VOL. 10, NO. 17, SEPTEMBER 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

ANALYSIS OF STATIC AND DYNAMIC MOTION ACCURACY FOR KINECT-VIRTUAL SENSEI SYSTEM

Ku Mohd Sabri Tuan Yusuf, Ahmad Faiz Ahmad Nazri, Ghazirah Mustapha and Jamaluddin Mahmud
Faculty of Mechanical Engineering, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia
E-Mail: 2010937585@isiswa.uitm.edu.my

ABSTRACT

Motion analysis has been widely adapted in research pertaining to biomechanics and used for many important applications such as injury reduction, sports performance enhancement and rehabilitation. Nevertheless, current available motion capture system such as the use of infrared cameras is very expensive. Microsoft Kinect has the potential to be used as an alternative low-cost motion analysis tool. Nevertheless, the standard procedure for measuring its accuracy and reliability has not been well established. Therefore, this study for the first time attempts to develop a standard procedure to assess and visualise the accuracy and repeatability of Microsoft Kinect. A single-camera system is used to capture static and dynamic motions of healthy volunteers. Adapting numerical and statistical tools, the data are analysed for the i) static motion capture (standing still with lateral hand lift) and ii) dynamic motion capture (simple lower arm movement), which are tracked by the sensor operated using open source Virtual Sensei Lite program. The variance and error value are then analysed to determine the accuracy of measurement. The study able to demonstrate average errors of less than 2% (static) and 5% (dynamic) accuracy respectively. The good results prove that the current study is important and could contribute a significant knowledge for further research in improving Microsoft Kinect functions and applications for motion analysis.

Keywords: microsoft kinect, virtual sensei, motion capture, motion analysis, accuracy, repeatability.

INTRODUCTION

The success of Microsoft in developing gaming console has led to the development of Kinect sensor technology (Corazza *et al*., 2006). Purely developed as Xbox360 gaming console accessory, researchers have discovered that the depth sensing technology of Kinect could be extended far beyond the gaming application as this technology offers much lower cost than traditional 3D cameras such as stereo, time-of-flight (TOF) and infrared cameras. Complementary nature of the depth and visual (RGB) information provided by Kinect, enables the motion sensor be widely used in a large number of applications, including close range 3D measurements (Molnár *et al.*, 2012). The high potential of this device can also be seen in ergonomics, coaching, sport and even security fields. Based on current studies, it appears that Kinect can be used to assess some motion in well-defined situation (Mustapha *et al.*, 2014a). The Kinect sensor has found itself in the mainstream of development of a low cost alternative for measurement and analysis of human motion kinematics (Timmi, 2011), (Bünger, 2013), (Mustapha *et al.*, 2014b).

Tracking Capability

Kinect sensor comprising of RGB camera, an infrared (IR) emitter and IR depth sensor (Dutta, 2012), which makes it capable to capture depth and colour images at 30 frames per second to generate three dimensional (3D) points from an infrared pattern projected on the subject. The IR emitter casts an IR speckle dot pattern into the 3D scene while the IR depth sensor captures the reflected IR speckles. The geometric relation between the IR emitter and IR depth sensor is obtained through an off-line calibration procedure. The IR emitter then projects a known light speckle into the 3D scene as this light speckle is invisible to the colour camera but can be viewed by IR depth sensor. Because of the uniqueness of each local pattern from the projected dots, the matching process between observed local dot patterns in the image with the calibrated projector dot pattern is feasible.

This technology of sensing and the software-based technology make it possible for Kinect be used for tracking the position of human body, even without the aid of handheld controllers or force platform (Galna *et al*, 2014). Moreover, it is able to track up to six people concurrently including two active players at the same time, featuring extraction up to 20 joints per player (Ott, 2013). The Microsoft Kinect is easy to operate and has the ability to contribute in dance gesture classification, pattern recognition and postural control assessment (Clark *et al.*, 2012). Furthermore, it may offer a portable 3D motion analysis system to overcome the limitation of the existing systems that require the laboratory experimental set up (Dutta 2012b). The device is cheaper than conventional motion capture system and has the capability of capturing data more easily.

Reliability

Kinect system could provide useful information, nonetheless, it is prone to variability in measurement due to the fact that all measurement contains error (Bonnechère *et al*., 2014), (Kazerouni, 2009). Thus, it is necessary to apply a suitable means of analysis to ensure that the image coordinates are correctly scaled to size (Canessa *et al.* 2014) and to determine the quality of measurement system (Kooshan 2012), (Mustapha *et al.* 2014). Exploration on

Kinect capability and accuracy is currently expanding. Many researchers seek on the accuracy before decided to explore further on its dynamic capability. This is confirmed by a number of discussion about the accuracy of Kinect for assessing 3D position in workplace environment, dancing gesture classification and human movement assessment (Clark *et al*., 2012), (Dutta, 2012), (Galna *et al*., 2014). As reported by Bonnechère *et al*. (2014), the validation of Kinect system in term of angular measurement and reproducibility during functional assessment has been done to overcome the lack of extensive validation related to functional evaluation and motion assessment using Kinect. Due to the reliability recognised to be so low, Bonnechère recommended further research must be performed again to get better insight on the topic. Study has also been conducted to compare the motion capture using Kinect with marker based motion analysis system (Schmitz *et al.,* 2014). However, this work considers only smaller scale motion focusing on hand and finger movement to be measured and compared. Prior to that, determination of an object (cubic) in field of work has been done for validation purposes only (Dutta, 2012). Referring to the above-mentioned literature, it is evident that Microsoft Kinect has the potential to be used as an alternative low-cost motion analysis tool. Nevertheless, the standard procedure for measuring its accuracy and reliability has not been well established. It may be impractical to implement the conventional motion capture system calibration technique due to the differences in terms of the principle of motion detection.

Therefore, this study for the first time attempts to develop a standard procedure to assess and visualise the accuracy and repeatability of Microsoft Kinect system as a measurement tool for static and dynamic motion. This is novel as no similar specific approach has been reported earlier.

METHODOLOGY

In general, this research comprises of experimental and numerical approaches that have been conducted in three (3) stages:

Stage 1: Experiments (System and parameters exploration for optimum experimental setup)

Stage 2: Static and Dynamic Motion Capture

Stage 3: Data tracking and Numerical Analysis (Accuracy and Repeatability)

The flow chart in Figure-1 illustrates the overall methodology of the study.

Experimental Setup

The measurement set up consists of a Kinect sensor connected to the USB port of a laptop running the Windows 7 operating system. The single camera-system is used to capture static and dynamic motion in a controlled lab environment (Figure-1). The captures provided position data of prescribed marker/points, where the marker/points movement are then tracked using an open source software Virtual Sensei Lite (version 0.4.2).

Virtual Sensei Lite is an innovative software for 3D sport motion analysis based on the Microsoft Kinect sensor. The software is capable to evaluate the level of kinetic energy during the motion analysis. It also can compute the speed of hand and feet in the recording. Furthermore, Virtual Sensei's have other capabilities that can communicate with the sensor via USB and allows auto calibration; as well as the ability to save the position data in *.csv file format.

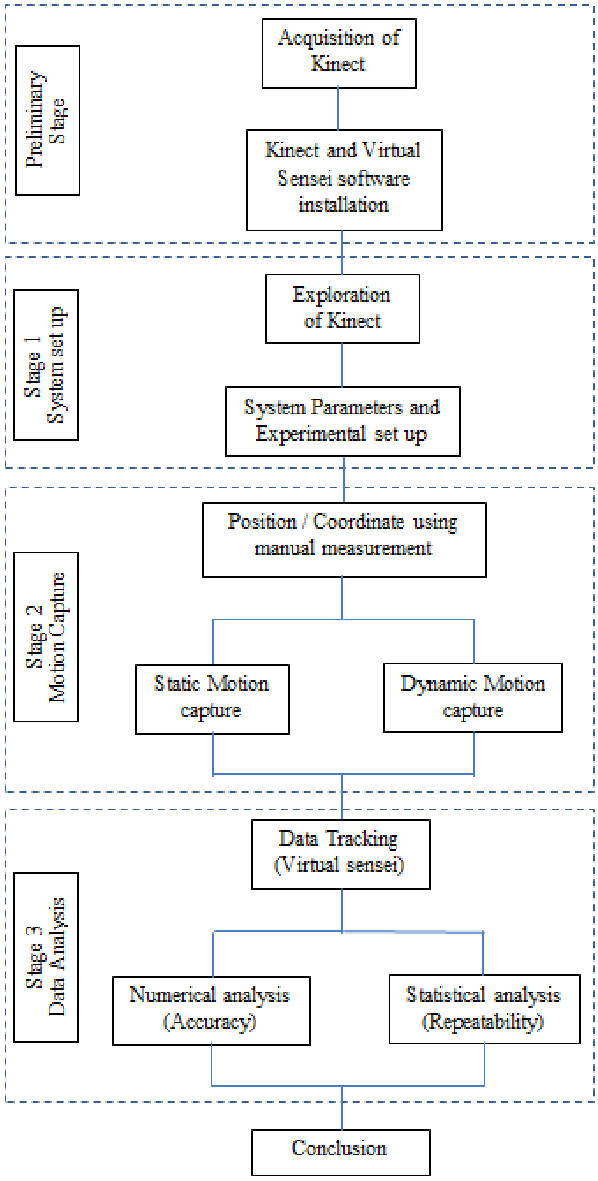

Figure-1. Overall process flow.

In order to run and use the Kinect for motion tracking, a few modules are installed first to run the Virtual Sensei software. To make it work, this study installed the essential modules based on the following order.

1. Microsoft Visual C++ 2008 Redistributable Package (x86)
2. Microsoft Visual C++ 2010 Redistributable Package (x86)
3. OpenNI for Windows x86 (32-bit);
4. PrimeSense NITE for Windows x86 (32-bit)
5. SensorKinectXXX-Bin-Win32-vX.X.X.X.msi

After Virtual Sensei installed successfully (the essential modules), the system is calibrated. Two forms of calibration are performed prior to. The current study uses Microsoft Xbox 360 (part of a gaming console) and thus it is calibrated accordingly. The console calibration process is performed using a simple game calibration card. This calibration method is simple, by just showing the card in front of the Kinect camera and try to complete the shape produces on the screen.

In general, this study adapts the procedure of Card Calibration using Xbox 360 game console.

1. The calibration card is downloaded for free as it did not come with the package we bought, consisting only Kinect sensor. A complete package of game and console system usually will provide the Kinect Game disc, and the calibration card is included.
2. The Guide button on the Xbox360 controller is pressed.
3. Next, Kinect Tuner is selected in the Settings menu.
4. Calibration is then selected.
5. The calibration card is showed in front of the Kinect camera
6. The card is moved around until it matched to the glasses on the screen. The edges of the calibration card should match the rectangle on screen.
7. The card is hold still until the green box is filled.
8. Step 6 and 7 are repeated until the Kinect sensor calibration procedure is completed.

The first calibration method is done only for the first time after the software successful installed. Then the second calibration method is applied for the next usage of the device. The second method is auto calibration performed using Virtual Sensei. The calibration is done automatically before the motion tracking is done. This calibration is to determine the range and to specify the length, height of recorded subject.

Initially, based on the tracking capabilities of Microsoft Kinect as proposed by Dutta (2012) the parameters, position, direction and axes of the single-camera system-subject are tested in a controlled lab environment and resulting the final set up of position (Figure-2); and effective range, depth limitation and angular field of views as shown in Table 1. The global coordinate of the origin and the directions of the axes (x, y and z) are synchronised between camera-tracking systems throughout the experiments in order to provide consistent data for numerical analysis.

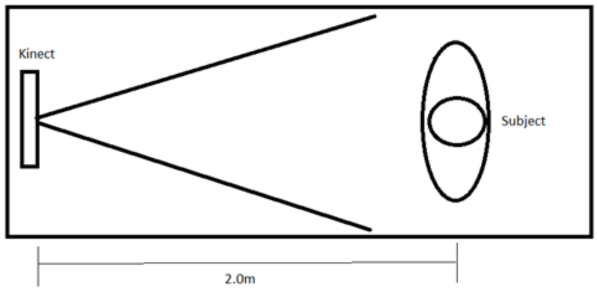

Figure-2. Schematic layout of the system position.

Table-1. General parameters set up for static/ dynamic motion capture.

Tracking Capabilities and Specification	
Effective Range	**1.5 m – 3.0 m**
Area	4 – 6 m^2
Depth Sensor Range	1.2 m – 3.5 m
Angular field of view	55° – 60° (horizontal), 40° – 45° (vertical)
Resolution	1.3 mm per pixel
Sensor speed	30 fps

Static and Dynamic Motion Capture

For better tracking, the healthy volunteers are required to wear tight-fitting clothing and stand 2 meter away from the Kinect Sensor to occupy the field of view. Before experiments, the age, gender and anthropometrical measurement, i.e. height, weight, length of arm and leg are recorded. Standard 15 marker/point set is employed in this study.

Since the current study aims to assess the accuracy of Microsoft Kinect system in detecting and measuring human motion for both static and dynamic conditions, two postural / body gesture tests are identified and performed. In the first stage for static motion capture, motionless images of the healthy volunteers are recorded. The posture of a straight arm is captured for approximately 15 seconds. The capture is repeated for five times.

The second stage focused on dynamic motion capture for a simple gesture. The healthy volunteers are requested to move their arms till reaching 90 degree to the elbow line. Once, the volunteers performed complete T-pose (both arm parallel with the floor), the cycle of the motion is completed. Each test is repeated for five times and recorded. The capture is conducted under controlled lab environment where for the complete cycle is controlled between 20 – 30 seconds/cycle.

ARPN Journal of Engineering and Applied Sciences

For a controlled lab environment, throughout the process of motion capturing, the following issues have been considered:

i. Shape of image control - Large clothing, skirts or long hair have been considered/ avoided as they could constitute problem for body tracking.
ii. Illumination control - Sunlight and reflection are avoided.
iii. Posture control - Arm tracking is less stable if it is very close to any body parts. Legs are open during capture to avoid loss of marker/points. The whole posture (head to toe) is in the field of view.
iv. Ambiance control - Empty space is used throughout motion capture.

Data Analysis for Accuracy

From the static and dynamic motion capture, the images are analysed and the marker/points are tracked using Vitrual Sensei. The tracked motion are organised into 3D coordinate of each marker/point for further analysis using numerical and statistical approach. The general idea is to use the data to compare with the actual value manually measured at the beginning of the experimental procedure. The accuracy of the Microsoft Kinect system in measuring the static and dynamic motion is determined from the error analysis (Equation 1).

$$\%\,Error = \frac{|Actual\ Reading - Measured\ Reading\ |}{Actual\ Reading}\ x\ 100 \qquad (1)$$

Angle Calculation for Dynamic Analysis

The angles of motion (arm gesture) are calculated using 3-points coordinate algebraic operation based on 2 equal lengths of lines as shown in Figure 3. In term of algebra, it is the sum of products of the corresponding value of 2 values. The angle of the arm gesture is determined using Equation 2. Comparing to the actual angle of 90°, percentage error is calculated to determine the minimum and maximum error.

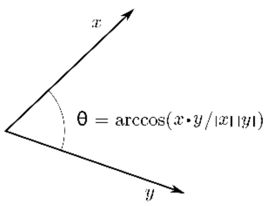

Figure-3. Vector angle.

$$\text{Angle, } \theta = arc\cos(x + y/|x||y|) \qquad (2)$$

Statistical analysis for Repeatability

The repeatability of the Microsoft Kinect system in measuring static and dynamic motion is determined using variance analysis. The concept of variance is to measure how far a set of numbers spreads out, where in this, study the points of measurement are taken from Kinect raw data. A zero measured variance means that all the data are close and fall on the same point/ curve. Only absolute value is considered (Modulus; non-negative). A very high variance value means that the data points are too wide-spread from the mean value. A low variance value means that the data points are close from the mean value.

The five captures for every posture/ motion are tabulated and the mean, standard deviation and variance (Equation 3) are computed.

$$s^2 = \frac{\sum(x - \bar{x})^2}{n - 1} \qquad (3)$$

In general, if the variance is found small (< 5%), it could be concluded that the process is repeatable.

RESULTS AND DISCUSSIONS

Figure-4 depicts a sample visualisation of the output images seen by Microsoft Kinect sensor, the tracked marker/points (images) by Virtual Sensei, raw data plotted and final data (mean position for 15-marker/point set) in spread sheet.

Figure-4(a) shows the images seen by Microsoft Kinect sensor where the images could be visualised based on the contour of images portraying shapes. Based on the algorithm provided in Virtual Sensei, the 15-marker points are shown with connecting lines. Once the markers are tracked using Virtual Sensei, a skeleton model could be developed in a 3-dimension motion space as shown in Figure 4(b). At this view, the axes are established and thus the motion could be observed more clearly. This simulation is very useful to visualise the tracked markers and movement. Virtual Sensei store the displacement of each marker and the data could be tabulated and transferred into a spreadsheet. Figure-4(c) shows an example of plotting the raw data consisting of x; y and z coordinate for each marker for a specific instance. For a specific duration of time frame, a video could replay the marker movement. These data are important and further processed. Figure-4(d) shows the mean position data (tracked marker/point) for each of the 15 points for visualisation, which are plotted from the coordinates tabulated in Table-2. The output shows that the data is consistent and the visual observation is then confirmed by the variance and calculated marginal error. For all measurement, the maximum variance computed is less than 5%, which can be considered acceptable in term of the Kinect accuracy.

For dynamic motion, the recorded motion is examined to identify the moment for initial and final position of arm motion during the test. Figure-5(a) shows four selected gesture output (Dynamic Motion) for tracked 15-marker/point set. The graph in Figure-5(b) also indicates the motion generated little kinetic energy and speed motion from the arms movements began at frame

ARPN Journal of Engineering and Applied Sciences

120 and ended at frame 250. Once the initial and final point has been identified, the angle between both hand segment (shoulder-elbow and elbow-wrist) is then calculated.

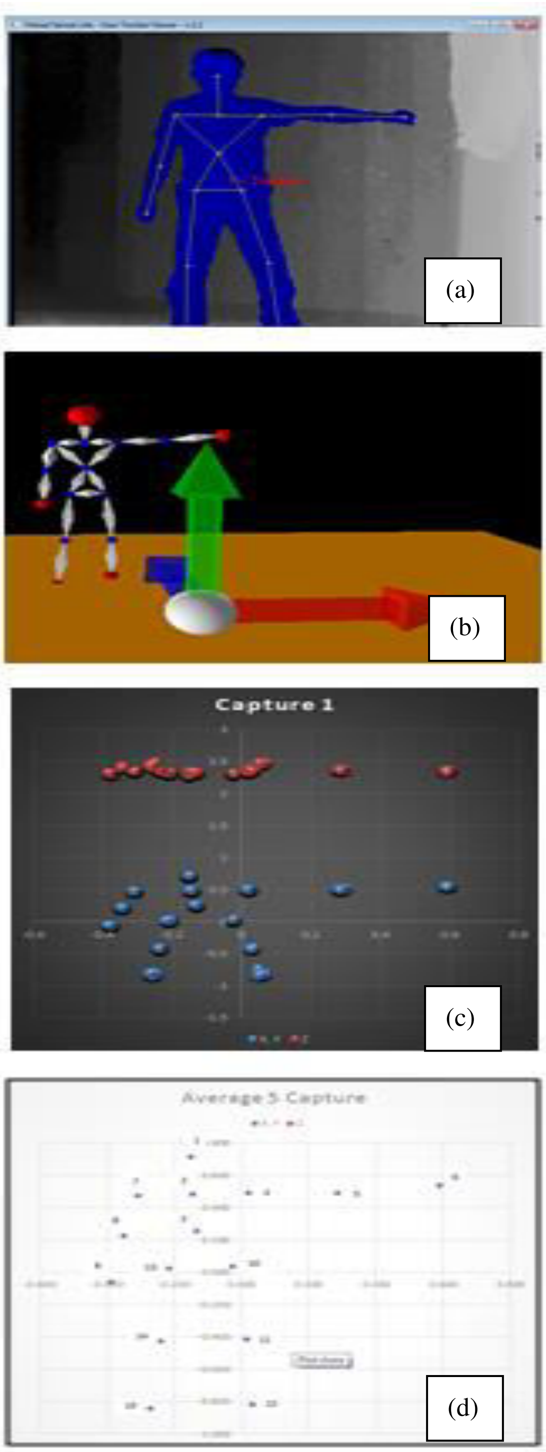

Figure-4. Outputs View: (a) The actual view from the motion capture (Kinect eye/ sensor); (b) View of tracked markers in Virtual Sensei; (c) Plotted raw data from Virtual Sensei; (d) Plotted final data (mean position for 15-marker/point set).

Table-2. Mean coordinate (x,y) of 15-marker/point set.

Point	x	y
(1)Head	0.151	0.714
(2)Neck	0.145	0.483
(3)Torso	0.134	0.257
(4)Left Shoulder	0.311	0.475
(5)Left Elbow	0.351	0.226
(6)Left hand	0.388	-0.061
(7)Right Shoulder	-0.02	0.492
(8)Right Elbow	-0.282	0.492
(9)Right hand	-0.588	0.538
(10)Left Hip	0.216	0.025
(11)Left Knee	0.242	-0.425
(12)Left Foot	0.272	-0.844
(13)Right Hip	0.029	0.035
(14)Right Knee	-0.015	-0.414
(15)Right Foot	-0.031	-0.816

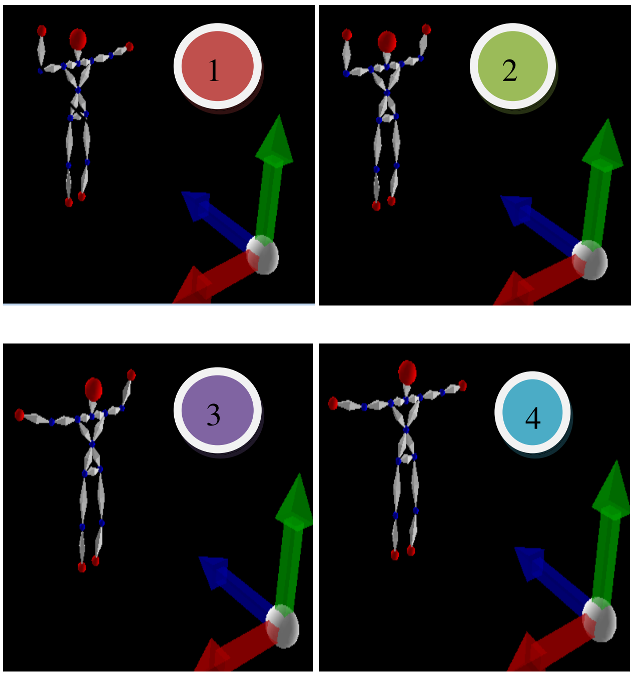

Figure-5(a). Four selected gesture output (Dynamic motion) for tracked 15-marker/point set.

VOL. 10, NO. 17, SEPTEMBER 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

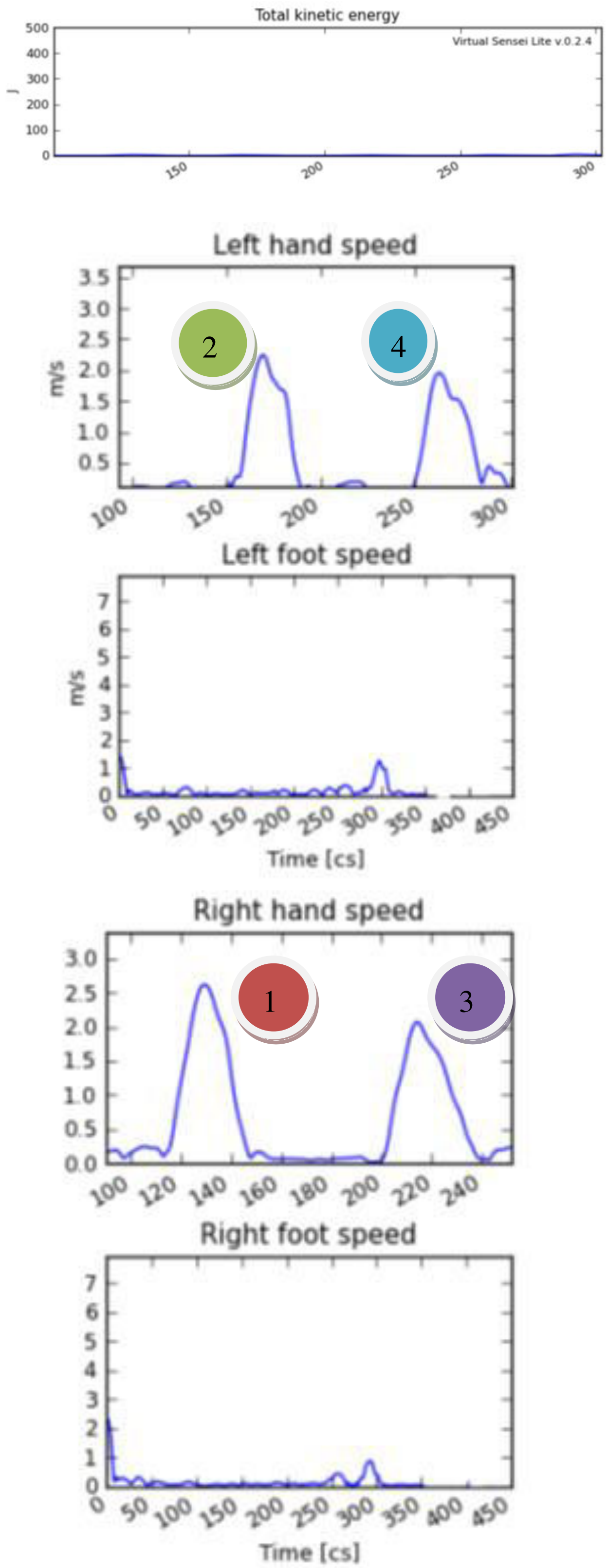

Figure-5(b). Graph generated by virtual sensei representing the kinetic energy and speed of motion.

Figure-6 shows the distribution of arm movement during the dynamic accuracy testing for five captures. The accuracy of Kinect sensor is then evaluated by observing the difference between the actual value (90°) and calculated angle value. Based on the plotted data in the Figure-6, it shows that the angle of 90 degree for each Elbow to Hand point does have a huge misalignment which been affected by the shape and length of the hand itself.

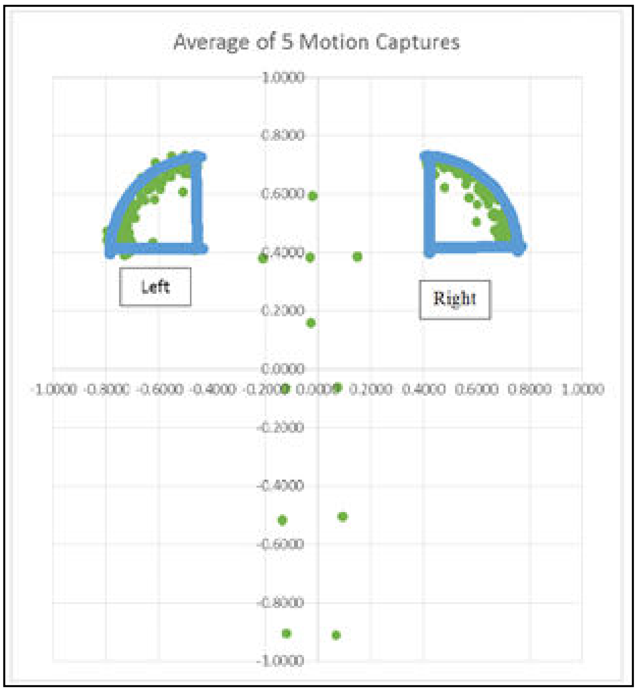

Figure-6. Average arm motion capture measurement.

For the Kinect tracking system, any small particular movement during the recording are recorded, so as the hand momentum motion. As an example if the motion is slightly faster than normal capture rate of the Kinect, the tracking point for the motion will difficult to follow and readjust to the exact point of the hand. This could cause inaccuracy in tracking especially during recording fast gesture. The results show that the percentage of error for right and left arm movement detection are 4.8 % and 5.1% respectively.

Based on the results from both tests (static and dynamic), the Kinect sensor performed better during static measurement. The greater error in dynamic motion is possibly due to the response of Kinect Sensor to any small particular movement and thus, this has affected the overall error for dynamic motion compared to static motion.

ISSUES AND CHALLENGES

During study, several issues are found related to the motion capture, recording, tracking and analysis that contributing to the overall accuracy of the system.

ARPN Journal of Engineering and Applied Sciences

Timing and Synchronisation

Initially, the study faces the challenge to synchronise between Kinect and Virtual Sensei for recording the data during motion capture. The data and recording are difficult to match since the output data generated from Virtual Sensei are not in the same frame as the Kinect's recording.

Capture Rate

Referring to the provided technical specification, the Kinect system used in this study is capable of capturing 30 frames per second under normal use. Based on this capability, it is expected that the system could record the maximum of 150 frames per second for duration of 5 sec recording. Nevertheless, the capture rates during live recording are not consistent and this made the comparison analysis between the captures became difficult. This inconsistency of the capture rates (fps) has contributed error as much as the timing in determining the accuracy.

Software

Another main issue arisen at the preliminary stage of the study is software installation. Setting up a Kinect Xbox360 (instead of Kinect for Windows), for running a motion analysis system on a PC is not an easy task since it is not designed for non-other works except for Xbox and gaming application. Furthermore, both type of Kinect devices run on a different architecture driver. Nevertheless, by getting a correct drivers and installation algorithm which involved a lot of time for preliminary study and software exploration, finally the Kinect Xbox360 system is able to perform as a motion capture system.

Hardware

During recording and later analysing the motion capture data, it is found that low specification hardware could affect the quality of data collection and analysis. Processing power (GPU process) is low when the computer ran at a low memory or simultaneously performed multi-tasking in the background. The low processing power caused capture rates drop and thus, timing and synchronising both Kinect capture and Virtual Sensei software is difficult and this became challenging in generating data and analysis.

CONCLUSIONS

This study aims to develop a standard procedure to assess and visualise the accuracy and repeatability of Microsoft Kinect and the results prove that this has been achieved. The study able to demonstrate average error less than 2% (static) and 5% (dynamic) accuracy test respectively. The data and results show that Microsoft Kinect system has the potential for a low cost motion capture and analysis application with reasonably accuracy. Nevertheless, further and rigorous research should be carried out for further assessment, including comparing Microsoft Kinect capabilities to other available tools and software. Other than that, the study also can consider to use different parameter of the volunteer with the different body size, different gender and different movement speed to implement in the future.

Nevertheless, it is not a straight forward comparison as existing motion capture tools such as infrared camera could detect actual marker/point position but currently available algorithm for Kinect Sensei such as used in Virtual Sensei approximate marker/point position from shape length. Thus, this opens a great opportunity for researchers to explore the function of Kinect sensor and develop computational algorithm for detecting the marker/point (a point) directly.

Therefore, it could be concluded that the current study is important and has contributed to enhancing knowledge about using Microsoft Kinect system as a measurement tool in motion analysis.

ACKNOWLEDGEMENT

This work is supported by Ministry of Education (MOE) Malaysia and Universiti Teknologi MARA (UiTM) under the Fundamental Research Grant Scheme (Grant no. FRGS/2/2013/TK01/UITM/03/3 and UiTM Ref. No. 600-RMI/FRGS 5/3 (76/2013)).

REFERENCES

[1] Timmi, A. 2011. Virtual Sensei Lite, URL http://www.virtualsensei.it/lite/, [last assessed: 25.5.2015].

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

[3] Bünger, M. 2013. Evaluation of Skeleton Trackers and Gesture Recognition for Interaction. Denmark: Aalborg University.

[4] Canessa, A., Chessa, M., Gibaldi, A., Sabatini, S.P., Solari, F. 2014. Calibrated depth and color cameras for accurate 3D interaction in a stereoscopic augmented reality environment. J. Vis. Commun. Image Represent. 25, pp. 227-237.

[5] Clark, R.A., Pa, Y.-H., Fortin, K., Ritchie, C., Webster, K.E., Denehy, L., Bryant, A.L. 2012. Validity of the Microsoft Kinect for Assessment of Postural Control. Gait Posture. 36, pp.372-377.

ARPN Journal of Engineering and Applied Sciences

[6] Corazza, S., Mundermann, L., Chaudari, A.M., Demattio, T., Cobelli, C., Aandriachi, T.P. 2006. A Markerless Motion Capture System to Study Musculoskeletal Biomechanics: Visual Hull and Simulated Annealing Approach. Ann. Biomed. Eng. 34, pp.1019-1029.

[7] Dutta, T. 2012. Evaluation of The kinect sensor for 3D Kinematic Measurement in The work place. Appl. Ergon. 43, pp.645-649.

[8] Galna, B., Barry, G., Jackson, D., Mhirippi, D., Olivier, P., Rochester, L. 2014. Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's Disease. Gait Posture. 39, pp.1062-1068.

[9] Kazerouni, A. 2009. Design and analysis of gauge R and R studies: Making decisions based on ANOVA method. World Academy of Science, Engineering and Technology. 52, pp.31-35.

[10] Kooshan, F. 2012. Implementation of Measurement System Analysis System (MSA): In the Piston Ring Company" Case Study. Int. J. Sci. Technol. 2(10), pp.749-761.

[11] Molnár, B., Toth, C.K., Detrekői, A., 2012. Accuracy Test of Microsoft Kinect for Human Morphologic Measurements. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, pp.544-547.

[12] Microsoft Xbox, URL http://www.xbox.com/en-US/, [last assessed: 27.5.2015].

[13] Mustapha, G., Aizat, M.S., Yahaya, N.H.M., Mahmud, J. 2014a. Measurement System Analysis of VS Lite. Applied Mechanics and Materials, 680, pp.447- 450.

[14] Mustapha, G., Sulaiman, W., Majeed, A.P.P.A., Yahya, N.H.M., Mahmud, J. 2014b. Kinematics and Efficiency Analysis of the Seni Silat Cekak Malaysia (Kaedah A). Applied Mechanics and Materials, 680, pp. 267-270.

[15] Mustapha, G., Zakaria, M., Sulaiman, W.R.W., Mahmud, J. 2014c. The Mechanical Aspects of Martial Arts : Total Time of Execution and Kinematics of Kaedah A. In: Adnan, R., Ismail, S.I., Sulaiman, N. (Eds.), Proceedings of the International Colloquium on Sports Science, Exercise, Engineering and Technology 2014 (ICoSSEET 2014). Springer, Singapore, pp.3-12.

[16] Ott, D. 2013. Kinect Analysis. Swiss Federal Institute of Technology in Zurich (ETHZ).

[17] Schmitz, A., Ye, M., Shapiro, R., Yang, R., Noehren, B. 2014. Accuracy and repeatability of joint angles measured using a single camera markerless motion capture system. J. Biomech. 47(2), pp. 587-591.