height of the ROI is obtained as in (2). In a similar way, the width of the ROI is calculated by multiplying  $w_{tl}$  with the safety factor as in (2). The safety factor accounts for positioning error and human error in the mapping process and should be determined experimentally. The four corners of the ROI can be calculated from  $h_{ROI}$  and  $w_{ROI}$  as in (3), and finally, the four points are projected on an image to generate the ROI.

$$h_1 = A_{+pitch} \times d + \frac{h_{tl}}{2}, \quad h_2 = -\left(A_{-pitch} \times d - \frac{h_{tl}}{2}\right)$$

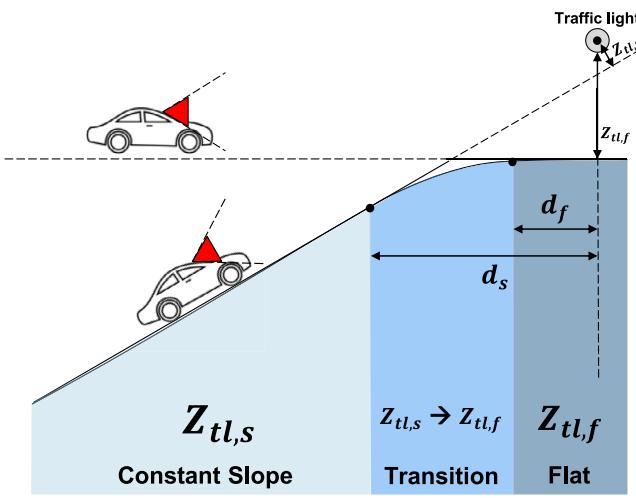
where  $A_{+pitch}$  : positive pitch [unit : radian]

$A_{-pitch}$  : negative pitch [unit : radian]

$d$  : distance from ego vehicle [unit : meter]

(1)

$$h_{ROI} = s_u \times (h_1 + h_2) \quad [\text{unit : meter}]$$



**Fig. 6.** Slope compensation for ROI generation.

$$w_{ROI} = s_u \times w_{tl} \quad [\text{unit : meter}]$$

where  $s_u$  : safety factor for uncertainties of positioning error and map

(2)

$$\begin{aligned} p_1 &= \left( x_{tl}, y_{tl} + \frac{w_{ROI}}{2}, z_{tl} + h_1 \right) \\ p_2 &= \left( x_{tl}, y_{tl} - \frac{w_{ROI}}{2}, z_{tl} + h_1 \right) \\ p_3 &= \left( x_{tl}, y_{tl} + \frac{w_{ROI}}{2}, z_{tl} + h_2 \right) \\ p_4 &= \left( x_{tl}, y_{tl} - \frac{w_{ROI}}{2}, z_{tl} + h_2 \right) \end{aligned} \quad (3)$$

However, this approach is invalid on slopes because it assumes flat road conditions. For instance, Fig. 6 represents the relationship between the ego-vehicle and TL when driving on an inclined road. The height of the TL from the ground is  $Z_{tl,f}$ . Assuming a constant slope for the inclined road,  $Z_{tl,s}$  is more suitable than  $Z_{tl,f}$  as the viewpoint of the ego-vehicle. In short, one reason why ROI generation cannot be applied on slopes is that  $Z_{tl,s}$  and  $Z_{tl,f}$  are entirely different. Therefore, the proposed method to compensate for road slope is as follows. First, the road is divided into three sections: constant slope, transition, and flat sections. The constant slope section uses  $Z_{tl,s}$  and the flat section uses  $Z_{tl,f}$ . The transition section changes height linearly from  $Z_{tl,s}$  to  $Z_{tl,f}$  according to the distance from the TL. All information related to the slope, such as the start point of the constant slope, the end point of constant slope, and the start point of flat road, is stored in the TL map. This method can be applied to decline roads as well by changing  $Z_{tl}$  in (3) to (4). The coefficients  $c_1$  and  $c_2$  are determined by  $d_s$  and  $d_f$ .

$$Z(d) = Z_{tl} + Z_c(d), \text{ where } Z_c = c_1d + c_2 \quad (4)$$

#### 4.2.2. Composition of recognition modules

Localization and map information can inform what type of TL is facing the ego-vehicle. Instead of a generalized approach for recognition, perception methods specialized for each type of TL can be applied to improve recognition rates individually. The recognition process involves the two steps of detection and classification. Several sub-modules are independently designed and implemented according to types of TLs for each step. Sub-modules are selected by indexes from the TL map in online mode.

This paper restricts the targets for TLR to three and four bulbs, since most TLs in South Korea can be categorized into these two types. Therefore, the recognition modules have two detectors and two classifiers as sub-modules. It is possible that the number of modules and the methodology for detection and classification could be modified and replaced for customization.

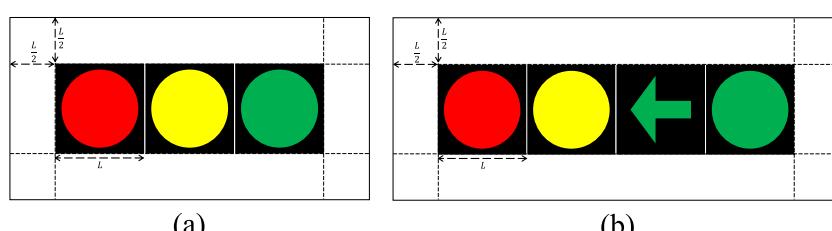
#### 4.2.3. Detection module

For detection, an Adaboost classifier based on Harr-like features is applied (Viola & Jones, 2004), which is a widely used approach for object recognition. There are three reasons for selecting the Adaboost classifier as the detector in recognition modules. First, this paper aims to show that a conventional algorithm can have high precision and recall rates by integrating localization and map information. Second, TLs are easily affected by various illumination and weather conditions, and appearances change slightly according to distance; therefore, applying this machine learning approach can overcome the limitations of color or shape-based approaches. Third, intensity differences between the TL box and the background, in addition to the brightness of an active bulb, are used as distinctive features for the reduction of false positives, as the Haar-like feature represents discrepancies between adjacent areas.

Fig. 7 depicts the specification of positive images for Adaboost training. Each sample for three bulbs and four bulbs has four corner margins with  $L/2$ , where  $L$  is the height of the TL box. Sample images for the three bulbs detector as listed in Table 2 are integrated for training sets that aim to detect all of the states of the three bulbs at once. Sample images for the four bulbs as listed in Table 3 are utilized in the same manner.

#### 4.2.4. Classification module

This paper proposes a cascade classifier to identify the state of the detected image in the previous step. There are two stages: one stage aims to distinguish TLs from the background, while the other stage seeks to determine the exact state of the active bulb as in Fig. 8. This cascade structure is useful for filtering out false positives. Two different features are adopted for each classifier. A histogram of oriented gradients (HOG) is composed of histograms to represent local directions of intensity differences; this can be useful for signifying the position of the active bulb and intensity differences between the black box and the background as shown in Fig. 9(a) (Dalal & Triggs, 2005). A histogram of HSV color for bulb regions as shown in Fig. 9(b) represents distributions of hue, saturation, and values in bulb areas as a feature. This feature can



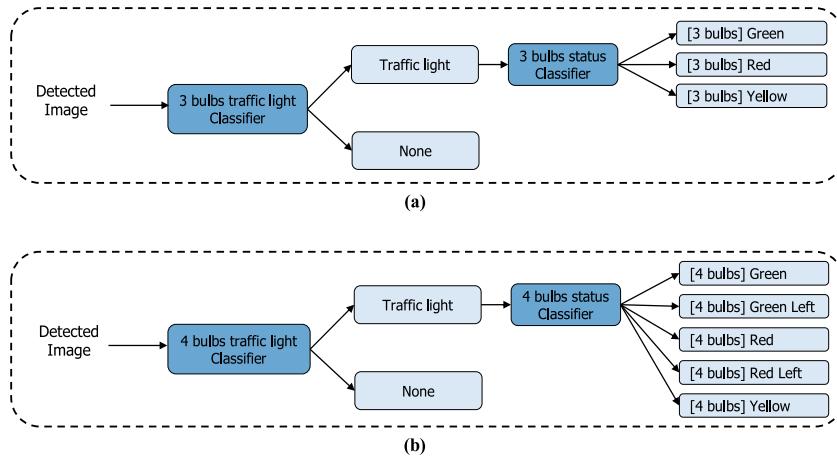
**Fig. 7.** Specification for training samples (a) three bulbs with margin, (b) four bulbs with margin.

**Table 2**  
Training samples for three bulbs detector.

Type	Green	Red	Yellow
Training image			

**Table 3**  
Training samples for four bulbs detector.

Type	Green	Green left	Red	Red left	Yellow
Training image					



**Fig. 8.** Cascade classifiers, (a) the specific classifier for a three bulbs traffic light, (b) the specific classifier for a four bulbs traffic light.

intensify the discernment of classifiers, and the two features are integrated and trained with support vector machines (SVM).

#### 4.2.5. Adaptive tracking with distance information

Kalman filter based multi-object tracking is applied for four states estimation that includes the top-left position, width, and height of a TL box in image coordinates. A static model is used for time updates, and a nearest neighborhood filter (NNF) is employed for data association. Tracks are managed using the three different states of tentative, confirmed, and terminated. Fig. 10 depicts the

system diagram. In general, the validation gate for the NNF is set as fixed. However, a fixed gate size can cause several problems. If the gate size is set as small, existing tracks cannot be associated with any measurements at short distances, whereas if the gate size is set as large, two different tracks can be merged into one track at long distances. This phenomenon is caused by the perspective effect of the camera. Fig. 11 shows positional discrepancies for the top-left corner between two consecutive frames. At long distances, the two positions are almost the same. In contrast, the difference is large at short distances.

To solve this problem, gate size should change according to the distance to TL. The relationship between gate size and distance is represented as a model containing non-linear characteristics in (5). This model compensates for the perspective effect during data association.

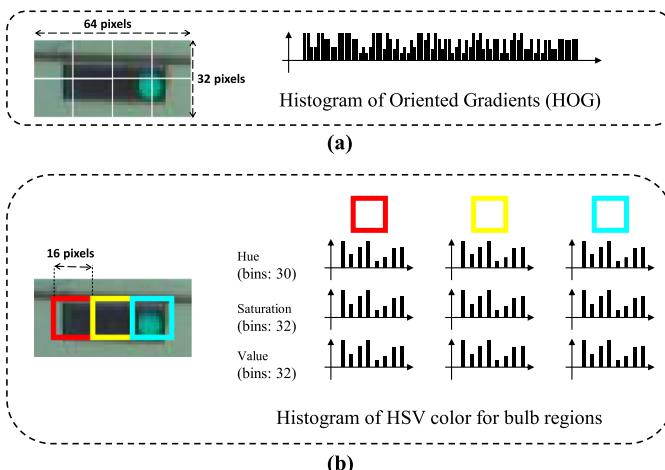
$$G_{\text{size}} = c_1 e^{c_2 d} + c_3 \quad (5)$$

$G_{\text{size}}$  : Gate size  
 $d$  : distance from ego vehicle to traffic light  
 $c_1, c_2, c_3$  : coefficients

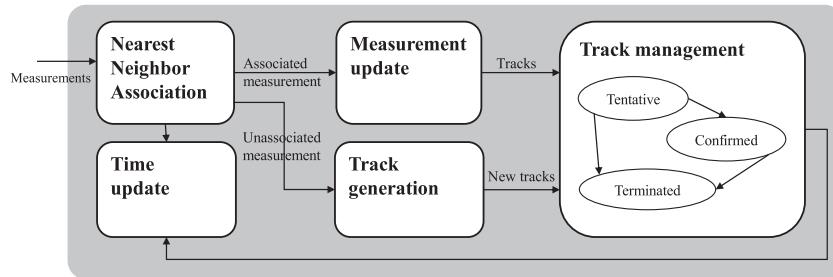
## 5. Preparation of recognition modules

### 5.1. Detection modules

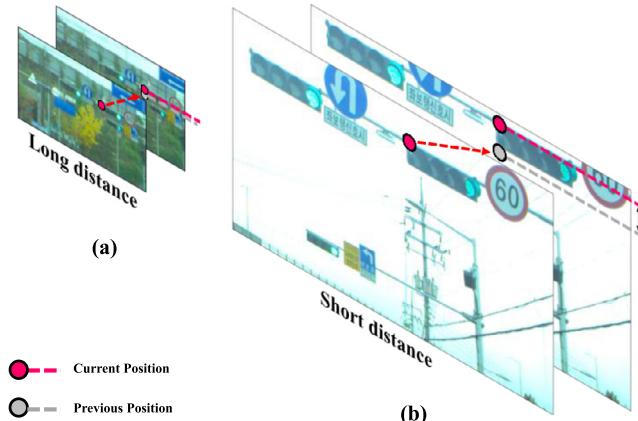
Adaboost classifiers for detection modules were trained with various positive samples as listed in Table 4. Half of the samples came from test sites and the remainder came from other locations. Furthermore, the samples were acquired during different times and weather conditions such as sunny, cloudy, rainy, and sunset, to reflect the diversity of driving and illumination conditions. For



**Fig. 9.** Features for classification, (a) Histogram of oriented gradients, (b) Histogram of HSV color for bulb regions.

**Fig. 10.** System diagram of Kalman filter based tracking with Nearest Neighbor Association.**Table 4**  
Training samples for detectors.

Type	3 bulbs			4 bulbs				Total
Status	Green	Red	Yellow	Green	Green left	Red	Red left	Yellow
samples	11079	14591	1921	7640	2603	4702	896	1080
Total	27591			16921				44512

**Fig. 11.** Position differences in the top-left corner of the traffic light between consecutive two frames at a long distance and a short distance. (a) Long distance, small difference (b) Short distance, large difference.

negative samples, images captured on roads without TLs, in addition to natural images, were used.

## 5.2. Classification modules

A specific classifier was selected by an index stored in the TL map according to the type of facing TL. The classification modules have two types of cascade classifiers, one for three bulbs and one for four bulbs. During the cascade process, the first step distinguished TLs from the background and the second step consequently identified the state of the active bulb. The training samples, composed of resultant images from detectors and the proportions for each state, are listed in Tables 5 and 6. During the process of SVM training, a jack-knife method was applied iteratively to achieve optimal cascade classifiers with high classification rates.

**Table 5**  
Training samples for the three states classifier for the three bulbs detector.

Class	3 bulbs		
	Green	Red	Yellow
Samples	13386	9297	499

## 6. Evaluation of the localization and map aided traffic light recognition system

### 6.1. Experimental environments

To evaluate the proposed system, a total of six experimental sites were selected and a TL map was constructed. The sites were carefully chosen using the following criteria: First, various types of TLs and different operation orders had to be included to show the effectiveness of the localization and map-aided approach. Second, road conditions consisting of flat roads, slanted roads, and uneven roads were needed to evaluate the proposed ROI generation method. Third, TLs in sites with complex backgrounds were utilized to assess the robustness of detection and classification. The types, road conditions, and environmental complexities of each location are explained in Table 7 and Fig. 12.

### 6.2. ROI Generation with slope compensation

To compensate for the pitch motions of the ego-vehicle,  $A_{-pitch}$  and  $A_{+pitch}$  in (1) should be analyzed in advance; however, it is hard to measure pitch precisely, and it depends on the type of vehicle and the vehicle's settings. Nevertheless, simulation analysis for pitch motions might be meaningful because it can provide insights related to the characteristics of the vehicle's dynamics. Carsim is a high-order model based vehicle simulator where pitch can be directly measured in various test scenarios. Two test scenarios were prepared using this simulator: one was a scenario involving full acceleration for maximum negative pitch, and the other was a scenario involving full deceleration for maximum positive pitch. Nine types of vehicles were involved, including a full size SUV, minivan, sedan, and hatchback, as shown in Fig. 13. The maximum pitch among them was 3.07° for the D-class minivan, whereas the minimum pitch was -1.13 for the E-class SUV. These pitch figures inform ROI generation; for example, in the case of a D-class SUV of a medium size, ROIs are generated according to distance changes as in Fig. 14. And the reason why the ratio of ROI height to TL box

**Table 6**  
Training samples for the five states classifier for the four bulbs detector.

Class	4 bulbs				
	Green	Green left	Red	Red left	Yellow
Samples	6621	2206	3412	295	525

**Table 7**  
Specification of traffic lights for evaluation.

No.	Bulbs	Operation	Road condition	Environment Complexity
1	3	Green → Yellow → Red	Flat	Normal
2			Inclined slope	Under the bridge, shade
3			Declined slope	Partial complex background
4			Declined slope	Complex background
5	4	Green Left → Yellow → Red	Flat	Normal
6		Green → Green Left → Yellow → Red	Uneven road	Complex background

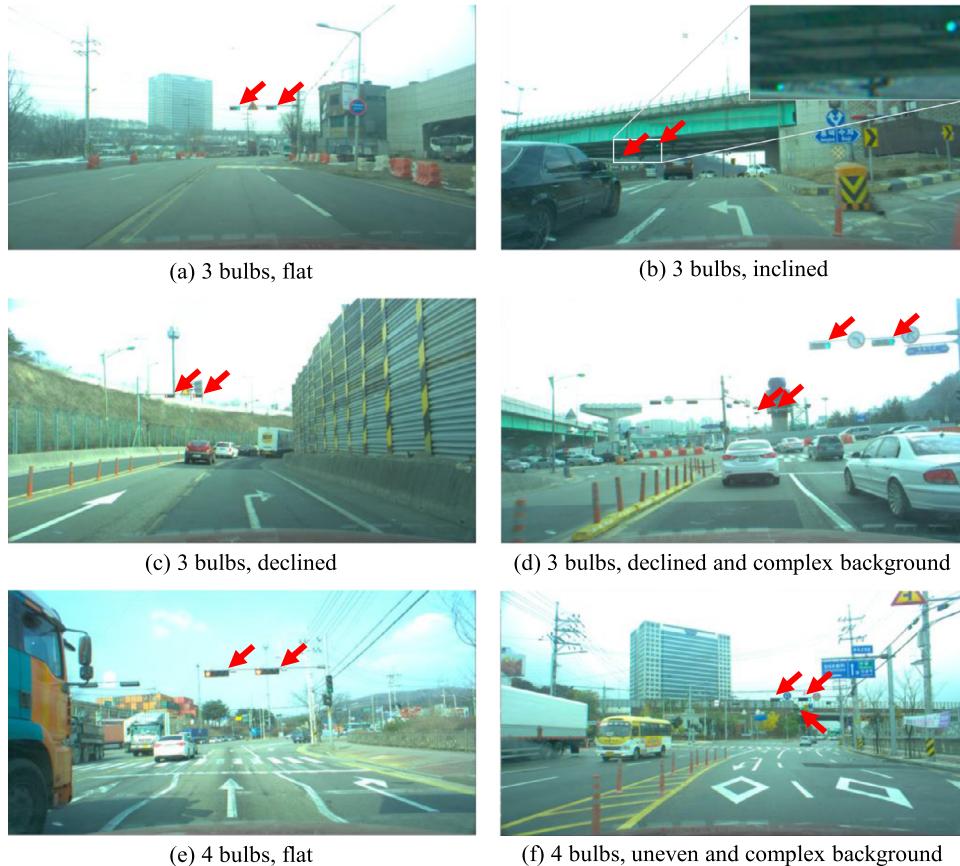


Fig. 12. Six intersections for evaluation.

height is larger at long distances than it is at short distances is that the effect of pitch is greater in images at long distances. This influence of perspective decreases gradually along with distance. Thus, in Fig. 14, the safety factor is set as 2. This safety factor should be obtained experimentally.

On slopes, ROI generation should be modified. For instance, TLs can be on a decline road as in Fig. 15(a). Although ROIs are gener-

ated with sufficient margins, TL boxes are placed at the top of the ROI. This approach may be successful, but only if there is no additional dynamic motion by the ego-vehicle. Next, Fig. 15(d) shows TLs on an inclined road. In this case, the TLs are placed out of ROIs due to a steep slope. It is impossible to detect the TL until the distance to the TL is small. This might cause rapid braking due to the sudden recognition of a red signal. To solve the problems posed by both cases, enlarged ROIs could be one solution; however, simple area expansion leads to increases in computation time and false detections against neighbor TLs.

As such, Fig. 15(b) and (e) show the results of the proposed slope compensation. First, ROIs are centered on the TL boxes and the constant slope and flat road heights that are used for the compensation are stored in the TL map as in Table 1. Then, during the transition section, the height is changed from  $Z_{tl,s}$  to  $Z_{tl,l}$  linearly by (4). This method is straightforward and efficient at compensating for the slope. Fig. 15(c) and (f) show the results of ROI generation at short distances. The two ROIs have identical sizes in Fig. 15(c) due to being at the same distance, whereas two different ROIs are generated according to longitudinal distances. Ideally, the horizontal centers of the TL box and ROI should be coincident; however, uncertainties from positioning errors and human errors

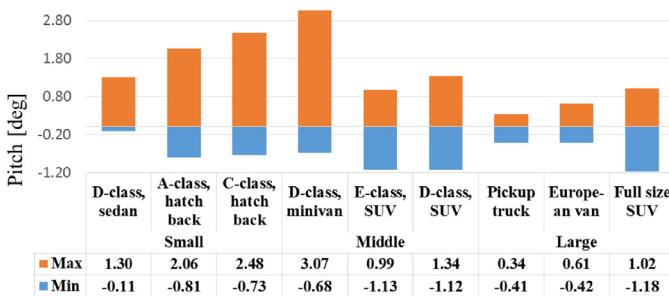
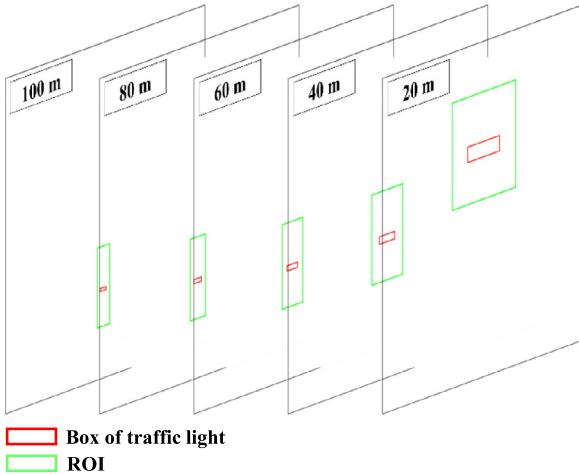


Fig. 13. Min and Max Pitch for full acceleration from 0 km/h and full deceleration from 100 to 0 km/h, respectively. Data is obtained from a vehicle simulator supporting high-order dynamic models, Carsim.



**Fig. 14.** ROI generation for D-class, SUV (min. pitch:  $-1.12^\circ$ , max. pitch:  $1.34^\circ$ , safety factor: 2) according to various distances.

during manual mapping cause discordance. The safety factor can absorb these uncertainties.

### 6.3. Test results and analysis

For performance evaluation of the proposed method, we recorded ten test sequences in the test sites, including 4054 frames with TLs and 3579 frames without TLs. The three performance indexes employed for detection and classifi-

cation were Precision=TP/(TP+FP), Recall=TP/(TP+FN), and Accuracy=(TP+TN)/((TP+TN+FP+FN)). For each sequence, the tables in Fig. 16 show the frame compositions and comparison results between the proposed method and the standalone vision method. Furthermore, the bottom graph for each sequence plots recognition results along with time. For the graphs, the gray marker represents the ground truth, the blue marker represents the response of the proposed method, and the red marker represents the results of the standalone vision method. The down-arrow with distance annotations stands for the initial recognition point of the proposed method. For calculating accurate recall, we set the maximum recognition distance of the proposed method as 80 m, which is the system's limit, and performed ground truth annotation for TLs within 80 m. If occlusion occurred, the annotation started from the moment of no occlusion. The y-axis labels for the graphs represent the seven classes of three bulb and four bulb TLs.

The proposed method resulted in a high precision rate (Detection precision: 99.68%, Detection + Classification precision: 98.68%) due to the task trigger that allows TLR to be executed in the limited regions and narrow ROIs in the image, as in Fig. 17 and Table 8. In contrast, the standalone vision method caused crucial false positives as shown in Figs. 16(ⓐ–①) and 18(a, b, d, e). The false positives in Figs. 16(ⓐ, ③–①) and 18(a, d) resulted from wrong detection and classification. Moreover, the false positives in Figs. 16(ⓑ, ④–①) and 18(b, e) show wrong recognition of TLs in other lanes. These faults represent one of the most challenging problems for the standalone vision method because it is hard to distinguish relevant TLs from irrelevant ones. In addition, for actual implementation, a low false positive rate is one of the most important factors for smooth driving and safety. Red false positives might cause



**Fig. 15.** Results of slope compensation for ROI generation in two different test sites, including incline and decline roads. No slope compensation: (a), (d). Slope compensation: (b), (e). Flat road: (c), (f).

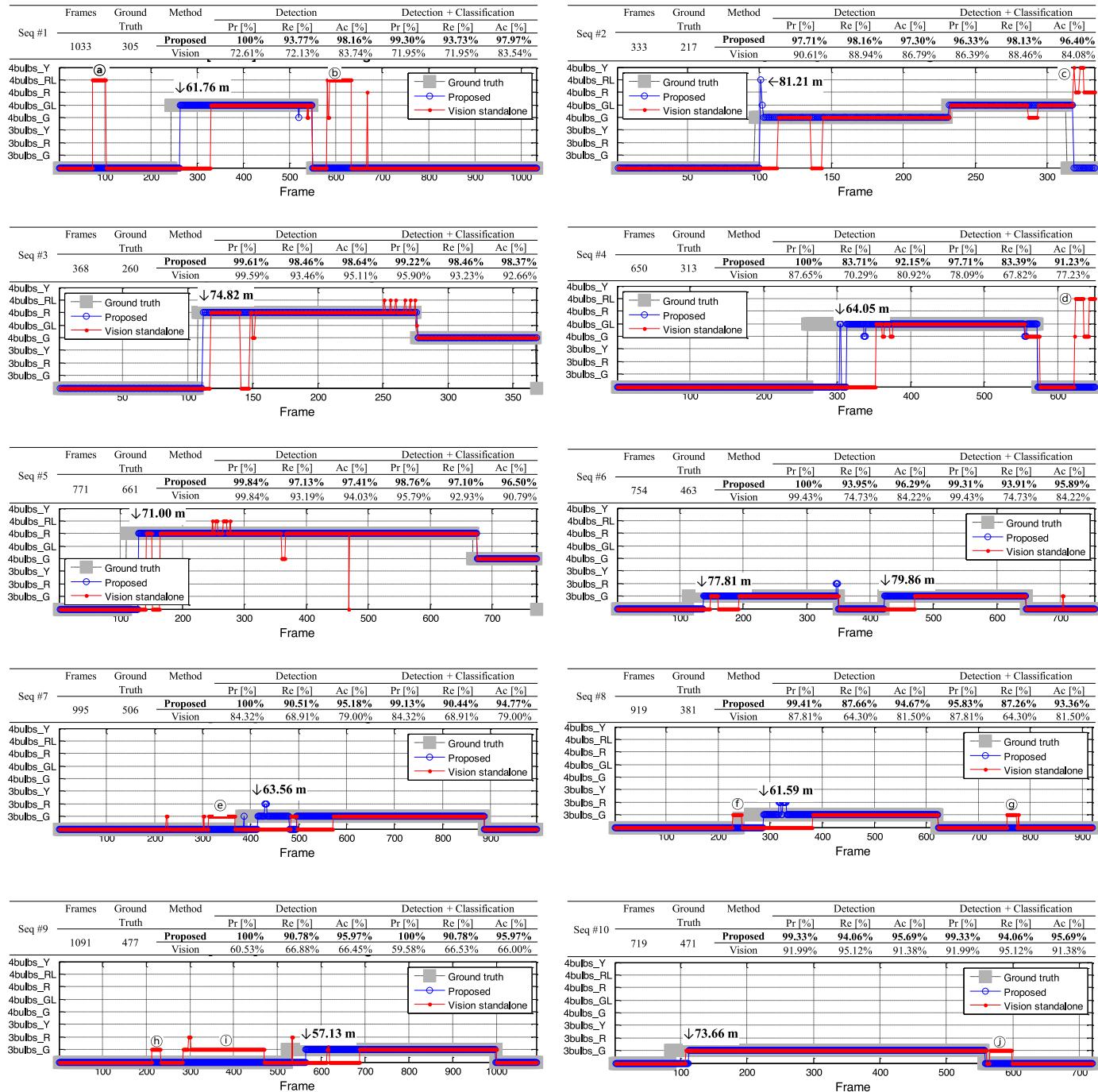


Fig. 16. Recognition results for ten test sequences.

**Table 8**  
Recognition rate.

Type	Frames	Ground truth	Method	Detection			Detection + Classification		
				Pr [%]	Re [%]	Ac [%]	Pr [%]	Re [%]	Ac [%]
3 bulbs + 4 bulbs	7633	3762	Proposed	99.68%	92.80%	96.02%	98.68%	92.73%	95.52%
			Standalone vision	86.88%	78.39%	82.24%	84.86%	77.99%	81.27%
3 bulbs	4478	2103	Proposed	99.76%	91.51%	95.53%	98.86%	91.44%	95.11%
			Standalone vision	83.05%	73.37%	78.65%	83.05%	73.37%	78.65%
4 bulbs	3155	1492	Proposed	99.58%	94.48%	96.70%	98.44%	94.42%	96.10%
			Standalone vision	91.65%	84.97%	87.32%	84.12%	76.28%	80.26%



**Fig. 17.** Correct recognition of localization and map-aided traffic light recognition. (a)-(c): 3 bulbs, (d)-(f): 4 bulbs. Blue boxes indicate ROI; red boxes represent recognition results for each traffic light. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 18.** False positives and missings. (a), (d): wrong detection and classification. (b), (e): wrong recognition of traffic lights in other lanes. (c), (f): missing due to the complex and dark background far away.

sudden emergency stops during driving, and green false positives might let the vehicle go despite red lights at intersections. Such abnormal operation can cause significant traffic accidents due to the response of other vehicles toward unexpected behaviors or violations of traffic regulations. The proposed method can solve these problems by integrating localization with the TL map and identifying which TLs are included in the ego lane.

The recognition manager chooses a specific detector and classifier according to the facing TL type. Table 8 shows the recognition analysis for each test sequence. The three rows represent the performance indexes for unified TLs, three bulbs, and four bulbs, respectively. Also, the detection column indicates detector results alone, whereas the detection + classification column represents the final recognition results. For unified TLs, the final recognition results are Precision 98.68%, Recall 92.73%, and Accuracy 95.52%. For three bulbs, the final recognition results are Precision 98.86%, Recall 91.44%, Accuracy 95.11%. For four bulbs, the final recognition results are Precision 98.44%, Recall 94.42%, Accuracy 96.10%. The proposed method reached high precision of over 98% due to the structure of the cascade classifiers with HOG and histograms of HSV. Complex and dark backgrounds reduced recall because the low-intensity differences between the box of the TL and the back-

**Table 9**

Execution times for the proposed method and the standalone vision method.

Step	Proposed [ms]	%	Standalone vision [ms]	%
ROI generation	0.36	3.18	0.0	0
Detection	7.83	69.92	170.26	95.23
Classification	2.97	26.49	8.45	4.73
Tracking	0.05	0.41	0.07	0.04
Total	11.20	100	178.78	100

ground in the far distance caused false negatives as in Fig. 18(c, f). However, if the distance came to within 50 m, recall reached 99% since the difference in intensity became distinguishable through the perspective effect. As a result, maximum distances varied from 57.13 m to 81.21 m, depending on the sequence. The experimental results mean that the proposed method cannot always guarantee a maximum distance of 80 m due to environmental effects. Therefore, the behavior of the ego vehicle at the intersections should be carefully determined after recognizing the facing TLs.

The proposed method was 16 times faster than the standalone vision method, as shown in Table 9. The narrow ROIs and limited scale range for the pyramid images sped up the detection process, which occupies the largest proportion of total execution time. The

**Table 10**

Comparison with previous studies.

Paper	Year	Dataset	Frames	Ground Truth	Resolution [Pixels]	TL types	TL states	Pr [%]	Re [%]	Ac [%]
Barnes	2015	Local	–	9,301 frames	6 × 1024 × 768	Vertical 3 bulbs	3 classes Red, Amber, Green, Red & Amber	92.3	99.0	–
John Levinson	2014	Local	3351	3351 TLs	–	Horizontal 3 bulbs	2 classes Red, Green	–	–	97.83
John Levinson	2011	Local	–	–	1280 × 1024	Vertical 3 bulbs	2 classes Red, Yellow, Green	81.46	77.98	92.85
Ziegler	2013	Local	–	–	–	Vertical 3 bulbs	3 classes Red, Yellow, Green	–	–	–
Lindner	2004	Local	–	–	–	Vertical 3 bulbs	3 classes Red, Yellow, Green	–	> 95	–
Fairfield	2011	Local	–	1,383 frames	2040 × 1080	Vertical	2 classes Red, Green	99.65	61.94	93.63
Hwang	2006	Local	–	–	720 × 480	Horizontal	2 classes Red, Green	–	–	–
Proposed	2016	Local	7995	4122 frames	2040 × 1086	Horizontal 3 bulbs	3 classes Green, Red, Yellow	98.68	92.73	95.52
						Horizontal 4 bulbs	4 classes Green, Green Left, Red, Red Left, Yellow			

time for classification was proportional to the number of hypotheses from the detection step. The standalone vision method required more time to classify hypotheses due to false positives. ROI generation worked only for the proposed method and required an average of 0.36 ms for two or three ROIs. This time includes image projection with camera calibration parameters and slope compensation. The proportion of tracking was very low. In addition to the reduction in execution time, the proposed method provides an advantage in resource utilization because the algorithm is run only at intersections or crosswalks.

A method that utilizes GPS position and high definition maps containing accurate prior knowledge of TL locations is categorized as an auxiliary detection method for TLR in Jensen et al. (2016), and is introduced as one effective way to reject false positives because extended digital maps provide sure locations for TLs. Table 10 shows the comparison results between the proposed method and previous studies of auxiliary detection methods. It was difficult to compare the results qualitatively and quantitatively because each proposition was evaluated using different databases and different configurations. In particular, the recognition performance of the auxiliary detection methods is tightly coupled to the accuracy of ego vehicle positioning and the maps. The proposed method offers high precision along with the capability to recognize seven TL states.

In contrast to the previous auxiliary methods, the proposed information fusion architecture makes it possible to use auxiliary information at every stage of recognition. The HD map not only contains accurate positions of TLs, but also their specific types; therefore, individual recognition modules can be designed and implemented separately. This modular architecture enables straightforward customization by replacing recognition modules according to the specific conventions of a country. Furthermore, we emphasize that TL tracking is very important for the driving stability of autonomous vehicles because track loss and unnecessary tracks might cause irregular operations such as unexpected braking or repetitive stop and go motions. To show this enhanced tracking performance, we draw recognition results on a timeline in comparison with standalone vision as in Fig. 16. It is noted that the proposed adaptive model based on distance from TLs in (5) compensates for the perspective effect of the camera.

The proposed method depends highly on prior information. In particular, the HD map should store the necessary geographical information of all TLs. If a TL location is changed or a new TL is installed, the HD map should be updated. This means that maintenance of the HD map is a critical factor in recognition reliability. Currently, due to the importance of an HD map for autonomous driving, several mapping companies have concentrated on map

construction; thus, we expect that mapping technologies will improve in terms of map accuracy and update frequency. Also, the positioning accuracy of ego vehicles should be guaranteed within 30 cm at intersections because this directly affects ROI generation and adaptive tracking, as they utilize distance to TLs. Although the positioning requirement can be satisfied with high precision RTK-GPS, which provides an error margin of 2 cm, it is not affordable for commercialization due to its cost. Thus, low cost GPS-based precise localization, such as in Jo, Jo, Suhr, Jung, and Sunwoo (2015) and Suhr, Jang, Min, and Jung (2017), is required.

## 7. Conclusion

This paper discussed an extended methodology for precise localization and map-aided TLR. The key factor of this paper is that the auxiliary information is utilized at every stage of recognition for improving the reliability, accuracy, and efficiency of TLR as follows:

- (1) In ROI generation, restriction of candidate areas on images results in a drastic reduction in the execution time for detection, going up to a 16x reduction, as compared to a full scan approach. This implies that high-resolution cameras can be utilized to cover long ranges and wide FOVs simultaneously in real time. Also, we proposed a new slope compensation method for ROI generation on incline or decline roads. The slanted road is divided into the three sections of constant slope, transition, and flat road, and for each section, different TL heights are calculated by using slope information from the map.
- (2) The recognition modules consist of two detection and two classification modules, and they are selected and applied to the retrieved indexes regarding facing TL types. This approach reduces the design complexity of recognition systems and allows for increased maintainability and flexibility. Different recognition modules can be designed and implemented not only for specific types, but also for various weather and time conditions. These added modules may improve the existing recognition system.
- (3) Adaptive tracking with distance information allows for the seamless continuity of tracks and minimizes the side effects of false alarms or missing problems. This advantage results in the exact estimation of the current state of the TL and smooth vehicle control without disturbances during autonomous driving.

Experimental results showed that the proposed system recognized traffic lights with 98.68% precision, 92.73% recall, and 95.52%

accuracy in the 10.02–81.21 m range. From the experiments, we realized that the maximum distance significantly depends on the intersection environment and TL conditions. For instance, complex backgrounds and low levels of TL brightness cause degradation of the recognition rate at remote distances. Such environmental factors might be hard to overcome by algorithmic evolution alone; thus, attempts at strategic resolution are necessary. For example, we can design a driving strategy that lowers vehicle speed before entering intersections and determines the behavior of the vehicle after the traffic lights are verified. This conservative approach might be preferable for safety in autonomous driving.

Furthermore, the proposed method requires a high precision TL map that requires considerable time and effort to construct. If either map information or location changes regarding facing TLs are unavailable, TLR will fail. However, intensive research in map generation is being performed, and the government is providing maps of limited areas in South Korea. If such maps become popular and such programs are expanded, the proposed method would become one of the most efficient ways of conducting TLR.

To enhance the current system, we plan to extend the maximum detection range to over 150 m by replacing recognition modules as the current system is limited to 80 m. We expect the extended range will help vehicle behavior planning at intersections. Furthermore, we will perform nighttime TLR with new recognition modules that will solve TL color saturation problems at long exposures. The recognition modules are changeable in the proposed architecture; thus, if we can detect night conditions, such nighttime recognition modules will be used for TLR at night instead of the daytime modules. Another challenging problem is TL map updating. By using the proposed recognition modules, the TL map might be offline-updated with high precision positioning data and locations through recorded images. Such an updated TL map could be used for online TLR in a circular fashion.

## References

- Almagambetov, A., Velipasalar, S., & Baitassova, A. (2015). Mobile standards-based traffic light detection in assistive devices for individuals with color-vision deficiency. *IEEE Transactions on Intelligent Transportation Systems*, 16(3), 1305–1320. doi:10.1109/TITS.2014.2361139.
- Angin, P., Bhargava, B., & Helal, S. (2010). A mobile-cloud collaborative traffic lights detector for blind navigation. *Paper presented at the mobile data management (MDM), 2010 eleventh international conference on*, 23–26 May 2010.
- Anh, V., Ramanandan, A., Anning, C., Farrell, J. A., & Barth, M. (2012). Real-time computer vision/DGPS-aided inertial navigation system for lane-level vehicle navigation. *Intelligent Transportation Systems, IEEE Transactions on*, 13(2), 899–913. doi:10.1109/TITS.2012.2187641.
- Barnes, D., Maddern, W., & Posner, I. (2015). Exploiting 3D semantic scene priors for online traffic light interpretation. *Paper presented at the 2015 IEEE intelligent vehicles symposium (IV)*.
- Bazzi, A., Zanella, A., & Masini, B. M. (2016). A distributed virtual traffic light algorithm exploiting short range V2V communications. *Ad Hoc Networks*, 49, 42–57.
- Borrman, J. M., Haxel, F., Nienhuser, D., Viehl, A., Zollner, J. M., Bringmann, O., & Rosenstiel, W. (2014). STELLAR – A case-study on systematically embedding a traffic light recognition. *Paper presented at the 2014 17th IEEE international conference on intelligent transportation systems, ITSC 2014*.
- Cai, Z., Gu, M., & Li, Y. (2012). Real-time arrow traffic light recognition system for intelligent vehicle. *Paper presented at the 2012 international conference on image processing, computer vision, and pattern recognition, IPCV 2012, Las Vegas, NV*.
- Chen, Z., Shi, Q., & Huang, X. (2015). Automatic detection of traffic lights using support vector machine. *Paper presented at the 2015 IEEE intelligent vehicles symposium (IV), June 28–July 1 2015*.
- Chin-Lun, L., Shu-Wen, C., & Jyh, S. (2009). An integrated portable vision assistant agency for the visual impaired people. *Paper presented at the control and automation, 2009. ICCA 2009. IEEE international conference on*, 9–11 December 2009.
- Chiu, Y. T., Chen, D. Y., & Hsieh, J. W. (2014). Real-time traffic light detection on resource-limited mobile platform. *Paper presented at the digest of technical papers - IEEE international conference on consumer electronics*.
- Chung, Y.-C., Wang, J.-M., & Chen, S.-W. (2002). A vision-based traffic light detection system at intersections. *Journal of Taiwan Normal University: Mathematics, Science and Technology*, 47(1), 67–86.
- Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. *Paper presented at the computer vision and pattern recognition, 2005. CVPR 2005. IEEE computer society conference on*, 25–25 June 2005.
- De Charette, R., & Nashashibi, F. (2009a). Real Time Visual Traffic Lights Recognition Based on Spot Light Detection and Adaptive Traffic Lights Templates. *IEEE Intelligent Vehicles Symposium, 2009* (pp. 358–363).
- De Charette, R., & Nashashibi, F. (2009b). Traffic light recognition using image processing compared to learning processes. *Paper presented at the intelligent robots and systems, 2009. IROS 2009. IEEE/RSJ international conference on*, 10–15 October 2009.
- Diaz-Cabrera, M., & Cerri, P. (2013). Traffic Light recognition during the night based on fuzzy logic clustering. In R. Moreno-Díaz, F. Pichler, & A. Quesada-Arencibia (Eds.). *In Computer aided systems theory - EUROCAST 2013: Vol. 8112* (pp. 93–100). Berlin Heidelberg: Springer.
- Diaz-Cabrera, M., Cerri, P., & Medici, P. (2015). Robust real-time traffic light detection and distance estimation using a single camera. *Expert Systems with Applications*, 42(8), 3911–3923. doi:10.1016/j.eswa.2014.12.037.
- Diaz-Cabrera, M., Cerri, P., & Sanchez-Medina, J. (2012). Suspended traffic lights detection and distance estimation using color features. *Paper presented at the 2012 15th international IEEE conference on intelligent transportation systems, ITSC 2012*.
- Diaz, M., Cerri, P., Pirlo, G., Ferrer, M. A., & Impedovo, D. (2015). A survey on traffic light detection. *Paper presented at the new trends in image analysis and processing-ICIAP 2015 workshops*.
- Fairfield, N., & Urmon, C. (2011). Traffic light mapping and detection. *Paper presented at the robotics and automation (ICRA), 2011 IEEE international conference on*, 9–13 May 2011.
- Fan, B., Lin, W., & Yang, X. (2012). An efficient framework for recognizing traffic lights in night traffic images. *Paper presented at the image and signal processing (CISP), 2012 5th international congress on*, 16–18 October 2012.
- Federal Highway Administration, W., DC, USA. (2015). *Manual on uniform traffic control devices* Retrieved from <http://mutcd.fhwa.dot.gov/>.
- Ferreira, M., & d'Orey, P. M. (2012). On the impact of virtual traffic lights on carbon emissions mitigation. *IEEE Transactions on Intelligent Transportation Systems*, 13(1), 284–295.
- Florin, R., & Olariu, S. (2015). A survey of vehicular communications for traffic signal optimization. *Vehicular Communications*, 2(2), 70–79.
- Gomez, A. E., Alencar, F. A. R., Prado, P. V., Osorio, F. S., & Wolf, D. F. (2014). Traffic lights detection and state estimation using Hidden Markov Models. *Paper presented at the IEEE intelligent vehicles symposium, proceedings*.
- Jang, C., Kim, C., Kim, D., Lee, M., & Sunwoo, M. (2014). Multiple exposure images based traffic light recognition. *Paper presented at the 2014 IEEE intelligent vehicles symposium proceedings*, 8–11 June 2014.
- Jensen, M. B., Philipsen, M. P., Bahnsen, C., Møgelmose, A., Moeslund, T. B., & Trivedi, M. M. (2015). Traffic light detection at night: comparison of a learning-based detector and three model-based detectors. In G. Bebis, R. Boyle, B. Parvin, D. Koracin, I. Pavlidis, R. Feris, T. McGraw, M. Elendt, R. Kopper, E. Ragan, Z. Ye, & G. Weber (Eds.), *Advances in visual computing: 11th international symposium, ISVC 2015, Las Vegas, NV, USA, December 14–16, 2015, proceedings, Part I* (pp. 774–783). Cham: Springer International Publishing.
- Jensen, M. B., Philipsen, M. P., Møgelmose, A., Moeslund, T. B., & Trivedi, M. M. (2016). Vision for Looking at traffic lights: issues, survey, and perspectives. *IEEE Transactions on Intelligent Transportation Systems*, PP(99), 1–16. doi:10.1109/TITS.2015.2509509.
- Jianhua, S. (2015). Arrow traffic lights real-time detection in urban environments for high resolution camera using geometry information. *International Journal of Earth Sciences and Engineering*, 8(1), 249–255.
- Jianwei, G., Yanhua, J., Guangming, X., Chaohua, G., Gang, T., & Huiyan, C. (2010). The recognition and tracking of traffic lights based on color segmentation and CAMSHIFT for intelligent vehicles. *Paper presented at the intelligent vehicles symposium (IV), 2010 IEEE*, 21–24 June 2010.
- Jo, K., Jo, Y., Suhr, J. K., Jung, H. G., & Sunwoo, M. (2015). Precise localization of an autonomous car based on probabilistic noise models of road surface marker features using multiple cameras. *IEEE Transactions on Intelligent Transportation Systems*, 16(6), 3377–3392. doi:10.1109/TITS.2015.2450738.
- John, V., Yoneda, K., Qi, B., Liu, Z., & Mita, S. (2014). Traffic light recognition in varying illumination using deep learning and saliency map. *Paper presented at the 2014 17th IEEE international conference on intelligent transportation systems, ITSC 2014*.
- Jongwon, C., Byung Tae, A., & In So, K. (2013). Crosswalk and traffic light detection via integral framework. *Paper presented at the frontiers of computer vision, (FCV), 2013 19th Korea-Japan joint workshop on*, January 30 2013–February 1 2013.
- Kim, H.-K., Shin, Y.-N., Kuk, S.-g., Park, J. H., & Jung, H.-Y. (2013). Night-Time traffic light detection based on SVM with geometric moment features. *Paper presented at the proceedings of world academy of science, engineering and technology*.
- Kim, Y. K., Kim, K. W., & Yang, X. (2007). Real time traffic light recognition system for color vision deficiencies Harbin.
- Levinson, J., Askeland, J., Dolson, J., & Thrun, S. (2011). Traffic light mapping, localization, and state detection for autonomous vehicles. *Paper presented at the robotics and automation (ICRA), 2011 IEEE international conference on*, 9–13 May 2011.
- Li, Y., Cai, Z. X., Gu, M. Q., & Yan, Q. Y. (2011). Traffic lights recognition based on morphology filtering and statistical classification Shanghai.
- Lindner, F., Kressel, U., & Kaelberer, S. (2004). Robust recognition of traffic signals. *Paper presented at the intelligent vehicles symposium, 2004*, 14–17 June 2004. IEEE.
- Michael, M., & Schlippling, M. (2015). Extending traffic light recognition: Efficient classification of phase and pictogram. *Paper presented at the proceedings of the international joint conference on neural networks*.
- Nienhuser, D., Drescher, M., & Zollner, J. M. (2010). Visual state estimation of traffic lights using hidden Markov models. *Paper presented at the intelligent*

- transportation systems (ITSC), 2010 13th international IEEE conference on, 19-22 September 2010.*
- Omachi, M., & Omachi, S. (2009). Traffic light detection with color and edge information. *Paper presented at the computer science and information technology, 2009. ICCSIT 2009. 2nd IEEE international conference on, 8-11 Aug. 2009.*
- Omachi, M., & Omachi, S. (2010). *Detection of traffic light using structural information* Beijing.
- Philipsen, M. P., Jensen, M. B., Mogelmose, A., Moeslund, T. B., & Trivedi, M. M. (2015). Traffic light detection: A learning algorithm and evaluations on challenging dataset. *Paper presented at the intelligent transportation systems (ITSC), 2015 IEEE 18th international conference on, 15-18 September 2015.*
- Philipsen, M. P., Jensen, M. B., Trivedi, M. M., Mogelmose, A., & Moeslund, T. B. (2015). Ongoing work on traffic lights: Detection and evaluation. *Paper presented at the advanced video and signal based surveillance (AVSS), 2015 12th IEEE international conference on, 25-28 August 2015.*
- Shi, Z., Zou, Z., & Zhang, C. (2015). Real-time traffic light detection with adaptive background suppression filter. *Intelligent Transportation Systems, IEEE Transactions on, PP(99), 1-11.* doi:10.1109/TITS.2015.2481459.
- Sooksatra, S., & Kondo, T. (2014). Red traffic light detection using fast radial symmetry transform. *Paper presented at the electrical engineering/electronics, computer, telecommunications and information technology (ECTI-CON), 2014 11th international conference on.*
- Suhr, J. K., Jang, J., Min, D., & Jung, H. G. (2017). Sensor fusion-based low-cost vehicle localization system for complex urban environments. *IEEE Transactions on Intelligent Transportation Systems, 18(5), 1078–1086.* doi:10.1109/TITS.2016.2595618.
- Tae-Hyun, H., In-Hak, J., & Seong-Ik, C. (2006). Detection of traffic lights for vision-based car navigation system. In *1st Pacific rim symposium on image and video technology, PSIVT 2006: Vol. 4319 LNCS, Hsinchu* (pp. 682–691).
- Trehard, G., Pollard, E., Bradai, B., & Nashashibi, F. (2014). Tracking both pose and status of a traffic light via an interacting multiple model filter. *Paper presented at the information fusion (FUSION), 2014 17th international conference on.*
- United Nations Economic Commission for Europe, G., Switzerland. (2006). *Vienna convention on road signs and signals* Retrieved from [www.unece.org/trans/conventn/signalse.pdf](http://www.unece.org/trans/conventn/signalse.pdf).
- Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International Journal of Computer Vision, 57(2), 137–154.* doi:10.1023/B:VISL.0000013087.49260.fb.
- Xuemei, L., Yanmin, Z., Minglu, L., & Qian, Z. (2012). POVA: Traffic light sensing with probe vehicles. *Paper presented at the INFOCOM, 2012 proceedings IEEE, 25-30 March 2012.*
- Yehu, S., Ozguner, U., Redmill, K., & Jilin, L. (2009). A robust video based traffic light detection algorithm for intelligent vehicles. *Paper presented at the intelligent vehicles symposium, 2009 IEEE, 3-5 June 2009.*
- Ying, J., Chen, X., Gao, P., & Xiong, Z. (2013). A new traffic light detection and recognition algorithm for electronic travel aid. *Paper presented at the intelligent control and information processing (ICICIP), 2013 fourth international conference on, 9-11 June 2013.*
- Younes, M. B., & Boukerche, A. (2016). Intelligent traffic light controlling algorithms using vehicular networks. *IEEE Transactions on Vehicular Technology, 65(8), 5887–5899.*
- Yu, C., Huang, C., & Lang, Y. (2010). *Traffic light detection during day and night conditions by a camera* Beijing.
- Zhang, L. T., Fu, M. Y., Yang, Y., & Wang, M. L. (2014). A framework of traffic lights detection, tracking and recognition based on motion models. *Paper presented at the 2014 17th IEEE international conference on intelligent transportation systems, ITSC 2014.*
- Ziegler, J., Bender, P., Schreiber, M., Lategahn, H., Strauss, T., Stiller, C., & Zeeb, E. (2014). Making bertha drive: An autonomous journey on a historic route. *Intelligent Transportation Systems Magazine, IEEE, 6(2), 8–20.* doi:10.1109/MITS.2014.2306552.
- Zixing, C., Yi, L., & Mingqin, G. (2012). Real-time recognition system of traffic light in urban environment. *Paper presented at the computational intelligence for security and defence applications (CISDA), 2012 IEEE symposium on, 11-13 July 2012.*
- Zong, W., & Chen, Q. (2014). Traffic light detection based on multi-feature segmentation and online selecting scheme. *Paper presented at the conference proceedings - IEEE international conference on systems, man and cybernetics.*