3D DISTANCE MEASUREMENT ACCURACY ON LOW-COST STEREO CAMERA

[1,3]M.F. Abu Hassan, [1]A. Hussain, [2]M.H. Md Saad, [1]K. Win
[1]Department of Electrical, Electronic & Systems Engineering,
[2]Department of Mechanical & Material Engineering,
Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia, Selangor, Malaysia
[3]Universiti Kuala Lumpur, Malaysia France Institute, Section 14, Jalan Teras Jernang, Bandar Baru Bangi, Selangor, Malaysia
For Correspondence; fadzil@siswa.ukm.edu.my

ABSTRACT: *This paper presents a comprehensive scheme for reconstructing a three-dimensional (3D) model from a stereo camera via multi-view calibration. In surveillance tasks, the depth information is useful so as to estimate the actual position of the target object. It is an essential parameter that can allow visualization from multiple perspectives. Despite the availability of various 3D sensors in the market, this technology still has its limitation. As such, in this paper, we afford a low-cost 3D sensor that uses a pair of mid-range webcam, Logitech HD Webcam C270. Both cameras were calibrated to obtain their intrinsic and extrinsic parameters by using two methods. Once done, the calibration method that produces the smallest mean re-projection error is selected. Next, the 3D model/image is reconstructed from a pair of two-dimensional (2D) images to visualize the perspective projection in a 3D space. To evaluate the 3D distance measurement accuracy of the sensor, an experiment was conducted by measuring the distance between the sensor and a target object within the range of 1 to 5.5 meters. The experimental results showed that the stereo camera is able to estimate a good distance measurement within an acceptable error range. The maximum and minimum measurement errors rates are 5.358 percent and 0.001 percent, respectively with mean error measurement of 1.717%.*

Keywords: stereo camera; 3D reconstruction; camera calibration; depth map

1. INTRODUCTION

Petty crimes such as purse snatching, pickpocketing, smash thefts and residential burglaries are the most common crimes committed against publics including tourists in Malaysia [1]. Noticeable increase of reported cases is forcing the security authorities to intensify surveillance activities especially within the hotspot area as means to curb such immoral acts. Various methods have been suggested and implemented to overcome such crimes and one of it is via surveillance system using the close-circuit televisions (CCTVs). To assist in the surveillance activities, video camera is being used as a monitoring tool to monitor people's behavior, activities, and other related information.

Typically, existing surveillance systems provide video capturing and recording function, while completely leaving the monitoring and evaluation task of threat detection to personal security or operators. It is agreed that such task requires high level of visual attention that is extremely demanding and incur high labor cost [2]. Thus, it has become an important research topic in the field of security and causes a paradigm shift from investigation to prevention of incidents.

Today, video surveillance technology has accelerated in term of growth because of the continuously decreasing price and better camera capability. The security services could harvest much better efficiency if the information is extracted from high- definition 3D images [3], instead of traditionally relied on 2D images. It is more advantageous if 3D techniques can solve the occlusion problem and distinguish multiple objects from the scene [4]. As such, in this work we propose a surveillance system that can provide real-time alert to respective personal as preemption. Our ultimate goal is to develop a surveillance system that is able to detect anomalous behavioral movement of human in the 3D space. The proposed research framework is as illustrated in Fig. 1, however only the blocks highlighted in blue will be discussed in this paper. At the preliminary stage, we will concentrate on the development of a low-cost 3D sensor. The intrinsic and extrinsic parameters are essential for 3D reconstruction process. Thus, camera calibration will be performed once the hardware development is completed. In this work, the estimated 3D coordinate reading from the system is evaluated to assess the performance of the hardware setup.

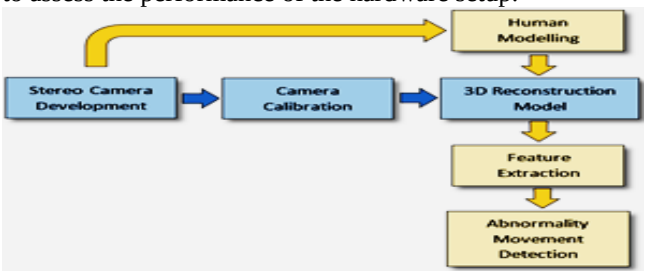

Fig. 1: Research framework

Several video-based methods have been proposed for movement analysis in recent years. Typically, the analyses use IR cameras and markers placed on the subject. These marker-based systems are accurate, but it is somewhat impractical unless use in laboratory settings and cause un-comfortableness to the subject under study. Additionally, passive or active markers must be correctly placed on the human body segments to ensure high sensing accuracy. Therefore, marker-less movement sensor is needed to overcome these issues.

Recently, many researchers demonstrate the effectiveness of off-the-shelf devices such as Microsoft Kinect Sensor and PlayStation Eye, wiimote, etc. for motion-sensing in several applications particularly in interactive video games, robotics, medical and also sports. [5] have used the Kinect Sensor to detect the abnormality of movement by analyzing the human joint. Besides, [6] did an investigation of Nintendo Wii controller in physical therapy interventions. The results showed that the subjects who performed the physical therapy

using Wii Fit training were able to perform independently without human intervention and improved their self-confidence and balance performance. Moreover, since the cost is reasonable and the sensing accuracy is within the acceptable limit, innovators are becoming more enthusiastic to use the end-user devices for video surveillance system. [7] proposed an algorithm to detect human aggressive behavior such as punching and kicking poses by using Kinect sensor. These type of "black boxes" have limitations and are suitable to certain application. The most critical limitation is the limitation involving the range of sensing range. Therefore, a sensing range capability of depth sensor should be considered in order to expand the existing motion sensing application.

A portable 3D reconstruction system with both hardware and software system designed by [8] was able to acquire the depth information from a pair of pictures and displayed the 3D reconstruction result. Unfortunately, the depth map accuracy and robustness of the system needed improvement. On the other hand, [9] presented a stereo-based obstacle detection system which was able to measure the distance between a car and an obstacle based on the depth information. The stereo vision allows 3D coordinate measurement and can be a cost-effective way of taking complex measurements. However, the estimated distance reading accuracy is still important especially in certain applications. A review on stereo vision accuracy using CCD cameras was made by [10] where he derives a limit for the accuracy based on the current state of the art technology. In this paper, we explain the development process of low-cost 3D sensor and calibration. Then, we construct a 3D model from the calibrated stereo camera and evaluate the accuracy of 3D distance measurement.

2. MATERIAL AND METHODS

Hardware System Setup

A 3D sensor can be constructed using a couple of low-cost cameras such as webcams [11, 12] while a stereo vision system can be constructed according to the specification of a particular application and requires a good calibration method to achieve the 3D point cloud to represent the external surface of an object. In this work, to develop the sensor, two units of the C270 Logitech® HD Webcams [13] were used in the set up and must be made capable of providing a maximum resolution of a resolution of 1280×720 video capture and taking sharp/crisp still 3Mp photos. Next, the USB 2.0 webcam which offers a relatively good flexibility and high definition quality image sensor was compared to other low-cost video recording devices. It provides view field of 60° and allows custom image parameters setting such as white balance, exposure, brightness, gain and contrast.

A customized enclosure was made by using a 3D printer to install the stereo camera, as shown in Fig. 2. The ABS typed material enclosure is formed to ensure a good strength to mount the stereo camera. A brass solid shaft holds the cameras in series and fastened to the enclosure using M3 screw. The cameras should be firmly placed in a stationary position to avoid external mechanical vibrations. This is highly important during camera calibration process where it will influence the calibration result. Consequently, the camera parameters will spoil to gain a good 3D point cloud. Moreover, to stabilize and elevate as well as able to rotate and tilt the stereo camera, the enclosure is attached with 35.5 to 106 cm height tripod.

Fig. 2: The proposed stereo camera hardware

Asynchronous Cameras

One of the most common disadvantages of low-cost USB webcams is that it is not equipped with hardware trigger and restricted it to an alternatively method that relies on software trigger. In principle, a pair of images is required to perform camera calibration and 3D reconstruction for stereo camera (explained in next subsection). For static scene and stationary cameras, the synchronous trigger is not essential but if when dealing with the dynamic background and foreground, the images should be captured from each camera simultaneously to avoid gross disparity in between the 2D images.

In the case of software trigger for Universal Serial Bus (USB) type webcam, it is impossible to get a matching time stamp for both acquisitions even though the webcam sensors are continuously exposed during the triggering process. One of the main reasons is that each time the software requests data transfer from the cameras, data transmission occurs in a serial form and the controller architecture is not allowed to do multiple data transmission in parallel. However, with the high bandwidth of USB 2.0 and frame frequency; at a maximum rate of 30 frame(s) per second (fps), the asynchronous time stamp problem could be minimized. The available physical frame frequency of camera for the selected image resolution should be evaluated to measure the performance.

The selected cameras are connected via USB 2.0 to a notebook computer with Intel i7 processor, Windows 10 OS and a 8 GB RAM. Using MATLAB® R2015a software, the timestamp correlation for 100 frames of each camera is evaluated and a delay is imposed before retrieving the data to accommodate data logging. Fig. 3 shows the timestamp for 800 x 600 images sampled at actual frame rate, F_{act} of 30 fps for 100 consecutive frames. We observed that it took 4.352 seconds and 4.938 seconds for right and left cameras, respectively to completely retrieve the 100th frames. Additionally, the average frame time stamp different, Δ_{ts} between them is 0.735 seconds, contributing a small error delay in time in synchronizing the retrieval time in general.

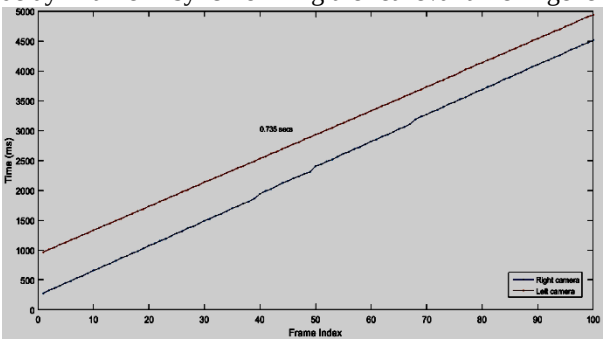

Fig. 3: Timestamp for right and left cameras

The time stamp difference between each frame index, T_{diffR} for right camera is calculated and we obtained the average of T_{diffR}, T_{aveR} is 41.6275 msec with standard deviation, SD_R of 10.36 msec as illustrated in Fig. 4. Therefore, the frame rate error, $Error_{fpsR}$ in between F_{act}=30 fps and experimental frame rate, $F_{expR}=1/T_{diffR}$ is 19.92%, corresponding to 6 frames/sec losses approximately. Whereas, we found the average time stamp difference for left camera, T_{aveL} is equal to 40.1123 msec, with standard deviation, SD_L of 8.8541 msec. Therefore, 16.90% was found to be the left camera frame rate error, $Error_{fpsL}$ which corresponds to 5 frames/sec losses approximately.

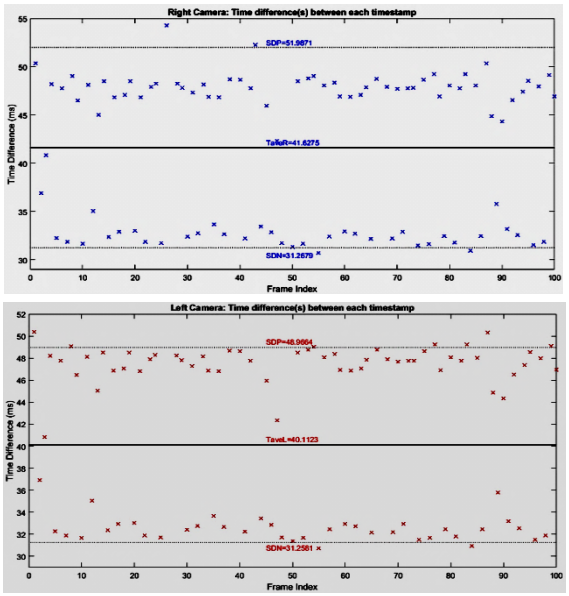

Fig. 4: Timestamp different for right (top) and left (below) cameras

The available frame frequency for each camera is evaluated individually within the available range of 5 to 30 fps with 0.5 fps sampling interval. Fig. 5 summarizes the experimental results in which the lowest frame frequency, that is, 5 fps contributed the highest transmission frame rate error. However, the system is able to retrieve the available frame at 25 fps with 4.48% for $Error_{fpsR}$ and 0.27% for $Error_{fpsL}$. The other available frame frequencies contributed to the gross error in frame rate transmission.

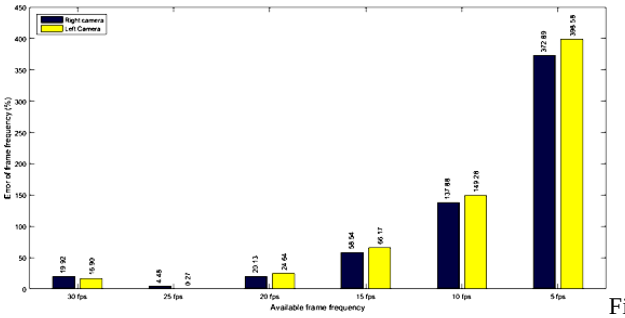

Fig. 5: Error percentage of frame frequency

To ensure a good result is obtained during the 3D reconstruction for motion scene, the highest frame frequency with the lowest error transmission is selected. Fig. 6 shows the transmission performance in 25 fps mode. The average retrieval time in between frames for right, $T_{avgTimeR}$ and left, $T_{avgTimeL}$ cameras are equivalent to 41.8763 msec and 40.1092 msec, respectively which correspond to 1.7777 msec delayed for each acquired frame session.

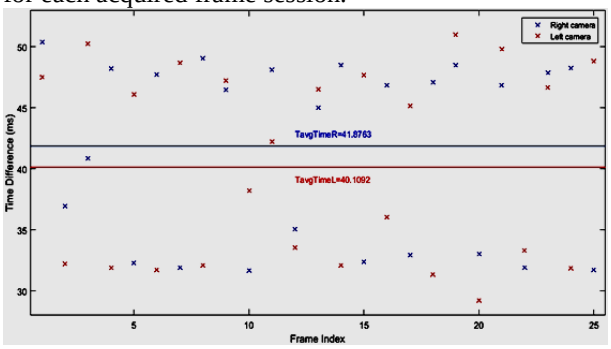

Fig. 6: Frame frequency for 25 fps performance

Calibration of the Stereo Camera

In computer vision, it is sometimes possible to acquire information from un-calibrated cameras, but camera calibration is absolutely necessary when dealing with metric information to obtain 3D information. In other words, the internal and external camera parameters obtained from the calibration process are required when constructing the 3D structure of a scene from the image pixel coordinates.

A. The Camera Model

A mathematical formulation which estimates the behavior of a camera is required. The extrinsic parameters are uniquely to identify the transformation between the known world reference and the unknown camera reference frames such as rotational and translation. Whereas, the intrinsic parameters are the internal characteristics of camera such as focal length, principal point and skew coefficient. Equation (1) derived by [14] represents a pinhole camera model that denote a projective mapping from the world coordinates (x, y, z) to pixel coordinates (u, v). The main objective of this transformation matrix is to find the relationship among the 3D points of the scene with their 2D projecting points in the plane image.

$$s\underbrace{\begin{bmatrix} u \\ v \\ 1 \end{bmatrix}}_{\tilde{\mathbf{m}}} = \underbrace{\begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & \mathbf{t} \end{bmatrix}}_{[\mathbf{R}\ \mathbf{t}]} \underbrace{\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}}_{\tilde{\mathbf{M}}} \quad (1)$$

Where,

α and β : Focal length
u_0 and v_0 : Principal point
γ : Skew parameter
r : Rotational vector
t : Translation vector

B.

The Calibration Methods

There are many different methods to compute the intrinsic and extrinsic parameters of camera [15]. In this work, two selective calibrating methods are being considered and evaluated. The first method is developed by [16] and it requires the two cameras to capture multiple planar chessboard pattern from different orientation views as shown in Fig. 7(a). This procedure consists of a closed-form solution and then a nonlinear refinement based on the maximum likelihood criterion. Simplicity and flexibility are the main reasons for choosing the calibration method since other advanced techniques require the use of expensive equipment and involve complex procedure.

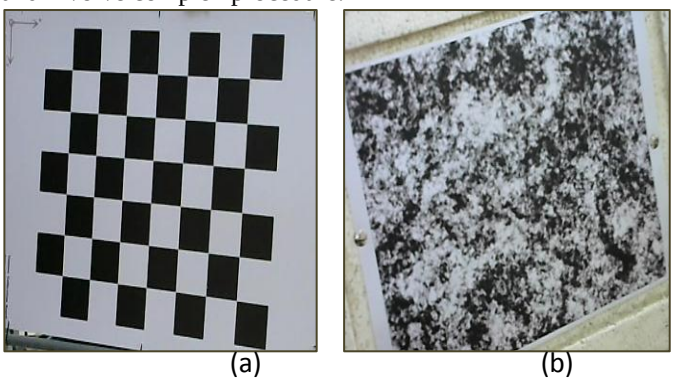

Fig. 7: Calibration patterns: (a) Chessboard (b) Feature Descriptor

The second calibration method considered is by [17] involving a novel calibration pattern as shown in Fig. 7(b) which consists of varying scales features. The features can be detected easily and guarantee a good calibration result for multiple cameras system. Unlike the other method, the calibration pattern does not have to be entirely within the field of stereo views.

C. Calibration Results

A 7 by 8, 74 mm^2 chessboard is used to calibrate the cameras using the first method while the second method uses the feature descriptor based pattern. Both printed patterns are mounted onto a rigid flat surface. Nineteen pairs of the calibration pattern of Fig. 7(a) and 20 pairs of the calibration pattern of Fig. 7(b) are captured by the neighboring cameras to obtain the unique feature point matches. The total number of feature points for the former method is 1596 while the latter has 2723 points.

The calibration results from the two methods are compared and it is observed that the estimated parameters are almost identical. However, the Mean Reprojection Error, MRE generated by the former method is much better than latter; with 0.1272 pixels and 0.6954 pixels, respectively. Nevertheless, both MRE results are still acceptable since the errors are less than one pixel. The resulting calibration results are as shown in Table 1.

Table 1: Stereo camera calibration result

	Method	1 (Former)	2 (Latter)
Criteria	**Pattern**	7x8 Chessboard	Feature Descriptor
	Image Size	480 x 640	480 x 640
	Images #	19 pairs	20 pairs
	Features # (a/ b)	798 / 798	1504/ 1219
Intrinsic Parameter	**Focal Length (α, β) [b]**	(832.2217, 831.6894)	(834.1979, 830.1998)
	Focal Length (α, β) [a]	(827.0566, 826.5095)	(828.5503, 824.7324)
	Principal Point (u0,v0) [b]	(306.0908, 240.3682)	(306.4708, 248.7132)
	Principal Point (u0, v0) [a]	(318.9395, 244.6339)	(305.1681, 248.8483)
	Skew (γ) [b]	0	0
	Skew (γ) [a]	0	0
Extrinsic Parameter	**Rotation vector (r) [a]**	[0.9999, -0.0076, -0.0084]	[0.9999, 0.0059, -0.0038]
	Translation vector (t) [a]	[75.2558, 0.7955, -2.7032] mm	[221.3464, 1.3795, 14.82282] mm
Accuracy	**Mean Reproduction Error, MRE**	0.1272 pixels	0.6954 pixels

[a]*Left camera*
[b]*Right camera*

D. Reprojection Error

The reprojection error is the geometric error distance in between the projected and the corresponding measured points in the pixel coordinate system. The chart in Fig. 8 shows the MRE calculated based on method 1 parameters. There are 42 detected points in each 19 images and the MRE is calculated individually, along with the overall mean error.

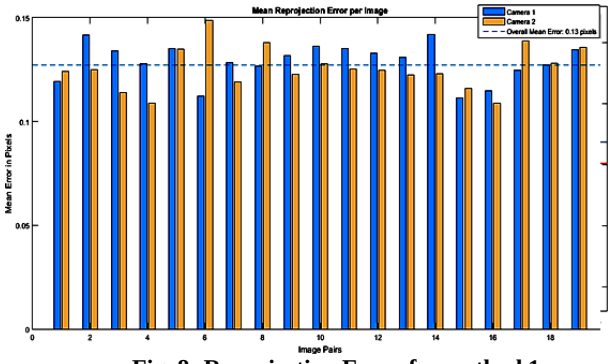

Fig. 8: Reprojection Error for method 1

3D Reconstruction

The actual shape and appearance of a physical scene can be reconstructed from stereo images [18]. The perspective projection in the *x* and *z*-axis of two cameras is as illustrated in Fig. 9 which portrays the fundamental of estimating the depth of the point of interest, P, in which it is considered as a single point in a 3D world.

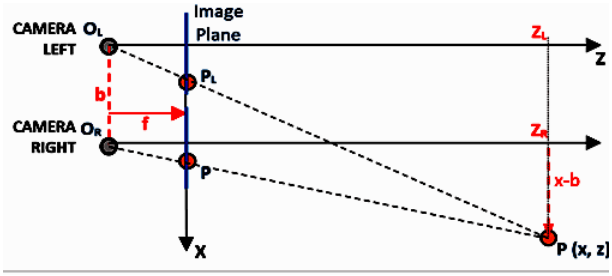

Fig. 9: Stereo camera geometrical projection

Note that, the *y*-axis is perpendicular to the plane. For an elementary description, the projection optic axes of two cameras are aligned in parallel. The points P_L and P_R are the projections of point P onto the image planes from projection centers, O_L and O_R respectively.

As can be seen, two major triangles (P, O_L, Z_L) and (P, O_R, Z_R) are formed from this stereo geometrical. Thus, the relation of corresponding sides can be derived as shown in Eq. (2).

$$Z / f = x / P_L \quad ; \quad Z / f = (x - b) / P_R \qquad (2)$$

Where,

$$Z = Z_R = Z_L$$

Based on this equation, the depth, *Z* of P can be computed as shown in Eq. (3) with respect to the known focal length, *f* and baseline, b which was computed in the previous subsection.

$$Z = (f \times b) / (P_L - P_R) \qquad (3)$$

Where,

$$\text{Disparity, } d = P_L - P_R$$

To find the disparity, *d* from the two images, the corresponding 2D coordinates (x, y) of P_L for P_R should be searched.

A. Image Rectification

Rectification is needed to minimized time exploration of the matching points by comparing each pixel in the images [19]. This process is done by horizontally aligned the images before the disparity of two points is calculated. Thus, the correlation routine searching of the corresponding point algorithm is applied for the *x*-axis only. Fig. 10 shows the red-cyan anaglyph from stereo images, where red represents image of the right camera and cyan represents image of the left camera.

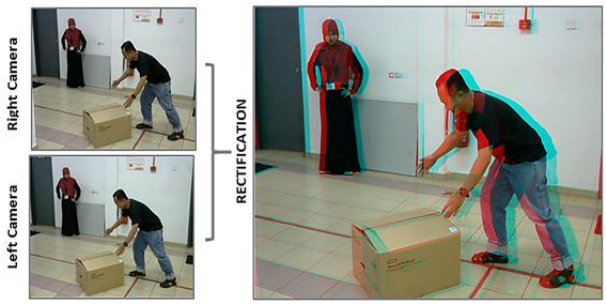

Fig. 10: Anaglyph image for rectified stereo images

B. Disparity

The correlation routine can be executed by computing disparity values for each pixel on the set of rectified stereo images as described in the previous subsection. The disparity map is composed by the computed depth value for each pixel in the 2D image as shown in Fig. 11.

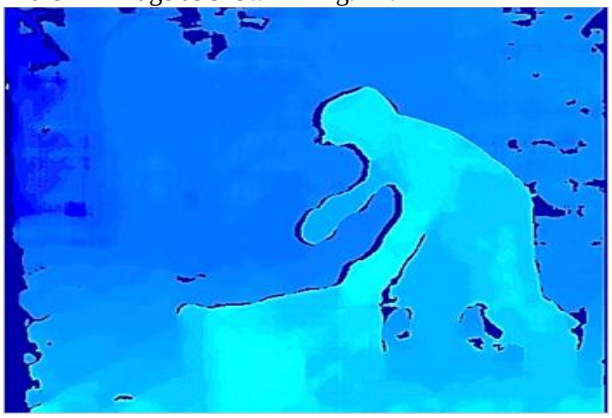

Fig. 11: Disparity map from the rectified 2D images

The darker intensities represent the objects that are far from the camera, whereas the brighter intensities correspond to objects closer to the camera. From the disparity map, each pixel corresponding to the (x, y, z) coordinates that describes its position in the 3D space can be re-projected into a set of data points in the coordinate system of a 3D point cloud. Fig. 12 and Fig. 13 represent the 3D scenes captured from the stereo camera.

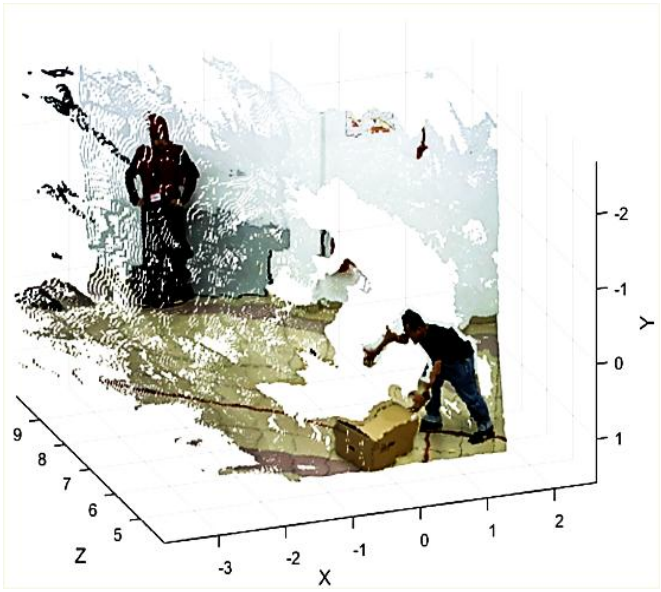

Fig. 12: 3D point cloud re-projected from stereo images

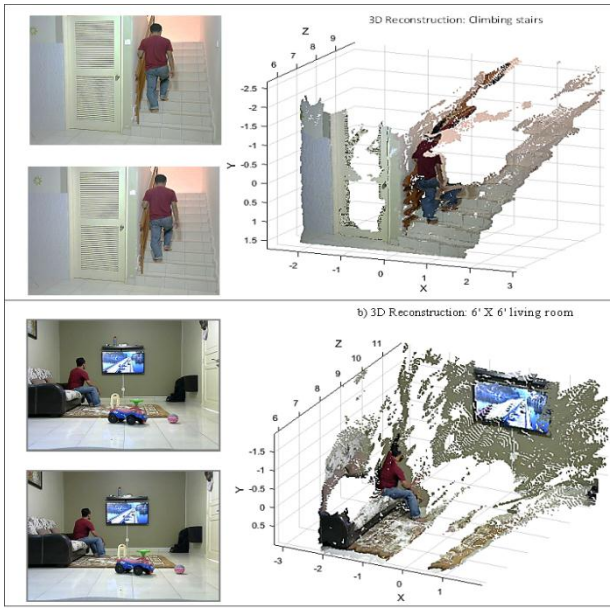

Fig. 13: 3D reconstruction from stereo camera images

3D Distance Measurement Accuracy

The two calibrating methods have been implemented on the stereo camera and based on the smallest MRE, the estimated parameters generated by method 1 are used to measure the 3D coordinates. The goal of this section is to compare the estimated distance extracted from 3D point cloud with the ground truth measurement. The distance in between the stereo camera and target board is the parameter to measure for evaluating the accuracy of 3D sensor reading.

The actual depth distance is measured by using a laser-type distance measuring device, Strait-Line that is perpendicularly mounted on the center top of the stereo camera enclosure. The checker board pattern is printed on a none glossy A4 size white paper and used as a target board. The stereo camera prototype and the target board are mounted on two separate tripods as shown in Fig. 14 during the measurement process, an assumption was made such that the camera and target board were perpendicularly to the flat surface of the floor. The four corners likelihood at each pixel are detected using the Corner Detection algorithm by [20].

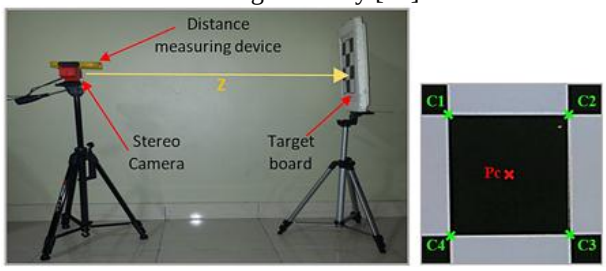

Fig. 14: Experimental hardware setup and target features

Then, the square center location is computed from the detected corner points diagonally (i.e C1 and C3) in 2D Euclidean space as shown in Eq. (4):

$$(x_p, y_p) = \left(\frac{x_{C1}+x_{C3}}{2}, \frac{y_{C1}+y_{C3}}{2}\right) \quad (4)$$

Fig. 14 shows the required features where the detected square corners are marked in green, and labelled as C1-C4 while the square center detection in marked in red as P_C. This process is imposed for both images captured by the stereo camera. Next, the depth, z distance from the square center location relative to the optical center of camera 1 are obtained from matching points in the images using stereo triangulation. The discrepancies between the ground truth value and i) the measurement obtained using the laser-type measuring device and ii) the estimated computed value is compared and defined as error measurements. In addition, to evaluate accuracy of the estimated 3D coordinate; x, y and z, the distance between point C1 and C2 is also computed in a 3D Euclidean space using Eq. (5).

$$D(C_1, C_2) = \sqrt{(x_{C1} - x_{C2})^2 + (y_{C1} - y_{C2})^2 + (z_{C1} - z_{C2})^2} \quad (5)$$

3. RESULTS AND DISCUSSION

The process above is iterated using different z distances; starting from 1000 mm to 5500 mm with an interval distance of 50 mm. As shown in Fig. 15, the stereo camera is able to give a good depth, z reading based on the estimated extrinsic and intrinsic parameters computed previously.

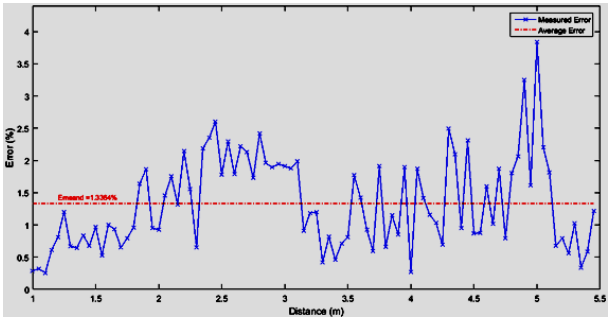

Fig. 15: Accuracy of depth, z reading from stereo camera

Based on this experiment, the maximum and minimum reading errors are 3.84% and 0.26%, respectively. The depth mean error measurement is 1.37% and the standard deviation is 0.71%; which is equivalent to 44.8 mm and 32.8 mm respectively. Unfortunately, the sensor was not able to give depth reading when the distance range was extended beyond 5500 mm. It was also observed that the corner detection algorithm was unable to detect square corner feature due to the limitation of image resolution captured by the two cameras. To resolve this, either a pre-processing step can be done on the images or a higher resolution camera are used.

Meanwhile, accuracy of the distance measurement in 3D Euclidean results in 5.36% and 0.001% of the maximum and minimum reading errors, respectively. The mean reading distance measurement error is 1.72% ± the standard deviation of 1.6%; which is equivalent to 44.8 mm and 32.8 mm, respectively.

4. CONCLUSION

This work affords the development of a low-cost 3D stereo vision system which consists of a pair of mid-range webcams

mounted inside a rigid enclosure. Two calibrating methods were reviewed, implemented on the custom-made 3D sensor and their performances evaluated. The first method being considered, that used the checker board pattern, outperformed the second method, which used the unique pattern. The former gives the smallest MRE of 0.1272 pixel and produces a good 3D image reconstruction representing the actual 3D physical scene. In conclusion, the developed 3D sensor was able to estimate a good 3D distance reading within an acceptable error percentage. For future work, the outcome of this work will be used to in an actual surveillance system to detect and identify anomalies in human action recognition tasks.

5. REFERENCES

[1] Kuala Lumpur Overseas Security Advisory Council, 2016. Malaysia. Crime and Safety Report, 2015. Retrieved from https://www.osac.gov/pages/ContentReportPDF.aspx?cid=17215

[2] Hampapur, A., Brown, L., Connell, J., Pankanti, S., Senior, A., & Tian, Y. Smart surveillance applications, technologies and implications. Information, Communications and Signal Processing, 1133–1138, 2003. DOI: 10.1109/ICICS.2003.1292637

[3] Müller, K., Smolić, A., Dröse, M., Voigt, P., & Wiegand, T. 3-D reconstruction of a dynamic environment with a fully calibrated background for traffic scenes. IEEE Transactions on Circuits and Systems for Video Technology, 15(4), 538–548, 2005. DOI: 10.1109/TCSVT.2005.844452

[4] Mishra, A. K., Ni, B., Winkler, S., & Kassim, A., 3D surveillance system using multiple cameras. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 6491, 12007. DOI: 10.1117/12.703449

[5] Abu Hassan, M. F., Zulkifley, M. A., & Hussain, A., Squat Exercise Abnormality Detection by Analyzing Joint Angle for Knee Osteoarthritis Rehabilitation. Jurnal Teknologi, 12, 133–139, 2015. DOI: 10.11113/jt.v77.6241

[6] D'Addio, G., Iuppariello, L., Gallo, F., Bifulco, P., Cesarelli, M., & Lanzillo, B. Comparison between clinical and instrumental assessing using Wii Fit system on balance control. IEEE MeMeA 2014 - IEEE International Symposium on Medical Measurements and Applications, Proceedings, 2014. DOI: 10.1109/MeMeA.2014.6860124

[7] Pang, J. M., Yap, V. V., & Soh, C. S. Human Behavioral Analytics System for Video Surveillance, (November), 28–30 , 2014

[8] Ma, J., Wei, Y., & Su, P. A Portable System for 3D Reconstruction Based on Stereo-View, 167–172, 2016 DOI: 10.1109/CYBER.2016.7574816

[9] Neethu, S., & Vinuchackravarthy, S. Object Detection Using Binocular Vision, 1558–1564, 2016.

[10] Schreve, K.. How accurate can a stereovision measurement be? 15th International Workshop on Research and Education in Mechatronics, REM 2014, (1), 2014. DOI: 10.1109/REM.2014.6920229

[11] Oleari, F., Rizzini, D. L., & Caselli, S. A Low-Cost Stereo System for 3D Object Recognition. IEEE International Conference on Intellgent Computer Communication and Processing, 127–132, 2013. DOI: 10.1109/ICCP.2013.6646095

[12] Huang, G., & Yu-Yong, T. Application of Stereo Vision 3D Target Recognition Using Camera Calibration Algorithm. In International Conference on Circuits and Systems (CAS 2015) (pp. 381–385) , 2015. DOI: 10.2991/cas-15.2015.91

[13] Logitech. (n.d.). HD Webcam C270 Specification. Retrieved from http://support.logitech.com/en_us/product/hd-webcam-c270

[14] Hartley, R., & Zisserman, A. Multiple view geometry in computer vision. Cambridge University Press, 2003.

[15] Salvi, J., Armangué, X., & Batlle, J. A comparative review of camera calibrating methods with accuracy evaluation. Pattern Recognition, 35(7), 1617–1635, 2001. DOI: 10.1016/S0031-3203(01)00126-1

[16] Bouguet, J. Y. Camera calibration toolbox for Matlab, 2004. Retrieved July 13, 2016, from http://www.vision.caltech.edu/bouguetj/calib_doc/.

[17] Li, B., Heng, L., Köser, K., & Pollefeys, M. A Multiple-Camera System Calibration Toolbox Using A Feature Descriptor-Based Calibration Pattern. IEEE RSJ International Conference on Intelligent Robots and Systems (IROS), 2013. DOI: 10.1109/IROS.2013.6696517

[18] Short, N. 3-D Point Cloud Generation from Rigid and Flexible Stereo Vision Systems, 2009.

[19] Loop, C., & Zhang, Z. Z. Z. Computing rectifying homographies for stereo vision. Proceedings of IEEE Conference on Computer Vision and Pattern Recog- Nition, 1, 125–131, 1999. DOI: 10.1109/CVPR.1999.786928.

[20] Geiger, A., Moosmann, F., Car, O., & Schuster, B. Automatic camera and range sensor calibration using a single shot. IEEE International Conference on Robotics and Automation, 3936–3943, 2012. DOI: 10.1109/ICRA.2012.6224570