Development and Performance Evaluation of Distance Measurements Detection Using 3 Axis Binocular Vision System

Mohamad Dzulhelmy Amari, Sukarnur Che Abdullah, Ng Shi Yen, Hanafiah Yusoff, Siti Aisyah Binti Jaafar

Faculty of Mechanical Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

Abstract
The world currently entering industry 4.0, which the industrial processes become automated and remotely control by computer with little input from human operators. Binocular vision is the result of signal sent to brain from eyes simultaneously which having advantage of depth perception. This project is about distance measurement of the object in 3-axis by using the advantages of binocular vision system. In this project, two webcams were used as binocular vision system to capture image and detect the distance measurement of the sample objects in 3-axis. The cameras must be calibrated before start to computing the distance measurement due to the fish eyes effect caused by the lens. In this project, the cameras were calibrated by using chessboard pattern with the size of 9x6 internal corners due to the simple geometry. Thedistance measurement algorithm will be executed once the cameras are calibrated, and the distance measurement is computed by using trigonometry concept.
Keywords: Binocular Vision, Camera Calibration, Measurement Detection

Introduction
The world currently is entered industry 4.0, which computers and automation will come together in an entirely new way with robotics connected remotely to computer systems equipped with machine learning algorithms that can learn and control the robotics with very little input from human operators. Vision sensor is a tool that consists of video camera, display and interface, and computer processor to automate industrial processes and decisions. It utilizes the images captured by a camera for determining the presence, orientation, and accuracy of parts. The vision sensor is widely used in many applications such as measurement, and observable characteristics regarding to the product quality.

Binocular vision is the result of the signal sent from the eye to brain, where the brain usually receive signal from both eyes simultaneously. The information contained in the signal for each eye is slightly different, and with the well-functioning binocular vision, which enables brain to judge the distance or the coordinates of the eye movement based on the differences. One of the benefits of binocular vision is the ability to judge the speed and depth of the objects, people with poor-functioning binocular vision might have difficulties to perform daily tasks such as pouring water into glass (Hamed et al., 2013).

Many researchers have been trying to simulate the binocular vision of human in robot applications by using vision sensor and software since the robot applications have been rising rap-idly in the society of well-developed technology. University of Washington and Microsoftdeveloped the wide baseline stereo vision system in 2003, the system used the same camera indifferent position of the Mars Reconnaissance Orbiter shoot images to determine the precise location of the camera with the non-linear optimization algorithm based on the landscape observed. Besides that, Massachusetts Institute of Technology had pro-posed a new sensor fusion algorithm for intelligence transportation to divide target location under high rates by combining the rough target depth from radar system and binocular vision, and with the improved fingerprint image segmentation method (Wang et al., 2017).

The computer vision for binocular vision system has been implemented in many robotic applications and safety purpose, where it concerned with the automatic extraction, analysis and understanding of useful information from a single image or a sequence ofimages. In addition, the development of theoretical and algorithm basis is involved in order to achieve automated visual understanding. Based on previous research, the trigonometry intersection points will be triggered by the combination of area of sight of the stereo vision in order to determine the distance of the object from the baseline. Furthermore, the measurementdetection algorithm is one of the issue that need to confront consequent since there are varioussort of calculations of measurement detection in the same field, however has unmistakable of use.

Methodology
Hardware Set Up
The vision sensor can be constructed with a couple of low-cost such as webcam (Abu Hassan et al., 2017), while a binocular stereo vision system can be constructed according to the particular application requirement. In this project, two units of Microsoft Lifecam HD-5000 were used in the set up as the cameras able to provide maximum output video resolution up to 1280 x 720 and maximum output still image resolution up to 1280 x 800. Furthermore, it also provides the diagonal field of view of 66° and allows to adjust image parameters setting such as brightness,exposure, sharpness, saturation, contrast, and focus (Microsoft, 2012). Figure 1 shows the experimental set-up of stereo cameras used in this project as the prototype of binocular vision system. The hardware was set up at Automation and Materials Lab in KDU University College.

Figure 1. Experimental set up for binocular vision

Camera Calibration

There are various methods to perform camera calibration for binocular vision system. According to research, there are two methods to do camera calibration with low cost stereo camera, which are using chessboard pattern, and feature descriptor pattern. The calibration using chessboard pattern was produced less mean re-projection error, which was better than using feature descriptor pattern. Furthermore, calibration with chessboard pattern required two cameras to capture multiple planar of chessboard pattern from different orientation view. In addition, this method consists of a closed-form solution and then a nonlinear refinement based on the maximum likelihood criterion (Abu Hassan et al., 2017).

Therefore, the chessboard pattern calibration method is chosen as the calibration method in this project due to the simple geometry. The procedure of stereo calibration process involved were taking few sets of raw image pair, execute calibration algorithm by OpenCV, process distortion images, process rectification and save the intrinsic and extrinsic camera parameters (Abdullah et al., 2015).
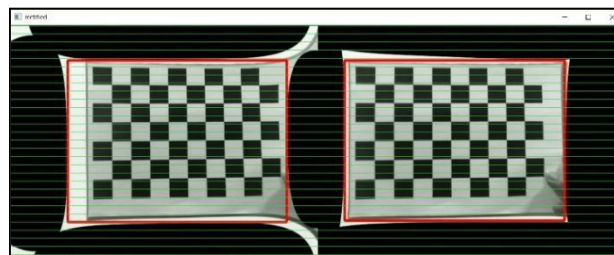

Figure 2. Rectified image pair by using chessboard pattern calibration

Table 1.
Stereo Calibration Result

Criteria	ChessboardPattern	9 x 6
	Image Size	640 x 480
Accuracy	Mean Reprojection Error	0.396611

Stereo Matching

Stereo matching is the process of matching a 3D point in the two different camera views, which can be computed only over the visual areas in which the views of the two cameras overlap (Bradski & Kaehler, 2008). This is the most crucial step in stereo vision domain as a lot of factors in thescene in a 3D space are composited and represented by single grey value when projected tothe in the 2D image. In addition, the factors include illumination, geometry shape of the object and physical property, noise effects and distortion, and the camera properties as well (Zou & Li, 2010). Furthermore, the disparity map will be generated by stereo correspondence and theinformation is used to determine the distance of the object in this project.

The block matching algorithm has three phases in order to create disparity map as known as depth map, which are pre-filtering phase, correspondence searching and computation, and post-filtering phase as shown in Figure 3.

Pre-filtering to Normalize Image Brightness and Enhance Texture → Correspondance Search Along Horizontal Epipolar Lines Using an SAD Window → Post-filtering to Eliminate the Bad Correspondane Matches
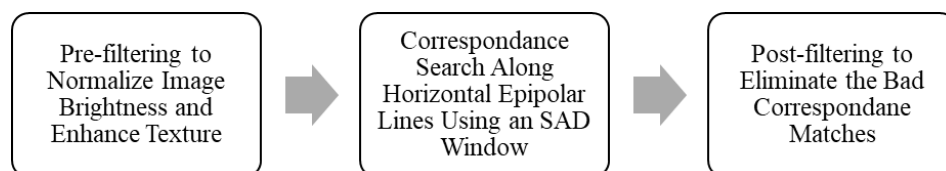
Figure 3. Phases in Block Matching [6]

Figure 4. Disparity Map

Distance Calculation

Based on the related research projects, the researchers used trigonometry concept to determinethe measurement for binocular vision system (Abu Hassan et al., 2017) (Huang & Cheng, 2013). Trigonometry concept is a method to determine the distance dimension in 3-axis. Figure 4 shows the analysis of binocular vision system. the binocular vision system mainly utilizes the differences of the targeted point in horizontal coordinates which are xl and xr from the two directed image from both left andright camera, and the disparity has an inverse relationship with distance Z from target point to imaginary plane as shown in (1), (2), and (3).

$$B/Z = (B - d)/(Z - f) \tag{1}$$

$$Z = f * B/d \qquad (2)$$

$$d = x_l - x_r \qquad (3)$$

Where,

B = The baseline distance between two cameras; Z = The depth distance;
d = The disparity of the images;
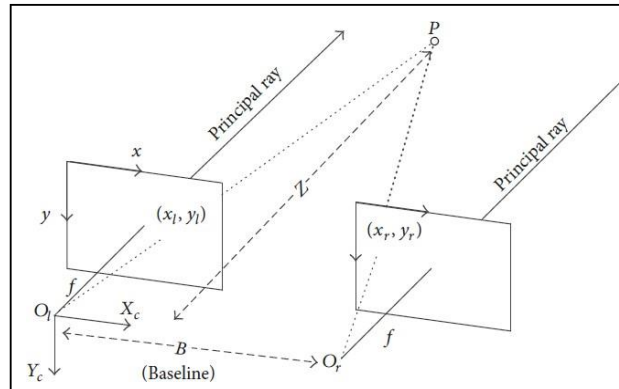f = the focal distance of the camera

Figure 5. Analysis of Binocular Vision System [8]

Furthermore, the distance measurement for X and Y axis can be determined after depth distance was determined by using the usual camera projective formula as shown in (4) and (5).

$$X = x_l * Z/f; \ X = B + x_r * Z/f \qquad (4)$$

$$Y = y_l * Z/f; \ Y = y_r * Z/f \qquad (5)$$

Results and Discussion
From the tabulated data, Table 2, Table 3, and Table 4 show the result of field measurement and distance measurement with binocular vision system in 3-axis of three primary colour objects. The possible reasons of errors occurred in the field measurement were parallax error and the scale on the ruler of field measurement only indicated one decimal place data unlike the binocular vision system the data can measure up to two or more decimal place data. Moreover, the distance detection results of this system are heavily relying on the calibration data, because the calibration algorithm would generate intrinsic and extrinsic parameters of the camera from the fixed set up, where the parameters were used to compute the distance of the object in 3-axis. If the setup was moved but the calibration data remain unchanged, then the disparity computed in block matching will not be accurate, which also affect the accuracy of distance measurement.

Table 2I.

Measured Distance in 3-axis with Binocular Stereo Vision System(Red Colour Object)

Field Measurement(x, y, z) (cm)	Camera (x, y, z) (cm)	Error (x, y, z) (%)
17.60, 9.50, 30.00	18.97, 9.17, 31.27	**7.78**, 3.47, 4.23
21.50, 21.80, 40.00	22.27, 21.64, 39.42	3.58, 0.73, 1.45
18.20, 26.30, 50.00	18.54, 25.82, 48.39	1.87, 1.83, 3.22
19.80, 34.50, 60.00	20.54, 33.87, 57.57	3.74, 1.83, 4.05
29.80, 45.30, 70.00	30.25, 41.85, 66.87	1.51, **7.62**, 4.47
27.50, 48.20, 80.00	27.99, 46.92, 76.04	1.78, 2.66, 4.95
29.30, 52.30, 90.00	29.87, 51.38, 85.55	1.95, 1.76, 4.94
29.00, 57.90, 100.00	30.36, 55.16, 94.96	4.69, 4.73, **5.04**

Table 3II.

Measured Distance in 3-axis with Binocular Stereo Vision System(Green Colour Object)

Field Measurement(x, y, z) (cm)	Camera (x, y, z) (cm)	Error (x, y, z) (%)
14.00, 15.80, 30.00	15.16, 15.26, 30.36	**8.29**, **3.42**, 1.20
17.30, 17.30, 40.00	18.45, 17.52, 40.23	6.65, 1.27, 0.57
15.00, 26.90, 50.00	15.98, 26.40, 49.05	6.53, 1.86, **1.90**
21.50, 36.90, 60.00	22.70, 35.68, 59.42	5.58, 3.31, 0.97
19.80, 40.80, 70.00	20.67, 40.35, 68.68	4.39, 1.10, 1.89
23.30, 48.20, 80.00	24.50, 48.65, 78.75	5.15, 0.93, 1.56
26.40, 54.50, 90.00	27.94, 54.90, 91.64	5.83, 0.73, 1.82
26.50, 58.60, 100.00	28.45, 58.51, 100.85	7.36, 0.15, 0.85

Table 4III.

Measured Distance in 3-axis with Binocular Stereo Vision System(Blue Colour Object)

Field Measurement(x, y, z) (cm)	Camera (x, y, z) (cm)	Error (x, y, z) (%)
11.50, 17.50, 30.00	12.49, 16.34, 30.81	8.61, 6.63, 2.70
16.70, 23.40, 40.00	17.07, 22.43, 39.42	2.22, 4.15, 1.45
13.40, 27.50, 50.00	15.07, 26.71, 50.50	12.46, 2.87, 1.00
14.40, 31.20, 60.00	15.10, 31.97, 60.05	4.86, 2.47, 0.08
16.70, 41.70, 70.00	17.92, 42.06, 70.38	7.31, 0.86, 0.54
25.70, 46.80, 80.00	28.38, 49.52, 85.77	10.43, 5.81, 7.21
29.30, 53.90, 90.00	33.26, 59.72, 98.09	**13.52, 10.80, 8.99**
28.80, 58.40, 100.00	32.67, 64.34, 108.71	13.44, 10.17, 8.71

Figure 6 shows the graph of percentage error of Z-axis measurement against the field measurement. Figure 7 and Figure 8 show the graph of percentage error of distance measurement while the objects were located in different depth location. There is no any significant trend that show in these figures and show any indication of having better measurement on certain colour. Hence, the colour of the object did not affect the distance measurement of binocular vision system. In addition, the system determines distance from disparity map in stereo matching, where the block that matched would appear on the disparity map regardless the object colour, and the matched block would appear in white, grey, or blackcolour depending on the object location.
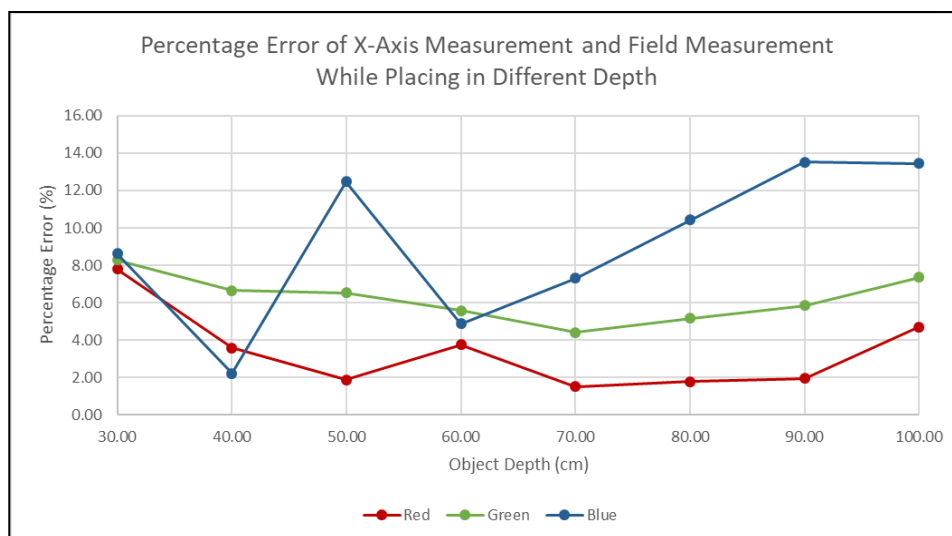

Figure 6. Percentage Error of Z-Axis Measurement Against Field MeasurementFigure 7. Percentage Error ofS X-Axis Measurement and Field Measurement While Placing inDifferent Depth
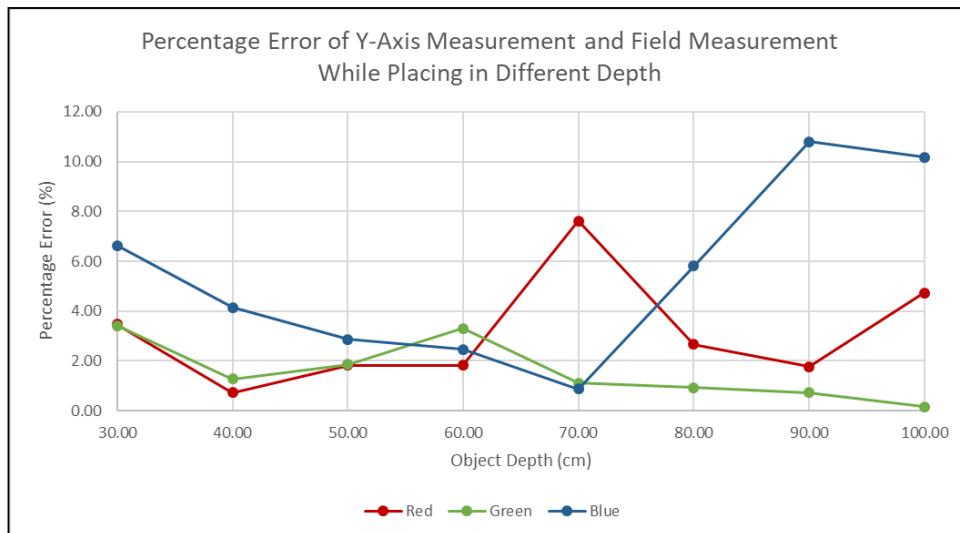
Figure 8. Percentage Error of Y-Axis Measurement and Field Measurement While Placing in Different Depth

Table 5 and Figure 9 show the average error of distance measurement in 3-axis when the object was located in different depth regrardless the object colour. The measurement of X-axis has highest average percentage error among the other two axis. Furthermore, the trend lines of the measurement of three axis show that the accuracy of the distance measurement were decreased as the further of the object located.

Table 5.
Average Error of Measurement in 3-Axis with Different Depth

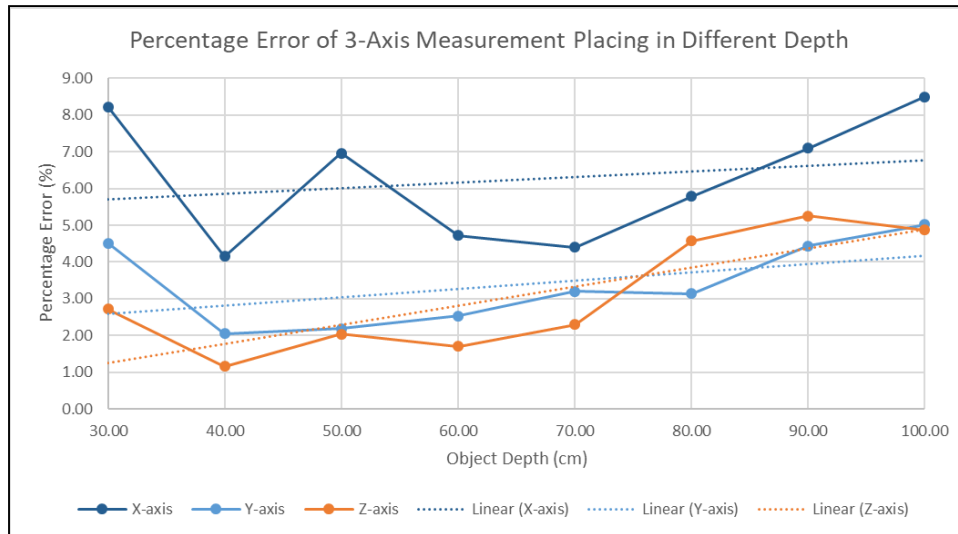| Z-axis Location (cm)	Average Error (%)		
	x	y	z
30	8.23	4.51	2.71
40	4.15	2.05	1.16
50	6.95	2.19	2.04
60	4.73	2.53	1.70
70	4.40	3.19	2.30
80	5.79	3.13	4.58
90	7.10	4.43	5.25
100	8.50	5.02	4.87

Figure 9. Percentage Error of 3-Axis Measurements Placing in Different Depth

The distance measurement sensors or application used in industry such as optical sensors, ultrasonic sensors, eddy current sensors, laser sensors could provide high accuracy output. Generally, these distance measurements send signal to the target surface and receive the reflected signal to generate the output. In addition, these sensors only able to detect the surface of the target and measure the distance in one dimensional axis.

Although the distance measurement of binocular vision system might not as accurate as these sensors, this system has several advantages in the application. This system able to detect the colour and the geometry of the object, which allows the system to determine the target object with the coordinates in the applications. Furthermore, this system is able to measure the distance of the object from the camera in 3-axis in once, which also has the advantages of time saving compared to other application

Conclusion

The proposed program algorithms for binocular vision system to detect 3-axis location of the RGB object was created successfully despite the accuracy might lower than the other applications. Furthermore, this system was able to filter the colours and detect the RGB objectand the accuracy of this system was evaluated with laser distance meter.

Acknowledgements

References

Abdullah, S. C., Yusof, Y., Rashid, H., & Ma'arof, M. N. (2015). Rectification Y-coordinate of epipolar geometry in calibration of binocular eye system. *Procedia Computer Science*, 76, 494-499.

Bradski, G., & Kaehler, A. (2008). *Learning OpenCV: Computer vision with the OpenCV library*. "O'Reilly Media, Inc.".

Hamed, M. M., David, A. G., & Marzieh, E. (2013). The relationship between binocular vision symptoms and near point of convergence. *Indian Journal of Ophthalmology*, 61(7), 325.

Hassan, M. A., Hussain, A., Saad, M. M., & Win, K. (2017). 3D distance measurement accuracy on low-cost stereo camera. *Science International*, 29(3), 599-599.

Huang, G. S., & Cheng, C. E. (2013). 3D coordinate identification of object using binocular vision system for mobile robot. *2013 CACS International Automatic Control Conference (CACS)* (pp. 91-96). IEEE.

Microsoft. (2012) Microsoft LifeCam HD5000 Technical Data Sheet. pp. 3–4.

Wang, T., Liu, B., Wang, Y., & Chen, Y. (2017). Research situation and development trend of the binocular stereo vision system. *AIP Conference Proceedings* (Vol. 1839, No. 1, p. 020220). AIP Publishing LLC

Zou, L., & Li, Y. (2010). A method of stereo vision matching based on OpenCV. *2010 International Conference on Audio, Language and Image Processing* (pp. 185-190). IEEE.