International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 16 (2017) pp. 5946-5951

Evaluation of 3D-Distance Measurement Accuracy of Stereo-Vision Systems

Ali Sophian[1], Wahju Sediono, Muhammad Ridzwan Salahudin, Mohd Saatarie Mohd Shamsuli
Dayang Qurratu'aini Awang Za'aba

Mechatronics Engineering Department, Faculty of Engineering, International Islamic University Malaysia, Malaysia

[1]Orcid :0000-0003-4991-6362

Abstract

Many applications would benefit from a cost effective 3D position or distance measurement systems. Stereo vision systems may offer this functionality with the extra benefits when the images are used for other purposes as well, such as object recognition. In this paper, an accuracy evaluation method of two stereo-vision systems, which use a coordinate measuring machine (CMM) and a reference block, has been presented, and the results have also been presented for systems using infra red cameras and webcams. Following a calibration process, the two systems were used in determining the dimensions of the reference block. The results show that the method could evaluate the two systems. While the evaluation results show that the webcams have a better accuracy and precision, the leap motion controller can be used when shorter measurement distance are required, thanks to their shorter lens distance and wide-angle lenses.

Keywords: Stereo vision; Machine vision; 3D Position Measurement; Measurement Evaluation; Leap motion controller

INTRODUCTION

Contactless solutions for 3D position or distance measurements are useful for various applications, both industrial and non-industrial, such as shaft misalignment [1], robot navigation [2], medical applications [3], manufacturing [4] etc. These solutions exploit various technologies, such as infrared [5], laser [1] and ultrasonic [6]. Each technique has their own strengths and weaknesses, and they are characterized by parameters, such as accuracy, measurement range, resolution and measurement rates. Their range of measurement and accuracy vary. Another alternative technique~~s~~ is based on stereo vision which may offer lower costs in addition to the convenient possibility of using the same system for other visual functions, such as object recognition, although their measurement performance maybe lower in accuracy.

The use of stereo vision in extracting 3D coordinates of a given point by using two planar projections has been widely exploited and reported. Other applications of stereo vision include robot navigation [7], inspection system, and human-computer interaction (HCI), which shows the versatility of the technique.

The suitability of a stereo-vision system for deployment in a given application depends on, among other things, the accuracy of the system. Surprisingly, there has not been much works reported in the determination of accuracy or uncertainties in stereo vision systems [8]. Sankowskli et al proposed an approach that used a general-purpose laser distance meter and a calibration board for determination of these uncertainties [8]. In this work, a simple evaluation method of stereo vision systems in determining a distance between two points have been presented. The method uses a reference block whose dimensions have been measured by using a CMM. The results of using the evaluation method in evaluating two different stereo vision systems for short-distance measurements of sub-1-m range have also been presented. Such systems may potentially be used for non-industrial applications, such as home-based hand rehabilitation [9], which is the central project of this work.

In the next section, the principles of 3D position measurement by using triangulation will be presented. The proposed evaluation method is then described. Subsequently, two different stereo vision systems used in the study are described. Then, in sections IV and V, the experimental setup and results are discussed. Finally, the conclusion is extracted.

3D-DISTANCE MEASUREMENT BY USING STEREO VISION

The principle of position measurement by using stereo vision can be found in many academic literatures, such as [5]. Figure 1 shows the procedural steps in determining the 3D position of a point in the real 3D world. The 3D position of a given pixel in the real 3D world is derived by using the principle of triangulation through finding the disparity, which is the difference between the positions of the matching pixels in the stereoscopic images. The corresponding pixels can be found more easily if the two images have been projected onto a common plane parallel to the line between the principal axis of the two lenses. This projection of the stereoscopic images is generally referred as image rectification, which is an important

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 16 (2017) pp. 5946-5951

step as misalignment between the two cameras or lenses always exists. Following the image rectification, the matching pixels can then be searched and found in the epipolar line and, in turn, their disparity can be identified as illustrated in Figure 2.

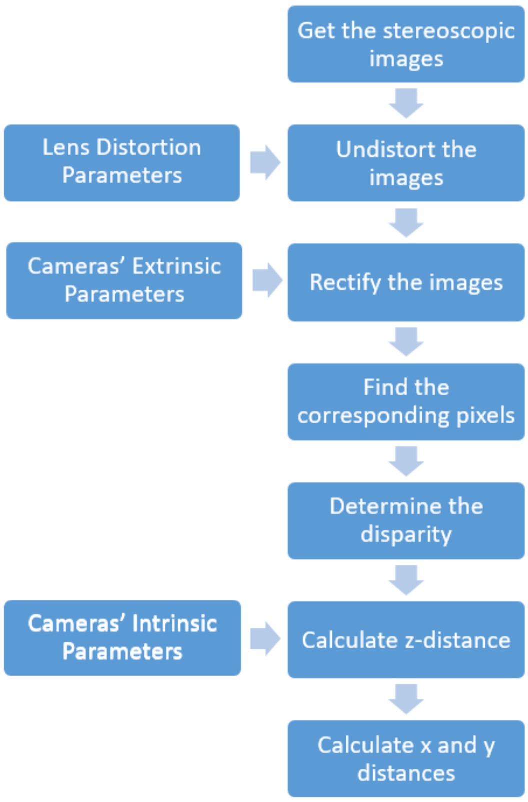

Figure 1: 3D Position Measurement Algorithm by Using Stereo Vision

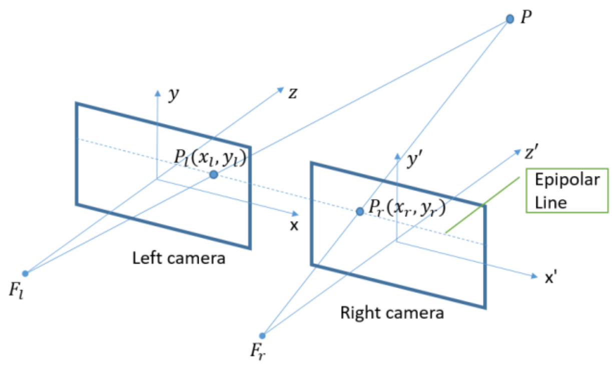

Figure 2: Epipolar line

Following the finding of the disparity, the z-distance between the cameras and the relevant point or object represented by the pixel, can then be determined, by using the following equation:

$$Z = \frac{fb}{d} \tag{1}$$

where: Z = distance along the the z-axis, f = focal length (in pixels) and b = distance between the axes of the two cameras and d is the disparity.

The distances in the x and y-axis, x and y respectively can then be determined by employing the following equations:

$$X = \frac{uZ}{f} \tag{2}$$

$$Y = \frac{vZ}{f} \tag{3}$$

where: X = distance along the the X-axis and Y = distance along the the Y-axis.

The length of the block can then be calculated by using the following equation:

$$l_{block} = \sqrt{\Delta X^2 + \Delta Y^2 + \Delta Z^2} \tag{4}$$

where: l_{block} = length of the block.

A. *Camera Model*

The foundation of the stereo vision by using cameras is the camera model, which defines the projection of a 3D world view onto a 2D image plane. Matlab and its relevant tool boxes have been used in this work, which uses the camera model which was presented by Jean-Yves Bouguet [3]. It is based on the pinhole camera model complemented with lens distortion. The pinhole camera model is explained by a camera matrix *P*, which transforms a 3-D world scene to a 2-D image plane and is defined as

$$P = K[R\ t] \tag{5}$$

where: *K* = calibration matrix or intrinsic matrix and $[R\ t]$ = extrinsic matrix that is comprised of rotation *R* and translation *t*.

It can be seen from the equation 5, that the model is comprised of intrinsic and extrinsic parameters and distortion coefficients. The intrinsic parameters include focal lengths (in pixels) and principal points, as shown below

$$K = \begin{bmatrix} f_x & 0 & 0 \\ s & f_y & 0 \\ c_x & c_y & 0 \end{bmatrix} \tag{6}$$

where: *fx* and *fy* = focal lengths, (cx, cy) = principal point, and *s* = skew parameter.

B. Camera Calibration

To use the camera model and to enable both the elimination of the image distortion and rectification, cameras' parameters are obtained through calibration. Calibration is performed for both the individual camera and the stereo configuration. For the individual camera, intrinsic and extrinsic parameters are obtained, in addition to the distortion coefficients. The relative positions and orientations of each camera are also established through the calibration, including the baseline, which is the distance between the two cameras.

The calibration was performed by using multiple images of a checkerboard pattern, which provides corresponding 2-D image points for 3-D world points. The size of each square in the checkerboard was 29.2 mm x 29.2 mm. Once the parameters are obtained, evaluation is performed on the accuracy on the parameters. If needed, more images of the calibration pattern can be obtained. Figure 3 shows a typical set of calibration checkerboard-pattern images.

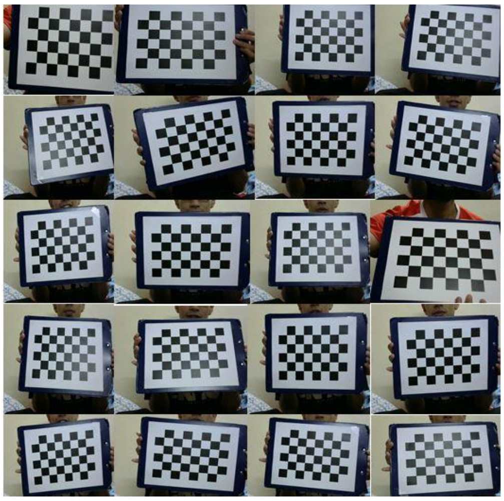

Figure 3: Checkerboard Pattern Captured at Different Distances and Orientations for Camera Calibration Purposes

THE STEREO VISION SYSTEMS

Two different stereo systems have been studied. The first one is based on commercial HD webcams (Logitech C270H) that have a resolution of 1280 x 720 pixels. The two cameras are arranged side-by-side as close as possible, as illustrated in Figure 4. The distance between the centres of the lenses is 70 mm. The field of view of the device is 60° and its focal length is 4 mm.

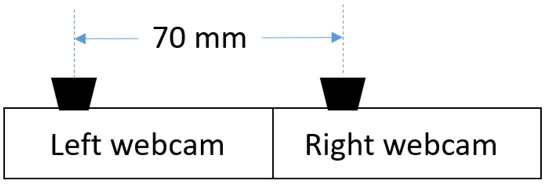

Figure 4: Diagram of the Webcam-based Stereo Vision

The other system is a Leap Motion Controller (LMC) which is a device that has been designed to detect the positions of fingertips, finger joints and palms. The device has two built-in infrared (IR) cameras, in addition to three IR LEDs, as illustrated in the diagram in Figure 5. The cameras capture the reflected IR beams and generate a pair of grayscale images with resolution of 320x120 pixels each. The images generated are much distorted due to the characteristics of the lenses used in the device.

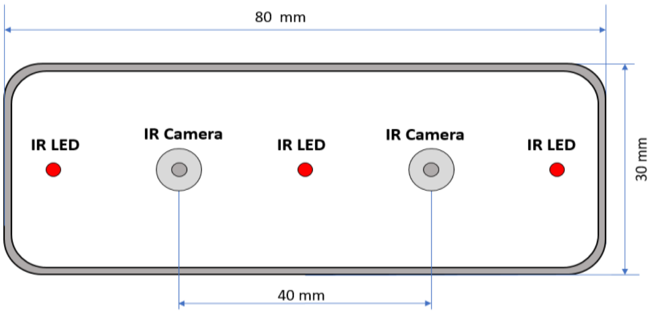

Figure 5: Leap Motion Controller

EXPERIMENTAL TESTS

A coordinate measuring machine (CMM) was used in this work to measure the dimensions of the workpiece that was going to be used in the evaluation of the stereo-vision systems. The CMM, which is Mitutoyo BH-V507, measures the xyz coordinates with a resolution of 0.001 mm and its maximum workpiece dimension is X = 500 mm, Y = 700 mm, Z = 400 mm. The workpiece used in this work is a rectangular metal block. Its length was measured by using the CMM repeatedly and its average was found to be 151.324 mm. Stereoscopic images of the reference workpiece were then taken at different z-distances, which were set-up as shown in the illustration in Figure 6.

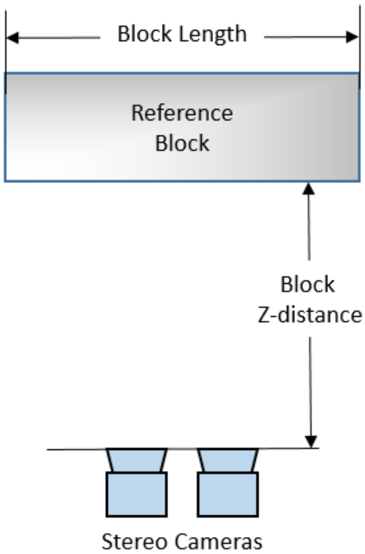

Figure 6: Experimental setup

The images were taken by using both the webcams and LMC, although the set of distances were different for each one as they have different optical characteristics and setups. Figure 7 shows typical examples of the distorted grey-scale images taken by using the LMC.

(a)

(b)

Figure 7. Examples of Stereo Images Taken by Using the Leap Motion Controller at a distance of 100 mm, (a) Left Image and (b) Right Image

Two sets of stereoscopic images were taken for determining the length of the reference block. For determining the length, both the top left hand side and top right hand side pixels were manually determined on both left and right images, as illustrated in Figure 8. In practical circumstances, the determination of the pixels would be done using some algorithm. By using the 'ground truth', the maximum potential of the systems in position measurement can be established. Then, the respective dimension can be determined by using the equations described in section 2 and the parameters obtained from the calibration, namely the focal lengths and the distances between the lenses.

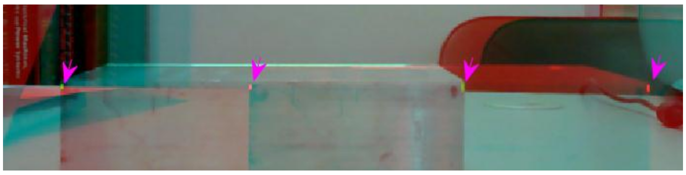

Figure 8: Four Pixels Identified on the Concatenated Rectified Stereo Images to be used for the Distance Measurement

RESULTS AND ANALYSIS

The positions and orientations of the checkerboards captured during the calibration process of the stereo webcam and the LMC are shown in Figure 9 and Figure 10 respectively. The figures show that the ranges of z-distances were different for the two systems. It should be highlighted that the LMC managed to capture the checkerboard pattern even when it was positioned less than 100 mm away from the cameras, whilst at the distance has to be more than 400 mm for the webcams. More than twenty pairs of images were taken for each calibration.

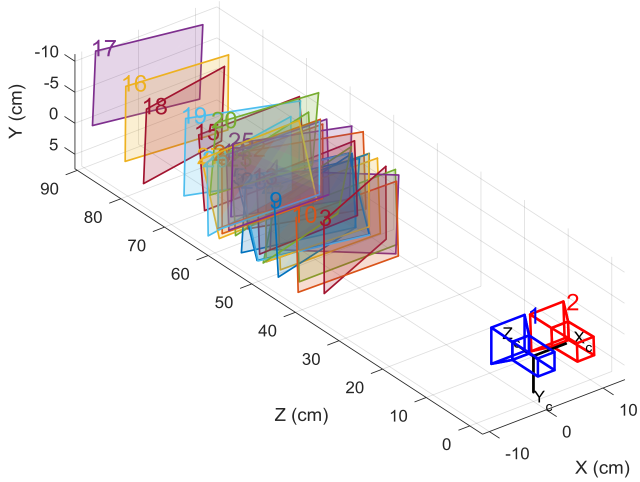

Figure 9: Positions and Orientations of the Checkerboard pattern in the Calibration of the stereo webcams

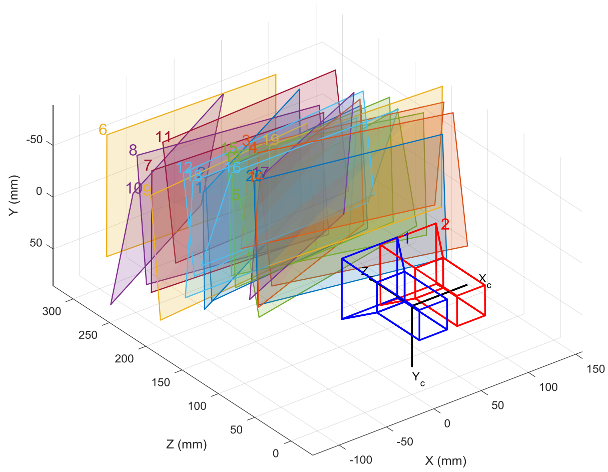

Figure 10: Positions and Orientations of the Checkerboard pattern in the Calibration of the LMC

The calibration outcomes display interesting results, especially for the LMC. The results reinforce the wide-angle characteristics of the lenses used in LMC, as can be seen from the significant difference in the approximated focal lengths f_x and f_y. This explains why LMC can capture the checkerboard pattern wholly at very short distance. Due to its wide-angle lenses, three coefficients had to be used for modelling the radial distortions of the LMC, instead of 2 used generally. Table 1. shows the parameters resulting from the calibration of both systems.

Table 1: Calibration Results

	Webcams	LMC
Focal lengths, f_x and f_y (pixels)	Cam 1: [812.6 813.2] Cam 2: [808.6 809.2]	Cam 1: [134.6 67.4] Cam 2: [134.3 67.3]
Principal point	Cam 1: (337.5, 230.7) Cam 2: (311.3 260.8)	Cam 1: (327.8 120.3) Cam 2: (323.2 123.4)
Radial distortion coefficients	Cam 1: [-0.016 0.728] Cam 2: [-0.029 0.526]	Cam 1: [0.16 -0.11 0.02] Cam 2: [0.17 -0.12 0.03]
Skew coefficient	Cam 1: 0 Cam 2: 0	Cam 1: 0 Cam 2: 0
Distance between the optical axis (mm)	69.8	39.9

Having obtained the camera parameters from the calibration step, the 3D positions of the upper-left and upper right-corners of the reference block can then be approximated by using the model, and subsequently the measured length of the block can be derived. The results of the measurement by using the webcams and LMC are shown in Table 2. The analysis of the measurement errors are also plotted in the graph shown in Figure 11. The results show that in general both systems can be used for approximation of the block length. However, they do differ in the performance. The webcams are generally better, shown by their lower errors, which mean better accuracy, with % error is always less than 1%. Their error standard deviations are also lower, signifying better precision. However, the advantage of the LMC can be highlighted from the measurement range when ranges less than 300 mm are required. This characteristic is possible due to its wide-angle lenses and the closer distance between the stereo lenses, i.e. 39.9 mm, instead of the 69.8-mm distance for the webcam.

Table 2: Measurement Evaluation Results

Stereo Vision	Block Distance (mm)	Averaged Measured Block Length (mm)	Abs Error (mm)		Error (%)
			Mean	Standard Deviation	
Webcams	350	151.2	0.12	1.26	0.08
	400	152.06	0.32	0.64	0.21
	500	150.72	0.6	1.69	0.40
	600	152.74	1.41	0.02	0.93
	700	149.99	1.34	0.01	0.89
LMC	75	153.06	1.74	8.22	1.15
	100	153.06	1.74	3.72	1.15
	150	152.5	1.17	1.75	0.77
	200	148.41	2.92	3.21	1.93
	300	148.82	1.6	4.00	1.06
	400	147.2	4.13	0.00	2.73

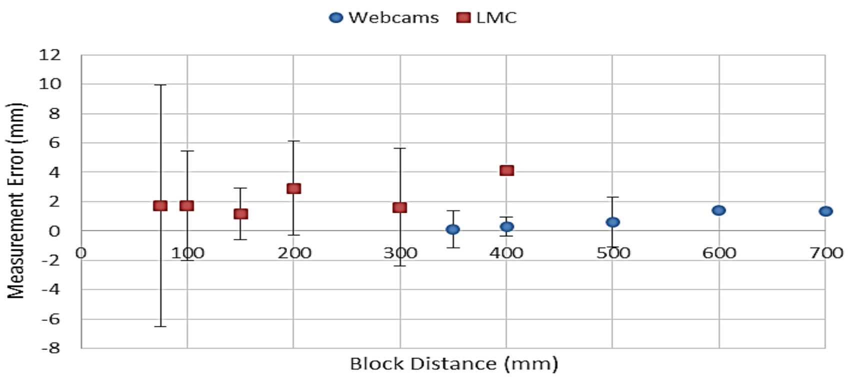

Figure 11: Plots of the Measurement Error vs Measurement Distance for Both Systems. The error bars signify the standard deviation of the error. Some standard deviations are too small to be seen in the plots.

CONCLUSION

There are many applications that would benefit from a cost effective 3D position or distance measurement systems. Stereo vision systems may offer this functionality with the extra benefits when the images are used for other purposes, as well such as object recognition. However, the performance of the stereo vision must be known to ensure its suitability for a given application. In this paper, an accuracy evaluation method of stereo-vision systems has been presented, and the results have also been presented and discussed for two different systems. One is based on commercial HD webcams, while the other is the leap motion controller, which uses IR cameras. Following a calibration process, the two systems were used in determining the dimensions of a metal block that has previously been measured using a CMM. The results show that the webcams generally have better performance in both accuracy and precision, however the LMC has some potential when shorter ranges are required, thanks to their shorter lens distance and wide-angle lenses.

ACKNOWLEDGMENT

The authors would like to thank the Ministry of Higher Education of Malaysia for financially sponsoring the work under the grant RIGS15-151-0151.

REFERENCES

[1] A. Simm, Q. Wang, S. Huang, and W. Zhao, "Laser based measurement for the monitoring of shaft misalignment," *Measurement*, vol. 87, pp. 104–116, 2016.

[2] M. A. Mahammed, A. I. Melhum, and F. A. Kochery, "Object Distance Measurement by Stereo Vision," *Int. J. Sci. Appl. Inf. Technol.*, vol. 2, no. 2, pp. 5–8, 2013.

[3] N. Saaidon, W. Sediono, and A. Sophian, "Altitude Tracking Using Colour Marker Based Navigation System for Image Guided Surgery," in *Proceedings - 6th International Conference on Computer and Communication Engineering: Innovative Technologies to Serve Humanity, ICCCE 2016*, 2016, pp. 465–469.

[4] L. Uriarte, M. Zatarain, D. Axinte, J. Yagüe-Fabra, S. Ihlenfeldt, J. Eguia, and A. Olarra, "Machine tools for large parts," *CIRP Ann. - Manuf. Technol.*, vol. 62, no. 2, pp. 731–750, 2013.

[5] G. Benet, F. Blanes, J. E. Simó, and P. Pérez, "Using infrared sensors for distance measurement in mobile robots," *Rob. Auton. Syst.*, vol. 40, no. 4, pp. 255–266, 2002.

[6] T. J. Licznerski, J. Jaroński, and D. Kosz, "Ultrasonic system for accurate distance measurement in the air," *Ultrasonics*, vol. 51, no. 8, pp. 960–965, 2011.

[7] N. Q. Ann, M. S. H. Achmad, L. Bayuaji, M. R. Daud, and D. Pebrianti, "Study on 3D scene reconstruction in robot navigation using stereo vision," *2016 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)*. pp. 72–77, 2016.

[8] W. Sankowski, M. Włodarczyk, D. Kacperski, and K. Grabowski, "Estimation of measurement uncertainty in stereo vision system," *Image Vis. Comput.*, vol. 61, pp. 70–81, 2017.

[9] A. Sophian, D. Q. Awang Za'aba, W. Sediono, H. Md. Yusof, and S. Sudirman, "Markerless detection of fingertips of object-manipulating hand," *ARPN J. Eng. Appl. Sci.*, vol. 11, no. 23, pp. 13716–13720, 2016.