Electronic Preprint for Journal of Information Processing Vol.24 No.2

Regular Paper

Traffic Light Detection Considering Color Saturation Using In-Vehicle Stereo Camera

Hiroki Moizumi[1,†1,a)] Yoshihiro Sugaya[1,b)] Masako Omachi[2,c)] Shinichiro Omachi[1,d)]

Received: June 29, 2015, Accepted: December 7, 2015

Abstract: One of the major causes of traffic accidents according to the statistical report on traffic accidents in Japan is the disregard of traffic lights by drivers. It would be useful if driving support systems could detect and recognize traffic lights and give appropriate information to drivers. Although many studies on intelligent transportation systems have been conducted, the detection of traffic lights using images remains a difficult problem. This is because traffic lights are very small as compared to other objects and there are many objects similar to traffic lights in the road environment. In addition, the pixel colors of traffic lights are easily over-saturated, which renders traffic light detection using color information difficult. The rapid deployment of the new LED traffic lights has led to a new problem. Since LED lights blink at high frequency, if they are captured by a digital video camera, there are frames in which all the traffic lights appear to be turned off. It is impossible to detect traffic lights in these frames by searching the ordinary color of traffic lights. In this paper, we focus on the stable detection of traffic lights, even when they are blinking or when their colors are over-saturated. A method for detecting candidate traffic lights utilizing intensity information together with color information is proposed for handling over-saturated pixels. To exclude candidates that are not traffic lights efficiently, the sizes of the detected candidates are calculated using a stereo image. In addition, we introduce tracking with a Kalman filter to avoid incorrect detection and achieve stable detection of blinking lights. The experimental results using video sequences taken by an in-vehicle stereo camera verify the efficacy of the proposed approaches.

Keywords: intelligent transportation systems, traffic light detection, stereo camera, image processing

1. Introduction

Various kinds of driving support systems have been developed and the driving environment has been improved by the recent research and development of intelligent transportation systems. For implementation in these systems, many techniques for video processing have been proposed [1]. Applications that use images or videos include vehicle detection [2], [3], license plate recognition [4], traffic lane detection [5], road sign recognition [6], [7], pedestrian detection [8], driver fatigue detection [9], and so on.

According to the statistical report on traffic accidents, drivers who disregard traffic lights cause many of the traffic accidents in Japan [10]. Hence, a driving support system that detects traffic lights and gives appropriate information to drivers would be useful. However, as compared to the applications mentioned above not many studies exist that address traffic light detection schemes that use images or videos. It is difficult to detect a traffic light by using typical pattern matching techniques, such as simple template matching [11], color indexing [12], or key-point detection [13], since they are very small as compared to other objects and objects similar to the traffic lights abound in the road environment. Nevertheless, some attempts to detect and recognize traffic lights have been made. Lindner et al. proposed a system for traffic light detection that uses a color- and shape-based detector [14]. They introduced GPS, digital maps, and a second camera that allows stereo vision into the system to enhance its performance. Kimura et al. proposed a method for detecting traffic lights that uses a color histogram [15] and Yelal et al. [16] presented a method for tracking the color of traffic lights with contour tracking. Joo et al. proposed a method using rotated principal component analysis [17]. Park and Jeong's traffic light detection system uses color clustering and a circularity check [18]. Charette and Nashashibi used a technique for spot light detection and adaptive template matching [19]. Yu et al. introduced a color-based detection and rule-based verification of traffic lights [20]. Omachi and Omachi proposed a method for detecting a traffic light with the modified Hough transform using color [21]. They improved the detection accuracy by introducing the structural information of traffic lights [22]. Gong et al. proposed a method for detecting the candidate region of a traffic light using threshold segmentation and tracking the target by using the continuously adaptive mean shift method [23]. Premachandra et al. utilized the blinking characteristic of LEDs and proposed a method for detecting LED traffic lights using a high-speed camera [24]. Li et al. used a top-hat transform to detect a bright region [25]. Gómez et al. proposed a method for utilizing a hidden Markov model for state estimation to provide robust recognition [26].

Most of these existing methods attempt to detect traffic lights

1 Graduate School of Engineering, Tohoku University, Sendai, Miyagi 980–8579, Japan
2 Sendai National Collage of Technology, Natori, Miyagi 981–1239, Japan
†1 Presently with Hitachi Automotive Systems, Ltd.
a) moizumi@iic.ecei.tohoku.ac.jp
b) sugaya@ecei.tohoku.ac.jp
c) omachi@sendai-nct.ac.jp
d) machi@ecei.tohoku.ac.jp

by searching pixels of the typical colors of traffic lights, i.e., red, yellow, and green. In these methods, traffic lights are detected through combining the color with other information, such as shape or position. However, our preliminary experiments revealed that the colors of traffic lights captured by a digital video camera are easily over-saturated in various conditions. In this case, the color is very different from the ordinary one. An example is shown in **Fig. 1** (a), where it can be seen that most of the pixels of the yellow traffic light are over-saturated and changed to white.

According to the above consideration, in this paper, we propose a method for detecting traffic lights even if the pixel colors of the traffic lights are over-saturated. We utilize the intensity information together with the color and detect over-saturated and unsaturated regions simultaneously. These regions are combined and regarded as candidate traffic light regions. However, many regions that are not traffic lights are also detected by this type of approach. To exclude candidates other than traffic lights efficiently, we introduce an in-vehicle stereo camera to calculate the size of the candidate regions using the acquired stereo images. In addition, tracking with a Kalman filter is introduced for two reasons. One is to avoid incorrect detection of objects other than traffic lights. In general, traffic lights are detected stably in successive frames, and other objects tend to be detected only in one frame. Therefore, only those objects that can be tracked continuously over several frames are probably traffic lights. The second reason is to stably detect LED traffic lights. There are two types of traffic lights: ordinary light bulbs and LEDs. In the case of LED traffic lights, since the LED lights blink at a high frequency, there are frames in which it appears that all the traffic lights are turned off. In that case, color-based detection fails. However, tracking results can complement the frame in which the light appears to be turned off. An example of a blinking light is shown in Fig. 1 (b). These two images were extracted from successive frames.

Using a high dynamic range camera or setting the camera exposure at a low level may be an alternative solution for handling the over-saturation. In addition, the introduction of a high-speed camera in the detection system may be effective for detecting LED traffic lights. However, such a camera is very expensive, and setting the camera exposure restricts the use of other applications that utilize an in-vehicle camera. As mentioned, many driving support systems using images have been developed and some have already been put into practical use. Ideally, it is desirable that our proposed algorithm can coexist with other applications within an existing system. Therefore, we focus on developing an algorithm suitable for a general-purpose inexpensive camera using image processing techniques.

Two types of experiments were performed. The experimental results show the effectiveness of the method and tracking scheme for stable traffic light detection when over-saturated pixels occur.

(a) Over-saturated pixels

(b) Blinking light

Fig. 1 Examples of traffic lights.

2. Proposed Method

Figure 2 shows the flow of the proposed method. First, a stereo image is acquired by a stereo camera, and distortion correction and paralleled transformation are applied using camera calibration parameters obtained in advance. Then, traffic light candidate regions are detected. The size of each region is compared to the standard size of a traffic light, and reliable candidates are selected. Each candidate is tracked with the Kalman filter to avoid an incorrect detection, and the recognition result is issued according to the tracking result.

In the proposed method, the stopping distance of a vehicle is considered when determining the final results. When the vehicle is moving at 60 km/h, this distance is 37 m if the road is dry and 41 m if the road is wet. Therefore, the result of traffic light detection obtained at a distance from the traffic light that is greater than these distances is useful. Therefore, the traffic light candidate that is 40 m to 60 m from the camera is selected as the output.

2.1 Candidate Region Detection

Candidate regions of traffic lights are detected using color histograms [15] that represent the colors of traffic lights. To separate intensity and color information, the RGB values are transformed into L*a*b* components [27][*1]. The L* component is used for detecting over-saturated pixels and the a* and b* components are used for detecting pixels of the typical (unsaturated) colors of traffic lights, i.e., red, yellow, and green. For each color, using multiple images of traffic lights of that color, a histogram for the a* and b* components is constructed in advance by dividing the

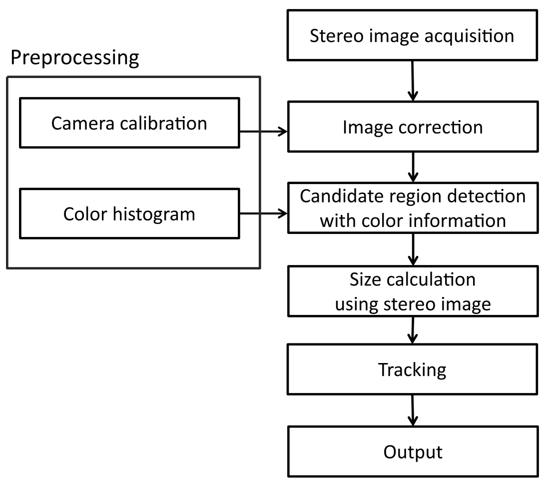


Fig. 2 Flow of the proposed method.

1 The values of L, a*, and b* are in the range 0 to 255.

a*-b* plane into 16×16 bins. The value of each bin is normalized from 0 to 1 by dividing by the mode.

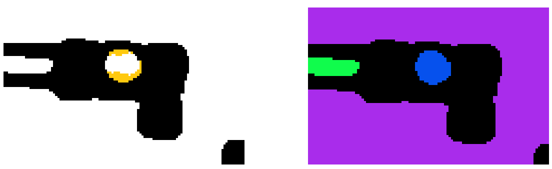
First, each pixel is tested as to whether it has one of the typical colors of traffic lights. When a pixel has a* and b* components, the histogram value of which is larger than a threshold θ_{ab}, it is judged to be a pixel of a traffic light of that color. **Figure 3** displays a result of pixel detection. Figure 3 (a) is the original image, and Fig. 3 (b) is the result of detecting pixels, the colors of which are one of the traffic light colors. Next, over-saturated pixels are detected. As mentioned above, the color of the pixels of the traffic lights are easily over-saturated and the color changes drastically, since the intensity of a traffic light is much greater than that of its surrounding regions. In the proposed method, we detect pixels having an L* value larger than a threshold θ_L to detect a traffic light with over-saturated pixels.

Then, adjacent pixels of the same color are unified into one region. At this time, if over-saturated pixels adjoin unsaturated pixels, they are unified into one region, the color of which is that of the unsaturated pixel. Hence, even if most of the pixels are over-saturated, a region can be detected as a traffic light candidate. An example is shown in **Fig. 4**. Figure 4 (a) displays the pixels detected in Fig. 1 (a). Yellow pixels are detected as yellow light, and white pixels are detected as over-saturated ones. Figure 4 (b) represents the result of labeling. The blocks of blue, green, and purple pixels are regarded as connected regions. This figure shows that the region of the yellow light is considered one connected region, even if many pixels are over-saturated.

To reduce the number of the detected regions, the number of detected pixels and the aspect ratio of the region including the pixels are used. A region that consists of fewer than $n_{\min}$ or more than $n_{\max}$ pixels, or the aspect ratio of which is more than r_a is excluded. Furthermore, candidate regions are selected by considering the structure of traffic lights. A traffic light device consists of three lights. In the traffic light systems used in Japan, when one light is turned on, the other two are turned off. Therefore, two regions next to the detected region are examined. First, the luminosity histograms of the target region and two neighboring regions are created. If the difference between the modes of the histogram of the target region and a neighboring region is larger than a threshold θ_n, the neighboring region is considered a light that is turned off. If both of the two neighboring regions are considered turned off lights, the target region is retained. A region that does not satisfy this condition is discarded.

(a) Input image (b) Detected pixels

Fig. 3 Result of detecting pixels.



(a) Detected pixels (b) Result of labeling

Fig. 4 Labeling.

In order to avoid incorrect detection, we calculate the actual size of the detected region using a stereo image and determine whether it is a traffic light or not by comparing its size with the standard traffic light sizes. In Japan, there are three standard traffic light diameters: Small (250 mm), medium (300 mm), and large (450 mm). In the proposed method, the standard size is regarded as 300 mm, which is the most common of the three standard sizes. The permissible error d_p mm is determined in advance, and the regions the sizes of which fall within the range $(300 - d_p)$ mm to $(300 + d_p)$ mm are selected as the candidate regions.

2.2 Tracking of Traffic Lights

In the process of candidate region detection, objects that are not traffic lights are sometimes incorrectly detected. In addition, in the case of LED traffic lights, sometimes no candidate region is detected, even if an image includes traffic lights, because the LED light is blinking. These types of failure can be avoided by tracking a traffic light. In the proposed method, a Kalman filter is assigned for each candidate region of a traffic light and tracked.

First, for each traffic light candidate region obtained in the first frame, a Kalman filter is created and tracking is initiated. For each frame, in order to obtain observation values for each Kalman filter, matching between Kalman filters and candidate regions is performed. After the matching process is completed, prediction and correction processes are performed using the state values at the previous time [28]. For each candidate region to which a Kalman filter has not been assigned, a new Kalman filter is created. On the other hand, if a Kalman filter has no corresponding candidate region, there exists no observation value. In this case, prediction values are used as observation values for these Kalman filters so that the tracking is not interrupted. When a Kalman filter has no corresponding candidate region continuously for three frames, the tracking is stopped and the Kalman filter vanishes.

The Kalman filter estimates the state by minimizing the error of the current value and the value estimated from the previous state [28]. Let the state at time k be $\boldsymbol{X}_k$. In general, the state is represented as

$$\boldsymbol{X}_k = A\boldsymbol{X}_{k-1} + B\boldsymbol{u}_{k-1} + \boldsymbol{w}_{k-1}, \tag{1}$$

where A is a matrix that represents the state transition from $k-1$ to k, B is a matrix that relates the control input $\boldsymbol{u}_{k-1}$ to $\boldsymbol{X}_k$, and $\boldsymbol{w}_{k-1}$ represents the process noise. In the proposed method, the state $\boldsymbol{X}_k$ is represented as

$$\boldsymbol{X}_k = [\begin{array}{cccc} x & y & \dot{x} & \dot{y} \end{array}]^T, \tag{2}$$

where (x, y) is the top left position of a candidate region and $(\dot{x}, \dot{y})$ is its velocity. Since there is no control input, Eq. (1) can be written as

$$\boldsymbol{X}_k = A\boldsymbol{X}_{k-1} + \boldsymbol{w}_{k-1}. \tag{3}$$

The matrix A is defined as

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{4}$$

since traffic lights actually do not move, and each candidate region moves approximately vertically in the frame according to the camera movement.

The observation value $\boldsymbol{Z}_k$ at time k is represented by the state $\boldsymbol{X}_k$ as

$$\boldsymbol{Z}_k = H\boldsymbol{X}_k + \boldsymbol{v}_k, \tag{5}$$

where H is a matrix that represents an observation model and $\boldsymbol{v}_k$ is measurement noise. In the proposed method, H is determined as

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \tag{6}$$

Process noise $\boldsymbol{w}_k$ and measurement noise $\boldsymbol{v}_k$ follow the Gaussian distribution, the covariance matrices of which are Q and R, respectively.

The state vector is updated by a predictor-corrector algorithm [28]. The prediction process is represented by

$$\hat{\boldsymbol{X}}_k^- = A\hat{\boldsymbol{X}}_{k-1}, \tag{7}$$

$$P_k^- = AP_{k-1}A^T + Q, \tag{8}$$

where $\hat{\boldsymbol{X}}_k^-$ is the *a priori* state estimate at time k and $\hat{\boldsymbol{X}}_{k-1}$ is the *a posteriori* state estimate at time $k-1$. P_k^- is the *a priori* estimate error covariance matrix at time k. The correction process is represented by

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1}, \tag{9}$$

$$\hat{\boldsymbol{X}}_k = \hat{\boldsymbol{X}}_k^- + K_k(\boldsymbol{Z}_k - H\hat{\boldsymbol{X}}_k^-), \tag{10}$$

$$P_k = (I - K_k H)P_k^-, \tag{11}$$

where K_k is the optimal Kalman gain, $\boldsymbol{Z}_k$ is the observation value, and P_k is the *a posteriori* error covariance matrix.

The observation values, which are necessary for updating the state values of the Kalman filters, are determined using the results of the candidate region detection. Let

$$F = \{f_1, f_2, ..., f_k\}$$

be a set of objects that the Kalman filters are tracking, where k is the number of objects. Each object f_i $(1 \leq i \leq k)$ has the following parameters: the position (x, y) and the velocity $(\dot{x}, \dot{y})$ in a frame, the width of the traffic light in a frame, the actual distance z from the camera, and the color of the traffic light. Let

$$M = \{m_1, m_2, ..., m_l\}$$

be a set of traffic light candidates, where l is the number of traffic light candidates. Each candidate m_j $(1 \leq j \leq l)$ has the following parameters: the position (x, y) in a frame, the width of the traffic light in a frame, the actual distance z from the camera, and the color of the traffic light. For each candidate m_j, the corresponding object f_i is assigned by the criterion

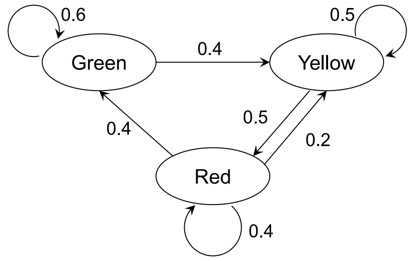


Fig. 5 Weights for color transition.

Table 1 Correction term considering color transition.

$W(m_j \mid f_i)$	m_j	f_i
0	green	green
0	yellow	yellow
0	red	red
w	yellow	green
w	red	yellow
$-2w$	green	red
$-w$	yellow	red

$$Cost(f_i, m_j) = \frac{\alpha D(f_i, m_j) + \beta Z(f_i, m_j)}{C(m_j \mid f_i)}. \tag{12}$$

In Eq. (12), $C(m_j|f_i)$ is the cost function that represents the color transition of traffic lights displayed in **Fig. 5**. In addition to the ordinary color transition, there is a path from red to yellow because sometimes a red light is recognized as a yellow one.

$$D(f_i, m_j) = \sqrt{(x_{fi} + W(m_j|f_i) - x_{mj})^2 + (y_{fi} - y_{mj})^2} \tag{13}$$

represents the distance between f_i and m_j in a frame, where (x_{fi}, y_{fi}) and (x_{mi}, y_{mi}) are the positions of f_i and m_j in a frame, respectively. $W(m_j|f_i)$ is the correction term considering the color transition from the color of f_i to the color of m_j, as shown in **Table 1**. In the table, w is the width of object f_i in a frame. For example, if the green light turns off and the yellow light turns on, the position of the light moves toward the right. $Z(f_i, m_j)$ is the difference between the actual distances [meters] from the camera, and α and β are the weighting factors. In the experiment, $\alpha = \beta = 1$ was used.

Note that Eq. (13) takes into account only horizontally arranged traffic lights. However, vertically arranged traffic lights also exist. In order to take these into account, Eq. (14), which has the vertical correction term $H(m_j|f_i)$, should be used instead of Eq. (13).

$$D(f_i, m_j) = \sqrt{(x_{fi} - x_{mj})^2 + (y_{fi} + H(m_j|f_i) - y_{mj})^2}. \tag{14}$$

3. Experiments

To verify the effect of the proposed method, we conducted experiments. Stereo images were acquired by an in-vehicle stereo camera. We used an Inrevium TD-BD-SCAMv2 stereo camera for the experiments. The baseline between the lenses is 50 cm. It can acquire a color stereo image of 640×480 pixels at 15 frames per second (fps). The proposed method was tested using Windows XP on an Intel Core i5-540M (2.53 GHz) with 2 GB of memory. The parameters defined in Section 2.1 were determined experimentally, as shown in **Table 2**.

Table 2 Parameters used in the experiment.

Parameter	Value
θ_{ab}	0.1
θ_L	236
$n_{\min}$	50
$n_{\max}$	1,200
r_a	1.1
θ_n	65
d_p	200

Table 3 Results of candidate region detection.

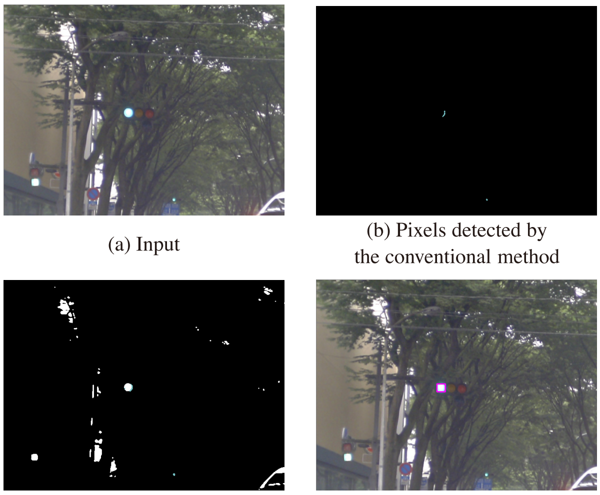
Type	Number of traffic lights	Method	Detected
Unsaturated	180	Proposed	179
		Conventional	179
Over-saturated	180	Proposed	174
		Conventional	34

3.1 Accuracy of Detecting Traffic Light Candidate Regions

The first experiment was performed to verify the accuracy of the candidate region detection. To show the effect of detecting pixels, the color of which is one of the traffic light colors, and over-saturated pixels simultaneously, the proposed candidate region detection method is compared to a method that does not use over-saturated pixels, as is common in most of the existing methods. In the experiment, six videos that include traffic lights with over-saturated pixels and six videos that include no over-saturated pixels were used. Thirty frames were used for each traffic light, and 360 frames were used in total.

Table 3 shows the number of traffic lights that were correctly detected (true positive) by the proposed method and the method that does not use over-saturated pixels (conventional method). The results show that both the methods can detect the traffic lights with unsaturated color. However, the accuracy of the proposed method for over-saturated color is much higher than that of the conventional method. These results show the effectiveness of the proposed approach. Note that it is important to detect as many traffic lights as possible in the candidate region detection step, in spite of the increase in false positives. As mentioned in Section 2.2, incorrectly detected objects are discarded by using tracking.

Figure 6 shows an example of a traffic light that was detected by the proposed method and was not detected by the conventional method. Figure 6 (a) is an input image and Fig. 6 (b) displays the detected pixels that have the color of a green traffic light. In this case, since the number of detected pixels was very small, they were not considered to be a traffic light. Figure 6 (c) shows the pixels detected by the proposed method. In addition to the green pixels, over-saturated pixels were also detected. As shown in Fig. 6 (d), the traffic light was correctly detected by the proposed method (surrounded by a square).



(a) Input

(b) Pixels detected by the conventional method

(c) Pixels detected by the proposed method

(d) Detected traffic light

Fig. 6 Traffic light detected by the proposed method.

Table 4 Accuracy of the proposed method.

Weather	Type of light	Precision	Recall	F-measure
Sunny	Light bulb	0.92	0.90	0.91
	LED	0.86	0.77	0.81
Cloudy	Light bulb	0.91	0.89	0.90
	LED	0.88	0.88	0.88
Average		0.89	0.86	0.88

Table 5 Accuracy of the existing method.

Weather	Type of light	Precision	Recall	F-measure
Sunny	Light bulb	0.74	0.47	0.58
	LED	0.60	0.22	0.32
Cloudy	Light bulb	0.67	0.39	0.49
	LED	0.74	0.28	0.41
Average		0.69	0.33	0.45

3.2 Total Performance

The second experiment was performed to verify the traffic light detection performance of the proposed method. Video sequences were acquired from an in-vehicle stereo camera in a driving car on sunny and cloudy days. We used video sequences of 13,721 frames in total, of which 3,527 frames included traffic lights. In our evaluation, we used traffic lights having a distance from the camera greater than 40 m and smaller than 60 m. Since our method uses L*a*b* color space, the method proposed by Yelal et al. [16], which also uses L*a*b* color space with contour tracking, was compared with the proposed method as an existing method.

The efficiency was evaluated by precision, recall, and F-measure, averaged for all the frames. Let the numbers of correctly detected, incorrectly detected, and undetected traffic lights be n_1, n_2, and n_3, respectively. If the detected region overlaps 30% of the region of the actual traffic light, the detection results are considered correct. The values of precision and recall are calculated as

$$Precision = \frac{n_1}{n_1 + n_2}, \tag{15}$$

$$Recall = \frac{n_1}{n_1 + n_3} \tag{16}$$

The F-measure is the harmonic mean of these values. The experimental results are shown in **Table 4** and **Table 5**. It was shown that the accuracy of the proposed method was much higher than that of the existing method.

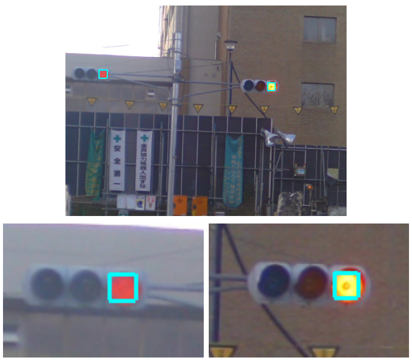
Examples of traffic lights correctly detected by the proposed method are shown in **Fig. 7**. An expanded image around each detected region is shown below the full image. The weather, color, type, and detected size of each traffic light are also shown. Some

(a) Cloudy, green bulb, 10 × 11 (b) Cloudy, yellow bulb, 13 × 14

Fig. 7 Results of detecting over-saturated traffic lights using the proposed method.

(a) Sunny, green LED, 17 × 18 (left) and green bulb, 8 × 8 (right)

(b) Sunny, yellow LED, 10 × 10 (left) and yellow bulb, 10 × 9 (right)



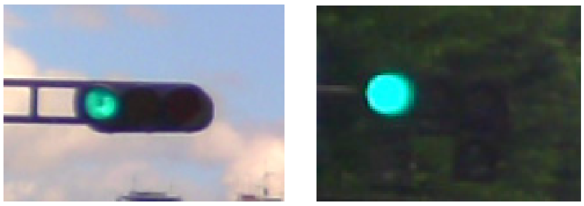
(c) Sunny, red LED, 20 × 20 (left) and red bulb, 18 × 19 (right)

Fig. 8 Results for detecting two types of traffic lights simultaneously using the proposed method.

(a) Cloudy, green LED, 8 × 8 (b) Cloudy, yellow LED, 11 × 11

(c) Cloudy, red LED, 12 × 12 (d) Cloudy, red bulb, 8 × 8

Fig. 9 Results of traffic light detection in various conditions using the proposed method.



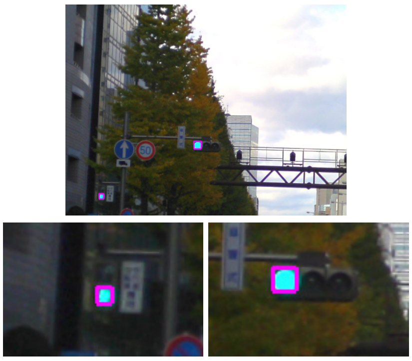
(a) Correctly detected (b) Not detected

Fig. 10 Detection by the existing method.

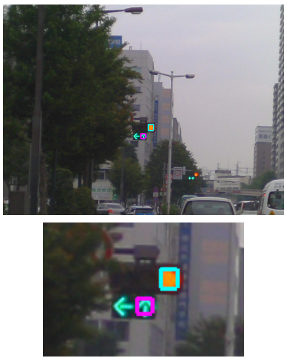
of the pixels of traffic lights were over-saturated in these examples. However, the proposed method correctly detected them by connecting the over-saturated pixels with unsaturated green pixels.

Figure 8 displays the results for images, including two sets of traffic lights. In every case, the traffic lights were correctly detected. In addition, it should be noted that in every case, one set of traffic lights consists of LEDs and the second set of ordinary light bulbs. As these figures show, the two types of traffic light were correctly detected simultaneously by the proposed method, which could not be achieved by the method designed only for LEDs. **Figure 9** displays the results of traffic light detection for combinations of weather, color, and type that are not included in Fig. 7 and Fig. 8. In every case, traffic lights were correctly detected.

Figure 10 displays the detection results produced by the existing method. The existing method correctly detected a traffic light if the contours were clear, as shown in Fig. 10 (a), since contour tracking was introduced. However, if the color of the background was similar to that of the traffic light device, as shown in Fig. 10 (b), the existing method failed to detect traffic lights, since clear contours could not be detected. In addition, in the existing method, the detection accuracy for LEDs was lower than that for ordinary light bulbs. This is because LED lights blink and it was difficult to detect the light in the frame where the LED light was turned off. On the other hand, the proposed method could interpo-

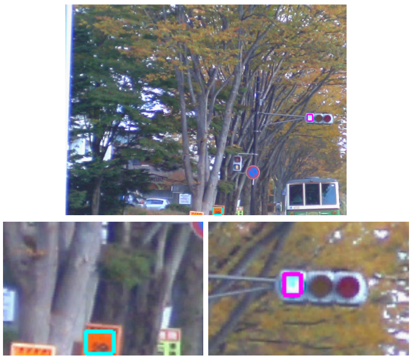


(a) Pedestrian traffic light (left)



(b) Arrow-shaped traffic light

(c) Red leaves (right)



(d) Signboard (left)

Fig. 11 Examples of incorrect detection.

late the detection by tracking, even if the traffic light could not be detected in one frame where the LED light was turned off. These results show the effect of tracking the traffic light itself instead of the contour.

Although its accuracy was high, the proposed method incorrectly detected objects other than traffic lights in some cases. **Figure 11** shows examples of incorrect detection. Figure 11 (a) shows a case where a pedestrian traffic light was detected. In this case, the color of the pedestrian traffic light was the same as that of the traffic light, and the neighboring regions were dark. Figure 11 (b) is a similar case where an arrow-shaped traffic light was detected. Since these objects were located near a traffic light, they appeared in the same frame. In order to distinguish them, it would be useful to take account of the shape of the object or its height from the road. In Fig. 11 (c) and Fig. 11 (d), red leaves and a red signboard were incorrectly detected. However, in every case, the traffic light was also detected correctly.

(a) Whole frame (b) Traffic light

Fig. 12 Motion blur.

Table 6 Processing time for one frame.

Process	Time [ms]
Image acquisition	2.00
Image correction	7.84
Candidate region detection	27.39
Stereo matching	1.39
Tracking	3.37
Image output	13.52
Total	55.51

A second reason for failure was the occurrence of motion blur. **Figure 12** displays an example. Motion blur occurs mainly when the car moves irregularly because of a distortion in the road. As shown in Fig. 12 (b), the shape of the traffic light changes drastically according to the blur, and it is difficult to detect traffic light regions or to track objects. Note that the Kalman filter can predict the position of the object to some extent, even if motion blur occurs. Therefore, the reason for failure is that no corresponding region existed continuously for three frames. One solution for this problem may be to increase the number of frames. However, an excessive number of frames may cause another misdetection. Therefore, some kind of image processing technique should be introduced to avoid this issue.

Finally, the averaged processing time of the proposed method for one frame is shown in **Table 6**. Each row shows the processing time for each process when it is run on a single-core CPU. The most time-consuming process was the detection of the traffic light candidate regions. In particular, when there were many over-saturated pixels, it took a longer time. Tracking also requires a considerable amount of time. When there were many candidate regions, it took a longer time since it was necessary to perform matching between Kalman filters and candidate regions. The total processing time for one frame was 55.51 ms, which is about 18 fps. The processing time of the existing method [16] was 58.81 ms. The proposed method yielded more accurate results than the existing method in almost the same processing time.

4. Conclusions

In this paper, we proposed a method for detecting traffic lights using an in-vehicle stereo camera. Traffic light detection using image recognition techniques remains a difficult problem because of the existence of similar objects in the environment and the over-saturation of pixels. In order to overcome these problems, we proposed a method for detecting connected regions that include over-saturated and unsaturated pixels. In this method, in order to avoid incorrect detection, the detection result is verified using the object's size by utilizing a stereo image. In addition, to stably detect blinking LED traffic lights, we introduced tracking with a Kalman filter.

Experiments were performed to verify the effect of the proposed method in two aspects. The effectiveness of the detection of traffic light candidate regions considering both over-saturated and unsaturated pixels was verified by the results of the first experiment. The results of the second experiment confirmed the effect of traffic light tracking and the total performance of the proposed method.

Although the performance of the proposed method was much higher than that of the existing method, there is room for improvement. It is important in future work to decrease the number of detections of objects other than traffic lights by utilizing the shape or height information. In addition, it is necessary to examine possible techniques for handling the detection and tracking failures caused by motion blur. The development of a driving support system that has a traffic light detection function is also an important future work.

Acknowledgments The authors would like to thank Dr. Tomo Miyazaki of Tohoku University for his fruitful comments. A part of this research was supported by JSPS KAKENHI Grant Numbers 22500164 and 25330216.

References

[1] Kastrinaki, V., Zervakis, M. and Kalaitzakis, K.: A survey of video processing techniques for traffic applications, *Image and Vision Computing*, Vol.21, No.4, pp.359–381 (April 2003).

[2] Sun, Z., Bebis, G. and Miller, R.: On-road vehicle detection: A review, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.28, No.5, pp.694–711 (May 2006).

[3] Sivaraman, S. and Trivedi, M.M.: Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking, and behavior analysis, *IEEE Trans. Intelligent Transportation Systems*, Vol.14, No.4, pp.1773–1795 (Dec. 2013).

[4] Anagnostopoulos, C.-N.E., Anagnostopoulos, I.E., Psoroulas, I.D., Loumos, V. and Kayafas, E.: License plate recognition from still images and video sequences: A survey, *IEEE Trans. Intelligent Transportation Systems*, Vol.9, No.3, pp.377–391 (Sep. 2008).

[5] Cheng, H.-Y., Jeng, B.-S., Tseng, P.-T. and Fan, K.-C.: Lane detection with moving vehicles in the traffic scenes, *IEEE Trans. Intelligent Transportation Systems*, Vol.7, No.4, pp.571–582 (Dec. 2006).

[6] Maldonado-Bascón, S., Lafuente-Arroyo, S., Gil-Jiménez, P., Gómez-Moreno, H. and López-Ferreras, F.: Road-sign detection and recognition based on support vector machines, *IEEE Trans. Intelligent Transportation Systems*, Vol.8, No.2, pp.264–278 (June 2007).

[7] Møgelmose, A., Trivedi, M.M. and Moeslund, T.B.: Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey, *IEEE Trans. Intelligent Transportation Systems*, Vol.13, No.4, pp.1484–1497 (Dec. 2012).

[8] Gerónimo, D., López, A.M., Sappa, A.D. and Graf, T.: Survey of pedestrian detection for advanced driver assistance systems, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.32, No.7, pp.1239–1258 (July 2010).

[9] Wang, Q., Yang, J., Ren, M. and Zheng, Y.: Driver fatigue detection: A survey, *Proc. 6th World Congress on Intelligent Control and Automation*, Vol.2, pp.8587–8591 (June 2006).

[10] International Association of Traffic and Safety Sciences, Statistics 2007 Road Accidents Japan (Aug. 2008).

[11] Duda, R.O. and Hart, P.E.: *Pattern Classification and Scene Analysis*, John Wiley & Sons, Inc., New York (1973).

[12] Swain, M.J. and Ballard, D.H.: Colour indexing, *International Journal of Computer Vision*, Vol.7, No.1, pp.11–32 (Nov. 1991).

[13] Lowe, D.G.: Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, Vol.60, No.2, pp.91–110 (Nov. 2004).

[14] Lindner, F., Kressel, U. and Kaelberer, S.: Robust recognition of traffic signals, *Proc. 2004 IEEE International Vehicle Symposium*, pp.49–53 (June 2004).

[15] Kimura, F., Takahashi, T., Mekada, Y., Ide, I. and Murase, H.: Recognition of traffic signals in various conditions for safety driving assistance, *Proc. 9th Meeting on Image Recognition and Understanding* (*MIRU2006*), pp.618–623 (July 2006). (in Japanese)

[16] Yelal, M.R., Sasi, S., Shaffer, G.R. and Kumar, A.K.: Color-based signal light tracking in real-time video, *Proc. IEEE International Conference on Video and Signal Based Surveillance* (Nov. 2006).

[17] Joo, S.-K., Kim, Y., Cho, S.I., Choi, K. and Lee, K.: Traffic light detection using rotated principal component analysis for video-based car navigation system, *IEICE Trans. Inf. Syst*. Vol.E91-D, No.12, pp.2884–2887 (Dec. 2008).

[18] Park, J.-H. and Jeong, C.-S.: Real-time signal light detection, *International Journal of Signal Processing, Image Processing and Pattern Recognition*, Vol.2, No.2, pp.1–10 (June 2009).

[19] de Charette, R. and Nashashibi, F.: Traffic light recognition using image processing compared to learning processes, *Proc. 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.333–338 (Oct. 2009).

[20] Yu, C., Huang, C. and Lang, Y.: Traffic light detection during day and night conditions by a camera, *Proc. 2010 IEEE 10th International Conference on Signal Processing*, pp.821–824 (Oct. 2010).

[21] Omachi, M. and Omachi, S.: Fast detection of traffic light with color and edge information, *The Journal of the Institute of Image Electronics Engineers of Japan*, Vol.38, No.5, pp.673–679 (Sep. 2009). (in Japanese)

[22] Omachi, M. and Omachi, S.: Detection of traffic light using structural information, *Proc. 2010 IEEE 10th International Conference on Signal Processing*, Vol.1, pp.809–812 (Oct. 2010).

[23] Gong, J., Jiang, Y., Xiong, G., Guan, C., Tao, G. and Chen, H.: The recognition and tracking of traffic lights based on color segmentation and CAMSHIFT for intelligent vehicles, *Proc. 2010 IEEE Intelligent Vehicles Symposium*, pp.431–435 (June 2010).

[24] Premachandra, H.C.N., Yendo, T., Tehrani, M.P., Yamazato, T., Okada, H., Fujii, T. and Tanimoto, M.: High-speed-camera image processing based LED traffic light detection for road-to-vehicle visible light communication, *Proc. 2010 IEEE Intelligent Vehicles Symposium*, pp.793–798 (June 2010).

[25] Li, Y., Cai, Z.-X., Gu, M.-Q. and Yan, Q.-Y.: Traffic lights recognition based on morphology filtering and statistical classification, *Proc. 2011 Seventh International Conference on Natural Computation* (*ICNC*), Vol.3, pp.1700–1704 (July 2011).

[26] Gómez, A.E., Alencar, F.A.R., Prado, P.V., Osório, F.S. and Wolf, D.F.: Traffic lights detection and state estimation using Hidden Markov Models, *Proc. 2014 IEEE Intelligent Vehicles Symposium*, pp.750–755 (June 2014).

[27] Gonzalez, R.C. and Woods, R.E.: *Digital Image Processing*, Second Edition, Prentice-Hall, Inc. (2002).

[28] Welch, G. and Bishop, G.: *An Introduction to the Kalman Filter*, UNC-Chapel Hill, TR 95-041 (July 2006).

Hiroki Moizumi received his B.E. and M.E. degrees from Tohoku University in 2012 and 2014, respectively. He joined Hitachi Automotive Systems, Ltd. in 2014.

Yoshihiro Sugaya received his B.E., M.E., and Doctor of Engineering degrees from Tohoku University, Sendai, Japan in 1995, 1997, and 2002, respectively. He has worked as a research associate at the Graduate School of Engineering, Tohoku University from 2000 to 2007, and since 2007, he has been an assistant professor. His research interests include the areas of pattern recognition, parallel processing, distributed computing, parallel architecture, and the smart grid. Dr. Sugaya is a member of IEICE.

Masako Omachi received her B.E., Master of Information Sciences, and Doctor of Engineering degrees from Tohoku University, Japan, in 1994, 1996, and 1999, respectively. From 1999 to 2010, she was with the Faculty of Science and Technology, Tohoku Bunka Gakuen University. Since 2010, she has been with the Advanced Course of Production System and Design Engineering, Sendai National College of Technology, where she is currently an associate professor. Her research interests include pattern recognition, character recognition, and image processing. She received the MIRU Nagao Award in 2007 and the IEICE Best Paper Award in 2012. Dr. Omachi is a member of IEICE.

Shinichiro Omachi received his B.E., M.E., and Doctor of Engineering degrees in Information Engineering from Tohoku University, Japan, in 1988, 1990, and 1993, respectively. He worked as a research associate at the Education Center for Information Processing at Tohoku University from 1993 to 1996. Since 1996, he has been with the Graduate School of Engineering at Tohoku University, where he is currently a professor. From 2000 to 2001, he was a visiting associate professor at Brown University. His research interests include pattern recognition, computer vision, image processing, image coding, and parallel processing. He received the MIRU Nagao Award in 2007, IAPR/ICDAR Best Paper Award in 2007, Best Paper Method Award of the 33rd Annual Conference of the GfKl in 2010, ICFHR Best Paper Award in 2010, and IEICE Best Paper Award in 2012. Dr. Omachi is a member of the IEEE, IEICE, among others.