Detection of Damaged Road Paints of Crosswalks by Focusing on Multi-layered Features

Takuma Ito 1) Kyoichi Tohriyama 2) Minoru Kamata 1)

1) The University of Tokyo, Institute of Gerontology
5-1-5 Kashiwanoha, Kashiwa, Chiba, 277-8589, Japan (E-mail: ito@iog.u-tokyo.ac.jp)
2) Toyota Motor Corporation
1-4-18 Koraku, Bunkyo-ku, Tokyo, 112-8701, Japan

Received on August 29, 2019

ABSTRACT: In this study, we aim to develop a detection system of damaged crosswalks as a basic component of a digital map localization system. In rural areas, because the road paints of crosswalks are sometimes damaged, the features on which the existing methods focus for detecting them, such as rectangular shapes with side edges, are not clear. Thus, we focus on multi-layered faint features: existence of the white-band bottom shape, distribution of the white-band bottoms, and shape of the white-band candidates. Through an experiment on public roads, we confirm the practical performance of the proposed system.

KEY WORDS: safety, intelligent vehicle, road environment recognition / damaged road mark, template matching [C1]

1. Introduction

Crosswalk detection is one of the beneficial clues for intelligent vehicles to understand the driving environment. As one of its potential applications, existing studies (1)(2)(3) have proposed the use of detected crosswalks as a supplemental information for preventing collisions with pedestrians near crosswalks. Similarly, detection of crosswalks as landmarks is useful for localization on a digital map. To these ends, the existing studies have proposed various methods for detecting crosswalks. From the perspective of the sensors for detection, existing methods are roughly classified into two approaches: methods using cameras and methods using light detection and ranging (LIDAR). As an example of the camera-based method, Suzuki et al. (1) proposed a method that focused on the feature called “cross ratio,” which is the constant rate of width of the white band to that of the blank space. Haselhoff et al. (2) focused on the alignment of the repeated rectangular shapes using an adaptive Haar-like filter. Sichelschmidt et al. (3) proposed a method that analyzed the horizontally repeated edges of the crosswalk band by fast Fourier transformation. Foucher et al. (4) identified a crosswalk from other road marks by focusing on the distribution of the repeated road-mark centroids. Zhai et al. (5) proposed a method that detected bands of crosswalks by elliptical fitting and determines them as a crosswalk by focusing on their distributions. On the contrary, as the example of a LIDAR-based method, Borcs et al. (6) proposed a method that detected a crosswalk by analyzing the intensity values of the LIDAR and its distributions. Hernández et al. (7) proposed a method that identified a crosswalk by focusing on the frequency of the reflected intensity. Although the existing studies have proposed various approaches to detect crosswalks, their methods have basically focused on two crosswalk features: rectangular shape of the white band with clear side edges, and regularly repeated distribution of the white bands. In crosswalks that are well maintained, as shown in Fig. 1, these features are effective for their detection. However, road marks including crosswalks are sometimes not well maintained on actual public roads, especially in rural areas, as shown in Fig. 2. Regarding the degree of damage of road paints, the Japan Contractors Association of Traffic Signs and Lane Markings (JCASM) (8) discussed a visual evaluation rank. They categorized the damage status into five ranks: rank 5 indicated the best and rank 1 indicated the worst. Based on their ranks, the status of the crosswalk in Fig. 1 belongs to either rank 5 or 4, while the status of the crosswalk in Fig. 2 belongs to either rank 2 or 1. In the abovementioned existing studies, the status of the target crosswalks could be categorized into either rank 5 or rank 4, although these categories are not global but domestic. Thus, the detection of crosswalks with rank of 3 or less has not been sufficiently discussed. In such situations, LIDAR is not suitable for detection because the reflected intensity is not stable for the damaged road paints. In addition, because crosswalks do not have clear white bands, the existing camera-based methods might not grasp the abovementioned features. For example, owing to the damaged road paint, these methods might detect neither the rectangular shape of the white bands nor their side edges. The feature values regarding repetition, such as cross ratio, FFT value, and frequency of the white bands, might not be constant in the situation of Fig. 2. Thus, other features are necessary to detect the damaged road paints of crosswalks on public roads. With these motivations, in this study, we aimed to develop a new detection system for damaged crosswalks on public roads.

The following parts of this paper are organized as follows: details of the system design are described in chapter 2. Offline evaluations of the proposed system using actual driving data on public roads are described in chapter 3. Conclusions and future works are described in chapter 4.
2. System Design

2.1. Concept Design

In our previous study (9), we developed a localization system with a digital map. In the previously proposed system, the intelligent vehicle detected stop lines and guardrails as landmarks for localization. Thus, in this study, we aimed to develop a crosswalk detection system as one of the basic components that enhanced our localization system. Figure 3 shows the conceptual schematic of a localization system based on crosswalk detection that the digital map supports. In this study, we aimed to develop the crosswalk detection system as a first part, on the assumption of the support by the digital map as a later part.

In the design of landmark detection systems, reducing the missed and false detections, and increasing the positive ones, is necessary. In this study, positive detection means that the system outputs the final detection results through the detection process where a crosswalk actually exists. Missed detection means that the system does not output the final detection results where a crosswalk actually exists. On the contrary, false detection means that the system outputs the final detection results where a crosswalk does not exist. In usual cases, there is a tradeoff between missed and false detections. Thus, determining their balance, considering purpose and usages, is important. The future usage of our system is a basic component for localization. Further, the digital map can exclude false detections where crosswalk data are not registered. Thus, the basic concept in this study is that reducing missed detections is more important than reducing the false ones. Based on this basic concept, we need to detect as many crosswalks as possible for increasing the chance of localization. Therefore, we need to focus on the faint features of crosswalks, such as bottom shape of a white band, for reducing missed detections, although focusing on these faint features has the possibility of provoking false detections. Additionally, the digital map cannot exclude a false detection near a true crosswalk, although it can exclude a false detection at a distance from a true crosswalk. Because there are usually other road marks around the crosswalks, such as arrow marks, we need to avoid mistaking them for the true crosswalk. Thus, we need to combine multi-layered faint features to discriminate other road marks from the crosswalks. For example, although the existence of bottom shape of a white band only shows a faint feature for detecting crosswalks, horizontally distributed band bottoms in addition to it give a bit probable clue. Further, if the shape of the white band candidate is adequate in addition to them, the probability becomes more confident. In this way, we focused on the multi-layered features of the crosswalk: existence of the band bottom, distribution of the band bottoms, and shape of a band candidate.

2.2. System Outline

Based on the basic concept discussed in Section 2.1, we developed the crosswalk detection system. Figure 4 shows the flowchart of the crosswalk detection algorithm. This flow is roughly separated into six steps from preprocess to confirmation of relative position, as shown on the left side of Fig. 4. Through several steps to extract these features, this algorithm detects the crosswalk. The details of each step are described in the following sections.

Copyright © 2019 Society of Automotive Engineers of Japan, Inc. All rights reserved
2.3. Preprocess

The first step is preprocess from “camera input” to “inverse perspective mapping” as shown in Fig. 4. Figure 2 shows an example of a camera input image. In our system, we equip the monocular camera, which is a GS3-U3-15S5C-C produced by FLIR, inside the windshield of the experimental vehicle. Because our system aims to detect various traffic elements such as road marks, other traffic participants, and traffic lights by only a single camera, we use a color camera with its resolution set at 1280 × 960. However, because the following process needs only grayscale information, we convert the input image from color to grayscale. Then, we conduct inverse perspective mapping and resize the image to 640 × 480. Figure 5 shows the example of the inverse perspective mapped image. Because the crosswalks are expected to be located in front of the vehicle, we set the region of interest (ROI) as shown in Fig. 5. By setting the ROI, we aim to speed up the calculation process.

Fig. 5 Example of inverse perspective mapped image.

Fig. 6 Template image for detecting the bottom shape of white bands.

2.4. Band Bottom Extraction

The second step starts with template matching. Because inverse perspective mapping can generate a pseudo ortho-image, we can use a template image, the size of which is fixed. Figure 6 shows the template image for detecting the bottom part of the white bands. We use the template matching method so that the detector calculates the zero-mean normalized cross correlation (ZNCC) as an index of similarity. Figure 7 shows the result of template matching. Brighter pixels indicate a region more similar to the template image. Then, we conduct the erosion and dilation process to reduce noise. After that, we exclude the pixels that, according to the similarity to the template image, are equal to or less than a threshold value. In this study, we assigned 0.7 to the threshold value. This threshold value was relatively relaxed one because we aimed to extract candidates of the white-band bottom with some damages. Figure 8 shows the result of the threshold process. The remaining white areas shown in Fig. 8 indicate the candidate positions of the white-band bottom.

Fig. 7 Result of template matching.

Fig. 8 Result of threshold process.

2.5. Crosswalk Bottom Extraction

The third step calculates the candidate ROI of crosswalks. Because the crosswalks consist of horizontally located white bands, a higher number of white-band candidates in a certain area indicates higher possibility of existence of a crosswalk. Based on this concept, we calculate the histogram of possibility scores of the crosswalk bottom positions. Figure 9 shows the conceptual schematic of the extraction process of the candidate ROI of a crosswalk bottom. For each candidate position of the white band bottom, as shown on the left side of Fig. 9, we pick up the point that has the local maximum value of similarity score for the template image within a contiguous area. For each point, we add the histogram with a weight based on the normalized Gaussian kernel multiplied by similarity score. The center part of Fig. 9 shows the result of this process. The vertical axis indicates the pixel height in the image while the horizontal axis indicates the possibility score. Larger possibility score indicates more number of horizontally distributed band bottoms around a certain pixel height; further, it indicates more probability of existence of a crosswalk.

Fig. 9 Conceptual schematic of extraction process of the crosswalk-bottom candidate ROI.
Then, we segmentalize the histogram, such as the segments A to D shown in Fig. 9. After that, we calculate the maximum value for each segment. If the maximum value of a segment is larger than a certain threshold value, we consider that segment as a candidate ROI of the crosswalk bottom. In this study, we assigned 1.06 to a threshold value. This process can select more than one ROI because on-board camera sometimes takes a picture of two crosswalks at a same time: one in front of a crossing and the other behind a crossing. If this process detects more than one ROI, the system executes the following process for each ROI, as shown in Fig. 4. In the case of Fig. 9, because only the maximum value of segment D is larger than a certain threshold value, we picked up only one ROI of the crosswalk bottom. The right side of Fig. 9 shows the ROI of the inverse perspective mapped image corresponding to the histogram. As shown in Fig. 9, this process adequately extracts the candidate ROI of the crosswalk bottom.

2.6. Band Extraction

The fourth step comprises the extraction of the valid bands, which are qualified as white bands from the perspective of the shape and the damage of road paint, and the main band, which is the most reliable of the valid bands. Figure 10 shows the conceptual schematic of extraction of the crosswalk valid bands. At first, we pick up the candidate positions of the white-band bottoms that are located inside the candidate crosswalk-bottom ROI. Based on the pixel that has the local maximum value in each candidate position of the white band bottom, we set the ROI, which are shown as orange rectangles in Fig. 10-1. The size of the ROI is the same as that of the template image. Then, by setting the ROIs to the inverse perspective mapped image, we calculate the threshold values for local binarization by the Otsu method \(^{(10)}\) around each ROI, as shown in Fig. 10-2. By confirming the threshold value for local binarization, we exclude the significantly damaged band bottom candidates. For example, because the road paint of the rightmost bounding box in the Fig. 10-2 is damaged too much, we exclude this bounding box from the candidates. After that, in the inverse perspective mapped image, we extract the contiguous areas starting from each ROI using the calculated threshold value. Figure 11 shows the supplemental schematic of this process. By confirming the size of the bounding box of the extracted areas, which are shown as blue rectangles in Figs. 10 and 11, we pick up the valid bands. If we cannot pick up more than one valid band, we stop the process and conclude that there is no crosswalk in the image. Otherwise, we extract the pixels of the valid bands, as shown Fig. 10-4. Additionally, we select the main band, which has the largest size among the valid bands.

2.7. Confirmation of Crosswalk Shape

The fifth step is the shape confirmation of the whole crosswalk. To confirm the shape of a whole crosswalk, we extrapolate the individual damaged bands. In this study, we call these extrapolated bands imaginary bands. Figure 12 shows the conceptual schematic for extrapolating the imaginary bands. First, we set the ROI based on the size of the bounding box containing the main band, as shown in left part of Fig. 12. Then, we extrapolate the imaginary bands based on the ROI of the main band as shown in the right of the Fig. 12. By taking the median of the bottom positions of these imaginary bands, we determine the vertical position of the crosswalk bottom. In addition, from the rightmost and leftmost imaginary-band positions, we determine the horizontal position of the crosswalk bottom. Owing to the imaginary bands, we can easily estimate the whole shape of the crosswalk.
2.8. Confirmation of Relative Position

The sixth step is a confirmation of relative position. To reduce the unexpected false detection, we compare the estimated distance to the crosswalk, which is calculated using inverse perspective mapped image, between previous and current frames. If the deviation of estimated distance between frames are larger than a certain threshold value, we consider the result as unstable false detection. Otherwise, we consider the detection result as valid one.

In some situations, we detect more than two crosswalks. For example, if the vehicle is located in front of a crossing, the system sometimes can detect two crosswalks: one in front of the crossing and the other behind the crossing. In this case, because the distance between two crosswalk candidates are large, we can consider these crosswalks as valid ones. On the contrary, if the system failed in detecting processes due to the damaged paint of the crosswalk, the system also detects more than one crosswalk candidate in a small area. In this case, because the distance between them are small, we can consider one of them as invalid one. In this way, by checking the distance between two crosswalk candidates, if exist, we can exclude the false detection.

Through the abovementioned processes, we detect a crosswalk. Figure 13 shows an example of the final detection results. The green line indicates the detection results that the system automatically draws.

3. Evaluation Experiment

As for the following experiment, we asked a member of the institutional review board (IRB) for human studies of the University of Tokyo about the necessity of the review by IRB regarding the experimental plan because we recorded the image data on the public roads. However, the member of the IRB answered that we did not need to be reviewed from a viewpoint of human studies because this experiment aimed to develop the system and not to evaluate the participants. Thus, we only took care of the privacy of the people who were visible in the recorded data from the viewpoint of the personal information protection.

3.1. Experimental Course

Although there are various elements needed to be considered in the evaluation of computer vision system, such as lighting conditions, the main focus of this study was detection of damaged crosswalks. Thus, to evaluate the performance of the proposed system from the viewpoint of the damage of road paint, we collected the driving data around various crosswalks on actual public roads. As for the lighting condition, we set up two conditions, which are explained later, as a first step. Figure 14 shows the aerial photographic map of the evaluation course. This figure is based on the map image published by Geospatial Information Authority of Japan (11). This course consisted of municipal roads and prefectural roads of Shizuoka. The surface of the road was not maintained as that of high-standard roads, such as highways. The total length was approximately 8 km. The evaluation targets were 22 crosswalks that the experimental vehicle drove over in a straight direction. Thus, for example, the crosswalks on both sides of the crossings were not evaluation targets.

![Aerial photographic map of evaluation course.](image)

**Fig. 14** Aerial photographic map of evaluation course.

In this course, we collected the image data of the monocular camera. To be more precise, we collected two data sets: one collected around midday and the other collected in the morning. Because the lighting condition owing to the solar elevation angle affects the contrast of the camera image, detection in the morning is more difficult than detection around midday. In addition, because this experiment was conducted in the end of autumn, the sizes of the shadows from the surrounding constructions were relatively large in the morning. To evaluate the negative effects of such factors for detecting crosswalks, we set the condition in the morning. Thus, we aimed to evaluate the basic performance from the results around midday, and grasp the limitations of performance from the results in the morning. Although there are much more various weather conditions, this conditions were examined as a first step.

3.2. Evaluation of Computational Speed

We evaluated the computational speed in two situations by measuring the average calculation speed from 100 frames for each situation. There was only one crosswalk in the first situation whereas there were two crosswalks in the second situation. Because the proposed method focuses on the multi-layered features, the computational speed may vary depending on the specific situations. In the first situation, the detection process took 0.0336 s (29.8 Hz) on average via a desktop PC (3.0 GHz CPU on the condition of hyper threading). On the contrary, in the second situation, it took 0.0385 s (26.0 Hz). The calculation speed
decreased owing to the increase in detected crosswalks because the system repeated the processes of band extraction and confirmation of crosswalk shape twice. However, the effect was not excessive. Because the framerate of the camera in our system is 15 fps, the proposed system realized real-time detection of crosswalks for both situations. Therefore, we confirmed that the processing speed of the proposed system was practical.

3.3. Detection Results

3.3.1. Summary of Detection Results

Table 1 presents the summary of detection results. There are three types of detection results: positive, missed, and false. Regarding the data set around midday, the proposed system correctly detected all crosswalks without false detections. Thus, the proposed system has a relatively good performance under basic lighting conditions. However, regarding the data set in the morning, the proposed system could not detect one crosswalk, and output one false detection. The following subsections explain the detailed examples of these results.

Table 1 Summary of detection results.

|                      | Around midday | Morning |
|----------------------|---------------|---------|
| Positive detection   | 22/22         | 21/22   |
| Missed detection     | 0/22          | 1/22    |
| False detection      | 0             | 1       |

In addition, to discuss the detecting performance from the viewpoint of damage of road paints, we calculated the ratio of valid pixels. Figure 15 shows the conceptual schematic of ROI for extracting valid pixels. We set the ROI, which is an orange rectangle shown in Fig.15, for the valid pixels extracted in the process discussed in the section 2.6. The height of the ROI corresponds to the height of main band whereas the width of the ROI corresponds to the width between rightmost and leftmost pixels of valid bands. The ratio of valid pixels is expressed as follows:

\[ R_{VP} = \frac{P_{V}}{2N-1} \]

(1)

where \( R_{VP} \) denotes the ratio of valid pixels, \( P_{V} \) denotes the number of valid pixels, \( P_{ROI} \) denotes the number of pixels of the ROI, and \( N \) denotes the number of the bands in the ROI. The ROI contains \( N \) white bands and \( N-1 \) blank spaces, and the width of a blank space is ideally equal to the width of a white band. Thus, we multiply the correction value \( N/(2N-1) \) to express the number of pixels of ideal white bands. Because the valid pixels are extracted by the proposed processes, \( R_{VP} \) is not an absolute index but an index how many pixels the proposed system used for detecting the crosswalk. Basically, smaller \( R_{VP} \) indicates more damaged crosswalk. Because \( N \) indicates the number of white bands including sufficiently damaged ones, which were not considered as valid and could not be counted systematically, we counted \( N \) visually.

We calculated the possibility score (PS) discussed in the section 2.5 and the abovementioned \( R_{VP} \) for all crosswalks in the data set around midday. Because both of these indices are time-series ones, we picked up the values when the PS got a maximum value in the time-series values around each crosswalk. Figure 16 shows the \( R_{VP} \) and PS of detected crosswalks. To explain the examples, we picked up the crosswalks A-D. Crosswalk A had the largest \( R_{VP} \) and Fig. 1 shows the corresponding appearance. As shown in the Fig. 1, the road paint of the crosswalk did not have any damages. Crosswalk B had the largest PS and relatively large \( R_{VP} \), and Fig. 17 shows the corresponding appearance. In contrast to the crosswalk A, because the crosswalk B had more number of white bands, PS of crosswalk B was larger than that of crosswalk A. As just described, crosswalks A and B are the examples of well maintained crosswalks. Thus, we think that even the existing methods can detect these ones easily. On the contrary, crosswalks C and D had small PS and \( R_{VP} \), and Fig. 18 shows the corresponding appearance. As shown in the figure, the road paints of these crosswalks had some damages. The \( R_{VP} \) of crosswalks C and D was 0.41 and 0.35, respectively. We think that these crosswalks could be categorized into rank 3 or less, which were not discussed in the existing studies, by the visual evaluation rank of JCASM. Thus, our proposed system could detect even such damaged crosswalks by using multi-layered features.
3.3.2. Example of Positive Detection

Figure 19 shows the damaged crosswalk that the proposed system could detect correctly. This crosswalk corresponds to the crosswalk C discussed in the section 3.3.1. The left figure shows an input image and the right one is an inverse perspective mapped image. The data set around midday contained this situation. As shown in Fig. 19, although the side edges of the white bands are not straight due to the damages of the road paint, our proposed system could detect the crosswalk owing to focusing on the multi-layered features of the crosswalks.

3.3.3. Example of Missed Detection

Figure 20 shows the example of damaged crosswalk that the proposed system could not detect. The data set in the morning contained this situation. Because our proposed system could detect the same crosswalk in the data set around midday, lighting condition was the reason of this missed detection. The frame-by-frame analysis indicated that the proposed system could detect the candidate ROI of crosswalk bottom, and could not continuously detect more than one valid band. Because the shapes of some white bands were changed due to the damage of road paints, the system did not consider such ones as valid. In addition, although there were some bands that might be detected under normal lighting conditions, the proposed system could not detect them owing to the dark lighting condition. On this point, adjusting the local brightness in the area of the shadow for detecting the white-band bottom considering the lighting conditions offers the possibility of solving this problem.

3.3.4. Example of False Detection

Figure 21 shows the example of the situation where the proposed system output a false detection. The data set in the morning contained this situation, where there were shadows near and far from the vehicle. In addition, there is a bridge around the center of the image. Around the bridge, the road surface is brighter than at other parts owing to the sunlight. In the brighter area, there are two straight shadow lines caused by the parapet of the bridge. In addition, the shape of the brighter areas is accidentally rectangular, and their size is similar to the white bands of a crosswalk. As a result of these lighting conditions, the proposed system finally output the false detection. Because the proposed system focuses on faint features, it is hard for it to avoid this type of false detections. However, as mentioned in Section 2.1, the proposed system includes the confirmation by a digital map in a later step. Thus, this false detection can be later excluded.

3.3.5. Prevention of False Detection

In this subsection, we explain the examples for preventing false detections. Figure 22 shows the example of the situation where the proposed system prevented the false detection of more than two crosswalks within the area of one crosswalk. As shown in the right image, there are three abnormal bands: A, B, and C. The length of band A is short owing to the existence of a grating cover. Because bands B and C are both damaged, the top of the bands can be considered as isolated small bands. As a result of these abnormal bands, the proposed system detects another crosswalk candidate, as the orange line indicates. However, the relative position confirmation, which is the final process of the proposed system as shown in Fig. 4, excluded this false detection because the distance between the bottom of this crosswalk candidate and the bottom of the true crosswalk was too short. As a result, the proposed system could prevent the false detection in this situation.
Fig. 22 Example of the situation where the proposed system prevented the false detection of more than two crosswalks within an area of one crosswalk.

Figure 23 shows the example of the situation where the proposed system prevented false detection around arrow marks. There are two arrow marks in the image. Although the bottom shapes of the arrow marks are similar to those of the white bands, the former are slimmer. Thus, because ZNCC scores of these parts were relatively small, the noise reduction process after template matching excluded these parts from the white-band candidates. Therefore, the proposed system could prevent the false detection in this situation.

Fig. 23 Example of the situation where the proposed system prevented false detection around arrow marks.

Figure 24 shows the example of the situation where the proposed system prevented false detection around the damaged road paint of Japanese characters. Owing to the damages of the road paints, some parts were accidentally similar to the template image from the perspective of shape and size. As a result, the proposed system extracted the candidate ROI of a crosswalk. However, the system excluded these parts in the process of valid-band extraction because these parts were not qualified as a white band from the perspective of whole shape. As a result, the proposed system could prevent false detection in this situation.

Fig. 24 Example of the situation where the proposed system prevented false detection around the damaged road paint.

As the abovementioned examples explain, our proposed system could prevent several false detections, except for the case of Fig. 21. Although focusing on the faint features of landmarks easily provokes false detections, the proposed system achieved prevention of them by confirming the multi-layered features of the crosswalks.

4. Conclusions

In this study, we developed a new detection system of damaged crosswalks on public roads. As a clue for detecting crosswalks, we focused on the bottom shape of the white bands; further, we extracted them by the template matching method. Then, we extracted a candidate ROI of a crosswalk by analyzing the distribution of the white-band bottoms. Furthermore, we checked the white-band shape to confirm the valid crosswalk white bands. Through these processes, the proposed system detected the damaged crosswalk. In addition, we conducted the experiment on actual public roads to evaluate the performance of the proposed system. As a result, we confirmed its practical performance and computational speed for detecting the actual crosswalk under basic lighting condition.

However, future work is still needed. Although the proposed system showed good performance with the data set around midday, the performance in the morning was not perfect. On this point, we need to further improve the system from the perspective of adaptive brightness adjustment depending on the area of the shadow. Based on such improvement, we need to verify the system in much more various lighting conditions. In addition, we need to improve the system to detect in the night and rain conditions. As for the night conditions, the system needs to detect crosswalks with using the headlight of the vehicle or roadside lamps. In contrast to the sunlight, brightness under such lights decreases according to the distance from them. Thus, more adaptive brightness adjustment will be necessary. On the contrary, as for the rain conditions, a drop of rain over the front windshield may prevent the detection of crosswalks. In addition, the headlight of the oncoming vehicle reflected on puddles may also prevent the detection. Thus, correction of the image considering such factors will be necessary. Furthermore, because the final goal of our system is to realize a localization system based on landmark detection, we need to develop the detection system of other landmarks, such as a damaged stop lines.

Acknowledgment

This research has been conducted as a part of the research project “Autonomous Driving System to Enhance Safe and Secured Traffic Society for Elderly Drivers” granted by Japan Science and Technology Agency (JST), S-Innovation (Strategic Promotion of Innovative Research and Development). The authors would like to thank the agency for providing financial support. The authors would also like to thank KPIT Technologies Ltd. for their contributions.

References

(1) Shigetaka Suzuki et al.: Sensor Fusion-Based Pedestrian Collision Warning System with Crosswalk Detection,
(2) Anselm Haselhoff et al.: On Visual Crosswalk Detection for Driver Assistance Systems, Proceedings of 2010 IEEE Intelligent Vehicles Symposium, pp. 883-888 (2010).

(3) Sebastian Sichelschmidt et al.: Pedestrian Crossing Detecting as a part of an Urban Pedestrian Safety System, Proceedings of 2010 IEEE Intelligent Vehicles Symposium, pp. 840-844 (2010).

(4) Philippe Foucher et al.: Detection and Recognition of Urban Road Markings Using Images, Proceedings of 2011 14th International IEEE Conference on Intelligent Transportation Systems, pp. 1747-1752 (2011).

(5) Yuqiang Zhai et al.: Crosswalk Detection Based on MSER and ERANSAC, Proceedings of 2015 18th International IEEE Conference on Intelligent Transportation Systems, pp. 2770-2775 (2015).

(6) Attila Börcs et al.: On board 3D Object Perception in Dynamic Urban Scenes, Proceedings of 4th IEEE International Conference on Cognitive Infocommunications, pp. 515-520 (2013).

(7) Danilo Cáceres Hernández et al.: Crosswalk Detection Based on Laser Scanning from Moving Vehicle, Proceedings of 2015 IEEE 13th International Conference on Industrial Informatics, pp. 1515-1519 (2015).

(8) Japan Contractors Association of Traffic Signs and Lane Markings: Road Paint and Traffic Safety, http://www.zenhyokyo.or.jp/common/pdf/romen_anzen_vol.10.pdf, (Accessed Aug. 1 2019) (in Japanese).

(9) Takuma Ito et al.: Novel Map Platform based on Primitive Elements of Traffic Environments for Automated Driving Technologies, International Journal of Automotive Engineering, Vol. 7, No. 4, pp. 143-151 (2016).

(10) Nobuyuki Otsu: A threshold selection method from gray-level histograms, IEEE transactions on systems, man, and cybernetics, Vol. 9, No. 1, pp. 62-66 (1979).

(11) Geospatial Information Authority of Japan: GSI Maps, https://maps.gsi.go.jp/, (Accessed Aug. 1 2019).