Accuracy and reliability of optimum distance for high performance Kinect Sensor

Conference Paper · March 2015

DOI: 10.1109/ICoBE.2015.7235927

5 authors, including:

Cheechin Lim
Universiti Malaysia Perlis
36 PUBLICATIONS 67 CITATIONS

Shafriza nisha Basah
Universiti Malaysia Perlis
76 PUBLICATIONS 564 CITATIONS

Sazali Yaacob
University of Kuala Lumpur
366 PUBLICATIONS 4,341 CITATIONS

ewe juan Yeap
Ministry of Health Malaysia
22 PUBLICATIONS 257 CITATIONS

2015 2nd International Conference on Biomedical Engineering (ICoBE), 30-31 March 2015, Penang

Accuracy and Reliability of Optimum Distance for High Performance Kinect Sensor

Lim Chee Chin, Shafriza Nisha Basah, Sazali Yaacob
School of Mechatronics
Universiti Malaysia Perlis
Ulu Pauh, Perlis, Malaysia
cheechin10@gmail.com, shafriza@unimap.edu.my, sazali@unimap.edu.my

Mohamad Yazid Din, Yeap Ewe Juan
Department of Orthopedic and Traumalogy
Hospital Tunku Fauziah
Kangar, Perlis, Malaysia
myaziddin@gmail.com, ejyeap@yahoo.com

***Abstract*— A depth/range camera is able to return images like an ordinary camera, however, instead of color, each pixel value represent a particular distance correspond to a point in a real world. Recently, the Microsoft Kinect Sensor is a popular measurement tool due to its ability to provide length measurement, volume measurement, and motion measurement in Biomedical Application. Every researcher wants to know accuracy, precision and reliability of Kinect sensor depth distance performance. This paper presents an investigation of the quality of depth data obtained by Kinect sensor. Based on Kinect Sensor's specification by Microsoft a theoretical accuracy, precision and reliability analysis is presented, which provides an insight into the factors influencing the accuracy of the data. Experimental results show that the absolute mean percentage error of depth measurement increases with increasing distance to the sensor, and ranges from a few millimeters up to about 40mm at the maximum range of sensor. The standard deviation at different distances of plane to the sensor is provided a similar quadratic curve in precision. The reliability of Kinect Sensor is evaluated by using Kuder-Richardson formula 20 shown that 0.76 in reliability. The accuracy, precision and reliability of depth distance is important and beneficial to biomedical application due to the depth perception is required to determine the 3D estimation pose in human motion application.**

Keywords*— Kinect Sensor, Kinect for Windows, Depth Distance, Default Range, Near Range.

I. INTRODUCTION *(HEADING 1)*

Nowadays 3D technologies are becoming more and more popular, which by itself increases the amount of 3D application in the biomedical application. There are a lot of range camera is developed in computer vision projects that applicable in biomedical applications such as Time-of-flight (TOF) and Microsoft Kinect Sensor.

The TOF camera emits modulated light to the scene and the reflection of the light is measured with the reference signal, which is correlated with the modulated light, in order to obtain the depth information [1][1]. The Kinect Sensor uses another kind of technique which consists of an infrared structured light projector and a monochrome CMOS camera to compute depth of the scene.

Now after the release of Microsoft Kinect Sensor and other depth cameras has provided 3D technologies in biomedical application. Kinect Sensor by Microsoft is one of the range cameras that developed for biomedical applications. The depth sensor is able to return images like an ordinary camera, but instead of color, each pixel value represents the distance to the point. Kinect Sensor is a researcher-like camera due to its specification. This paper is written to prove that this Kinect Sensor is able to provide the accuracy, precision and reliability of depth distance values same with actual distance. The accuracy of Kinect Sensor is important to analysis for the degree of closeness of measurements of depth distance to the actual distance. In Kinect Sensor's precision, the depth array has a precision of up to only 11 bits or 2^{11} = 2048 difference values. Thus, the Kinect Sensor's depth measurement will provided a non-linear function of distance [2]. However, there is no previous researchers review on reliability of its depth distance. But they are focus on its reliability of the Kinect within functional assessment activities in biomedical application [3], [4]. Moreover, the default range and near range of the distance from Kinect Sensor is evaluated is this paper.

I. LITERATURE REVIEW

The Kinect Sensor hardware specification has been described in many recent publications [3], [5]–[7]. The Kinect Sensor (Figure 1) has a 640 x 480 32-bit RGB color camera and a 320*240 16-bit depth camera both running at 30 frames per second. The field of view (FOV) of cameras are 57º horizontally and 43º vertically. It has a motorized angle driver which can tile the both camera 27º vertically and a microphone array which is used for detecting voice commands [8]. The Kinect Sensor hardware specification is able to apply in many applications such as medical, robotic, and gaming application. The working principle of Microsoft Kinect Sensor is that emission of IR pattern from Kinect depth sensor and simultaneously image capture of IR images with a (traditional) CMOS camera that is fitted with an IR-pass filter. The image processor of Kinect Sensor uses the relative positions of the dots in the pattern to calculate the depth displacement at each pixel position in the image. The algorithm for calculating the depth and accuracy of the Kinect Sensor has yet not released by Microsoft. However, few papers have computed the absolute accuracy of the depth map [9]–[11].The depth sensor from Kinect Sensor returns an 11-bit number raw data which needs further processing to get the true depth. The true depth is expressed as found in [9].

978-1-4799-1749-5/15/$31.00 ©2015 IEEE

Khoshelham *et al. [12]* provide an insight into the geometric quality of Kinect Sensor depth data based on analysing the accuracy and resolution of the depth signal. Experimental results show that the random error of depth measurement increases when the distance between the scene and the sensor increases, ranging from a few millimetres at close range to about 4 cm at the maximum range of the sensor.

Fig. 1. Kinect Sensor with external case removed showing the three primary sensors: IR structured light projector (left), RGB camera (center left), and IR detector (center right) [6].

Based on Microsoft Kinect Developer [13], the depth sensor has two depth ranges: the default range and the near range (shown in the Depth Range Enumeration). In Figure 2 illustrates that the sensor depth ranges in meters. The default range is available in both the Kinect for Windows sensor and the Kinect for Xbox 360 sensor; whereas the near range is only available in the Kinect for Windows sensor which shown in Figure 2. As such, the depth sensor can be seen as a device that returns (x, y, z)-coordinates of 3D objects. From Figure 1, default range has a blind spot from 0m to 0.8m, where it cannot generate raw depth data for anything in that region. It also cannot generate raw depth data for anything after 4 meters. On the other hand, the near range has a blind spot from 0m to 0.4m where it cannot generate any raw depth data for anything in that region. It also cannot generate any raw depth data for anything after 3 meters.

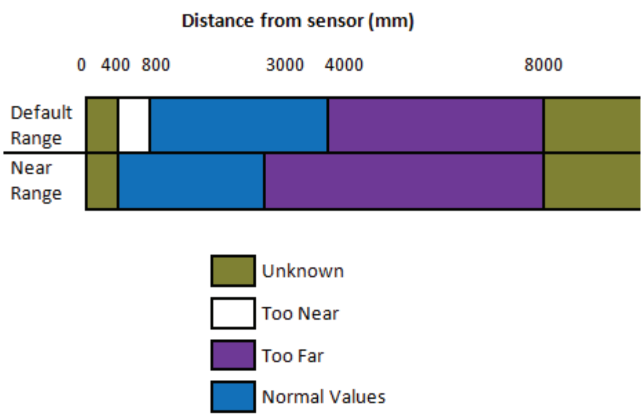

Fig. 2. Microsoft Kinect Sensor Depth Distance (Source: Microsoft Developer Network Officer Website[13]*)*

II. Design Methodology

An experiment had been conducted to evaluate the Microsoft Kinect Sensor Depth Distance for default range and near range. A Kinect Sensor is placed on the floor and also a long piece of ruler tape orthogonal to the Kinect's field of view. The measurements are marked along the piece of tape from 200mm to 4000mm from the sensor. A cardboard box is placed in the center of the Kinect's field of view for each hash; from 200mm all the way back to 4000mm.

The experiment layout is shown in the Figure 3. The actual distance between the Kinect Sensor and cardboard is varied along the piece of ruler tape. The actual distance and the depth distance from the Kinect Sensor are recorded for every varied 100mm. The purpose that actual distance and depth distance is recorded to identify its absolute accuracy of the depth map.

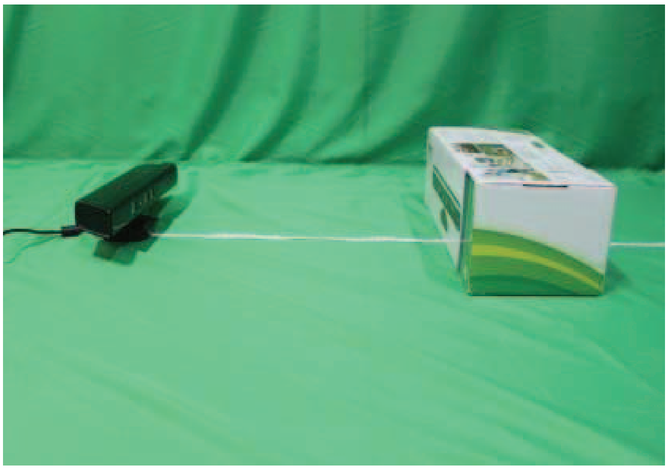

Fig. 3. Experiment Layout

A. Program Design

The Kinect Sensor is programmed by using Kinect for Window Software Development Kit (SDK) 1.8 with C# programming languages. The integrated developer toolkit includes sample applications with access to full source code, Kinect Studio, and resources to simplify and speed up application development. The SDK provides the tools and APIs, both native and managed.

An algorithm program had been designed to calculate the depth distance of the Kinect Sensor depicted in Figure 5. The Kinect depth sensor calculates distance based on the straight distance between the object and sensor. The distance calculates by an internally drawn line is perpendicular to the sensor in a diagonal views of pixel. In Kinect Sensor properties, depth image stream and depth stream frame classes have the *MaxDepth* and *MinDepth* properties, which return the maximum and minimum depth ranges for that particular stream or captured image frame, in millimeters. This range value returns the best range where depth distance can be measured properly.

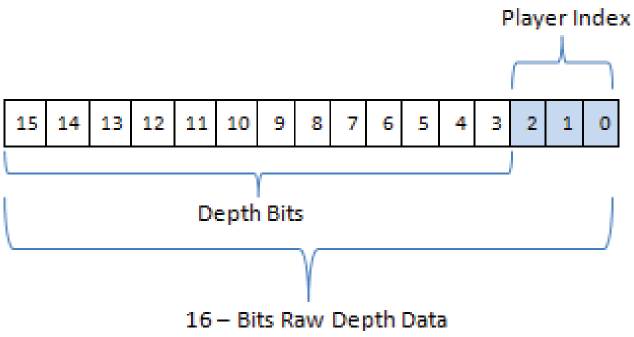

Fig. 4. Depth Frame bit pixel data in Kinect [14]

The Kinect Sensor returns 16-bit raw depth frame data in a row of stride in Figure 4. The first three bits are used to represent the identified players and the remaining 13 bits give you the measured depth distance in millimeters. In the Figure 4, the upper 13 bits, which represent the actual distance of the pixel value from the sensor and lower 3 bits represent player index. For depth distance calculation from the 13 bits, a bitwise shift operation (>>) need to be perform to move the bits to their correct position.

Algorithm Depth Distance
Start
1. Enabling the color and depth stream channel *2. Kinect Sensor is Started* *3. Attaching Color, Depth, and Skeleton Event handler* *4. Content of Mindepth and Maxdepth is get from the maximum depth ranges and minimum reliable depth ranges for the current frame* *5. Processing the color frames* 6. *Get the pixel data in byte array* *7. Copy the pixel data and assign bitmap image source into image* 8. *Calculate stride = imageFrame.Width * imageFrame.BytesPerPixel;* *9. Calculate pixelIndex from the X and Y coordinates of the selected position* *10. Distance calculation with depth data by using bitwise shift operation.11. Display depth data, and Coordinate X, Y in GUI for analysis.*
End

Fig. 5. Algorithm of calculating Depth Distance from Kinect Sensor

B. *Depth Distance Calculations*

In this experiment, there are 5 tests were done for each difference distance range for default range and near range. The distance range is taken from 200mm until 4000mm, with different of 100mm. The purpose to carry out 5 tests is to get the average distance. Thus, the average distance for 5 tests has been calculated by using Equation (1). For analysis its distance error, the absolute mean percentage error (*AMPE*) is calculated to identify the most accurate range for measurement in default range and near range application presented in Equation (2). Further analysis on its precision, the standard deviation is calculated to identify the precision of the depth data provide by Kinect for every distance. The reliability of Kuder-Richardson formula 20 is used for analysis the consistency of Kinect depth data result from the 5 testing experiments which are Test 1, Test 2, Test 3, Test 4, and Test 5.

$$Average, \bar{x}\ (mm) = \frac{(Test\ 1 + Test\ 2 + Test\ 3 + Test\ 4 + Test\ 5)}{5} \quad (1)$$

$$AMPE(\%) = \left|\frac{100\%}{n} \times \sum_{i=1}^{n} \left(\frac{Test\ i - Actual\ Distance}{Actual\ Distance}\right)\right| \quad (2)$$

$$Standard\ Deviation = \sqrt{\frac{\sum(Actual\ Distance - \bar{x})^2}{n}} \quad (3)$$

i is the number of each test, $i=1,2,3,4,5$
$\bar{x}$ is the average of each distance
$\sum$ is indicate to sum
n is the total number of test taken which $n=5$

$$r_{KR20} = \left(\frac{k}{k-1}\right)\left(1 - \frac{\sum pq}{\sigma^2}\right) \quad (4)$$

R_{KR20} is the Kuder-Richardson formula 20
k is the total number of test items
$\sum$ is indicate to sum
p is the distance of testing is pass ±5mm
q is the distance of testing is fail
σ^2 is the variation of entire test

III. Result And Discussion

This Experimental Design with a cardboard box is placed in the center of the Kinect's field of view for each hash; from 200mm all the way back to 4000mm. The design of measurement setting in the center area of depth images due to human motion tracking application is mostly focus on the center of view. In this experiment, the cardboard box is used instead of human bodies. The reasons that using cardboard box is to focus on depth distance instead of human bodies' detection. If using human bodies, it consists of curve surface, which will cause error due to the detection problem. However, researcher unable to get the effect of depth distance when using human bodies in the experiment.

The Graphical User Interface (GUI) is display from the Kinect SDK programming C# with visual studio as depicted in Figure 6. The reason that Microsoft SDK is chosen to be used is because it is the only available for Windows. Microsoft SDK able to provide human anatomical landmark data automatically determined in close to real-time by the machine-learning algorithm as presented in [15], [16]. Microsoft Kinect SDK supports skeleton tracking with depth to create animations of body movements for the avatar [17]. Whereas, the others software frameworks such as OpenNI and OpenKinect are multi-platforms and open-source [18]. The Microsoft SDK only delivers depth values up to approximately 4000 millimeters. This is likely to be a choice by Microsoft due to low resolution and noise resistance on greater distances. The disadvantage of OpenNI is its difficulties in calibration to provide good depth distance [8]. However, the OpenKinect framework is not linearized and therefore always has a resolution of 1 bit [18]. As the depth value is not linearized this does not mean that the OpenKinect can measure smaller depth differences. Thus, Microsoft SDK is used as software framework in the experiment.

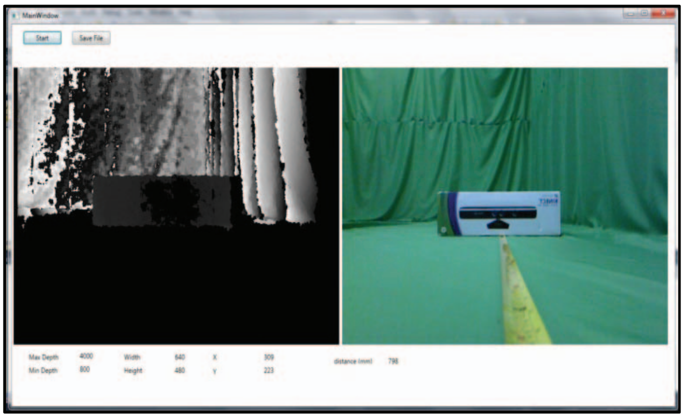

Fig.6. Graphical User Interface (GUI) of default range and near range

The Microsoft SDK is also limited to depths above 800 millimeters for default range and go down to 500 millimeters for near range [18]. However, we discovery that default range mode is able provide range measurement start at 500 millimeters until 4000mm which is better than near range as presented in Figure 7. Therefore, the default range is a function in Kinect for Windows that enables the depth camera to see objects as close as 500mm in front of the sensor and as far as 4000mm while maintaining accuracy, precision and reliability. In addition, the AMPE for the range in between 600mm until 2900mm is less the 1%. Thus, default range given the accurate depth values result in its own range and also near range as shown in Figure 8. The red line in the Figure 8 had shown the similar quadratic graph shape with increasing data from 40mm to maximum ranges of 4000mm. The overall error of depth distance is in tolerance of 1.5%.

Default range skeletal tracking measurement will return position data for all of the joints (20 joint types) for full human body in default mode [19]; whereas, near range skeletal tracking are only for 10 joints are tracked for upper body in seated mode. Near range skeletal tracking only focus on the user's head, torso, and arms are visible and tracked due to its limited field of view for closer application in skeletal tracking [19]. Thus, the lower 10 joints are returned as untracked. Therefore, default range is more applicable and researcher-like for most of the computer vision application. Default range used in a lot of application such as face detection [20], physical rehabilitation [10], 3D human pose estimation [21]–[24], robotic application [25] and so on.

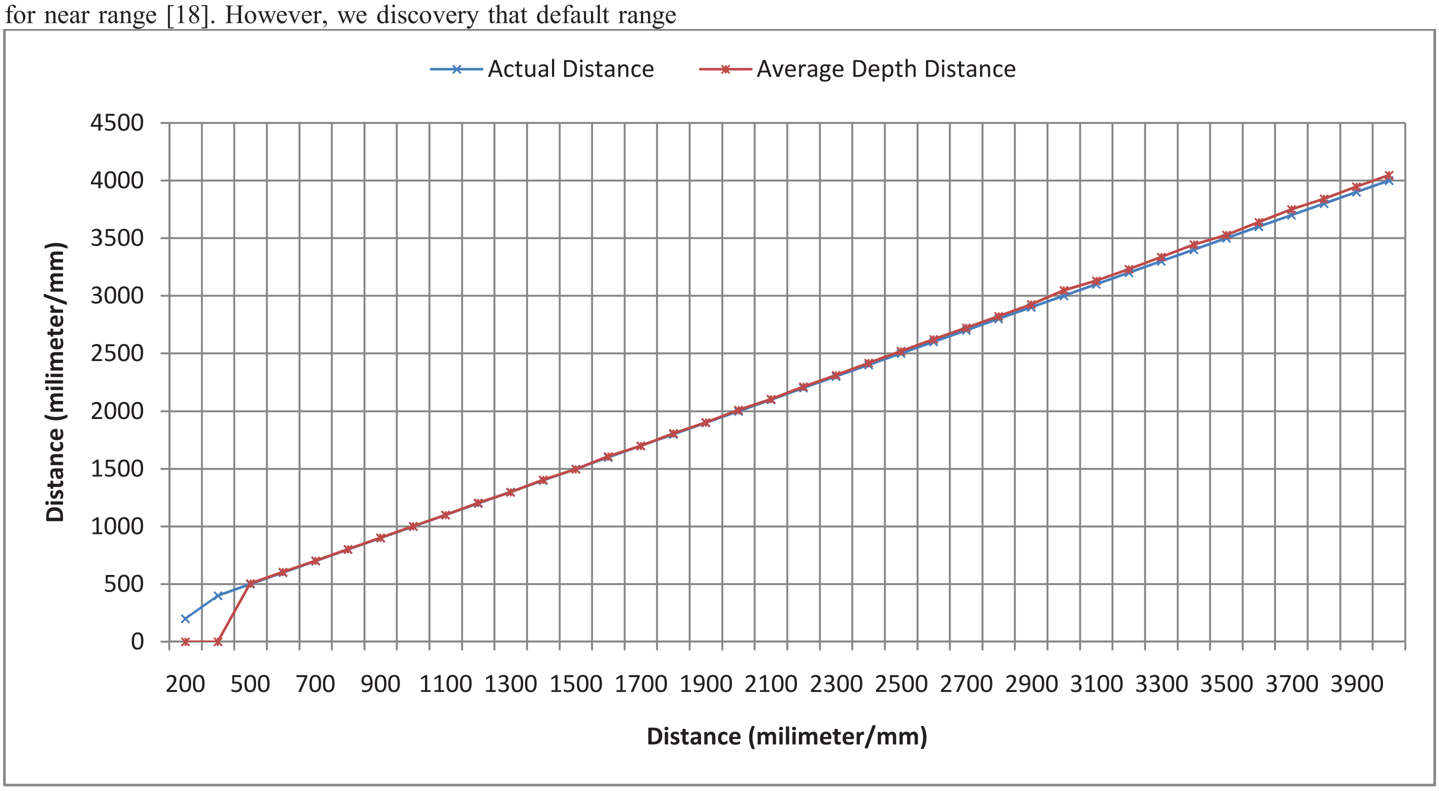

Fig.7. Comparison Average Depth Distance vs. Actual Distance.

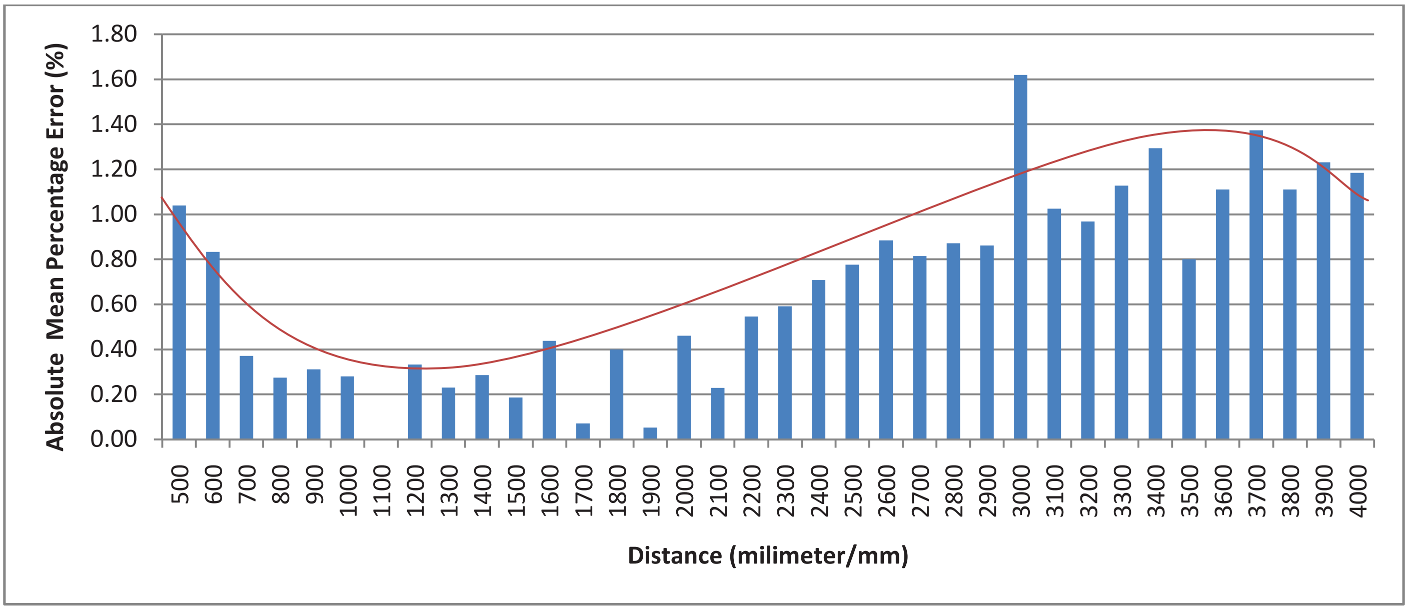

Fig. 8. For Accuracy Analysis with Absolute Mean Percentage Error vs. Distance

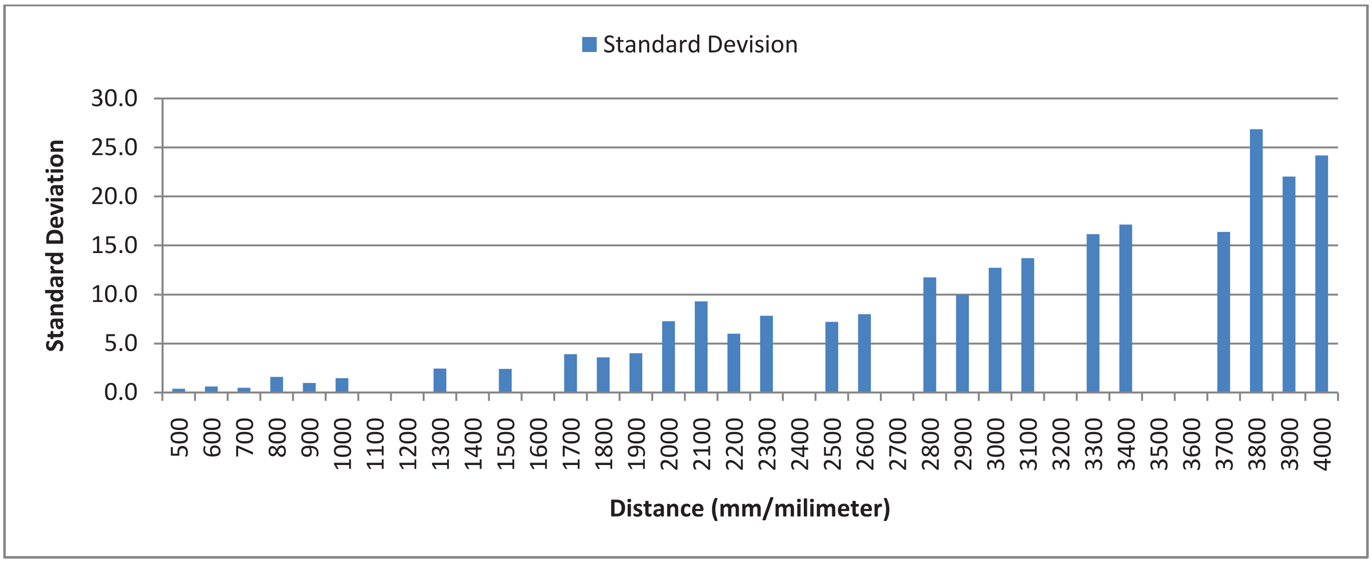

Fig. 9. For Precision Analysis with Standard Deviation vs. Distance

IV. CONCLUSION

In conclusion, the depth distance of default range from the experimental results show that Kinect Sensor is able to provide data with good accuracy precision and reliability of depth distance values compared to the actual distances. From the results the following general conclusions can be drawn:

- The Kinect Sensor does not contain large systematic errors thought its depth distance. This error of depth distance increase with at the low range of 600mm and high range of 2900mm.
- The random error of depth distance measurements increases quadratically with increasing data from the sensor and reaches 40mm at maximum ranges
- The consistency of depth distance results is prerequisite to validity of the Kinect sensor in all distance range.

In general, for biomedical applications the data should be acquired within 600mm ~ 3000mm distance to the sensor. At the too far and too near distances, the quality of the data is degraded by the noise and low resolution of the depth distance measurements.

ACKNOWLEDGEMENT

We would like to thank to the Universiti Malaysia Perlis for their support of our projects. Thanks are also giving to RAGS Funds for providing Funding support for project operations.

REFERENCES

[1] J. U. Kuehnle, Z. Xue, M. Stotz, J. M. Zoellner, A. Verl, and R. Dillmann, "Grasping in depth maps of time-of-flight cameras," in *ROSE 2008 - IEEE International Workshop on Robotic and Sensors Environments Proceedings*, 2008, pp. 132–137.

[2] N. Crock, "Kinect Depth vs. Actual Distance," 2011. [Online]. Available: http://mathnathan.com/2011/02/depthvsdistance/. [Accessed: 03-Jan-2015].

[3] B. F. Mentiplay, R. a Clark, A. Mullins, A. L. Bryant, S. Bartold, and K. Paterson, "Reliability and validity of the Microsoft Kinect for evaluating static foot posture.," *J. Foot Ankle Res.*, vol. 6, p. 14, 2013.

[4] B. Bonnechère, B. Jansen, P. Salvia, H. Bouzahouene, L. Omelina, F. Moiseev, V. Sholukha, J. Cornelis, M. Rooze, and S. Van Sint Jan, "Validity and reliability of the Kinect within functional assessment activities: Comparison with standard stereophotogrammetry," *Gait Posture*, vol. 39, pp. 593–598, 2014.

[5] Microsoft, "Kinect for Windows," 2012. [Online]. Available: http://www.microsoft.com/en-us/kinectforwindows/. [Accessed: 03-Jan-2015].

[6] Open Source Bobotic Foundation, "kinect - ROS Wiki," 2013. [Online]. Available: http://wiki.ros.org/kinect. [Accessed: 03-Jan-2015].

[7] Microsoft Developer Network, "Kinect for Windows Sensor Components and Specifications," 2014. [Online]. Available: http://msdn.microsoft.com/en-us/library/jj131033.aspx. [Accessed: 03-Jan-2015].

[8] Microsoft Developer Network, "Skeleton Tracking Mode Enumeration," 2014. [Online]. Available: http://msdn.microsoft.com/en-us/library/microsoft.kinect.skeletontrackingmode.aspx. [Accessed: 03-Jan-2015].

[9] K. D. Mankoff and T. A. Russo, "The Kinect: A low-cost, high-resolution, short-range 3D camera," *Earth Surf. Process. Landforms*, vol. 38, pp. 926–936, 2013.

[10] N. Kitsunezaki, E. Adachi, T. Masuda, and J. I. Mizusawa, "KINECT applications for the physical rehabilitation," in *MeMeA 2013 - IEEE International Symposium on Medical Measurements and Applications, Proceedings*, 2013, pp. 294–299.

[11] J. Han, L. Shao, D. Xu, and J. Shotton, "Enhanced computer vision with Microsoft Kinect sensor: A review," *IEEE Trans. Cybern.*, vol. 43, pp. 1318–1334, 2013.

[12] K. Khoshelham and S. O. Elberink, "Accuracy and resolution of kinect depth data for indoor mapping applications," *Sensors*, vol. 12, pp. 1437–1454, 2012.

[13] Microsoft Developer Network, "Coordinate Spaces," 2014. [Online]. Available: http://msdn.microsoft.com/en-us/library/hh973078.aspx#Depth_Ranges. [Accessed: 03-Jan-2015].

[14] A. Jana, *Kinect for Windows SDK Programming Guide*, First. Birmingham B3 3PB, UK: Packt Piblishing Ltd., 2012, p. 392.

[15] R. A. Clark, Y. H. Pua, K. Fortin, C. Ritchie, K. E. Webster, L. Denehy, and A. L. Bryant, "Validity of the Microsoft Kinect for assessment of postural control," *Gait Posture*, vol. 36, pp. 372–377, 2012.

[16] K. Labelle, "EVALUATION OF KINECT JOINT TRACKING FOR CLINICAL AND IN-HOME STROKE REHABILITATION TOOLS," 2011.

[17] K. Aitpayev and J. Gaber, "Creation of 3D Human Avatar using Kinect," vol. 01, no. 05, pp. 3–5, 2012.

[18] T. G. and P. A. M.R. Andersen, T. Jensen, P. Lisouski, A.K. Mortensen, M.K. Hansen, "KINECT DEPTH SENSOR EVALUATION FOR COMPUTER VISION APPLICATIONS," 2012.

[19] Microsoft Developer Network, "Tracking Skeletons in Near Depth Range," 2014. [Online]. Available: http://msdn.microsoft.com/en-us/library/hh855353.aspx. [Accessed: 03-Jan-2015].

[20] F. Guimbretière and C. Nguyen, "Bimanual marking menu for near surface interactions," *Proc. 2012 ACM Annu. Conf. Hum. Factors Comput. Syst. - CHI '12*, p. 825, 2012.

[21] J. Shotton, R. Girshick, A. Fitzgibbon, T. Sharp, M. Cook, M. Finocchio, R. Moore, P. Kohli, A. Criminisi, A. Kipman, and A. Blake, "Efficient human pose estimation from single depth images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, pp. 2821–2840, 2013.

[22] R. Girshick, J. Shotton, P. Kohli, A. Criminisi, and A. Fitzgibbon, "Efficient regression of general-activity human poses from depth images," in *Proceedings of the IEEE International Conference on Computer Vision*, 2011, pp. 415–422.

[23] K. Lai, J. Konrad, and P. Ishwar, "A gesture-driven computer interface using Kinect," in *2012 IEEE Southwest Symposium on Image Analysis and Interpretation*, 2012, pp. 185–188.

[24] L. Chen, H. Wei, and J. Ferryman, "A survey of human motion analysis using depth imagery," *Pattern Recognit. Lett.*, vol. 34, pp. 1995–2006, 2013.

[25] R. A. El-Laithy, J. Huang, and M. Yeh, "Study on the use of Microsoft Kinect for robotics applications," in *Record -*

IEEE PLANS, Position Location and Navigation Symposium, 2012, pp. 1280–1288.