Improving The Stereo Distance Measurement Accuracy on The Barelang-FC Humanoid Robot

Winarti [1], Susanto[2], Riska Analia[3] , Eko Rudiawan Jamzuri[4]

{ winarti0899@gmail.com [1], susanto@polibatam.ac.id [2], riskaanalia@polibatam.ac.id[3] , ekorudiawan@polibatam.ac.id [4] }

Department of Electrical Engineering, Politeknik Negeri Batam, Batam, Indonesia

Abstract. Distance estimation is essential in developing humanoid soccer robots. Accurate distance measurement can minimize an error while the robot is maneuvering, chasing a ball, or passing the ball to the proponent robots. Currently, stereo vision and feature matching is the conventional method to estimate the distance. Distance is estimated based on the disparity value between detected features on the stereo image. However, the matching process needs high cost computationally. Furthermore, the estimated distance based on feature matching is less accurate. Therefore, in this work, the distance estimation based on the object coordinates detected using the YOLOv3 has been proposed. Additionally, a linear regression algorithm added to improve the measurement accuracy. Several experiments have been done to verify this proposed method in real-time applications. As a result, our proposed method successfully improves the distance measurement accuracy from 86.58% to 98.01%.

Keywords: humanoid robot, object detection, stereo distance measurement, YOLOv3, linear regression.

1 Introduction

A humanoid robot is a robot that popularly competes in robotics contests such as Robocup. In this contest, robots are competed to play football. The robot must recognize the ball, goal, opponent, and proponent. Furthermore, from the recognized objects, the robot must dribble the ball, kick the ball, and determine the target kick to the goal opponent. Every humanoid robot has an object recognition system and a distance estimation to overcome this situation. Distance measuring is essential to ensure how far the robot is to the ball or the proponent robots. Active and passive methods can be utilized for estimating the distance. Actively measuring distance can be done by sending a signal directly to the object. This measurement can be performed using a laser or a radio signal. Meanwhile, passive object measurement utilizes images from the camera sensor, which usually uses a stereo camera [1]. The stereo vision technique has several advantages, such as high levels of precision, efficiency, and automation [2]. Estimating the distance using a camera sensor is the right preference because, in addition to measuring distances, the camera sensor can also identify specific objects. The camera that can be used to

ICAE 2022, October 05, Batam, Indonesia

DOI 10.4108/eai.5-10-2022.2327753

measure distance is a stereo camera. This camera produces two angles of view from the right and left lenses. So need a way to combine images or find the middle point of the image [3].

Determining the distance is needed so the robot can estimate the distance of the object to be addressed. Determining the distance can be done if the final result of the detected object with specific characteristics has been obtained on both cameras. In a study conducted, an image matching algorithm was used to recognize specific patterns [4]. However, this usage requires a long computation time and extensive memory usage. In addition, this study states that the proposed method can only be used in the same environment and camera parameters. A deep neural network model has been proposed [5] to simultaneously predict the semantic information and depth image to improve the traditional stereo matching algorithm. In other research [6], the overlapping area is measured by two cameras when the object was located outside the optical axes to measure the object distance. In this work, they also employed the matching method to understand each image's angle and verified it in simulation, and in [7], they utilized a frequency domain from the captured image and implemented it on a stereo camera to enhance the object distance estimation. Moreover, to estimate the distance for autonomous tomato harvesting [8], they implemented the YOLO (You Only Look Once) deep learning method for detecting the object and utilized the OpenCV library Stereo SGBM algorithm to estimate the distance of each tomato.

In contrast with previous works mentioned above, a simple linear regression is proposed algorithm to help the stereo camera improve the distance estimation. The ZED Mini stereo camera has been used in this work which is mounted on the humanoid robot called Barelang-FC. The distance measured is the ball distance towards the robot. First, detect the ball position and collect the coordinate using YOLOv3 [9]. Then, this coordinate will be used to calculate the distance afterward.

2 Materials and Method

This research is not only focused on how to improve the distance measurement from the stereo camera in terms of calculating the real-time distance from the robot to the object (ball) but also detect the ball in a parallel process. The whole block diagram system can be seen in **Fig. 1**, where the ZED Mini stereo camera employed as the robot vision sensor in this work. The ZED Mini stereo camera includes two HD720-resolution frames (right and left) cameras.

At a glance, when the ZED camera captured the ball with its two cameras, object detection and distance optimization performed afterward so it could detect the ball and measure its distance of the ball at the same time. The YOLOv3 has been used as the object detection method, then a linear regression equation added to optimize the distance estimated from the ZED Mini camera. And then, this detection dan measurement results will be implemented on the humanoid Barelang-FC. This section explains the hardware design, object detection, and distance optimization into three sub-sections.

2.1 Object Detection

The YOLOv3 [10] deep learning has been used as the object detection method in this work. The detection architecture is illustrated in **Fig. 2**. After the vision system captured the ball on the field, the object detection method is performed to detect the ball. Seven convolution layers and six max-pooling layers has been the object detection method performed to detect the ball. In addition, this compressed network can detect objects from up to 80 different classes [11]. The output prediction of this object detection will be the bounding box of the ball detection and

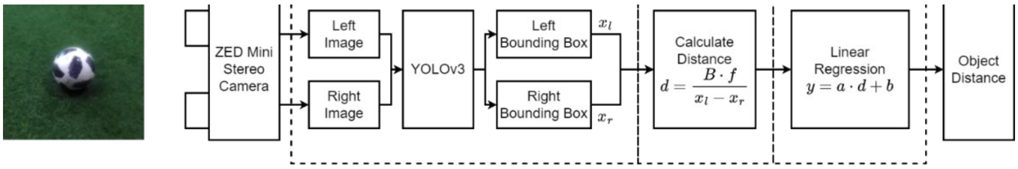


Fig. 1. Block diagram of ball detection and distance measurement.

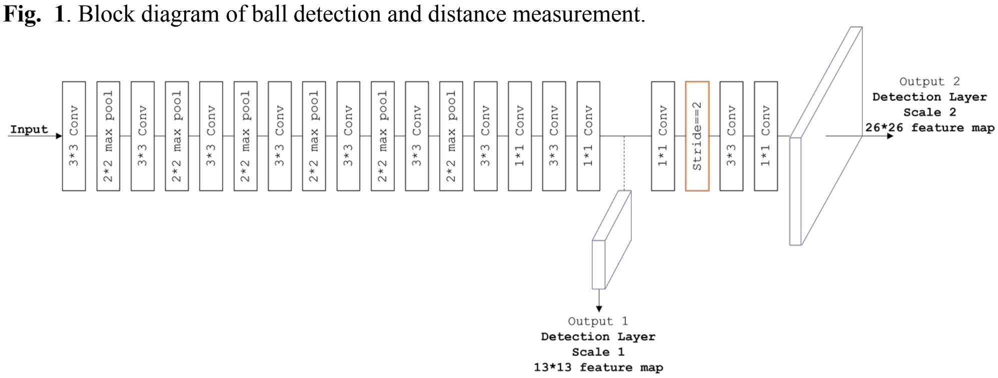
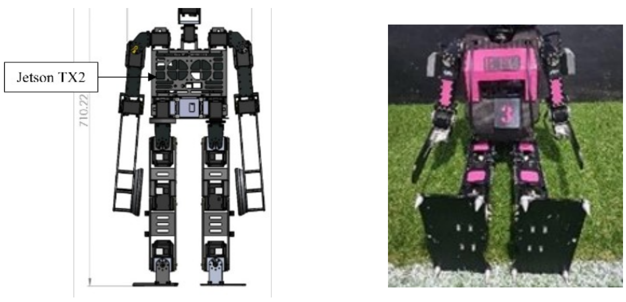


Fig. 2. The YOLOv3 architecture design.



(a) (b)

Fig. 3. BarelangFC robot. (a) the hardware design, and (b) the prototype.

the coordinate of the ball. After detecting the ball, it is used to estimate the distance between the robot and the ball in real-time.

2.2 Hardware Design

The design of Barelang-FC is represented in **Fig. 3**, where **Fig. 3** (a) describes the mechanical design of the robot and **Fig. 3** (b) the prototype of Barelang-FC. The robot design

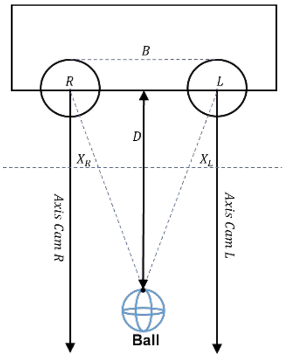


Fig. 4. The geometric model principle for distance estimation of ZED Mini.

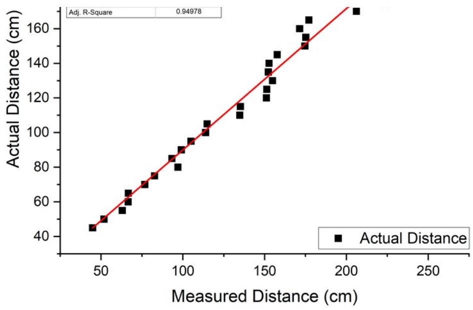


Fig. 5. The linear regression graph.

has been changed in about three years [12-14]. Where in our brand-new design, the ZED Mini stereo camera as the robot vision is utilized, and the robot height and width are about 71 cm × 26 cm. the ZED Mini stereo camera has 1280×720 resolution with a Field of View of 54° for vertical and 85° for horizontal and the focal length of 700 pixels for both cameras. This robot is also equipped with NVIDIA Jetson TX2 and robotic servo 106 for the actuators. The servos configuration generated 20 DOF in total; 12 DOF for the legs joint, 6 DOF for the arms, and 2 for the neck joint. As for the robot body parts, aluminum is used and produced them using the CNC machine.

2.3 Distance Measurement and Linear Regression Fitting

The principle of distance estimation from the ZED Mini camera is illustrated in **Fig. 4**. The distance estimation can be determined from the depth point D calculated from Equation (1). B for the baseline, f for the focal length, and $X_L - X_R$ for the left and right camera horizontal coordinates.

$$D = \frac{Bf}{X_L - X_R} \tag{1}$$

Estimating the distance using a stereo camera depends on the distance between the camera and the object. When the object is too close to the camera, the distance measurement cannot estimate correctly. Therefore, in this work, linear regression is added for optimizing the distance calculation to lessen the error results while detecting dan measuring the ball.

In order to determine the regression equation, several experiments to collect the distance between the actual distance and the distance result from the stereo camera has been done. The actual distance value with the distance from the stereo camera to measure the distance from 45 cm to 200 cm with about 35 sample data are used for comparison. The linear regression graph is illustrated in **Fig. 5**, where the regression equation can be determined using equation (2). The Y here represented the output of our proposed distance estimation, while D was the distance prediction from the stereo camera.

$$Y = 0.816344563000542 \cdot D + 8.26903136660302 \tag{2}$$

3 Experimental Results

All the experiments in this work were conducted in real-time application with the vision embedded in the Barelang-FC. The first experiment was to verify the object detection method to detect the ball. In this experiment, the robot detects a white ball on a field several meters away. As presented in **Fig. 6**, the robot could detect the ball accurately and produce the coordinate of the ball directly while detecting the ball.

The other experiment, the robot is allowed to understand the distance from it to the ball. In this experiment, the robot placed about 159 cm toward the ball and ordered the system to estimate the distance directly. The result can be seen in **Fig. 7**, where in this figure, the result estimation from the stereo camera and our improvement strategy was available by the system. As shown in **Fig. 7**, the stereo camera estimates the distance at about 183 cm, and our system produces approximately 158 cm to estimate the distance in real-time application.

In order to make sure that our distance measurement has improved compared to the original stereo camera results, several experiments has been done to compare the result of these two and plot them on the graph. In this experiment, the robot was let detect the ball and estimate its distance, ranging from 45 cm to 200 cm. The result of this experiment is represented in **Fig. 8**, where the red one is the result of our proposed improvement, and the grey one is the result of the stereo camera. The addition of this linear regression algorithm increases the accuracy of the distance measurement results from 86.58% using stereo vision to 98.01% when optimized use linear regression. As shown **Fig. 9**, the prediction results are generated by our proposed distance optimization almost close to the actual value. The error comparison is described in **Fig. 9**, where the blue line are proposed system, and the stereo camera results are orange. As can be see from the error comparison result, the stereo camera generated a more significant error than our proposed system to predict the ball distance in a real-time application.

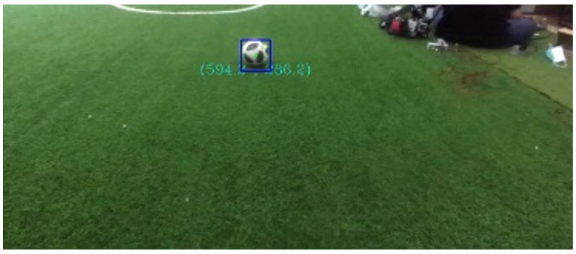


Fig. 6. The results of detecting the ball.

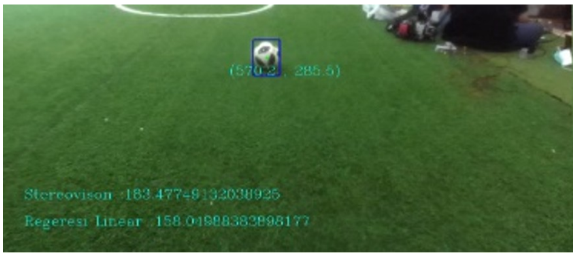


Fig. 7. The result of distance estimation from the robot to the ball.

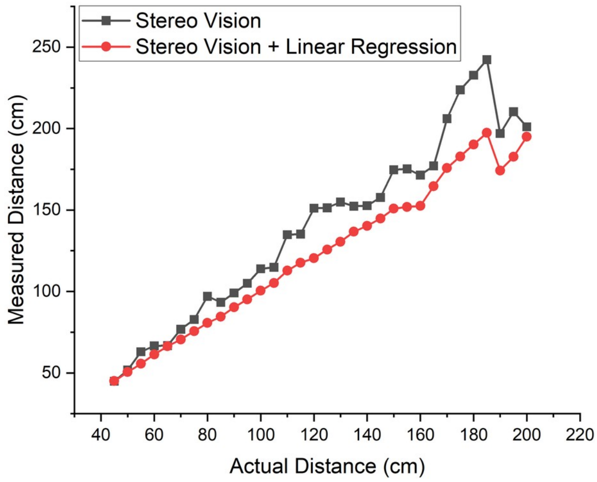


Fig. 8. The distance measurement comparison results.

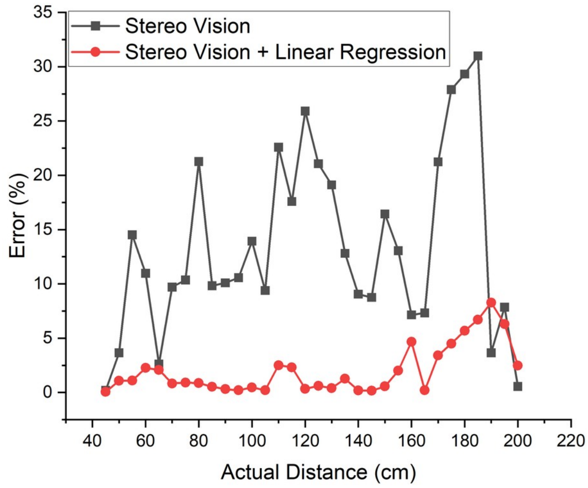


Fig. 9. The error comparison produced by the distance estimation.

4 Conclusion

This paper presented a modest equation to improve the distance prediction from the stereo camera. The linear regression equation has been used to cop the distance estimation. Around 35 sample data were collected to estimate the distance from the stereo camera and compare it to the actual distance value to get the linear regression equation. This work also used the stereo camera to detect the ball using the YOLOv3 deep learning. To verify our proposed system, several experiments has been done in real-time applications. From the experiment, it shows that the proposed system could estimate the distance almost the same as the actual distance and precisely detect the ball simultaneously. In the future, we will use this proposed method to estimate the distance between two robots so that each robot can pass the ball to its team member.

References

[1] J. Mrovlje and D. Vran, "Distance measuring based on stereoscopic pictures," *9th Int. PhD Work. Syst. Control Young Gener. Viewp.*, vol. 2, no. October, pp. 1–6, 2008.

[2] J. Yang, J. Qian, J. Yang, and Z. Fu, "Research on computer aided surgical navigation based on binocular stereovision," *2006 IEEE Int. Conf. Mechatronics Autom. ICMA 2006*, vol. 2006, pp. 1532–1536, 2006, doi: 10.1109/ICMA.2006.257402.

[3] E. Dandil and K. K. Cevik, "Computer Vision Based Distance Measurement System using Stereo Camera View," *3rd Int. Symp. Multidiscip. Stud. Innov. Technol. ISMSIT 2019 - Proc.*, pp. 6–9, 2019, doi: 10.1109/ISMSIT.2019.8932817.

[4] N. Q. Ann *et al.*, *SKF-based image template matching for distance measurement by using stereo vision*, vol. 0, no. 9789811087875. Springer Singapore, 2018. doi: 10.1007/978-981-10-8788-2_39.

[5] H. Wu, H. Su, Y. Liu, and H. Gao, "Object Detection and Localization Using Stereo Cameras," *ICARM 2020 - 2020 5th IEEE Int. Conf. Adv. Robot. Mechatronics*, pp. 628–633, 2020, doi: 10.1109/ICARM49381.2020.9195365.

[6] H. S. Baek, J. M. Choi, and B. S. Lee, "Improvement of distance measurement algorithm on stereo vision system(SVS)," *2010 Proc. 5th Int. Conf. Ubiquitous Inf. Technol. Appl. CUTE 2010*, no. 1, pp. 4–6, 2010, doi: 10.1109/ICUT.2010.5678176.

[7] I. S. Yoo and S. W. Seo, "Object distance estimation based on frequency domain analysis using a stereo

camera," *2015 IEEE Int. Conf. Consum. Electron. ICCE 2015*, pp. 343–344, 2015, doi: 10.1109/ICCE.2015.7066437.

[8] Z. Buribayev, T. Merembayev, Y. Amirgaliyev, and T. Miyachi, "The Optimized Distance Calculation Method with Stereo Camera for an Autonomous Tomato Harvesting," *IEEE Access*, 2021, doi: 10.1109/SIST50301.2021.9465944.

[9] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," 2018, [Online]. Available: http://arxiv.org/abs/1804.02767

[10] W. Fang, L. Wang, and P. Ren, "Tinier-YOLO: A Real-Time Object Detection Method for Constrained Environments," *IEEE Access*, vol. 8, pp. 1935–1944, 2020, doi: 10.1109/ACCESS.2019.2961959.

[11] J. Gong, J. Zhao, F. Li, and H. Zhang, "Vehicle detection in thermal images with an improved yolov3-tiny," *Proc. 2020 IEEE Int. Conf. Power, Intell. Comput. Syst. ICPICS 2020*, pp. 253–256, 2020, doi: 10.1109/ICPICS50287.2020.9201995.

[12] Susanto, R. Saputra, and R. Analia, "A Control Strategy to Estimate the Robot Position Of Barelang-FC Striker," *IEEE Access*, p. 5, 2019, doi: 10.1109/ICAE47758.2019.9221736.

[13] S. Susanto, F. A. Putra, and R. Analia, "XNOR-YOLO: The high precision of the ball and goal detecting on the barelang-FC robot soccer," *Proc. ICAE 2020 - 3rd Int. Conf. Appl. Eng.*, 2020, doi: 10.1109/ICAE50557.2020.9350386.

[14] S. Susanto, J. Suroto, and R. Analia, "The ROS: Kinetic Kame for Humanoid Robot BarelangFC," *J. Integr.*, vol. 13, no. 1, pp. 68–77, 2021, doi: 10.30871/ji.v13i1.2686.