

Performance of Dual Depth Camera Motion Capture System for Athletes' Biomechanics Analysis

*Wee Chang An*¹, *Mohd Zamani Ngali*^{1, *}, *Zulhilmi Kaharuddin*¹, and *Siti Badriah Khairu Razak*¹

¹Faculty of Mechanical and Manufacturing Engineering Universiti Tun Hussein Onn Malaysia, Batu Pahat, Malaysia

Abstract. Motion capture system has recently being brought to light and drawn much attention in many fields of research, especially in biomechanics. Marker-based motion capture systems have been used as the main tool in capturing motion for years. Marker-based motion capture systems are very pricey, lab-based and beyond reach of many researchers, hence it cannot be applied to ubiquitous applications. The game however has changed with the introduction of depth camera technology, a marker-less yet affordable motion capture system. By means of this system, motion capture has been promoted as more portable application and does not require substantial time in setting up the system. Limitation in terms of nodal coverage of single depth camera has widely accepted but the performance of dual depth camera system is still doubtful since it is expected to improve the coverage issue but at the same time has bigger issues on data merging and accuracy. This work appraises the accuracy performance of dual depth camera motion capture system specifically for athletes' running biomechanics analysis. Kinect sensors were selected to capture motions of an athlete simultaneously in three-dimension, and fused the recorded data into an analysable data. Running was chosen as the biomechanics motion and interpreted in the form of angle-time, angle-angle and continuous relative phase plot. The linear and angular kinematics were analysed and represented graphically. Quantitative interpretations of the result allowed the deep insight of the movement and joint coordination of the athlete. The result showed that the root-mean-square error of the Kinect sensor measurement to exact measurement data and rigid transformation were 0.0045 and 0.0077291 respectively. The velocity and acceleration of the subject were determined to be 3.3479 ms^{-1} and -4.1444 ms^{-2} . The result showed that the dual Kinect camera motion capture system was feasible to perform athletes' biomechanics analysis.

*Corresponding author: zamani@uthm.edu.my

1 Introduction

Since the capturing of the motion of an athlete requires highly accurate system, both camera based (e.g. Vicon, UK) or inertia sensor based (e.g. XSens, Holland) were dubbed as the gold standard equipment for the purpose. However these types of systems are expensive; make it outside the reach of most users. Microsoft Kinect (Microsoft Corp., Redmond, WA, USA) sensor was then considered as the low-cost alternatives to the pricey systems. Previous study concluded that the Kinect sensor “showed good performance versus the clinical gold standard (Vicon and AMS/Visual-3D)” [1]. Another significant feature of the Kinect sensor is that it does not required any marker to be attached onto the subject – a marker-less motion capturing system. The Vicon system on the other hand requires reflective markers to be attached on the subject’s body; the markers are definitely a huge disadvantage. Thus with the application of Kinect sensors as the dual depth camera motion capture system, it could facilitate the set-up process and phase out complex equipment set-up. The fast, low cost, portable system for capturing these biomechanical measures would facilitate the biomechanists or researcher to bring their work task to the athlete instead of bringing the subject into the lab. Biomechanics analysis is used to assess the performance of athletes in order to gain a greater understanding of human movement. Thus, appropriate quantitative solutions could be developed to improve athlete’s techniques and thereby increase the performance, minimizing injuries and in turn, promoting career longevity. A biomechanics researcher uses quantitative biomechanical analysis methods to discover new techniques, which communicated to the teachers and coaches who will implement them [2].

2 Related work

There were numerous works have been done on motion analysis via depth cameras. Each of them yields outstanding results in term of performance of the Kinect sensor.

In the study conducted by Dutta [3], Kinect was used to capture three-dimensional image of a scene with cubes positioned at varying distances. The result of the study was compared to the gold standard motion capture system (Vicon), and the root-mean-square error (RMSE) in the investigation were reported to be 0.0169m, 0.0348m and 0.0141m in the x, y and z directions respectively, which relatively small and accurate.

As claimed by Seung-kook Jun [4], without any post-processing to the raw data captured by the Kinect, the application of such system tend to be limited. However with suitable data post-processing, such limitation was unbarred. The post-processing as suggested was to perform kinematic calibration and apply Kalman filter to the raw data captured by Kinect. As a result, it showed a good performance when compared to the gold standard equipment (Vicon and AMS/Visual-3D).

In Caon’s [5] research paper, the writer employed multiple Kinect sensors to detect subject’s posture and gesture estimation, in three dimensional smart environments. The system developed was able to track multiple users and reconstructs its situational information such as posture and coordinates. The system was used to smartly interact with objects around it, such as turning a lamp on and off, turning a TV on and off and so on.

Based on the literature review, most of the research works were only focusing on the subject’s posture and the limitation and conflict of the usage of Kinect system. Quantitative analysis on how good dual Kinect sensor system can be applied as motion capture system for sports and biomechanics analysis was still questionable. Therefore, this work was initiated to obtain numeric measurement of dual depth camera system feasibility in these fields.

3 Methodology

This research was subjected to a systematic sequence. The process started with the data acquisition from each Kinect sensor, so that the skeletons data from the sensors could be retrieved. Two Kinect sensors were projected onto the subject. Each sensor streams its own skeletal data in its field of view (FOV). The Microsoft Kinect Software Development Kit (SDK) used in this study allows the skeletal tracking and data acquisition into the system. With the three coordinate spaces: the colour spaces; the depth spaces and; the skeletal spaces, the three-dimensional (3D) position data of the 20 predefined skeletal joints of human body were retrieved. Each sensor streamed the 3-dimensional space coordinates of the tracked skeleton into a custom formatted readable text file which later used for skeletons merging. The system however required a few calibration steps before fused into a single skeleton model. The calibration method is described in Figure-1 below.

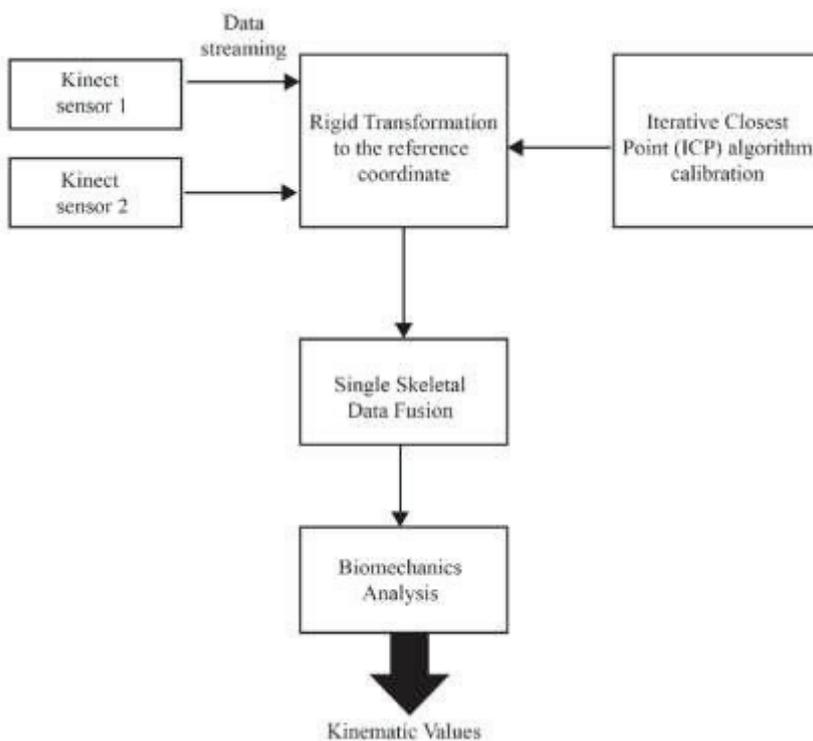


Fig. 1. Calibration method.

The overall research flow was described in Figure-2, from the beginning of data acquisition to data analysis. As illustrated in this figure, the skeletal data from each sensor has to be transformed to a common reference coordinate system before data processing were carried out. From the system, one sensor is selected as a reference sensor (principal camera, denoted by K_B) so as to provide a reference frame for other sensor (secondary camera, denoted by K_A). Next, the usage of the Iterative Closest Point (ICP) [3] algorithm was introduced to the K_A . It was fed into the ICP to estimate the rigid transformation (or the Rotation-Translation) that minimized the distance between the transition positions in the reference frame.

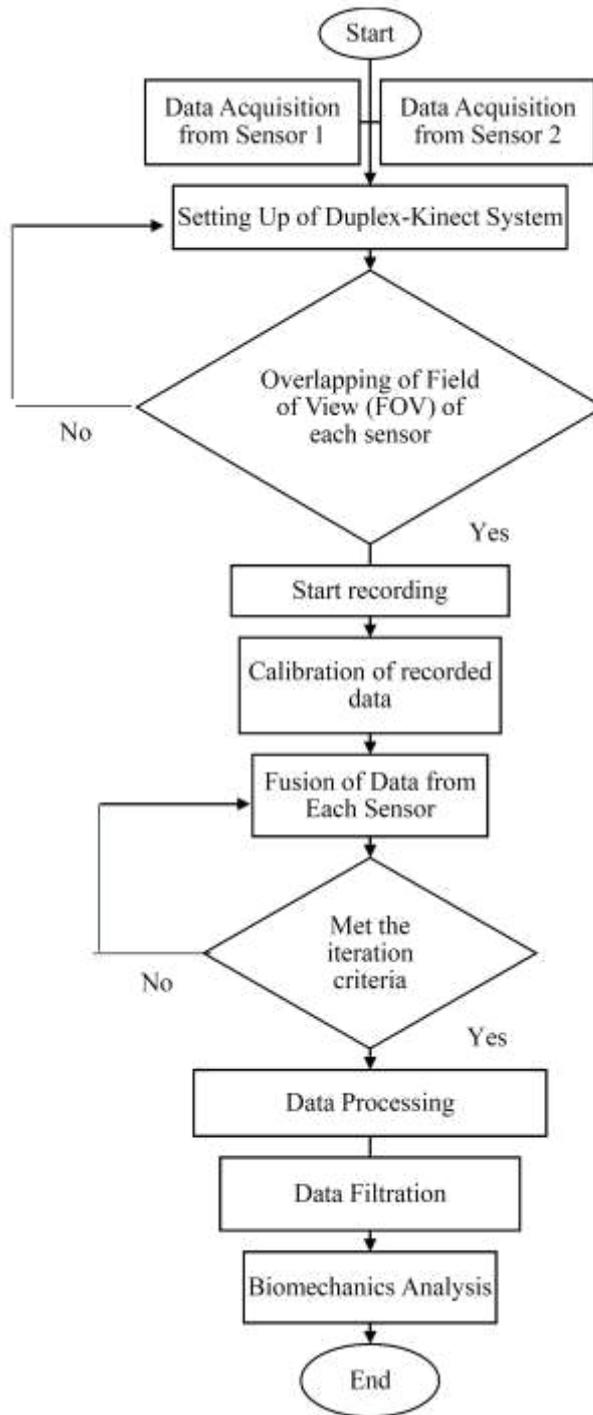


Fig. 2. Flowchart of research.

The skeleton was sliced into three sections, namely upper section, middle section and lower section. From each section, a joint was selected. The three selected joint was then used to compute the rigid transformation. The transformation obtained in term of rotation (R) and translation (t) was then used to register the skeletons acquired from each sensor in the reference coordinate system. The formula used to calculate the R and t based on the Eqn. (1).

$$B = R * A + t \tag{1}$$

Where R , t are the transforms applied to dataset A to align it with dataset B, as best as possible. To solve the equation, it involved three steps: computation the centroids of both dataset; moving of both dataset to the origin and then evaluate the optimal rotation, (matrix R) and; calculation of the translation, t . Figure-3 described the process graphically.

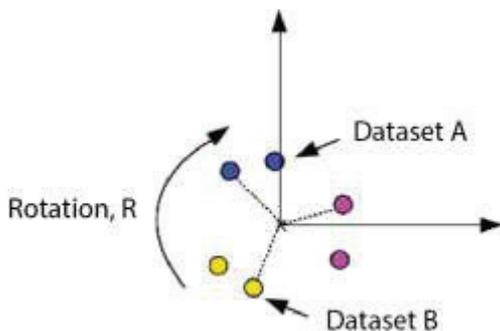


Fig. 3. Moving data set A and B to origin and rotate the data.

Once all the prerequisite conditions were met, which was the root-mean-square error (RMSE) of the rigid transformation was lower than the previous iteration in a number of given iterations; the merging of both skeletons was started. The merging of skeleton was illustrated in Figure-4.

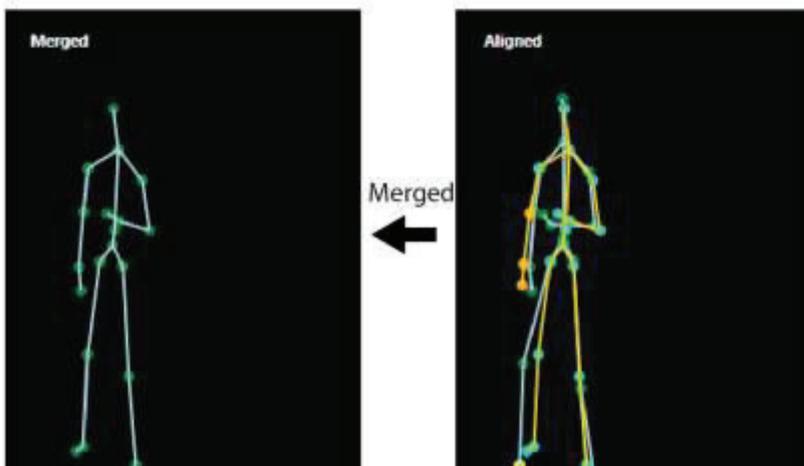


Fig. 4. Merged skeleton (left) and transformed skeleton (right).

Next, the data processing was done on the fused skeleton by extracting the joints coordinates in x, y, and y positions. Once the coordinates were successfully retrieved, biomechanics analysis was proceeded. However prior to analysis, Butterworth filter was applied to the data. Most system that measure physical quantity using a sensor return errors and noise, this includes the Kinect sensor used in this research. By using such filter, it had minimized the jittering and hence stabilized the joint positions over time. After several tuning were done on the recorded data, a 3Hz cut-off frequency of 5th order low pass Butterworth filter was chosen. Since a filter will delay the signal, thus, the data was filtered twice, in forward and backward direction to produce a zero lag filter [4]. After the filtration, linear and angular kinematics motion of an athlete were analysed using the fundamental equations, Eqn. (2), Eqn. (3), Eqn. (4) and Eqn. (5).

$$displacement, d_i = \sqrt{d_{x,i}^2 + d_{y,i}^2 + d_{z,i}^2} \tag{2}$$

$$velocity, v = \frac{\Delta d}{\Delta t} \tag{3}$$

$$acceleration, a = \frac{\Delta v}{\Delta t} \tag{4}$$

$$\mathbf{A} \cdot \mathbf{M} = |\mathbf{A}||\mathbf{M}|\cos\theta \tag{5}$$

4 Results and discussion

Analysis on the athlete was done using Mathworks' Matlab. The subject chosen was a long distance running athlete. The raw data showed that although the subject was standing statically in front of the Kinect sensor, it was observed that the joints of the skeleton jittered. This gave an unstable movement of the joints to be analysed. Figure-5 however showed both the data before and after the application of Butterworth filter.

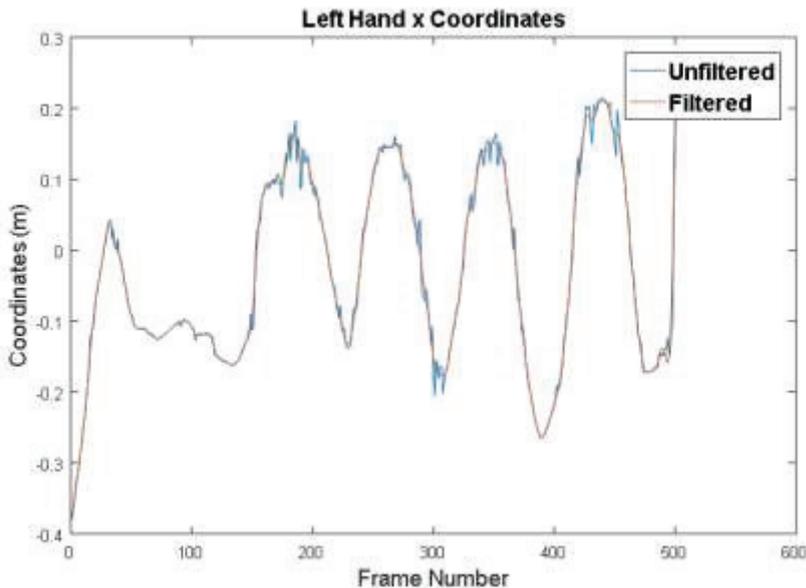


Fig. 5. Noise filtration.

All joints data were filtered as mentioned previously in methodology section with 5th order low pass Butterworth filter at cut-off frequency of 3Hz. The dual Kinect system used in this research was compared with the use of single Kinect sensor. Three types of basic human movements were tested in both systems, the flexion in sagittal plane, lateral in frontal plane and twist in transverse plane. The flexion result was illustrated in Figure-6.

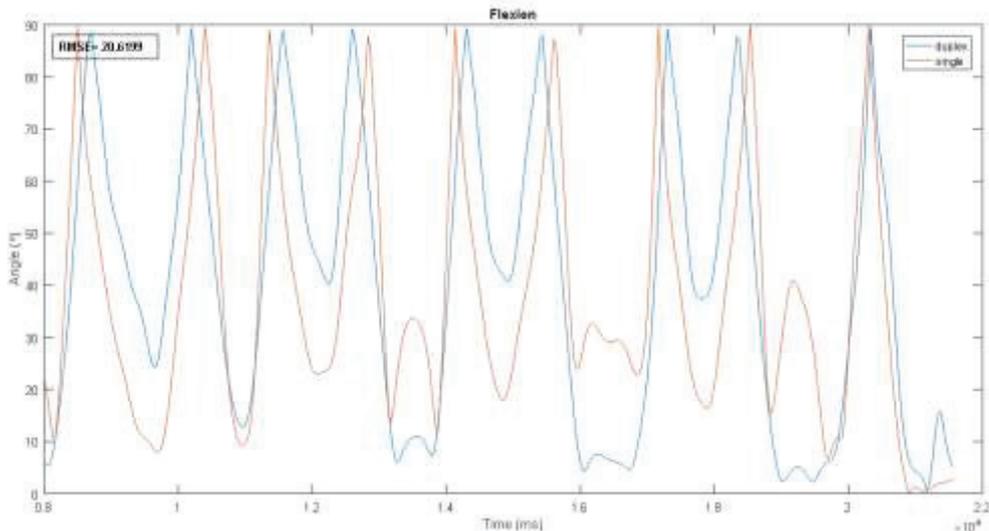


Fig. 6. Flexion angle of dual and single Kinect system.

After the comparison, it returned root-mean-square error (RMSE) of 20.8199, which is numerically large. The flexion angle for single Kinect system show there was a large angle jumping between one instance of time to another instance time, however for dual Kinect system, the differences was smaller. For lateral, the comparison is shown in Figure-7.

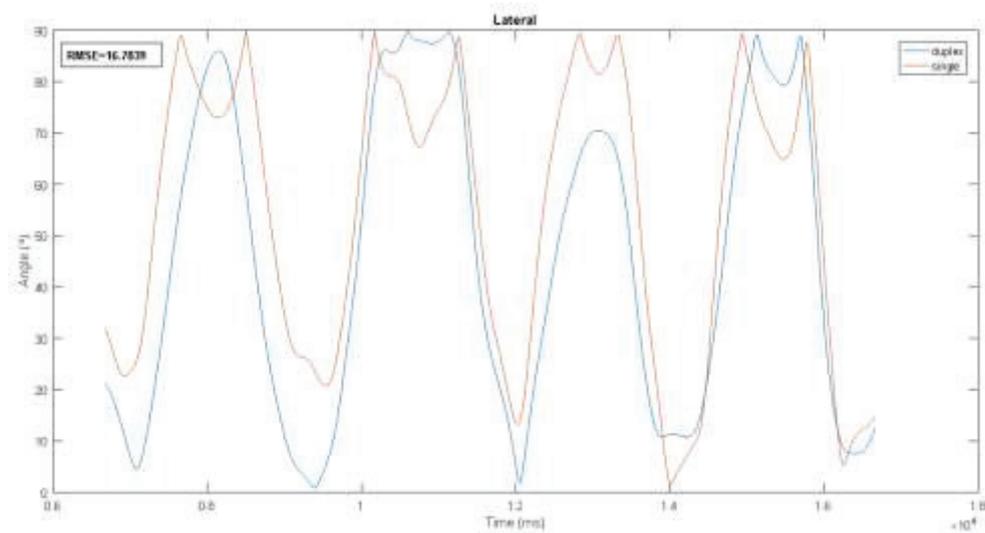


Fig. 7. Lateral angle of dual and single Kinect system.

The error for lateral movement for single Kinect system had a smaller RMSE compared to the flexion in sagittal plane. The RMSE was only 16.7839. This set of data had proved that the single Kinect system was taking the ‘inferred’ or ‘not tracked’ data as the actual data. The most significant feature of dual Kinect was that it could reduce occlusion problem. The twisting motion of subject is shown in Figure-8.

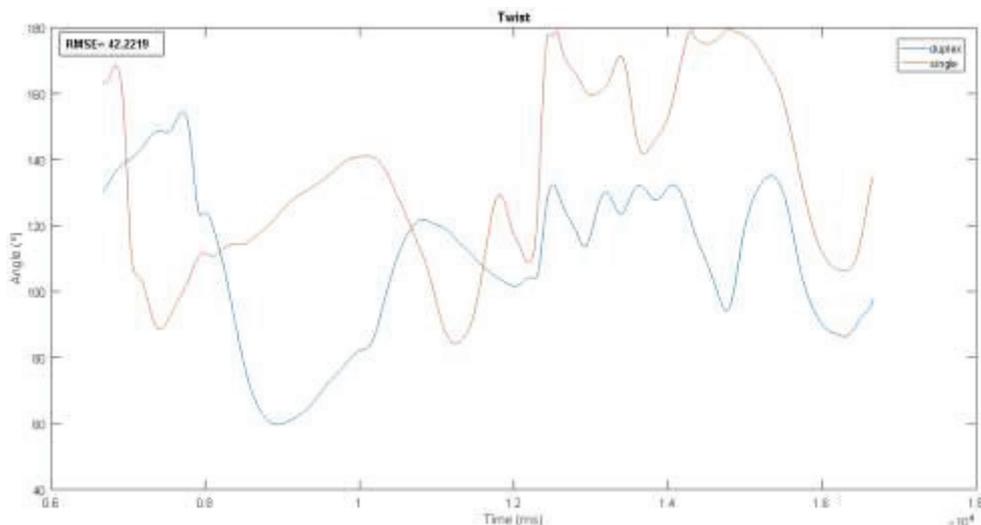


Fig. 8. Twist angle of dual and single Kinect system.

The RMSE error for single Kinect system in measuring twist angle was greatly deviated from the dual system with a high of 42.2219. As the subject twisting, the single Kinect was unable to track the joints that were occluded. However, dual system made it possible by considering the data from the other sensor. Next, the rigid transformation of the athlete running, return a RMSE of 0.0077291. Such transformation considered to be comparably accurate since the value is approximately close to 0. However, the transformation sometimes could be erroneous and a resolution was taken by increasing the number of iterations and set the maximum acceptable error to accommodate various set-up and recording.

By using the transformed data, linear and angular kinematics were analysed next. The centre of mass (CM) of a vertically standing athlete was assumed to be approximately at navel, which was estimated to be at the centre hip joint [5]. The linear kinematics motion of the athlete, the displacement, the velocity and the acceleration was described in Figure-9, Figure-10 and Figure-11 respectively.

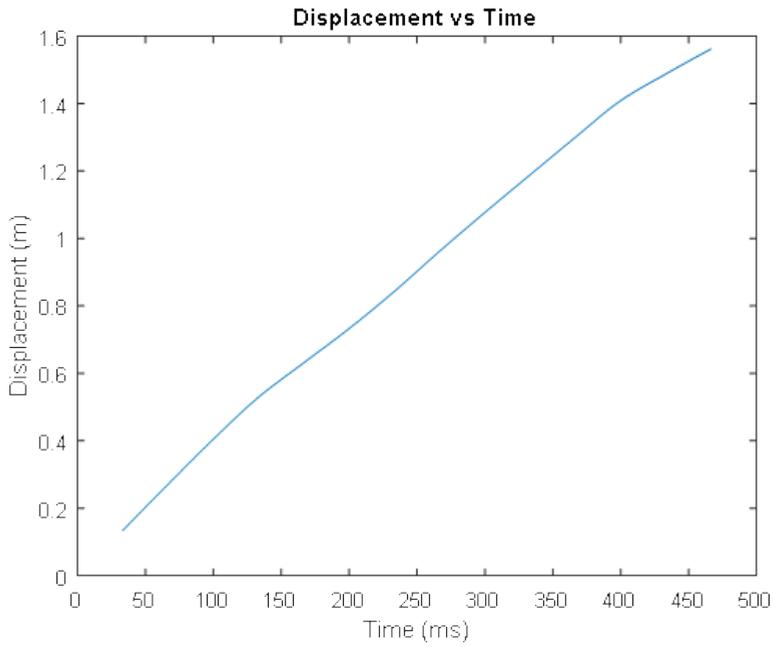


Fig. 9. Displacement of an athlete versus time.

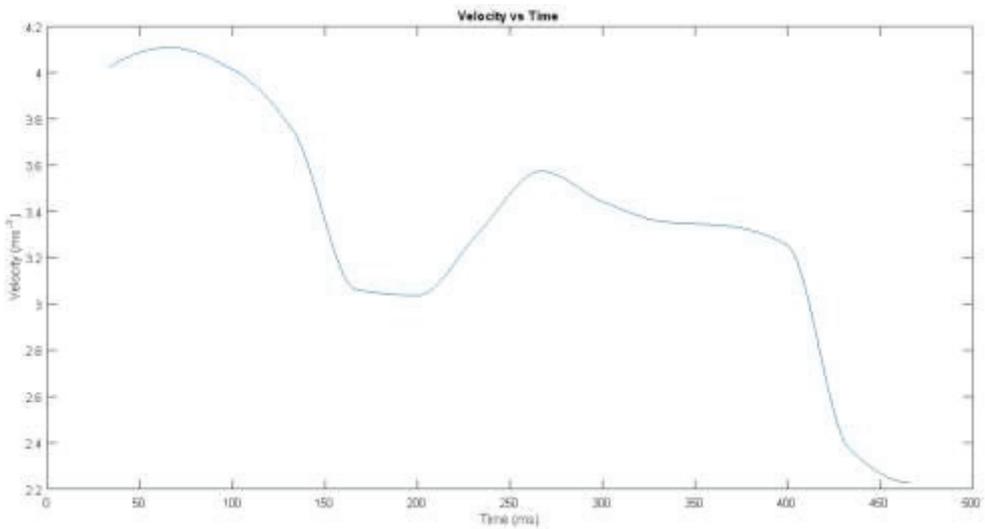


Fig. 10. Velocity of the athlete.

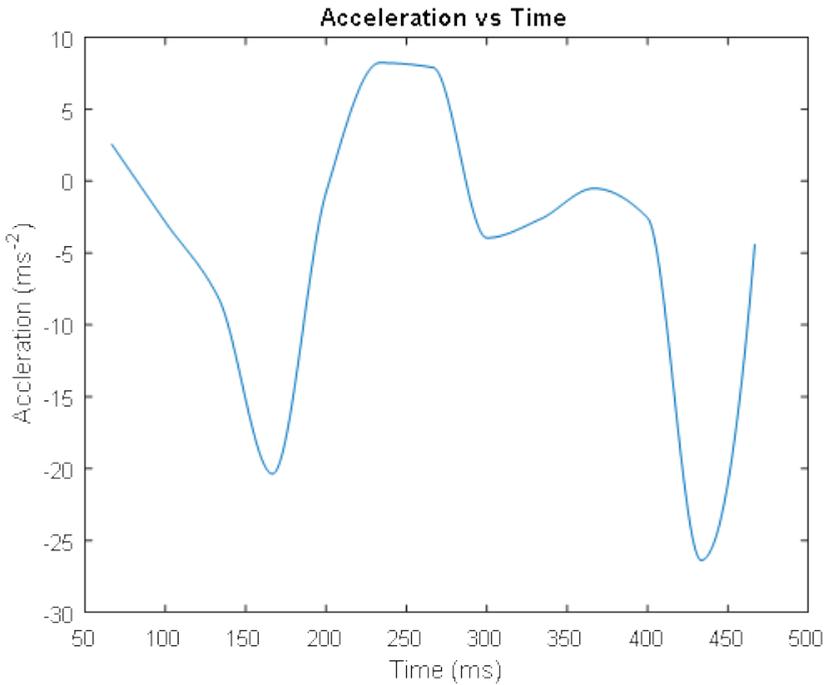


Fig. 11. Acceleration of the athlete.

The average velocity and acceleration of the athlete was calculated to be 3.3479ms^{-1} and -4.1444ms^{-2} respectively. Figure-12 below depicted the left knee angle over time when the athlete was running.

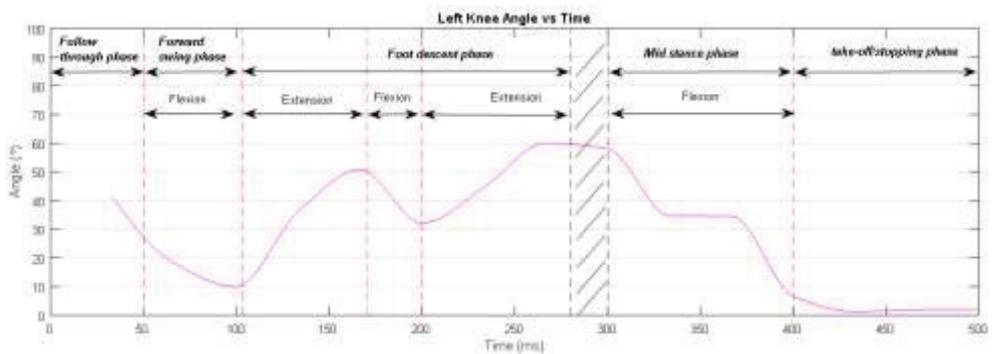


Fig. 12. Flexion and extension of left knee with respective phase.

The key events that mainly were analysed in this angular motion were the foot strike, mid-support and take-off, follow through, forward swing and foot descent, as described in Figure-13.



Fig. 13. The analysed stages of a running.

Foot strike, also known as the initial contact absorbs the impact, mid-support or mid-stance maintains the forward momentum and support the body weight, take-off critically responsible for the acceleration of the body, follow through or initial swing decelerating the running process for the recovery leg, forward swing also known as mid-swing coordinating the recovery leg movement, and foot descent or terminal swing prepare the foot for active landing. All the key events were identified in the Figure-12 for knee and Figure-14 below for hip.

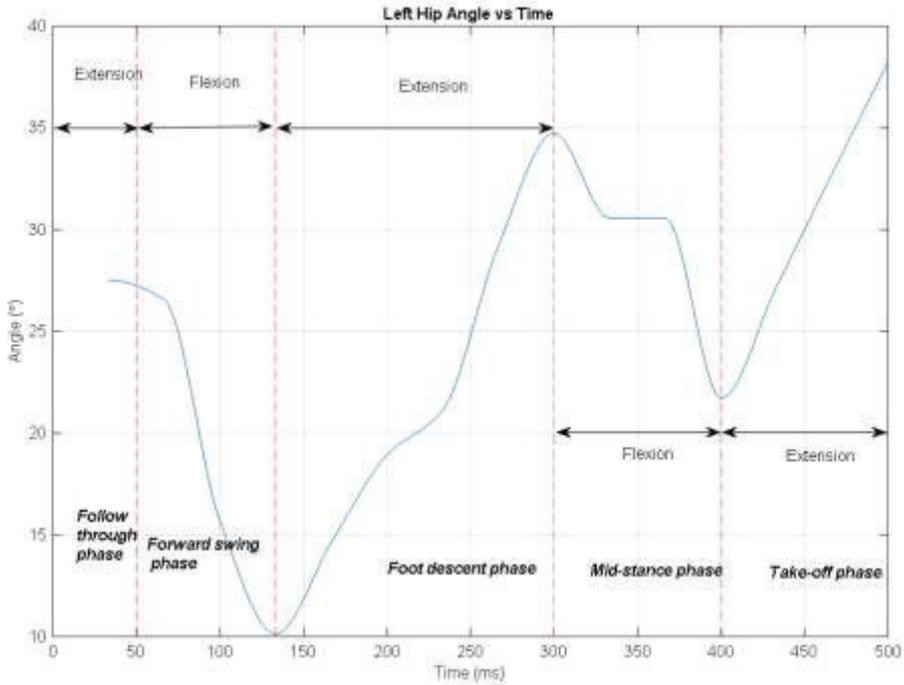


Fig. 14. Left hip angle during the running.

However to further provide insight into the interrelationship between joints when running, an angle-angle diagram (also known as cyclogram) easily interpreted the relation. The coordinations of the joints were plotted in Figure-15 for the hip and knee.

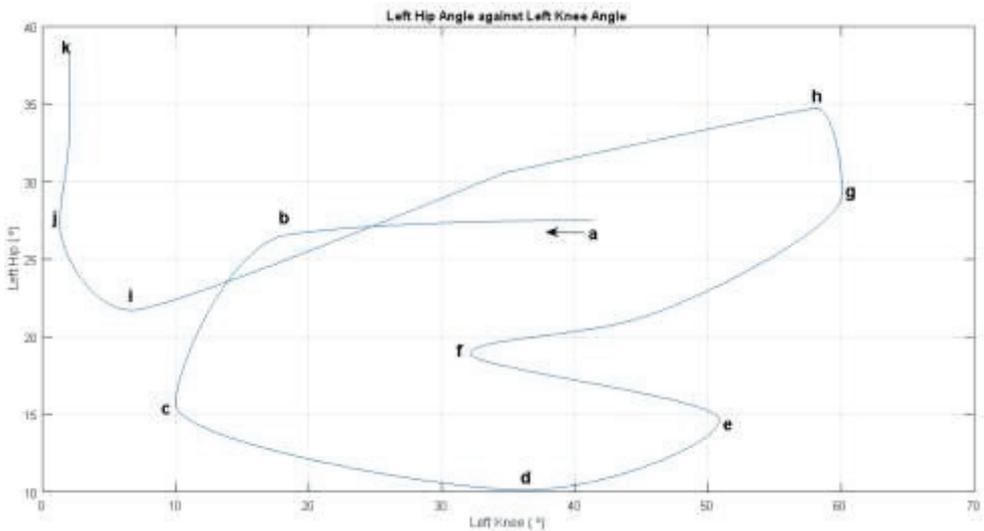


Fig. 15. Running angle-angle diagram performed by the athlete.

From Figure-15, the arrow mark indicated the movement direction of the hip and knee, with 'a' as starting point of the running. Next a continuous relative phase (CRP) was plotted for the differences of phase for both joints, as shown in Figure 16 below.

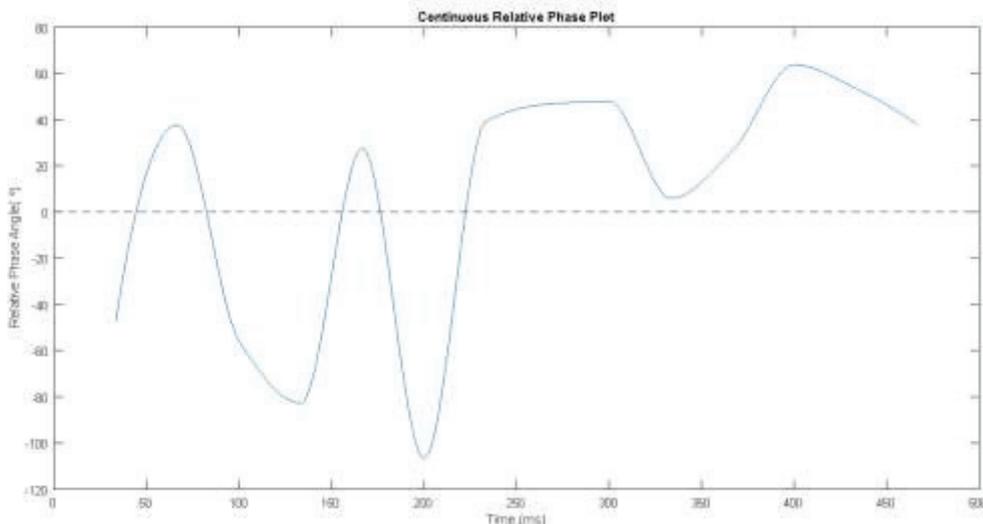


Fig. 16. Continuous relative phase of the runner between the hip and knee angles.

The two joints being in-phase when the CRP is equal to 0° , which was the instance at 44th, 82nd 155th, 176th and 222nd milliseconds. At these moments, the hip and knee undergoes a same action (either flexion or extension) at the same time. Referring to Figure-12 and Figure-14, both joints flexed at 44th and 82nd, extended at 155th and slightly flexed at 176th and back to extension at 222nd milliseconds. The anti-phase of joints occurred at 180° . However, from the graph, there were zero anti-phase of the hip and knee, but the phase is approaching the anti-phase and back once again (sinusoidal shape).

5 Conclusion

The depth camera motion capture system (Kinect) allows retrieval of skeleton joints coordinates of athletes. With the obtained coordinates, biomechanics analysis on the athlete can be done quantitatively to improve their performance while reducing the risk of injury. The biomechanics analysis includes linear and angular kinematics that consists of displacement, velocity, acceleration, angular velocity and angular motion (flexion, extension). This work verifies that the dual Kinect system is capable of performing the biomechanics analysis on athletes. The dual Kinect system is able to track the athlete's motion in a wider coverage as compared to single Kinect system. The system is portable, uncomplicated set-up, economical, and does not require substantial amount of time and effort for data acquiring and processing. In this work, linear kinematics analysis of running biomechanics which involves the centre of mass (CM) of the athlete were successfully computed. Displacement, velocity and acceleration of the athlete were calculated using the joint coordinates retrieved from the Kinect system. The three parameters were presented graphically. Angular kinematics analyses, which were much more complicated as compared to linear motion, were successfully calculated too using the joint coordinates and vectors. Angular velocity and joint coordination quantitatively defined the athlete's movement in

terms of angle-angle relation and in time series. The flexion and extension of joints were interpreted in term of time and their interrelationships were represented in a comprehensive continuous phase planes plot. Foremost, the system omits the need of markers attached to the subject's body, which allows more comfort and a lot of time saving in setting up the system, compared to conventional marker-based motion capture system which is expensive and lab-based. Referring to the performance of the dual depth camera system, higher number of depth cameras are expected to give better features to the motion caption system.

Acknowledgement

This research is supported by the Fundamental Research Grant Scheme (FRGS), Grant No: FRGS/1/2015 Vot 1543 by Ministry of Education, Malaysia.

References

1. J. Seung-kook, X. Zhou, D. K. Ramsey, V. N. Krovi, "A Comparative Study of Human Motion Capture and Analysis Tools A Comparative Study of Human Motion Capture and Computational Analysis Tools," in *The 2nd International Digital Human Modeling Symposium*, Michigan, USA (2003)
2. P. McGinnis, *Biomechanics of Sport and Exercise With Web Resource and MaxTRAQ 2D Educational Software Access* (3rd ed.), Illinois: Human Kinetics (2013)
3. D. T., "Evaluation of the Kinect Sensor for 3D Kinematic Measurement in the Workspace," *Applied Ergonomics* **43**(4), pp. 645-49 (2011)
4. J. Seung-kook, X. Zhou, D. K. Ramsey, V. N. Krovi, "A Comparative Study of Human Motion Capture and Analysis Tools A Comparative Study of Human Motion Capture and Computational Analysis Tools," in *The 2nd International Digital Human Modeling Symposium*, Michigan, USA (2003)
5. M. Caon, J. Tscherrig, E. Mugellini, O. A. Khaled, "Context-Aware 3D Gesture Interaction Based on Multiple Kinects," *First International Conference on Ambient Computing, Applications, Services and Technologies (AMBIENT)*, pp. 7-12 (2011)
6. P. J. Besl, N. McKay, "A Method for Registration of 3-D Shapes.," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. **14**(2), pp. 239-256 (1992)
7. D. B. J., . D. J. Robertson, "Response Characteristics Of Different Butterworth Low-Pass Digital Filters.," in *Proceedings of the 2nd North American Congress on Biomechanics*, (pp. 105–106) (1992)
8. G. L. Moir, *Strength and Conditioning*, Burlington, Massachusetts, United States: Jones & Bartlett Learning (2016)