IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 9, 2020, accepted July 20, 2020, date of publication July 23, 2020, date of current version August 4, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3011430

Detection and Tracking of Moving Objects at Road Intersections Using a 360-Degree Camera for Driver Assistance and Automated Driving

CHINTHAKA PREMACHANDRA, (Member, IEEE), SHOHEI UEDA, AND YUYA SUZUKI
Department of Electronic Engineering, Shibaura Institute of Technology, Tokyo 135-8548, Japan
Corresponding author: Chinthaka Premachandra (chintaka@sic.shibaura-it.ac.jp)

This work was supported in part by the Branding Research Fund of the Shibaura Institute of Technology.

ABSTRACT Many complicated road intersections are seen while driving. In some, blind spots make it difficult for drivers or automated vehicles to discern moving objects coming from certain directions, possibly confusing drivers or autonomous vehicles wishing to cross or to turn at the intersection. To address this problem, we investigate detection and tracking of all moving objects at an intersection using a single 360-degree-view camera (3DVC). Through experiments, we develop methods allowing a 3DVC to capture the entirety of a four-way intersection when installed at one corner. This paper also presents image processing algorithms for detecting and tracking moving objects at intersections by processing images from the installed 3DVC. Experiments under varied conditions demonstrate that the proposed detection algorithm has a very high detection rate. We also confirm the tracking ability for moving objects detected using the proposed algorithm.

INDEX TERMS 360-degree camera (omnidirectional camera), driver assistance, autonomous driving, moving object detection and tracking, image conversion.

I. INTRODUCTION

Traffic accidents are reported around the world daily. In many countries, the majority of accidents occur at road intersections. Objects such as pedestrians, vehicles, cyclists, and wheelchairs move differently at intersections, and this complexity can confuse drivers trying to comprehend the intersection while driving through it or turning. If drivers can understand the conditions of an intersection—including its blind spots—a few seconds before arriving there, they can better handle their vehicle while traversing the intersection. In the case of automated driving, moving objects in blind spots of intersections are particularly confusing for automation systems without proper understanding of objects approaching from different directions. Such situations can cause traffic accidents in both manual and autonomous driving. Automatic detection of moving objects and real-time transfer of moving object information to vehicles are extremely important for solving such intersection traffic problems. In this study, we apply use of a 360-degree-view camera (3DVC) to tackle the problem of detecting moving objects at intersections. We experimentally verified that installing a 3DVC at an intersection corner can capture the entire intersection area, as described in the next section. In this paper, we mainly focus on the detection and tracking of vehicles, pedestrians, cyclists, wheelchairs and other moving objects at intersections by processing images from a 3DVC installed at an intersection corner. This study targeted detection of moving objects within 30 m from the camera because we experimentally confirmed that detection of moving objects within this range is sufficient to support drivers a few (2–5) seconds before arriving at the intersection by informing them of moving object distributions there. Furthermore, we believe this detection range is sufficient to achieve turning in the case of automated driving, since the movement of autonomous vehicles is nearly identical to manual driving.

There have been many previous studies of moving object detection at intersections through use of two-dimensional (2D) cameras [1]–[7], but it is difficult for 2D cameras to capture all moving objects in an intersection. Even so, some approaches focus on comprehensive moving vehicle detection using a single 2D camera installed high above the ground [8], [9]. Moving vehicles are detected by processing

The associate editor coordinating the review of this manuscript and approving it for publication was Shadi Alawneh.

images from the camera. When the camera is installed high above the intersection, it is possible to capture the entire area. However, it remains difficult to capture pedestrians and bicycles, which do not clearly appear in camera images. Furthermore, installing such cameras at a sufficient height can be difficult and costly.

Other studies have thus attempted moving object detection throughout intersections by installing multiple cameras at complex intersections [10]. Multiple cameras can capture images from different intersection areas, allowing moving object detection by image processing. Moving objects detected by all cameras can be merged to grasp the behavior of moving objects situation throughout the intersection. However, the required image data processing increases when multiple cameras are used, potentially making real-time detection difficult.

In this paper we target to detect all the moving objects around an intersection. We installed the 3DVC 3m above the ground at a corner of an intersection, as shown in Fig. 2. We experimentally found that this camera installation is sufficient for capturing images of the entire intersection, including all nearby moving objects. In the other word, an image from 3DVC includes the entire intersection area, when the 3DVC is fixed at an intersection corner following the above mentioned installation procedure. For example, Fig. 4 shows an image with this camera installing at an intersection. In the image, four roads of the intersection can easily be found when you focus on the four road crossings. In this intersection, there is a crossing on each road. Furthermore, the moving objects on roads as well as near the intersection region can also be confirmed.

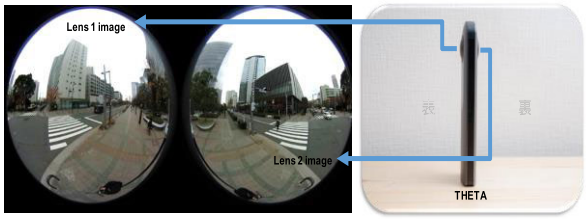

FIGURE 1. Structure of an image captured by a Theta 3DVC.

In this study, we attempt detection and tracking of all moving objects in a road intersection using a single 3DVC camera. Generally, a 3DVC has two lenses, each providing a 180-degree circular image like those shown in Fig. 1. Two 180-degree view circular images combined provide a 360-degree view. These 3DVC cameras have been used in studies related to topics such as virtual reality [11], mapping [12], [13], 360-degree imaging systems [14], video coding [15], depth estimation [16], and navigation behavior analysis [17]. To the extent of our knowledge, however, there have been no studies regarding their application to detection of traffic situations at an intersection. In early stages of this work, we attempted detection of moving objects in the entirety of an intersection by processing the original circular images [18]. However, we found that tracking moving objects is difficult with these images, as explained in the next section. In this study, therefore, we generated rectangular images from the original circular images and performed moving object detection and tracking using those rectangular images.

Experiments conducted to confirm the effectiveness of the proposed detection algorithm showed high detection rates. Experiments also demonstrated that the proposed tracking algorithm can track detected moving objects with high accuracy. This detection and tracking approach is therefore practically applicable to development of driver assistance systems capable of autonomous driving at complex intersections with blind spots.

The remainder of this paper is organized as follows. Section 2 describes 3DVC installations at intersections. Section 3 presents the proposed moving object detection method, and Section 4 describes it in detail. Section 5 presents and discusses the experimental results of this work. Finally, Section 6 concludes the paper.

II. APPLICATION OF 3DVC TO DETECTING MOVING OBJECTS AT INTERSECTINS

In this study, we used a Ricoh Theta 360-degree camera as the 3DVC. This section describes the camera installation at an intersection performed for this work and the image conversion necessary for detecting and tracking moving objects. The Theta has two lenses, each generating a 180-degree circular image like those shown in Fig. 1. Combined, these circular images provide a 360-degree view.

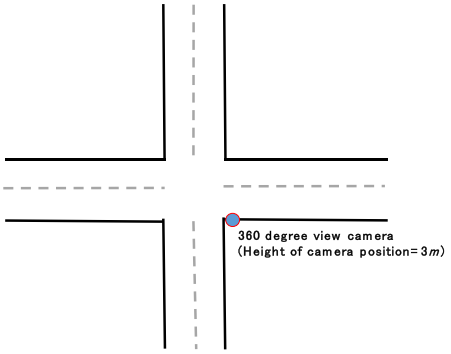

FIGURE 2. 3DVC installation at a road intersection.

A. CAMERA INSTALLATION

We installed the 3DVC 3m above the ground at a corner of an intersection, as shown in Fig. 2. We experimentally found that this camera installation is sufficient for capturing images of the entire intersection, including all nearby moving objects.

As mentioned above, full 3DVC images combine a circular image from each of its lenses. Figure 3 shows an image from a 3DVC installed at an intersection. Note that the brightness might be poor in some images, according to the lighting environment at the time (Fig. 3).

FIGURE 3. A 3DVC-captured image.

B. BRIGHTNESS IMPROVEMENT

As preprocessing, we apply gamma correction to improve the brightness of original images [19]. In gamma correction, an original camera image with 8-bit pixel intensity levels (I_i) is converted into levels in the range [0, 1]. Gamma conversion is then performed following Eq. (1), and a gamma-converted image I_o of I_i is generated following Eq. (2).

$$I_g = (\frac{I_i}{255})^{\frac{1}{\gamma}} \tag{1}$$

$$I_o = I_g * 255 \tag{2}$$

Figure 4 shows gamma conversion of the image in Fig. 3. This brightness correction can improve detection of moving objects that are far from the camera.

FIGURE 4. Gamma correction of the image shown in Fig. 3.

C. CONVERSION TO RECTANGULAR IMAGES

In early stages of this work, we attempted to detect moving objects at intersections using the original circular images [18]. We could achieve moving object detection to some extent, but it was difficult to track moving objects using these images, owing to the difficulty of tracking crossover from one image to the other. In this study, therefore, we performed moving object detection and tracking after converting the circular images to rectangular ones. Specifically, we continually convert pairs of source circular images to combined rectangular images.

To convert the original images to rectangular images, we use equirectangular conversion, a simple map projection method [20]. In this study, we perform projections from spherical coordinates to planar coordinates by Eqs. (3) and (4).

$$\mathrm{x} = (\lambda - \lambda_0)\cos\vartheta_1 \tag{3}$$

$$\mathrm{y} = \vartheta - \vartheta_1 \tag{4}$$

In the above equations, x and y respectively denote horizontal and vertical coordinates of the projected location coordinate, λ and λ_0 are respectively the longitude and central meridian of the location to project, and ϑ and ϑ_1 denote the latitude of the location to project and standard parallels.

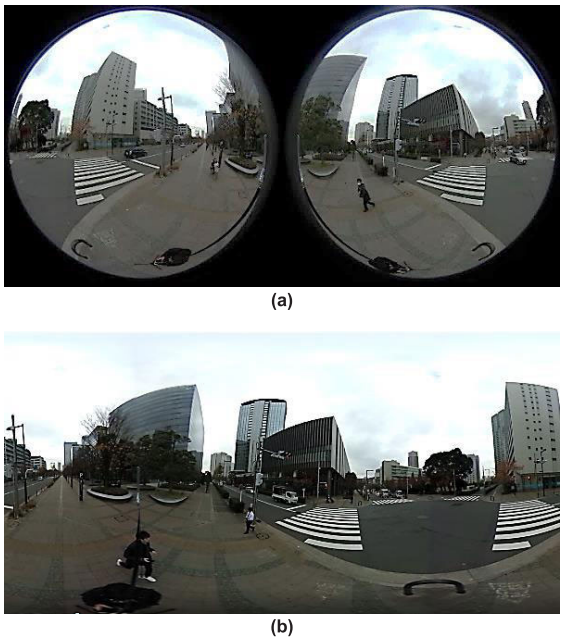

FIGURE 5. Example equirectangular conversion (a) Original image, (b) Equirectangular conversion.

Figure 5(a) shows an original image from the camera and Fig. 5(b) shows its equirectangular conversion.

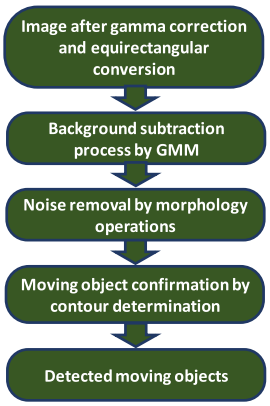

FIGURE 6. Major steps in the moving object detection algorithm.

III. MOVING OBJECT DETECTION AND TRACKING

We use images after brightness correction and equirectangular conversion to perform moving object detection. Our proposed algorithm for moving object detection is sufficiently lightweight to allow real-time detection. Figure 6 shows the

major steps in the algorithm, as described in the following sub-sections.

A. MOVING OBJECT CANDIDATE EXTRACTION BY GAUSSIAN MIXTURE MODEL BASED BACKGROUND SUBTRACTION

Frame and background subtraction can be used to detect moving objects in images [21]–[32]. In frame subtraction, moving objects are detected by calculating the image subtraction between two or a few consecutive frames [21], [23], [28]. Moving objects must have some minimal speed for detection by frame subtraction. In this study, the camera is static and some targeted moving objects, such as wheelchairs, have very low speeds. Note that objects like pedestrians waiting at the intersection cannot be detected with frame subtraction, so we decided not to use frame subtraction to extract moving objects. Instead, we use background subtraction, which generally performs subtraction between a current frame and a previously prepared background image [22], [27], which must be updated as lighting conditions at the intersection change. We use a background subtraction method that follows a Gaussian mixture model (GMM) to extract moving object candidates from the images [24]–[26], [46], [47] because this method can automatically update background pixels in the image. The light condition of the outdoor environment slightly changes as sunlight changes. Sometimes it changes sharply following the changes of cloud conditions. In GMM, multi Gaussian models (K) are generated and updated following the pixel value variation over the time, regarding each pixel in the image. Then, if the current pixel value is away from the multi Gaussian models, it is picked up as a pixel of a moving object. The same process is applied to all the pixels in the image to extract the pixels of moving objects. GMM is robust against above-mentioned light changes in the outdoor environment because a pixel of moving objects are extracted by comparing with several Gaussian models, those are generated following it's value variation over the time. Thus, the GMM is adaptive enough for detecting moving objects from consecutive images from a static camera in an outdoor environment. This method models each background pixel by a mixture of K Gaussian distributions ($K = 3 \ldots 5$). The mixture weight denotes the time interval over which those colors remain in the scene. The probability of each pixel to be in the foreground or background is calculated using its intensity in RGB color space as

$$\mathrm{P}\,(X_t) = \sum_{i=1}^{K} \omega_{i,t} \cdot \tau(X_t, \mu_{i,t}, \sum_{i,t}), \tag{5}$$

where X_t is the current pixel in frame t, K is distributions in the mixture, $\omega_{i,t}$ is the weight of the k-th distribution in frame t, $\mu_{i,t}$ is the mean of the k-th distribution in frame t, and $\sum_{i,t}$ is the standard deviation of the k-th distribution in frame t. $\tau(X_t, \mu_{i,t}, \sum_{i,t})$ is a probability density function with general structure

$$\tau\,(X_t, \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}}\,|\Sigma|^{\frac{1}{2}}} exp^{-\frac{1}{2}(X_t-\mu)\Sigma^{-1}(X_t-\mu)} \tag{6}$$

Not every RGB is correlated with others [25], [26], so differences in intensity can be assumed to have uniform standard deviations. The covariance matrix can thus be expressed as

$$\Sigma_{i,t} = \sigma_{i,t}^2 I. \tag{7}$$

Every Gaussian exceeding a defined threshold (th) is extracted as background, and others are extracted as foreground. When a pixel matches at least one K Gaussian in the mixture, ω, μ, and σ are updated. To clearly indicate this classification, pixels selected as foreground are colored white, while pixels selected as background are colored black, as

$$I_{GMM}(i,j) \begin{cases} 255\ (if\ every\ Gaussian > th) \\ 0 \end{cases} \tag{8}$$

Figure 7 shows an image after gamma correction and equirectangular conversion. This image was captured by an 3DVC installed at an intersection under the above-described installation conditions. Figure 8 shows its background subtraction following the GMM method, demonstrating that candidate moving objects can be extracted as white blobs. Note that these candidates do not have a uniform shape, and some noise is present. As mentioned above, in this paper the camera is static, so these subtraction methods can easily be used to detect moving objects from consecutive images from the camera rather than color based [40]–[43] and learning based [44], [45], [48] methods.

FIGURE 7. Image after gamma correction and equirectangular conversion.

FIGURE 8. Background subtraction by GMM.

B. SHAPING MOVING OBJECT CANDIDATES AND NOISE REMOVAL

To remove noise from the subtraction results and recover the shape of candidate moving objects, we apply a morphology

operation process. Previous studies have addressed this problem by proposing a hole-filling method [26]. In consideration of the computational times required, however, we instead apply a morphology operation, in which the subtraction image is eroded m times and the resulting eroded image is dilated m times. Figure 9 shows the result of this morphology operation applied to the subtraction image shown in Fig. 8.

FIGURE 9. Results of the morphology operation process.

As Fig. 9 illustrates, the morphology operation process reduces noise and improves the shape of candidate moving objects. However, some noise unrelated to subtractions of moving objects still exists in the image, the object in the red circle being an example. To extract actual moving objects from among candidates in an image resulting from the morphology operation, we apply the contour-tracking method described in the next subsection.

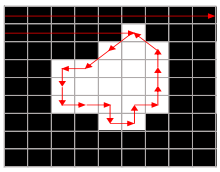

FIGURE 10. Results of raster scanning and contour tracking.

C. EXTRACTING MOVING OBJECTS BY CONTOUR TRACKING

We perform contour tracking to calculate the contour length of each candidate in images resulting from the morphology operation. Figure 10 shows this contour tracking operation, in which image pixels are raster scanned. If a white pixel is found during scanning, contour pixels of the object in which that white pixel is contained are tracked. This tracking is performed in the counterclockwise direction, as shown Fig. 10.

During the tracking process, we count the number of contour pixels (C_p). Candidate moving objects with C_p exceeding a predefined threshold (T_h) are extracted as moving objects. In this paper, $T_h = 7$.

Figures 11 and 12 respectively show moving objects detected in images resulting from the morphology operation and in a rectangularized source image. We calculate the circumscribing rectangle for each detected moving object (green rectangles in Figs. 11 and 12), the middle point of which is used as the object's position (p_t^{det}). In this paper, $p_{t(i)}^{det}$ is the number of detected moving objects in a frame, where i is counting order of them.

FIGURE 11. Moving objects detected in an image resulting from the morphology operation.

FIGURE 12. Moving objects detected in a rectangularized original image.

IV. TRACKING OF DETECTED OBJECTS

Detected moving objects are tracked to determine their movement direction, an important factor on the vehicle side. We plan to create a system that can send detected moving object information to vehicles approaching the intersection from all directions. In manual driving, drivers can more easily perform decision-making at intersections if moving object information can be received a few seconds before arriving. In particular, providing information about moving objects in blind spots can make drivers much more comfortable at intersections and allow computers to better control autonomous vehicles.

A. KALMAN FILTER-BASED MOVING OBJECT POSITION ESTIMATION

Object tracking is an interesting problem in computer vision, and there are many previous studies of this topic [33]–[35]. In this paper, we perform tracking based on a Kalman filter [36], which is less time-consuming and has been applied to similar tracking problems [37], [38]. We estimate the position of a moving object in a current frame (p_t^{est}) by following its position in the previous frame (p_{t-1}^{est}), following Eqs. (9) and (10). In these equations, v_t^{est} is the estimated moving object velocity following the velocity value estimated in the previous frame (v_{t-1}^{est}). In addition, a_t is acceleration of the moving object, and w_p and w_v are respectively noises

regarding the moving object's position and velocity.

$$p_t^{est} = p_{t-1}^{est} + v_{t-1}^{est} dt + \frac{1}{2} a_t (dt)^2 + w_p \tag{9}$$

$$v_t^{est} = v_{t-1}^{est} + a_t dt + w_v \tag{10}$$

Equations (9) and (10) can be summarized as

$$\begin{bmatrix} p_t^{est} \\ v_t^{est} \end{bmatrix} = \begin{bmatrix} 1 & dt \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p_{t-1}^{est} \\ v_{t-1}^{est} \end{bmatrix} + \begin{bmatrix} \frac{1}{2}(dt)^2 \\ dt \end{bmatrix} [a_t] + \begin{bmatrix} w_p \\ w_v \end{bmatrix}, \tag{11}$$

By substituting into Eq. (11) the definitions in Eqs. (12)–(15), we derive Eq. (16), which gives the estimated position of a moving object in the current frame (p_t^{est}).

$$f_t = \begin{bmatrix} 1 & dt \\ 0 & 1 \end{bmatrix} \tag{12}$$

$$b_t = \begin{bmatrix} \frac{1}{2}(dt)^2 \\ dt \end{bmatrix} \tag{13}$$

$$\boldsymbol{u}_t = [a_t] \tag{14}$$

$$\boldsymbol{w}_t = \begin{bmatrix} w_p \\ w_v \end{bmatrix} \tag{15}$$

$$p_t^{est} = f_t \boldsymbol{p}_{t-1}^{est} + b_t \boldsymbol{u}_t + \boldsymbol{w}_t \tag{16}$$

In this paper, p_t^{est} is calculated assuming that $\boldsymbol{w}_t = 0$. In addition, when tracking starts after detecting a moving object, we initialize the estimation assuming that the moving object's estimated positions in a few previous frames are the same as the current detected position.

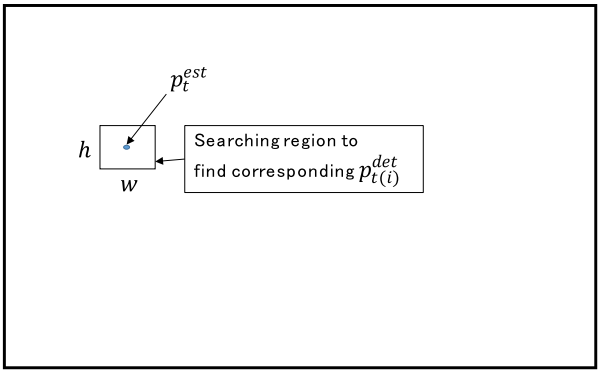

FIGURE 13. Definition of a searching region for finding a corresponding $p_{t(i)}^{det}$.

B. CONNECTING ESTIMATED MOVING OBJECT POSITIONS WITH POSITIONS BY THE DETECTION PROCESS

After estimating a current position for a moving object, it is compared with the positions of moving objects in the current frame detected by the detection process ($p_{t(i)}^{det}$) described in the previous section. Here, i is an index for detected moving objects in the current frame. We define a region keeping p_t^{est} as the center point, as illustrated in Fig. 13. The $p_{t(i)}^{det}$ located within that region is selected as the corresponding position of the currently tracked object. At an intersection, however, different types of moving objects move in different directions. Multiple detected positions can thus appear within the defined region, leading to tracking errors. In such situations, we focus on the directions of p_t^{est}. We calculate the moving direction of an object by following its estimated positions in a few previous frames. By comparing that direction value, we select the $p_{t(i)}^{det}$ corresponding to p_t^{est}. After confirming the relation between detected positions $p_{t(i)}^{det}$ and estimated positions $p_{t(i)}^{est}$, estimated $p_{t(i)}^{est}$ values can be updated by $p_{t(i)}^{det}$.

FIGURE 14. Tracking results, with red dots indicating detected moving object positions ($p_{t(i)}^{det}$) and white crosses indicating estimated positions (p_t^{est}).

The image in Fig. 14 shows an example of this tracking method. In that image, green color rectangles are circumscribing rectangles for detected moving objects, with red dots at their center. As mentioned above, we consider these dots as the detected positions ($p_{t(i)}^{det}$) of moving objects. White crosses denote estimated moving object positions ($p_{t(i)}^{est}$).

V. EXPERIMENTS

We conducted experiments to evaluate the proposed moving object detection and tracking method. All experiments were conducted installing the 3DVC camera at several intersections.

FIGURE 15. Theta camera installation at an intersection corner.

A. EXPERIMENTAL ENVIRONMENT

The 3DVC was installed at corners of the intersections used in experiments to evaluate the proposed detection and tracking methods. Figure 15 shows an example camera installation. We conducted experiments only at intersections where two

roads with four lanes intersect. In each intersection, there is a pedestrian crossing across each road segment near the intersection (Fig. 15). Furthermore, they are complicated intersections having blind spots from drivers' perspectives.

Configuration of personal computer used for experiments: 3.40 GHz (Core i7, 8 GB RAM). Dimensions of came images: 640 × 480 pixels.

We evaluated moving object detection at intersections for targets such as vehicles, pedestrians, wheelchairs, and cyclists. The detection rate was calculated for 400 objects at each through multiple experiments under sunny and cloudy conditions. Each experimental time (total video length) was 110 min. Human-pushed baby strollers and wheelchairs, as well as powered wheelchairs, were all categorized as ''wheelchairs.'' In some cases, moving behavior by multiple pedestrians walking together resulted in detection as a single object. We counted such clusters as a pedestrian. During the experiments, intersections were not very crowded and traffic flow was normal without traffic jams. The maximum vehicle speed on these roads was 70 km/h.

B. RESULTS

As Fig. 14 shows, a single 3DVC could detect objects moving from different directions at intersections.

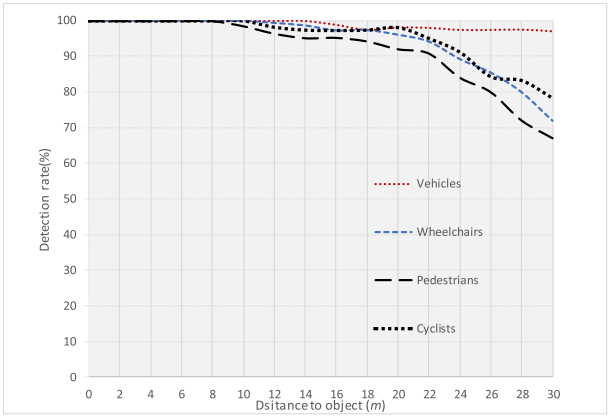

FIGURE 16. Detection rate variation by distance.

Figure 16 shows detection rate variations for each moving object, which suggest that all moving object types had very high detection rates at distances within 25 m. At greater distances, however, detection rates generally showed gradual decreases. However, vehicles maintained high detection rates even when moving 30 m from the camera because they produce relatively many pixels in the images. The average false positive rate for all moving object types was 5.8%.

The tracking algorithm was capable of tracking all types of detected moving objects. Figure 17 shows differences between estimated and detected positions for a moving vehicle and a cyclist. Figure 18 shows differences between estimated and detected positions for a pedestrian and a wheelchair.

Across all tracking experiments, the mean difference between estimated and detected positions was 1.67 pixels and

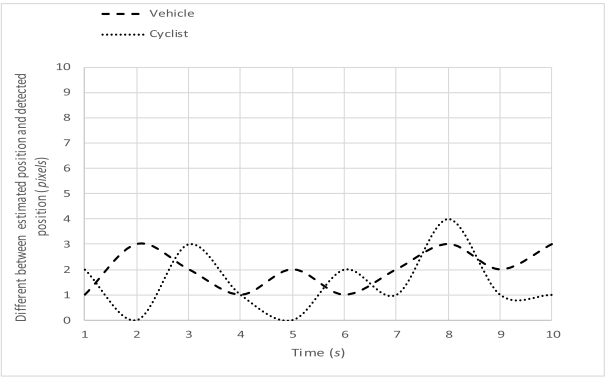

FIGURE 17. Differences between estimated and detected positions for a vehicle and a cyclist.

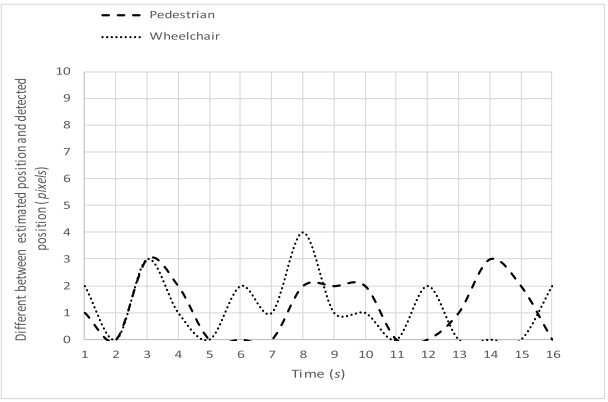

FIGURE 18. Differences between estimated and detected positions for a pedestrian and a wheelchair.

mean error rates when matching detected and estimated positions was 1.8%. The tracking method thus has high accuracy.

The average processing time was 28 msec/frame, so detection and tracking can be performed nearly in real time. According to the experiments, online processing is possible when the frame rate of the camera is less than 25 fps.

According to the overall experimental results, the tracking and detection algorithms both worked well when distances to moving objects were less than 30 m. Performance of the proposed method did not change with traffic flow. As mentioned above, there were some false detections in the experiments, but not at rates that would become a significant issue when applying the proposed methods to driver assistance and automated driving. The performance demonstrated above thus shows high potential for applying the proposed method to supporting drivers and for achieving safe automatic vehicle control at intersections. (See the submitted video to confirm these results.)

C. COMPARISONS WITH CONVENTIONAL METHODS

We compared the proposed moving object detection method with three conventional methods from the literature, denoted as Con 1 [22], Con 2 [27], and Con 3 [39]. Specifically, we compared detection rates for all moving objects in 110 min

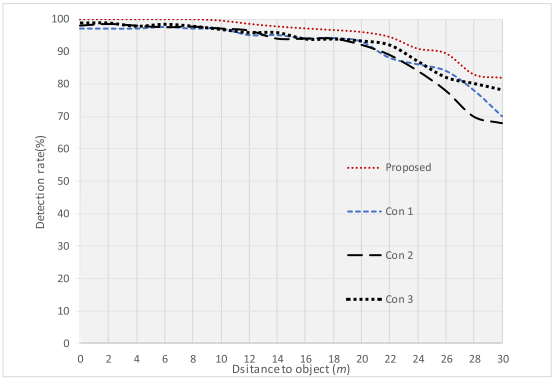

FIGURE 19. Comparison of the proposed moving object detection method with some conventional methods.

TABLE 1. Comparison of false positive rates.

Method	False Positive Rates
Con 1	11.3%
Con 2	10.3%
Con 3	6.2%
Proposed	5.8%

of the above-described video data, thereby providing similar experimental environments. Figure 19 shows the results, which confirm that the proposed moving object detection method based mainly on a GMM provided the best results. Table 1 summarizes false positive rates for each method. The proposed method resulted in fewer false positives than did the conventional methods.

VI. CONCLUSION

We proposed moving object detection and tracking methods for application at road intersections using a 3DVC for driver assistance and automated driving. We experimentally confirmed that a 3DVC installed at an intersection corner can image the entire intersection area. We also developed algorithms for detecting and tracking moving objects at the intersection using those 3DVC images. Experiments performed under varied conditions demonstrated that the developed algorithms showed high performance for detecting and tracking moving objects.

The proposed methods performed well when the distance to moving objects was within 25 m. The proposed methods can thus be applied at intersections between intersecting four-lane roads. Application of the proposed methods at larger intersections will require improvements to the detection and tracking algorithms, which we plan to investigate in future work. In addition, all experiments in this study were conducted during daytime under cloudy or sunny conditions. We are working to achieve similar performance under rainy and snowy conditions. By conducting tests with different cameras, we plan to achieve moving object detection and tracking during nighttime as well. Future work will also investigate estimations of detected moving object directions and development of a driver support system.

REFERENCES

[1] J. Eguchi and H. Koike, "Discrimination of an approaching vehicle at an intersection using a monocular camera," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2007, pp. 618–623.
[2] S. G. Ebrahimi, N. Seifnaraghi, and E. A. Ince, "Traffic analysis of avenues and intersections based on video surveillance from fixed video cameras," in *Proc. IEEE 17th Signal Process. Commun. Appl. Conf.*, Apr. 2009, pp. 848–851.
[3] F.-F. Xun, X.-H. Yang, Y. Xie, and L.-Y. Wang, "Congestion detection of urban intersections based on surveillance video," in *Proc. 18th Int. Symp. Commun. Inf. Technol. (ISCIT)*, Sep. 2018, pp. 495–498.
[4] Q. Baig, O. Aycard, T. D. Vu, and T. Fraichard, "Fusion between laser and stereo vision data for moving objects tracking in intersection like scenario," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2011, pp. 362–367.
[5] S. Sonoda, J. K. Tan, H. Kim, S. Ishikawa, and T. Morie, "Moving objects detection at an intersection by sequential background extraction," in *Proc. 11th Int. Conf. Control, Autom. Syst.*, Oct. 2011, pp. 1752–1755.
[6] R. A. Bedruz, E. Sybingco, A. Bandala, A. R. Quiros, A. C. Uy, and E. Dadios, "Real-time vehicle detection and tracking using a mean-shift based blob analysis and tracking approach," in *Proc. IEEE 9th Int. Conf. Humanoid, Nanotechnol., Inf. Technol., Commun. Control, Environ. Manage. (HNICEM)*, Dec. 2017, pp. 1–5.
[7] S. Wang, K. Ozcan, and A. Sharma, "Region-based deformable fully convolutional networks for multi-class object detection at signalized traffic intersections: NVIDIA AICity challenge 2017 track 1," in *Proc. IEEE SmartWorld, Ubiquitous Intell. Comput.*, Aug. 2017, pp. 1–4.
[8] J.-P. Jodoin, G.-A. Bilodeau, and N. Saunier, "Tracking all road users at multimodal urban traffic intersections," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 11, pp. 3241–3251, Nov. 2016.
[9] L. Juntao, L. Bingwu, and H. Lingyu, "Multiple objects segmentation and tracking algorithm for intersection monitoring," in *Proc. 3rd IEEE Conf. Ind. Electron. Appl.*, Jun. 2008, pp. 1413–1416.
[10] P. Babaei and M. Fathy, "Abnormality detection and traffic flow measurement using a hybrid scheme of SIFT in distributed multi camera intersection monitoring," in *Proc. 7th Iranian Conf. Mach. Vis. Image Process.*, 2011, pp. 1–5.
[11] H. Huang, J. Chen, H. Xue, Y. Huang, and T. Zhao, "Time-variant visual attention in 360-degree video playback," in *Proc. IEEE Int. Symp. Haptic, Audio Vis. Environ. Games (HAVE)*, Sep. 2018, pp. 1–5.
[12] T. Sawa, T. Yanagi, Y. Kusayanagi, S. Tsukui, and A. Yoshida, "Seafloor mapping by 360 degree view camera with sonar supports," in *Proc. MTS/IEEE Kobe Techno-Oceans (OTO)*, May 2018, pp. 1–4.
[13] L. Li, Z. Li, M. Budagavi, and H. Li, "Projection based advanced motion model for cubic mapping for 360-degree video," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 1427–1431.
[14] H. Abdelhamid, W. DongDong, C. Can, H. Abdelkarim, Z. Mounir, and G. Raouf, "360 degrees imaging systems design, implementation and evaluation," in *Proc. Int. Conf. Mech. Sci., Electr. Eng. Comput. (MEC)*, Dec. 2013, pp. 2034–2038.
[15] M. Budagavi, J. Furton, G. Jin, A. Saxena, J. Wilkinson, and A. Dickerson, "360 degrees video coding using region adaptive smoothing," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 750–754.
[16] K. Wegner, O. Stankiewicz, T. Grajek, and M. Domanski, "Depth estimation from stereoscopic 360-degree video," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 2945–2948.
[17] F. Duanmu, Y. Mao, S. Liu, S. Srinivasan, and Y. Wang, "A subjective study of viewer navigation behaviors when watching 360-degree videos on computers," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2018, pp. 1–6.
[18] C. Premachandra, S. Ueda, and Y. Suzuki, "Road intersection moving object detection by 360-degree view camera," in *Proc. IEEE 16th Int. Conf. Netw., Sens. Control (ICNSC)*, May 2019, pp. 369–372.
[19] F. Ebner and M. D. Fairchild, "Development and testing of a color space (IPT) with improved hue uniformity," in *Proc. 6th Color Imag. Conf.*, 1998, pp. 8–13.
[20] P. J. Snyder, "Flattening the Earth: Two thousand years of map projections," Univ. Chicago Press, Street Chicago, IL, USA, Tech. Rep., 1993.
[21] C. Premachandra, M. Otsuka, R. Gohara, T. Ninomiya, and K. Kato, "A study on development of a hybrid aerial terrestrial robot system for avoiding ground obstacles by flight," *IEEE/CAA J. Automatica Sinica*, vol. 6, no. 1, pp. 327–336, Jan. 2019.
[22] Y. Xia, S. Ning, and H. Shen, "Moving targets detection algorithm based on background subtraction and frames subtraction," in *Proc. 2nd Int. Conf. Ind. Mechatronics Autom.*, May 2010, pp. 122–125.

[23] C. Premachandra, D. Ueda, and K. Kato, "Speed-up automatic quadcopter position detection by sensing propeller rotation," *IEEE Sensors J.*, vol. 19, no. 7, pp. 2758–2766, Apr. 2019.
[24] P. KaewTraKulPong and R. Bowden, "An improved adaptive background mixture model for realtime tracking with shadow detection," *Proc. 2nd Eur. Workshop Adv. Video Based Surveill. Syst.*, vol. 1, Sep. 2001, pp. 1–5.
[25] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Dec. 1999, pp. 23–25.
[26] A. Nurhadiyatna, W. Jatmiko, B. Hardjono, A. Wibisono, I. Sina, and P. Mursanto, "Background subtraction using Gaussian mixture model enhanced by hole filling algorithm (GMMHF)," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 2013, pp. 4006–4011.
[27] L. Chen, P. Zhu, and G. Zhu, "Moving objects detection based on background subtraction combined with consecutive frames subtraction," in *Proc. Int. Conf. Future Inf. Technol. Manage. Eng.*, Oct. 2010, pp. 545–548.
[28] C. Premachandra, R. Gohara, and K. Kato, "Fast lane boundary recognition by a parallel image processor," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2016, pp. 947–952.
[29] Y. Okamoto, T. U. o. S. Department of Electrical EngineeringGraduate School of Engineering, C. Premachandra, and K. Kato, "A study on computational time reduction of road obstacle detection by parallel image processor," *J. Adv. Comput. Intell. Intell. Informat.*, vol. 18, no. 5, pp. 849–855, Sep. 2014.
[30] C. Premachandra, R. Gohara, T. Ninomiya, and K. Kato, "Smooth automatic stopping for ultra-compact vehicles," *IEEE Trans. Intell. Veh.*, vol. 4, no. 4, pp. 561–568, Dec. 2019.
[31] Y. Yamazaki, C. Premachandra, and C. J. Perea, "Audio-Processing-Based human detection at disaster sites with unmanned aerial vehicle," *IEEE Access*, vol. 8, pp. 101398–101405, 2020.
[32] P. Chinthaka, C. Premachandra, and S. Amarakeerthi, "Effective natural communication between human hand and mobile robot using raspberry-pi," in *Proc. IEEE Int. Conf. Consum. Electron.*, Jan. 2018, pp. 1–3.
[33] K. Kale, S. Pawar, and P. Dhulekar, "Moving object tracking using optical flow and motion vector estimation," in *Proc. 4th Int. Conf. Rel., Infocom Technol. Optim. (ICRITO)*, Sep. 2015, pp. 1–6.
[34] I. Tchikk, "A survey on moving object detection and tracking techniques," *Int. J. Eng. Comput. Sci.*, vol. 6, no. 6, pp. 5212–5215, Apr. 2016.
[35] S. Shantaiya, K. Verma, and K. Mehta, "Multiple object tracking using Kalman filter and optical flow," *Eur. J. Adv. Eng. Technol.*, vol. 2, no. 2, pp. 34–39, 2015.
[36] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME J. Basic Eng.*, vol. 5, pp. 35–45, Mar. 1960.
[37] H. A. Patel and D. G. Thakore, "Moving object tracking using Kalman filter," *Int. J. Comput. Sci. Mobile Comput.*, vol. 10, pp. 326–332, Apr. 2013.
[38] X. Li, K. Wang, W. Wang, and Y. Li, "A multiple object tracking method using Kalman filter," in *Proc. IEEE Int. Conf. Inf. Autom.*, Jun. 2010, pp. 1862–1866.
[39] J. Kim, D. Yeom, and Y. Joo, "Fast and robust algorithm of tracking multiple moving objects for intelligent video surveillance systems," *IEEE Trans. Consum. Electron.*, vol. 57, no. 3, pp. 1165–1170, Aug. 2011.
[40] C. Premachandra, D. N. H. Thanh, T. Kimura, and H. Kawanaka, "A study on hovering control of small aerial robot by sensing existing floor features," *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 4, pp. 1016–1025, Jul. 2020.
[41] M. Tamaki and C. Premachandra, "An automatic compensation system for unclear area in 360-degree images using pan-tilt camera," in *Proc. Int. Symp. Syst. Eng. (ISSE)*, Oct. 2019, pp. 1–4.
[42] I. Yuki, C. Premachandra, S. Sumathipala, and B. H. Sudantha, "HSV conversion based tactile paving detection for developing walking support system to visually handicapped people," in *Proc. IEEE 23rd Int. Symp. Consum. Technol. (ISCT)*, Jun. 2019, pp. 138–142.
[43] M. Tsunoda, C. Premachandra, H. A. H. Y. Sarathchandra, K. L. A. N. Perera, I. T. Lakmal, and H. W. H. Premachandra, "Visible light communication by using LED array for automatic wheelchair control in hospitals," in *Proc. IEEE 23rd Int. Symp. Consum. Technol. (ISCT)*, Jun. 2019, pp. 210–215.
[44] R. C. Dilip, T. Lucas, A. Saarangan, S. Sumathipala, and C. Premachandra, "Making online content viral through text analysis," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2018, pp. 1–6.
[45] A. S. Winoto, M. Kristianus, and C. Premachandra, "Small and slim deep convolutional neural network for mobile device," *IEEE Access*, vol. 8, pp. 125210–125222, 2020.
[46] Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," in *Proc. 17th Int. Conf. Pattern Recognit.*, 2004, pp. 28–31.
[47] Z. Zivkovic and F. van der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," *Pattern Recognit. Lett.*, vol. 27, no. 7, pp. 773–780, May 2006.
[48] S. S. Thenuwara, C. Premachandra, and S. Sumathipala, "Hybrid approach to face recognition system using PCA & LDA in border control," in *Proc. Nat. Inf. Technol. Conf. (NITC)*, Oct. 2019, pp. 9–15.

CHINTHAKA PREMACHANDRA (Member, IEEE) was born in Sri Lanka. He received the B.Sc. and M.Sc. degrees from Mie University, Tsu, Japan, in 2006 and 2008, respectively, and the Ph.D. degree from Nagoya University, Nagoya, Japan, in 2011.

From 2012 to 2015, he was an Assistant Professor with the Department of Electrical Engineering, Faculty of Engineering, Tokyo University of Science, Tokyo, Japan. From 2016 to 2017, he was an Assistant Professor with the Department of Electronic Engineering, School of Engineering, Shibaura Institute of Technology, Tokyo. In 2018, he was promoted to an Associate Professor with the Department of Electronic Engineering, School of Engineering/Graduate School of Engineering and Science, Shibaura Institute of Technology, where he is currently the Manager of the Image Processing and Robotic Laboratory. His laboratory conducts research in two main fields image processing and robotics. His former research interests include computer vision, pattern recognition, speed up image processing, and camera-based intelligent transportation systems, while latter interests include terrestrial robotic systems, flying robotic systems, and integration of terrestrial robot and flying robot.

Dr. Premachandra is a member of IEICE, Japan, SICE, Japan, and SOFT, Japan. He received the FIT Best Paper Award and the FIT Young Researchers Award from IEICE and IPSJ, Japan, in 2009 and 2010. He has served many international conferences and journals as a Steering Committee Member and an Editor. He is the Founding Chair of the International Conference on Image Processing and Robotics (ICIPRoB).

SHOHEI UEDA received the B.Sc. degree in electronics engineering from the Shibaura Institute of Technology, Tokyo, Japan, in 2019.

His research interest includes moving object tracking and their applications in robotics, automated driving, and intelligent transportations systems.

YUYA SUZUKI received the B.Sc. degree in electronics engineering from the Shibaura Institute of Technology, Tokyo, Japan, in 2018.

His research interest includes moving object detection and their applications in automated systems.