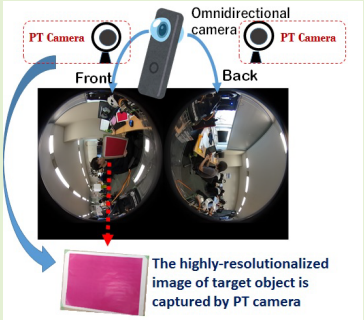
A Hybrid Camera System for High-Resolutionization of Target Objects in Omnidirectional Images

Chinthaka Premachandra, *Senior Member, IEEE*, and Masaya Tamaki

***Abstract*—Recent improvements in cameras and image processing techniques have advanced research on object detection and classification. The cameras capable of capturing omnidirectional images are advantageous for detecting surrounding objects, we focus on their usage method. Omnidirectional camera images generally have lower resolution than available cameras, leading to difficulty in identifying objects farther from the camera. To address this problem, we propose a hybrid camera system that first detects indistinct target regions during omnidirectional viewing, then uses a pan–tilt camera to capture a high-resolution image of the target. We also propose an algorithm for reducing the number of required complementary shots to realize efficient shooting.**



***Index Terms*—Omnidirectional camera, hybrid camera system, high-resolution image capturing, PT camera control, indistinct image region.**

I. Introduction

RECENT improvements in image processing techniques have advanced research on object detection and classification, and we have been engaged in extensive research in that field [1]–[10]. In these studies, a camera photographs an object and image processing is performed to obtain information about that object. Cameras are also used to ensure safety, such as facility surveillance cameras for monitoring vehicles and people. It is possible to process images captured by surveillance cameras to identify danger and then take appropriate measures.

Detection often uses image enhancement technologies to obtain desired information from surveillance camera images, and related research is being actively conducted in this field [11], [12]. These studies assume that the desired information is contained within the field of view of the installed camera. It is therefore necessary to install multiple cameras when the field of view or resolution are likely to be insufficient. Research done by Haifeng *et al.* [13] has attempted to mitigate this problem by proposing a cooperative observation algorithm using pan–tilt–zoom (PTZ) cameras supported by wide-angle cameras; PTZ cameras can capture high-resolution images while wide-angle cameras ensure a wide field of view.

For surveillance cameras, due to the importance of collecting information with a wide field of view, there has been increased interest in the use of 360-degree cameras to monitor surroundings. A 360-degree camera uses a fisheye lens to collect light across a wider range. Generally speaking, use of a fisheye lens can expand a finite field of view to 360 degrees (a hemisphere). Because the field of view of monocular cameras is wide, they are advantageous in that installation costs are low for surveillance systems in open places in comparison with multiple general-purpose cameras. A camera like the one in Fig. 1 having fisheye lenses sandwiching the camera body on either side is called an omnidirectional camera. Theoretically, processing two hemispherical images should make any blind spots extremely small, allowing most of the surrounding area to be captured.

Recent studies of omnidirectional cameras have examined topics such as object recognition in omnidirectional images, coding of omnidirectional images, and combining omnidirectional images [14]–[27], [30]–[35]. In one example of research on object recognition using omnidirectional cameras, Premachandra *et al.* [33] proposed a moving object detection

Manuscript received December 28, 2020; revised February 9, 2021; accepted February 9, 2021. Date of publication February 12, 2021; date of current version April 5, 2021. This work was supported in part by the Branding Research Fund of Shibaura Institute of Technology. The associate editor coordinating the review of this article and approving it for publication was Prof. Aime Lay-Ekuakille. *(Corresponding author: Chinthaka Premachandra.)*

Chinthaka Premachandra is with the Department of Electronic Engineering, School of Engineering/Graduate School of Engineering and Science, Shibaura Institute of Technology, Tokyo 135-8548, Japan (e-mail: chintaka@shibaura-it.ac.jp).

Masaya Tamaki is with the Department of Electronic Engineering, Shibaura Institute of Technology, Tokyo 135-8548, Japan (e-mail: ma18066@shibaura-it.ac.jp).

This article has supplementary downloadable material available at https://doi.org/10.1109/JSEN.2021.3059102, provided by the authors.

Digital Object Identifier 10.1109/JSEN.2021.3059102

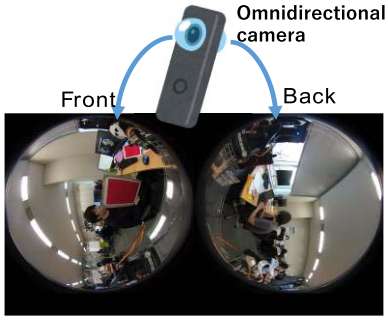


Fig. 1. An omnidirectional camera.

system for road intersections and demonstrated the advantages of using omnidirectional cameras for traffic surveillance. However, they also reported that distant objects are captured at low resolution, making recognition difficult, and that there is thus need for mechanisms preventing erroneous detection.

Increasing camera resolution is one method for utilizing the wide 360-degree field of view of an omnidirectional camera. Budagavi *et al.* pointed out that utilization of the wide field of view in 360-degree images would require increasing video resolutions to at least 4K (3840 × 2160 pixels), consequently increasing bitrate requirements and creating a need for efficient compression technologies [16].

Images captured by omnidirectional cameras are highly affected by lens distortion, so processing of the raw images and videos is difficult. Therefore, 360-degree 3D images are generally converted and projected onto 2D images to facilitate image processing. Li *et al.* [14] proposed addition of a new motion model to the High Efficiency Video Coding (HEVC) models for handling shape distortions in cubic projections, which are often used to develop 360-degree images from omnidirectional cameras, thereby realizing efficient projections.

We noted the need to increase video resolutions to at least 4K in order to mitigate the problem of resolution degradation due to wide fields of view when using omnidirectional cameras for monitoring. However, continuous processing under the high computational loads incurred by tasks such as moving object detection and tracking, not to mention converting 360-degree video at 4K or higher resolutions into 2D images, is infeasible in terms of real-time performance and implementation costs. We therefore considered that a system using images from a conventional omnidirectional camera while using a separate camera for capturing high-resolution images of far-off objects would prevent misidentifications in object detection in monitoring systems using omnidirectional cameras. Some studies in the literature have also proposed camera platforms by combining single omnidirectional camera and single 2D camera, and have studied necessary calibration techniques as well [19], [20].

In this study, we constructed a hybrid camera platform comprising two pan–tilt (PT) cameras in combination with a monocular omnidirectional camera. While we have already published some of the basic ideas behind the proposed platform [21], here we describe its implementation, usefulness, and effectiveness. The developed camera platform identifies regions where recognition is difficult due to low resolution in the omnidirectional image, then points the PT cameras toward regions of interest to capture high-resolution images.

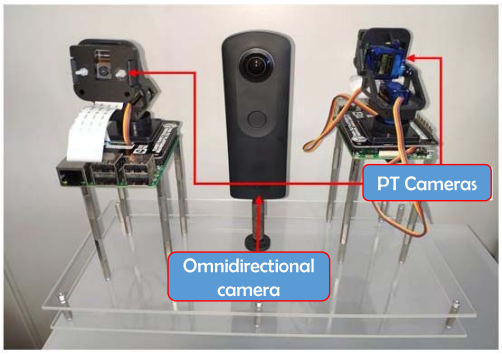


Fig. 2. The fabricated camera platform.

The 2D expansion needed for spherical image processing is normally not performed; rather, the PT cameras capture high-resolution images, allowing highly accurate object identification while keeping computational loads low.

In this paper, we propose a system that allows dynamic processing from recognition in low-resolution regions to capture of high-resolution images. Through multiple imaging experiments, we also confirm that the proposed method effectively acquires high-resolution images of target objects.

This paper comprises several sections to describe this research project in detail. Section 2 describes details of the hardware configuration in our proposed hybrid camera system, and Section 3 describes object recognition in omnidirectional images. Section 4 describes the method by which we capture complementary images with the PT cameras, and Section 5 describes experiments for confirming the effectiveness and utility of the proposed hybrid camera system. Section 6 concludes this paper.

II. Hybrid Camera Platform

Figure 2 shows the fabricated camera platform, which comprises a monocular omnidirectional camera and two PT cameras. The PT cameras are used to allow capture of 180-degree ranges on either side of the omnidirectional camera. All cameras are installed along the same line, with spacers used to position them at the same height. Cameras are situated 20 cm above the platform base, and the PT cameras are each 10 cm to the right and left of the central omnidirectional camera.

We used a Ricoh THETA S (hereinafter, THETA) as the omnidirectional camera, and the PT cameras were Raspberry Pi Camera Modules v2.1 mounted on a Pan-Tilt HAT by Pimoroni Ltd., connected to and controlled by a Raspberry Pi 3 Model B (hereinafter, Raspberry Pi).

The fabricated camera platform is connected to a personal computer for overall control. Figure 3 shows connections between each camera and the computer. We used Python 2.7.13 and OpenCV 2.4.9.1 as the operational environment for control software.

A. System Operation

Figure 4 shows the operational flow for the proposed camera platform. First, we process an omnidirectional image captured by THETA to extract target regions. Barycentric coordinates of the target region are then converted into angle information and transmitted to the corresponding Raspberry Pi. Here, the angle information is the pan and tilt angles necessary

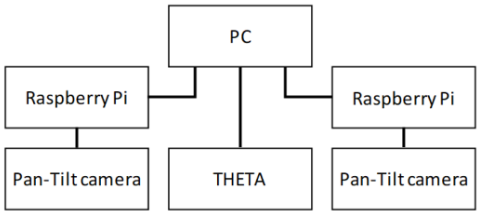


Fig. 3. Connections between each camera and the computer.

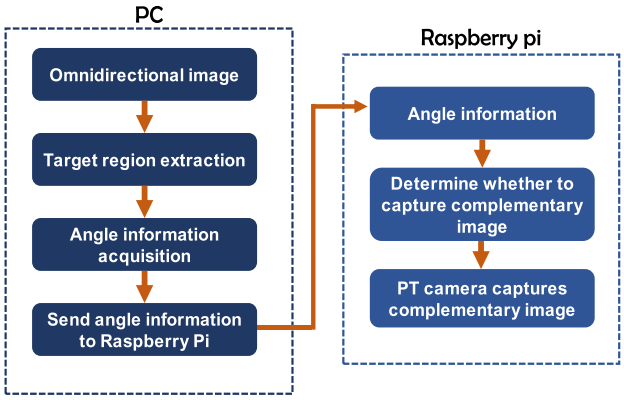


Fig. 4. Acquisition process for complementary images.

Fig. 5. Colored rectangles.

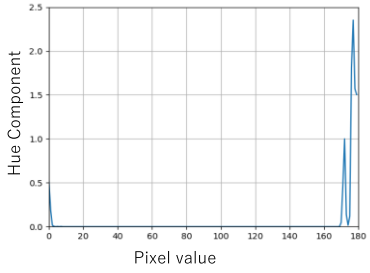
for PT camera operations to capture the target region with the corresponding PT camera within the target region as viewed from THETA. After receiving this angle information, the corresponding Raspberry Pi determines whether to capture a complementary image of the target region, controlling each connected PT camera based on the angle information. The following section describes these processes.

III. Extraction of Objects in Omnidirectional Images

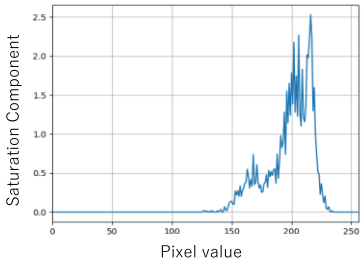
A. Setting the Target Region

When the proposed camera platform is actually used, the detection target will differ according to the application. To facilitate experiments, therefore, we applied it to detection of colored rectangles as the target region for extraction. As detection targets, we used A4- and A3-sized copy paper on which were printed pink rectangles (Fig. 5). The colored rectangles had dimensions 18.5 × 25.5 cm on the A4 paper and 27.2 × 25.5 cm on the A3 paper.

Fig. 6. Color samples.



(a) Hue Component variation



(a) Saturation Component variation

Fig. 7. Combined histogram.

Detection of colored rectangles is performed by converting the RGB omnidirectional images captured by THETA into HSV images. Expression as an HSV color space is intuitive and easy for humans to understand and is suited to detecting single-colored objects as in these experiments. We used Eqs. (1)–(3) to perform RGB-to-HSV conversions [28], [29], [36].

$$\mathrm{H} = \begin{cases} 60 \times \dfrac{G-B}{max-min}, & (max = R) \\ 60 \times \dfrac{B-R}{max-min} + 120, & (max = G) \\ 60 \times \dfrac{R-G}{max-min} + 240, & (max = B) \\ 0 & (R = G = B) \end{cases} \tag{1}$$

$$S = \frac{max - min}{max} \tag{2}$$

$$V = max \tag{3}$$

We next prepared five images of colored rectangles captured by THETA, then extracted regions showing colored rectangles as color samples (Fig. 6).

We then used Eqs. (1)–(3) to convert these color samples to HSV. Next, we converted the output images into histograms for each color component, normalized them, and combined all histograms. Figure 7(a) shows the hue component of the combined histogram, and Fig. 7(b) shows the

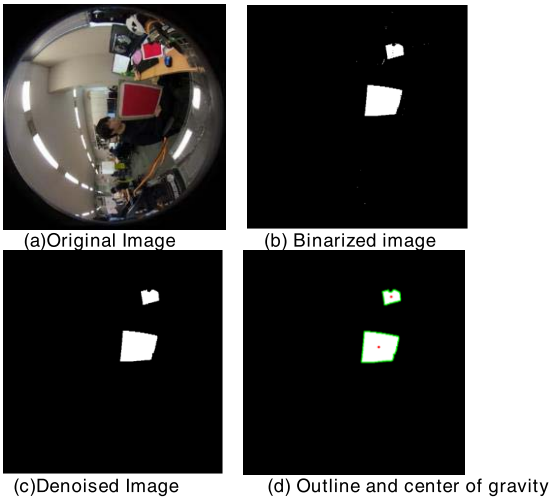


Fig. 8. Process for colored rectangle detection.

saturation component. Note that the brightness component depends on the illumination environment and thus was not used as a target component.

B. Colored Rectangle Detection

The process for detecting colored rectangles is not the focus of this research and thus is only briefly described below.

First, we binarize and display colored rectangle regions from the omnidirectional image making a threshold setting [37]. We then use morphological operations to remove noise components from the binarized image. Next, colored rectangle regions in the denoised image are displayed as outlines. Finally, we apply the image moment function for a binarized image given by Eq. (4) to the contour image, and calculate the center of gravity (C_x, C_y) and the area A by Eqs. (5)–(6). Figures 8(a)–(d) summarize the processing steps for detecting colored rectangles.

$$M_{mn} = \sum_{x,y} \left(x^m \cdot y^n\right) \tag{4}$$

$$(C_x, C_y) = \left(\frac{M_{10}}{M_{00}}, \frac{M_{01}}{M_{00}}\right) \tag{5}$$

$$A = M_{00} \tag{6}$$

IV. PT Camera Capture of Complementary Images

To capture complementary images, we obtained pan and tilt angles for orienting the PT camera toward the target region, with barycentric coordinates of the target object in the omnidirectional image converted into angular coordinates.

THETA in this camera platform uses an equidistant projection method [20], [29]. As shown in Fig. 9, in images captured by equidistant projection, ratios of angles of incidence on the lens and distances to points plotted on the screen (image) are the same.

Letting d be the distance from the screen center to the lens and letting θ be the angle of incidence, the position x on the screen can be calculated as

$$x = d \cdot \theta. \tag{7}$$

From Eq. (7), because the position of the point on the screen and its associated angle are in a correspondence relation,

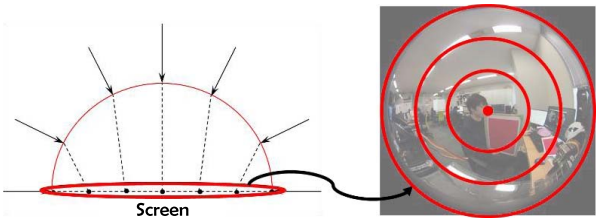


Fig. 9. Correspondence between positions and angles on the screen.

we can convert between. Thus, letting (x_0, y_0) be the origin in the 360-degree image, the angle conversion formulas for coordinate (p, q) from that origin are

$$\begin{cases} p = x - x_0 \\ q = y - y_0 \end{cases} \tag{8}$$

$$\begin{cases} \theta = q \cdot \dfrac{\pi}{2R} \\ \phi = p \cdot \dfrac{\pi}{2R}, \end{cases} \tag{9}$$

where θ is the pan angle, ϕ is the tilt angle, and R is the radius of the 360-degree image.

A. Imaging Priority When Multiple Target Objects Are Detected

The proposed camera platform can take only one photo on each side at a time, so when multiple target objects simultaneously appear, it is necessary to prioritize their imaging order. This prioritization utilizes a characteristic of the 360-degree image, namely, that distortion of the imaged object increases with distance from the center. Therefore, the distance d from the image center is calculated for each target region as

$$d = \sqrt{\theta^2 + \phi^2}, \tag{10}$$

and the imaging priority is arranged in order of increasing distance.

B. Same Region Imaging

In a dynamically operating system, same or nearby regions may be multiply detected as target regions after detecting target regions and before capturing a complementary image. Because it is unnecessary to continuously capture the same areas, in such cases the determination is made based on conditions for preventing redundant imaging of the same area. Specifically, a judgement is made by individually comparing data immediately after acquisition with previous image data. Here, we set m to the number of objects extracted from one image in data immediately after acquisition, and set n to the maximum number of objects fitting in an arbitrary time range for image data not extending too far into the past.

Figure 10 shows a flowchart for performing this judgment, which occurs in two stages: comparison according to the distance between coordinates and then comparison of differences in imaging times. We use D_{th} as the threshold for distance comparisons between coordinates and T_{th} as the threshold for comparisons of differences in shooting times. Individual between-coordinate distances d_i and time differences T_{di} are

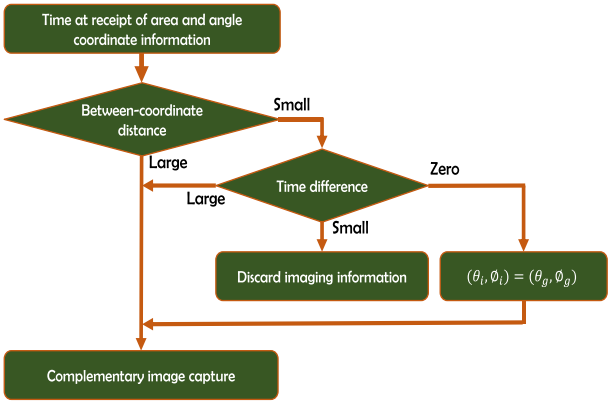


Fig. 10. Preventing repeated same-area imaging.

calculated as

$$d_i = \sqrt{\left|\theta_i - \theta_j^{his}\right|^2 + \left|\phi_i - \phi_j^{his}\right|^2} \tag{11}$$

$$T_{di} = T_i - T_j^{his}. \tag{12}$$

θ_i and ϕ_i indicate current pan and tilt values while θ_j^{his} and ϕ_j^{his} are same values in the previous image capturing. On the other hand, T_i indicates current time and T_j^{his} indicates previous image capturing time. Four patterns can result from following the flowchart in Fig. 10. The first is the case where the coordinates are sufficiently separated to indicate separate objects, so individual imaging is performed. In the second case, the distance between the coordinates is small and the time difference between imaging is large, suggesting that imaging times were different, so individual imaging is performed. In the third case, there are small differences between coordinates and imaging times, suggesting continuous imaging in the same region, so imaging is not performed and images are not recorded as past data. Finally, in the fourth case, there is a small between-coordinate difference and a time difference of zero, suggesting that the coordinates were extracted from images taken at the same time. Since the between-coordinate distance is small and the coordinates are dense, in this case, the dense target regions are collectively imaged by the following method.

When target regions are imaged together, the centers of gravity of each are integrated into one, and that integrated center of gravity is imaged. We can calculate the integrated center of gravity (θ_g, ϕ_g) as

$$\begin{cases} \theta_g = \dfrac{\sum_{i=1}^{m} A_i \cdot \theta_i}{\sum_{i=1}^{m} A_i} \\ \phi_g = \dfrac{\sum_{i=1}^{m} A_i \cdot \phi_i}{\sum_{i=1}^{m} A_i}, \end{cases} \tag{13}$$

where A is the area of each targeted region.

C. Capturing Complementary Images

After making a determination to perform complementary imaging according to the imaging conditions, the corresponding Raspberry Pi controls a servo according to the received pan and tilt angles to orient the camera toward the target region. During continuous imaging, however, the entire image will be blurred if the camera orientation changes too quickly. It is therefore necessary to establish a delay time before exposure immediately after turning the camera. We experimentally set this delay time to 0.5s.

Fig. 11. Camera platform used in the experiments.

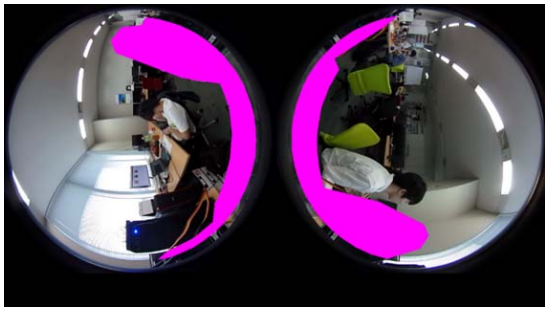


Fig. 12. Field of view of the camera platform.

V. Experiments

We performed four main experiments to confirm performance in four different aspects of the fabricated camera platform, each with the camera platform mounted on a tripod. In addition to the four main experiments, we conducted different experiments to confirm the image capturing performance of the platform by setting the target objects in different locations from the platform (called as *overall capturing experiments*). The camera height above the floor was 115 cm. Figure 11 shows the camera platform used in these experiments, during which we used imaging resolutions of 1280 × 720 pixels for THETA and 640 × 480 pixels for each PT camera.

A. Experiment 1

In experiment 1, we installed a camera platform, captured images using THETA, and confirmed regions in the captured images in which objects could be detected. Using the captured images, we processed blind-spot regions due to the camera platform housing and PT cameras, and calculated the area coverage of the camera platform from pixel counts. Figure 12 shows the field of view of the camera platform, and Table I shows the calculation results. In the Fig. 12, blind-spot regions due to the camera platform housing and PT cameras are shown in pink.

B. Experiment 2

In Experiment 2, we photographed separate complementary images while simultaneously detecting multiple target regions.

TABLE I
REGION COVERAGE BY THE CAMERA PLATFORM

	Formula	Measurement value [pixels]
THETA FOV	FOV	525,700
Blind-spot region	B	99,579
Coverage	$1-\frac{B}{FOV}$	81.1%

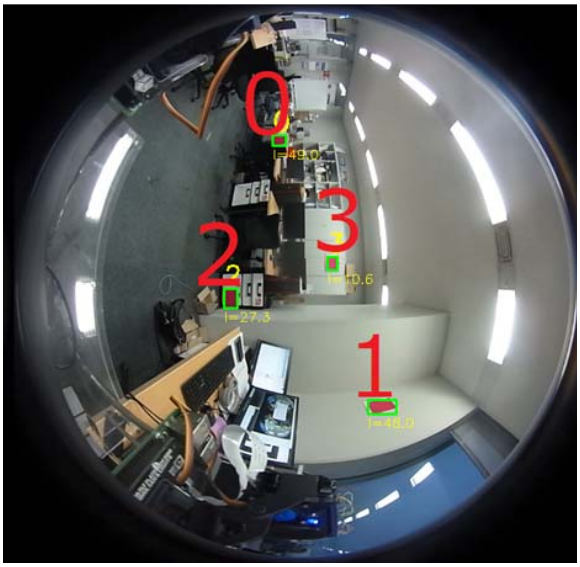
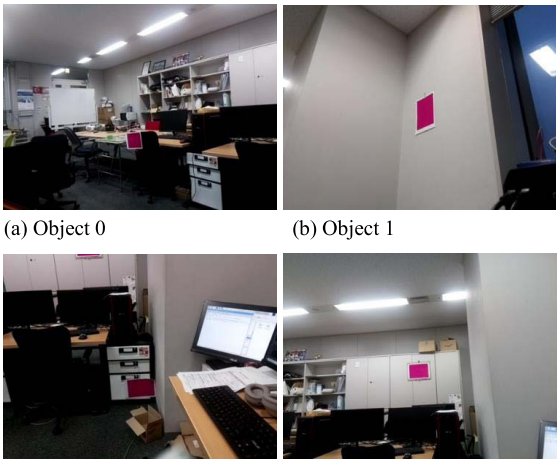


Fig. 13. Detection of multiple target regions.



(a) Object 0 (b) Object 1

(c) Object 2 (d) Object 3

Fig. 14. Complementary images for multiple target regions.

We prepared four colored rectangles, installed them at sufficient separations 2–4 m from the camera and 0.05–2 m above the floor, and captured complementary images.

The results of this experiment showed that THETA detected multiple target regions, and immediately after that, the PT cameras successfully captured individual complementary images at high resolutions. Figure 13 shows the results of detecting multiple target regions, and Figs. 14(a)–(d) show complementary images individually captured by the PT cameras. The submitted supplementary video also shows this capturing situation.

Fig. 15. Image captured by THETA.

(a) (b)

Fig. 16. Center-of-gravity images for multiple target regions. (a) Individual centers of gravity (b) Integrated center of gravity.

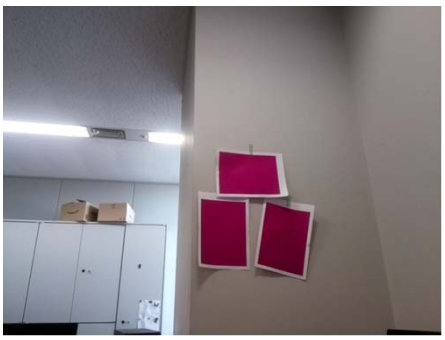


Fig. 17. Complementary captured image by integrating centers of gravity.

C. Experiment 3

In experiment 3, we simultaneously captured grouping of target regions. We prepared three colored rectangles and installed them 2 m from the camera and sufficiently close to each other to form a group. The results of this experiment demonstrated summarized complementary image capturing. Namely, while three captures would have been required for separate complementary images, we successfully captured a complementary image in a single shot. Figure 15 shows an image from THETA, Fig. 16(a) shows centers of gravity for individual target regions, and Fig. 16(b) shows the integrated center of gravity for the three target regions. Figure 17 shows a complementary image from a PT camera taking the integrated center of gravity as the coordinate for the target region.

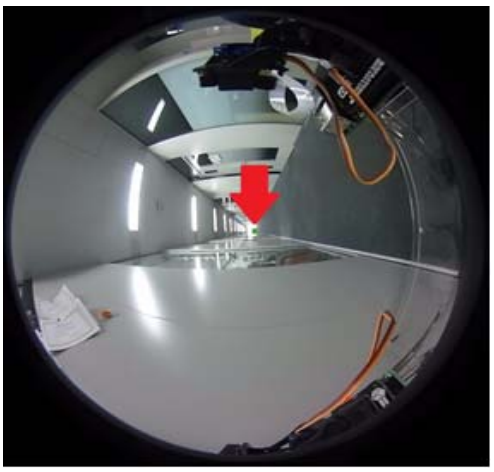


Fig. 18. Example image from THETA.

Fig. 19. Complementary image from a PT camera.

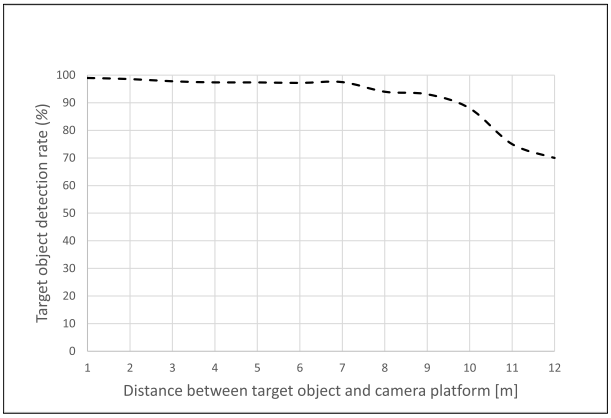


Fig. 20. Target object detection results.

TABLE II
EXPERIMENTAL EVALUATION OF INCREASE IN IMAGE QUALITY

Exp. No.	Object resolution in THETA Image [pixels]	Object resolution in PT camera image [pixels]	Resolution increment by PT camera [pixel]	Ratio of increment [%]
1	36	234	+198	+550.0
2	12	198	+186	+1550.0
3	24	216	+192	+800.0
4	18	207	+189	+1050.0
5	15	204	+189	+1260.0
6	16	204	+188	+1175.0
7	26	219	+193	+742.3
8	31	228	+197	+635.5
9	17	206	+189	+1111.8
10	15	202	+187	+1246.7

D. Experiment 4

In experiment 4, we evaluated the extent to which the proposed camera platform improves image quality. We used the assumed application environment as the experimental environment, with a 10m distance to five target regions placed at $\pm 90^\circ$ from each side of the camera platform, taking the forward direction as 0°. We thus captured ten images in total. As in the previous experiments, camera resolutions were 1280×720 pixels for THETA and 640×480 for each PT camera. Figure 18 shows an example image from THETA, and Fig. 19 shows the complementary image from a PT camera captured at the same time..

E. Overall Target Object Capturing Experiments

In the overall target object capturing experiments, we set the target objects (colored rectangle) in a different location from the camera platform and capturing was conducted. In some cases, capturing was done setting few target objects near each other like in Fig.17.

First, the detection of the target object from the THETA images was evaluated changing the distance between the camera platform and the target objects. Fig. 20 illustrates the results of detection rate. The proposed target object detection method showed a high detection rate when the distance between target object and camera platform is less than 10m. In addition to that, the average false positive rate of the target object detection is 3.9%. These minor false positives occurred due to availability of similar type objects in the environment.

Then, table II illustrates the results of ten complementary images captured by the camera platform in detail. To investigate the extent of image quality improvement, we extracted color components from colored rectangles in each image, counted the actual number of pixels, and calculated the rate of increase in image quality. In Table II, column two illustrates the counted pixel numbers of the original objects in the images from the THETA while column three illustrates counted pixel numbers of their corresponding complementary images, captured by the PT cameras. The column four of Table II shows the pixel number increment while the column five of Table II shows the ratio of the increment. These results show that the average area of target regions in THETA images was 21 pixels, while the average area of target regions in complementary images from PT cameras was 211.8 pixels, indicating average increases of 190.8 pixels, or 1101.2% in terms of average ratio of increase. We therefore demonstrated that high-resolution images of objects appearing at low resolution in omnidirectional images can be obtained when capturing target regions.

Finally, we evaluate the operation of the proposed camera platform by generating a receiver operating characteristic (ROC) curve of overall image capturing experiments, as shown in Fig. 21. This ROC curve was generated regarding

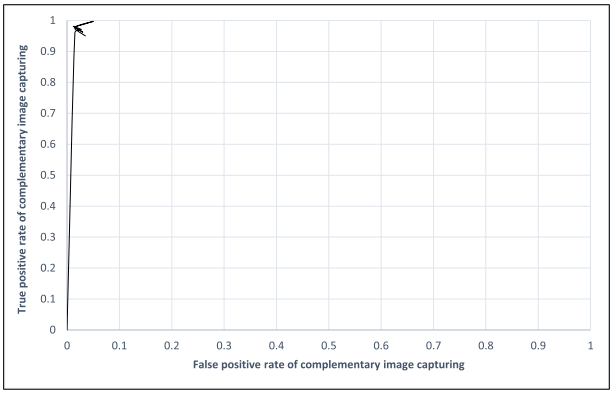


Fig. 21. Receiver Operating Characteristic (ROC) curve of complementary image capturing operations.

only the image capturing operation of the target objects that are detected by proposed object detection method. In the generated ROC curve, false positive rate varies from 0% to 3.5%. Therefore, the false positive rate of complementary image capturing was very low. We believe that these minor false positives occurred due to mechanical errors of servos that move PT cameras.

VI. Conclusion

In this study, sought to address the difficulty of object recognition due to low resolutions of omnidirectional cameras used in the context of object detection. We accomplished this by fabricating a hybrid camera platform consisting of a single omnidirectional camera and two PT cameras. Experimental results demonstrated that use of this platform to capture target regions allowed acquisition of higher-resolution images compared with use of a single omnidirectional camera. A remaining issue is that when a moving object is to be captured as the target region, captured complementary images may shift due to the delay time required for image acquisition by the PT camera. Potential countermeasures for addressing this include introduction of optical flow or a Kalman filter in object detection to predict future coordinates of the target object when capturing images.

Other topics for further development of this research include allowing more dynamic use of the system by enabling automatic object identification in captured complementary images, developing applications that extend the types of targets that can be detected as target regions, and adopting cameras with zoom functionality to further improve image quality.

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] C. Premachandra, R. Gohara, T. Ninomiya, and K. Kato, "Smooth automatic stopping for ultra-compact vehicles," *IEEE Trans. Intell. Vehicles*, vol. 4, no. 4, pp. 561–568, Dec. 2019.

[6] C. Premachandra, M. Murakami, R. Gohara, T. Ninomiya, and K. Kato, "Improving landmark detection accuracy for self-localization through baseboard recognition," *Int. J. Mach. Learn. Cybern.*, vol. 8, no. 6, pp. 1815–1826, Dec. 2017.

[7] H. W. H. Premachandra, C. Premachandra, C. D. Parape, and H. Kawanaka, "Speed-up ellipse enclosing character detection approach for large-size document images by parallel scanning and Hough transform," *Int. J. Mach. Learn. Cybern.*, vol. 8, no. 1, pp. 371–378, Feb. 2017.

[8] 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.

[9] 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.

[10] 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.

[11] N. Chumuang, M. Ketcham, and T. Yingthawornsuk, "CCTV based surveillance system for railway station security," in *Proc. Int. Conf. Digit. Arts, Media Technol. (ICDAMT)*, Feb. 2018, pp. 7–12.

[12] M. Sodanil and C. Intarat, "A development of image enhancement for CCTV images," in *Proc. 5th Int. Conf. IT Converg. Secur. (ICITCS)*, Aug. 2015, pp. 1–4.

[13] L. Haifeng, W. Lufeng, and S. Jieqiong, "Collaborative observation algorithm for multiple moving targets using a PTZ camera assisted by a wide-angle camera," in *Proc. 27th Chin. Control Decis. Conf. (CCDC)*, May 2015, pp. 778–782.

[14] 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.

[15] 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.

[16] 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.

[17] K. Wegner, O. Stankiewicz, T. Grajek, and M. Domański, "Depth estimation from stereoscopic 360-degree video," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 2945–2948.

[18] 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.

[19] Y. Bastanlar, "A simplified two-view geometry based external calibration method for omnidirectional and PTZ camera pairs," *Pattern Recognit. Lett.*, vol. 71, pp. 1–7, Feb. 2016.

[20] Y. Bastanlar, A. Temizel, Y. Yardimci, and P. Sturm, "Multi-view structure-from-motion for hybrid camera scenarios," *Image Vis. Comput.*, vol. 30, no. 8, pp. 557–572, Aug. 2012.

[21] 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.

[22] 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.

[23] A. K. Mulya, F. Ardilla, and D. Pramadihanto, "Ball tracking and goal detection for middle size soccer robot using omnidirectional camera," in *Proc. Int. Electron. Symp. (IES)*, Sep. 2016, pp. 432–437.

[24] G. Pudics, M. Z. Szabo-Resch, and Z. Vamossy, "Safe robot navigation using an omnidirectional camera," in *Proc. 16th IEEE Int. Symp. Comput. Intell. Informat. (CINTI)*, Nov. 2015, pp. 227–231.

[25] E. Acevedo *et al.*, "An implementation of a monocular 360-degree vision system for mobile robot navigation," in *Proc. Int. Conf. Mechatronics, Electron. Automot. Eng. (ICMEAE)*, Nov. 2018, pp. 21–25.

[26] K. Bhongale and S. Gore, "Design of robot navigation monitoring system using image feature analysis and omnidirectional camera images," in *Proc. 2nd Int. Conf. for Converg. Technol. (I2CT)*, Apr. 2017, pp. 405–409.
[27] Y. Sasaki and M. Ishii, "A study of mobile robot localization using omnidirectional images," in *Proc. SICE Annu. Conf. (SICE)*, Aug. 2012, pp. 1035–1039.
[28] X.-N. Zhang, J. Jiang, Z.-H. Liang, and C.-L. Liu, "Skin color enhancement based on favorite skin color in HSV color space," *IEEE Trans. Consum. Electron.*, vol. 56, no. 3, pp. 1789–1793, Aug. 2010.
[29] L. Yan, H. Ye, Y. Wang, and Y. Chang, "A lip localization method based on HSV transformation in smart phone environment," in *Proc. Int. Conf. Signal Process. (ICSP)*, Oct. 2014, pp. 1285–1290.
[30] Y. Wang, Y. Li, D. Yang, and Z. Chen, "A fast intra prediction algorithm for 360-degree equirectangular panoramic video," in *Proc. IEEE Vis. Commun. Image Process. (VCIP)*, Dec. 2017, pp. 1–4.
[31] W. Yang, Y. Qian, J.-K. Kamarainen, F. Cricri, and L. Fan, "Object detection in equirectangular panorama," in *Proc. 24th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2018, pp. 2190–2195.
[32] S. Ono and C. Premachandra, "Generation of panoramic images by two hemispherical cameras independent of installation location," *IEEE Consum. Electron. Mag.*, early access, Oct. 14, 2020, doi: 10.1109/MCE.2020.3031090.
[33] C. Premachandra, S. Ueda, and Y. Suzuki, "Detection and tracking of moving objects at road intersections using a 360-degree camera for driver assistance and automated driving," *IEEE Access*, vol. 8, pp. 135652–135660, 2020.
[34] H. C. Karaimer, I. Baris, and Y. Bastanlar, "Detection and Classification of Vehicles from Omnidirectional Videos Using Multiple Silhouettes," *Pattern Anal. Appl.*, vol. 54, no. 7, pp. 424–426, Aug. 2017.
[35] I. Cinaroglu and Y. Bastanlar, "A direct approach for object detection with catadioptric omnidirectional cameras," *Signal, Image Video Process.*, vol. 10, no. 2, pp. 413–420, Feb. 2016.
[36] Y. Ito, C. Premachandra, S. Sumathipala, H. W. H. Premachandra, and B. S. Sudantha, "Tactile paving detection by dynamic thresholding based on HSV space analysis for developing a walking support system," *IEEE Access*, vol. 9, pp. 20358–20367, 2021.
[37] M. Demirhan and C. Premachandra, "Development of an automated camera-based drone landing system," *IEEE Access*, vol. 8, pp. 202111–202121, 2020.

Chinthaka Premachandra (Senior 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. He is currently the Manager of the Image Processing and Robotic Laboratory. His laboratory conducts research in two main fields: image processing and robotics. The former field includes AI, computer vision, pattern recognition, image processing, and camera-based intelligent transportation systems, while the latter field includes mobile robotics, aerial robotics, and integration of mobile robotics and aerial robotics.

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, respectively. He has served many international conferences and journals as a Steering Committee Member and an Editor, respectively. He is the Founding Chair of the International Conference on Image Processing and Robotics (ICIPRoB), which is technically co-sponsored by the IEEE.

Masaya Tamaki received the B.S. degree in electronic engineering from the Shibaura Institute of Technology, Tokyo, Japan, in 2017, where he is currently pursuing the M.S. degree with the Graduate School of Engineering and Science.

His research interests include home automation systems, image/video processing, and hybrid camera systems.