Object Detection Using Omnidirectional Camera

Hendawan Soebhakti[1], Haidar Afnan Fakhruddin[2], Rizqy Pratama Singarimbun[3] and Nawwafal Dzaki Musthofa[4]

{hendawan@polibatam.ac.id[1], hdrafnan@gmail.com[2], mirzanabil34@gmail.com[3], d.mdzaki1222@gmail.com[4]}

Barelang Robotics Artificial and Intellegence Lab (BRAIL), Department of Electrical Engineering, Politeknik Negeri Batam, Batam, Indonesia[1,2,3,4]

Abstract. Object detection is a technology for detecting objects in images or videos. In this research, we implemented an object detection system using an omnidirectional camera as a crucial source of information in object detection systems, especially in real-time applications. Color spaces, thresholding, and morphological transformations were used as object detection methods, and the omnidirectional camera configuration transformed the view into a panorama. The evaluation results demonstrate that this system can detect objects and their distance in real-time. The conclusion of this study is that the object detection system with an omnidirectional camera, which transforms its view into a panorama, provides high accuracy and wide viewing flexibility.

Keywords: object detection, omnidirectional camera, panoramic

1 Introduction

Object detection is a crucial field in image processing, especially in the context of robot control. However, traditional cameras used in object detection face challenges such as image distortion, limited field of view, and difficulty in detecting objects in hard-to-reach areas. Omnidirectional cameras offer a solution to address these issues [1]. Omnidirectional cameras can capture images from all directions, reducing image distortion and expanding the field of view. Additionally, they can be used in challenging environments like small rooms or obstacle-filled spaces [2].

In order for a robot to engage in activities like playing soccer, it must be able to detect objects using an omnidirectional camera, as it allows the robot to have a 360° field of view without changing the camera's position [3]. In the previous method of detecting objects we used the YOLO V3 real-time object detection algorithm. This method is also used to detect objects such as balls, obstacles, goals. This method has advantages in detecting objects because of its ability to detect in realtime and is able to recognize objects well because it uses three different grid levels to detect objects at different

ICAE 2023, November 07, Batam, Indonesia
Copyright © 2024 EAI
DOI 10.4108/eai.7-11-2023.2342944

scales in the image, so it can detect diverse objects. However, in actual play, the robot moves quite quickly, resulting in vibration of the camera and sometimes making the view on the camera a little blurry which makes objects difficult to detect. In addition, efficiency and computational factors also affect this because, to get good results, it requires a lot and varied data labeling. The computational process in the YOLO method is also quite heavy because it uses the GPU to detect objects in realtime, making battery power consumption decrease quickly. Therefore, the right solution in this case is using the color segmentation with a wide visual perception system. This method is used because the computation process is quite light and efficient and also makes the robot can better understand the positional state of objects compared to a regular camera view. Therefore, in this study, we modify the initial radial or omnidirectional visual perspective into a panoramic view [4].

2 Method

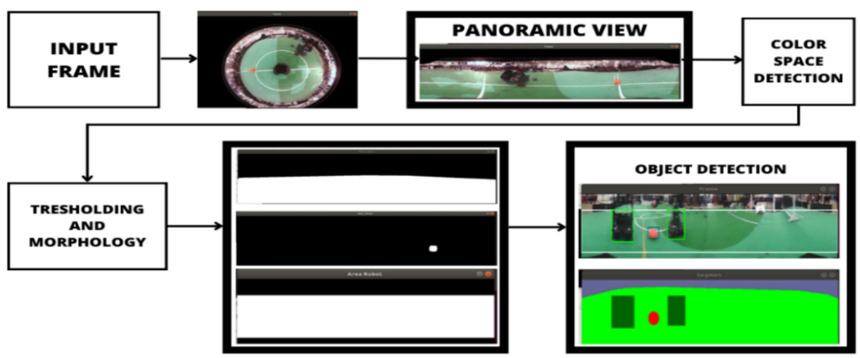

Fig. 1. The framework of the system movement flow.

Image processing is carried out by inputting frames using an omnidirectional camera, followed by transforming the view into a panorama with the aim of determining the position and distance of detected objects. Subsequently, color space values are processed for each detected object. There are differences in the color space transformations for the field area, the ball object, and the robot object. For the field area, a good transformation is achieved in the HSV color space [5], while for the ball object, a good transformation is achieved in the HSL color space [6]. As for the robot object, a good transformation is achieved in the RGB color space [7]. The color space output for the ball object then undergoes thresholding and morphological transformation filtering stages to produce optimal binary values, and the same applies to the robot object [8], [9].

2.1 Omni directional view to panoramic

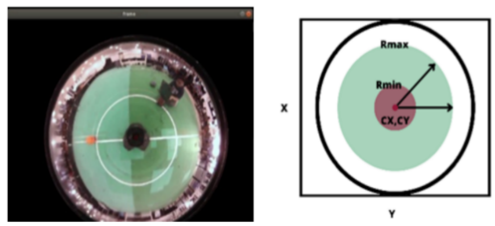

Fig. 2. Input frame, Implementation function in frame.

Cameras play a crucial role in the use of robots as visual sensors to perceive their surroundings. With omnidirectional cameras, they can utilize them as one of the visual sensors that can ensure a 360-degree field of view [10], as shown in Figure 2a. Changes in the camera's perspective are made to encompass the detection of positions, leveraging comprehensive detection range. Thus, the transformation of frames into a panorama shape becomes more optimal by applying formulas that can adjust pixel states as desired [4]. Therefore, in **Fig. 2**. on the right side, the implementation of the function naming to be modified is shown, along with the formula for transforming the shape into a panorama.

$$X_{pano} = 2.0 \times \left(\frac{R_{max} + R_{min}}{2}\right) \times \frac{22}{7} \tag{1}$$

$$Y_{pano} = R_{max} + R_{min} \tag{2}$$

$$R = \left(\frac{Y}{Y_{pano}}\right) \times (R_{max} + R_{min}) + R_{min} \tag{3}$$

$$\theta = \left(\frac{X}{X_{pano}}\right) \times 2.0 \times \frac{22}{7} \tag{4}$$

$$X_{equel} = Cx + R \ \times sin\ sin\ \theta \tag{5}$$

$$Y_{equel} = Cy + R \ \times cos\ cos\ \theta \tag{6}$$

Fig. 3. After applying panoramic.

Determining the position of the frame size value to be transformed involves defining equations 1 and 2, resulting in the height and width of a new frame. In equation 3, the radius area y is generated between R_(max). Meanwhile, in equation 4, the radius area x is obtained from what is generated between R_(min). Thus, from these equations, the values of X_equal and Y_equal can be determined, which will undergo a frame condition change [4], leading to frame transformation after going through functions from 1 to 6, as shown in **Fig. 3.**

2.1 Object detection

Determining the detected objects using color space serves as a method for selecting colors to define, create, and visualize individual colors. The color space process is subjective and inherently variable; its purpose is to describe colors between different hues and standardize them [7], [11]. There is also a thresholding method that achieves sensitivity in color space outcomes for the transformed frame values. Thresholding is a technique that distinguishes objects from the background by detecting differences in darkness or brightness [9], [11]. Additionally, there is a method called morphological transformation to assist in identifying target regions. Morphological transformation is a binary image process that alters the image's perception as a processing tool. Morphological transformation can change pixel values, pixel value comparisons, and alter the original structure of the frame [12]. Morphological transformation consists of several parts, namely, erosion and dilation [11]. Erosion is the process of removing object boundaries to make them part of the background based on the texture element used; this function reduces the size of objects by eroding binary values. Meanwhile, dilation is the process of merging background points into an object based on the texture element used; this function enlarges the size of objects by adding binary values [13].

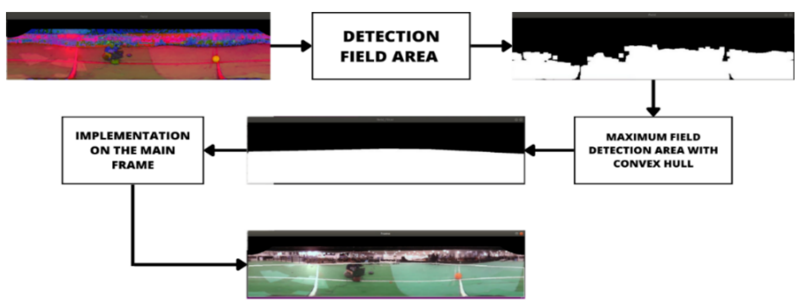

Fig. 4. Convert to HSV frame for field detection, Biner field detection, Maximize the result obtained from Thresholding, and Morphological Transformation in the field detection, The result implemented on the frame.

Field detection involves identifying the robot's object detection coverage area using the HSV color space method to locate green color. This generates binary values corresponding to **Fig. 4.** area. However, using HSV alone seems inadequate, as some expected values are missing. To address this, thresholding and morphological transformations adjust binary values, converting large to small and vice versa. These values are then transformed into a convex hull, enhancing object detection in the selected area **Fig. 4**.

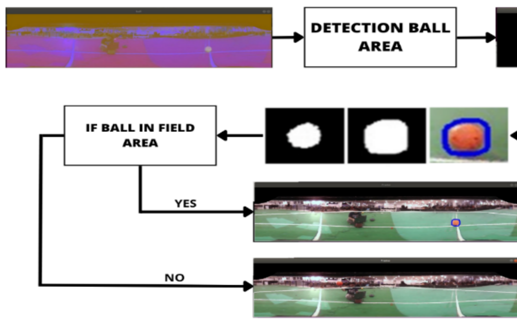

Fig. 5. Convert to HSL frame for ball detection, Biner ball detection, The result biner of the ball, Maximize the result obtained from Thresholding, and Morphological Transformation in the ball detection, Result on frame, The position of the ball in the field, The position of the ball outside the field.

In this phase, the ball's position within the frame is determined using the HSL color space method, chosen for superior results. Maximizing binary values in the HSL color space corresponds to the area in **Fig. 5.** Adjusting binary values through thresholding and morphological transformations enhances sensitivity, transitioning between large and small values. The maximized ball detection combines with the field area, setting the boundary for ball detection as the maximum ball area. If the ball falls outside the field area, it remains undetected, as illustrated in **Fig. 5**.

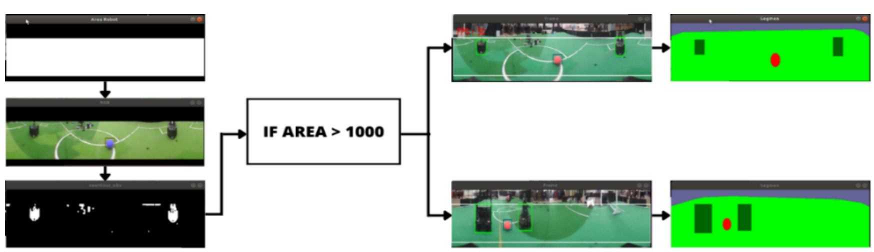

Fig. 6. Robot object detection flowchart.

In robot detection, the RGB color space excels, maximizing binary values as depicted in Fig. 5. These values undergo enhancement via thresholding and morphological transformations, akin to ball object detection. Specific boundaries are set to exclude robot look-alikes, as shown in Fig. 6. Objects under 1000 in area are considered undetected. Combining robot and field area boundaries forms the ultimate detection area, as in Fig. 6.

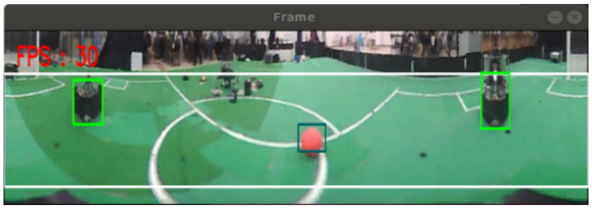

Fig. 7. Frames per second (FPS) on all object detection processing.

3 Results and Discussion

Fig. 8. (a) The size of the x and y frames as a reference for object distance, (b) Determines the value of the detected object against the pixel value.

Testing an object detection system with an adjustable field-of-view omnidirectional camera in panoramic mode has significantly improved accuracy and overall effectiveness. After object detection, additional steps are crucial for the robot to determine the position and distance of detected objects. To establish a practical global point distance metric, the frame's area approximates the pixel frame value (Fig. 7a). Data from detected objects is condensed into a single value, correlated with the pixel frame (Fig. 7b), estimating object distance relative to the pixel frame. This ensures precise distance estimation by comparing pixel frame data with actual distances. Ball object detection spans 60 centimeters (camera's minimum) to 700 centimeters, while robot object detection ranges from 60 to 300 centimeters due to detection area constraints.

Table 1. Estimated ball detection

Actual Distance (cm)	Pixel Frame	Ball Distance Regression (cm)	Estimate Ball Distance Error (cm)	Detect or No Detect
0	0	0	0	No Detect
60	15	64,07	4,07	Detect
100	20	105,12	5,12	Detect
150	25	152,25	2,25	Detect
200	31	205,12	5,12	Detect
250	36	253,5	3,50	Detect
300	41	304,64	4,64	Detect
350	45	350,24	0,24	Detect
400	50	402,55	2,55	Detect
450	55	459,27	9,27	Detect
500	59	505,39	5,39	Detect
550	64	540,24	-9,76	Detect
600	68	598,46	-1,54	Detect
650	71	656,43	6,43	Detect
700	75	707,25	7,25	Detect

Table 2. Estimated Robot Detection

Actual Distance (cm)	Pixel Frame	Robot Distance Regression (cm)	Estimate Robot Distance Error (cm)	Detect or No Detect
0	0	0	0	No Detect
60	15	64,87	4,87	Detect
100	20	105,12	5,12	Detect
150	25	152,25	2,25	Detect
200	31	196,12	-3,88	Detect
250	36	243,61	-6,39	Detect
300	41	294,22	-5,78	Detect

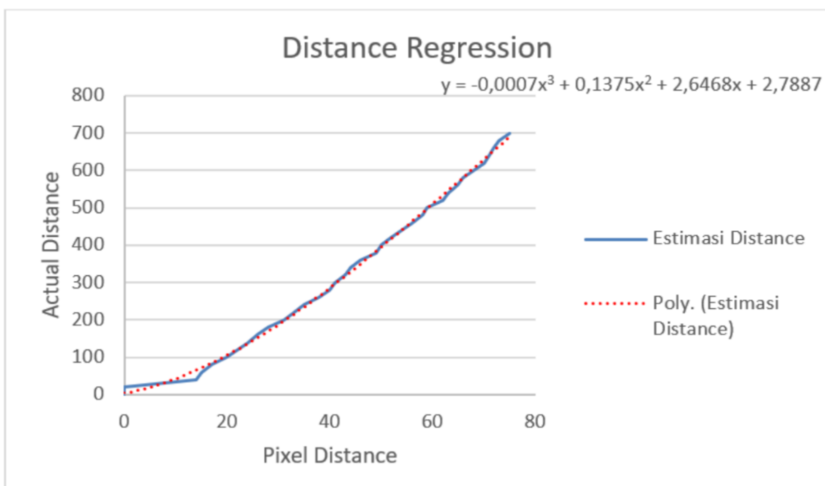

Fig. 9. Comparison graph of the actual distance and the pixel distance of the object.

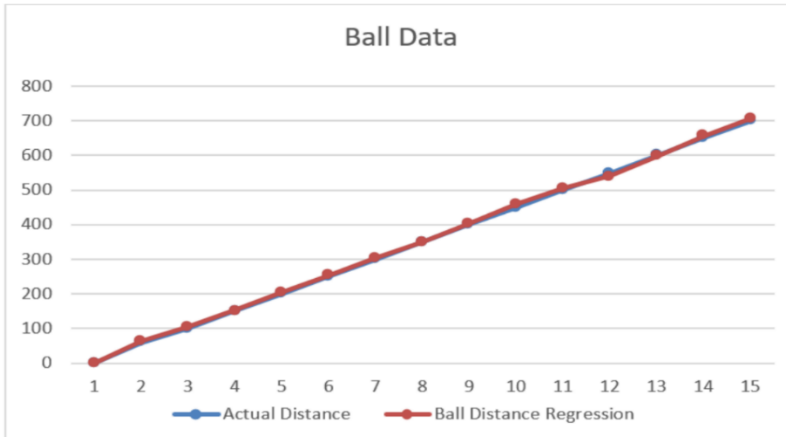

Fig. 10 . Comparison chart of actual distance and ball distance.

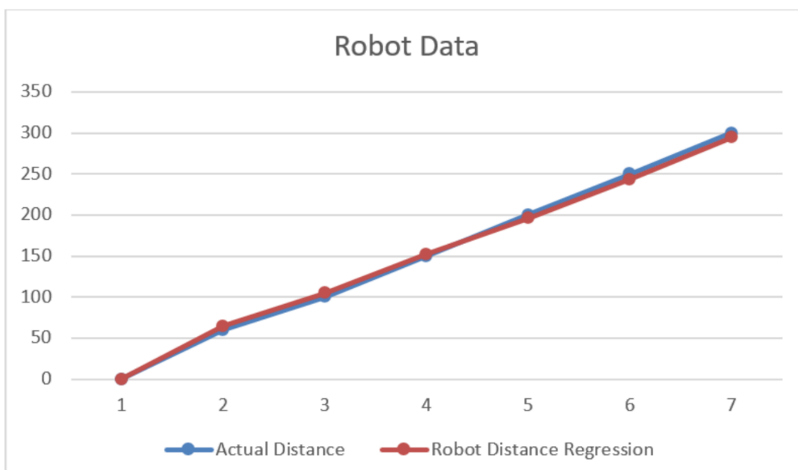

Fig. 11. Comparison chart of actual distance and robot distance.

4 Conclusion

The results obtained from testing an object detection system using an omnidirectional camera and employing color space methods, thresholding techniques, and morphological transformations indicate that this system exhibits a high level of accuracy and achieves rapid processing with a frame rate of up to 30 FPS (Frames Per Second) during image processing. Moreover, this system demonstrates robust performance in mitigating the issues of false positives and false negatives, which are often encountered in object detection systems.

References

[1] C. Premachandra and M. Tamaki : A Hybrid Camera System for High-Resolutionization of Target Objects in Omnidirectional Images, IEEE Sens J, vol. 21, no. 9, pp. 10752–10760, May 2021, doi: 10.1109/JSEN.2021.3059102.

[2] E. A. Nurrohmah, B. Sena Bayu, M. M. Bachtiar, I. Kurnianto Wibowo, and R. Adryantoro : Detecting Features of Middle Size Soccer Field using Omnidirectional Camera for Robot Soccer ERSOW, in Proceeding - ICoSTA 2020: 2020 International Conference on Smart Technology and Applications: Empowering Industrial IoT by Implementing Green Technology for Sustainable Development, Feb. 2020. doi: 10.1109/ICoSTA48221.2020.1570615971.

[3] IEEE Staff, 2019 International Conference on Electrical, Electronics and Information Engineering (ICEEIE). IEEE, 2019.

[4] Guo li ji nan guo ji da xue, IEEE Consumer Electronics Society, IEEE Consumer Electronics Society. Taipei Chapter, and Institute of Electrical and Electronics Engineers, 2016 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW).

[5] H. Zhang, S. Qu, H. Li, J. Luo, and W. Xu : A Moving Shadow Elimination Method Based on Fusion of Multi-Feature, IEEE Access, vol. 8, pp. 63971–63982, 2020, doi: 10.1109/ACCESS.2020.2984680.

[6] Z. Zhong, T. Wang, K. Zeng, X. Zhou, and Z. Li : White Blood Cell Segmentation via Sparsity and Geometry Constraints, IEEE Access, vol. 7, pp. 167593–167604, 2019, doi: 10.1109/ACCESS.2019.2954457.

[7] S. Zhi, Y. Cui, J. Deng, and W. Du : An FPGA-based simple RGB-HSI space conversion algorithm for hardware image processing, IEEE Access, vol. 8, pp. 173838–173853, 2020, doi: 10.1109/ACCESS.2020.3026189.

[8] G. Boato, D. T. Dang-Nguyen, and F. G. B. de Natale: Morphological Filter Detector for Image Forensics Applications, IEEE Access, vol. 8, pp. 13549–13560, 2020, doi: 10.1109/ACCESS.2020.2965745.

[9] A. A. Ewees, M. Abd Elaziz, M. A. A. Al-Qaness, H. A. Khalil, and S. Kim : Improved Artificial Bee Colony Using Sine-Cosine Algorithm for Multi-Level Thresholding Image Segmentation, IEEE Access, vol. 8, pp. 26304–26315, 2020, doi: 10.1109/ACCESS.2020.2971249.

[10] Winarno, A. S. Agoes, E. I. Agustin, and D. Arifianto : Object detection for KRSBI robot soccer using PeleeNet on omnidirectional camera, Telkomnika (Telecommunication Computing Electronics and Control), vol. 18, no. 4, pp. 1942–1953, 2020, doi: 10.12928/TELKOMNIKA.V18I4.15009.

[11] H. Li, X. Zhao, A. Su, H. Zhang, J. Liu, and G. Gu : Color Space Transformation and Multi-Class Weighted Loss for Adhesive White Blood Cell Segmentation, IEEE Access, vol. 8, pp. 24808–24818, 2020, doi: 10.1109/ACCESS.2020.2970485.

[12] Z. Gong, N. Qin, and G. Zhang : Three-Dimensional Generalized Discrete Fuzzy Number and Applications in Color Mathematical Morphology, IEEE Access, vol. 9, pp. 25405–25421, 2021, doi: 10.1109/ACCESS.2021.3057623.
[13] S. Zhao, C. Chen, and Y. Luo : Probabilistic Principal Component Analysis Assisted New Optimal Scale Morphological Top-Hat Filter for the Fault Diagnosis of Rolling Bearing, IEEE Access, vol. 8, pp. 156774–156791, 2020, doi: 10.1109/ACCESS.2020.3019638.
[14] X. Wei, J. Li, B. Long, X. Hu, H. Wu, and H. Li : Method for Vortex Shape Retrieval and Area Calculation Based on Convex Hull Algorithm, IEEE Access, vol. 9, pp. 2706–2714, 2021, doi: 10.1109/ACCESS.2020.3046651.