Accuracy of joint angles tracking using markerless motion system ☆

Nazeeh Alothmany[1], Afzal Khan[1], Majdi Alnowaimi[2,*], Ali H. Morfeq[1,*], Ehab A Hafez[4,*]

Abstract

Human motion analysis is a widely accepted diagnostic tool in the field of medicine, sports, and biomechanics. Clinicians and Rehabilitators use this technology to identify the fundamentals of human gait and posture deviations. Various motion analysis methods exist, some track body motion using markers placed on the body while others track motion without using markers. In this paper, the performance of Microsoft Kinect as a marker-less system to track Sit-to-Stand movement is compared with the performance of a Myomotion system that uses inertial sensors to track the same motion. Four healthy adult persons performed sit to stand movement and Joint angles of hips, knees and ankles were collected at real time using Kinect and Inertial sensors.

A correlation analysis was performed on the angles calculated from the data collected using the Kinect and Inertial sensors. Results showed a high correlation for kinematics of knee and lower values for hips and ankles. RMSE (Root mean square error) was calculated for joint angles. The RMSE was found to be 21.1 for left knee and 24.4 for right knee which is lower than hip and ankle. This indicates that Kinect systems can be used to diagnose the motion of knee joints directly facing the camera

Keywords: Gait, Inertial sensors, Joint Angles, Kinect, Motion analysis, Markerless, XBOX

1. Introduction

Motion analysis is a standard technology used to examine behavior of neuro-musclo-skeletal system over various health issues like neurological, orthopedic disorders [1, 2, 3] and fall prediction [4]. Finding reliable and accurate low cost Motion analysis systems is a challenge for biomechanics. There are generally two kind of motion analysis techniques: Kinematics and Kinetics. Kinematics work with actual motion and its associated properties whereas Kinetics is concerned with Newtonian forces experienced by body. Both the fields are driven by motion capture systems providing necessary tools in estimating the properties of motion [4].

Various methods had been proposed for motion analysis most notably Marker-based and Marker-less systems. Marker-based systems use multiple infra-red (IR) cameras and reflective markers placed on subjects. Marker-based systems are accurate but not widely available in many hospitals and institute because these systems are expensive and not portable as a result they are limited to clinical settings. Recent research had shown the use of wearable sensor as inertial sensors [5] being less expensive, mobile and light weight and can be implemented in clinical settings. Wearable sensors do have some disadvantages, including the artifacts arising due to signal drift and noise [6] and each sensor is limited to measure a few properties hence more than one sensor is required and need regular updating.

Marker-less motion tracking technologies have advanced drastically in recent years. Segen and Kumar [7] used light source, a camera, and shadows, to achieve real time motion tracking of the hand, however shadow variations caused inconsistencies due to distance of the hand from the background. Marker-less motion tracking system developed by Cheung et al used shapes from silhouette images [8]. The technology used human kinematic models for tracking motion of person but their system did not work in real time. was accomplished by Breuer et al. [9]. The system was not able to track different movements of each individual finger. Soutschek et al. [10] tracked hand

☆This project was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah under grants/Project no: MG/34/16 Dated: 14/7/2013 supported in part by the U.S. Department of Commerce under Grant "BS123456".

*Corresponding author

Email address: malnowaimi@kau.edu.sa (Ehab A Hafez)

[1]Electrical and Computer Engineering Department, Faculty of Engineering, King Abdulaziz University, PO Box 80204, Jeddah 21589, Saudi Arabia.

[2]Nuclear Engineering Department, Faculty of Engineering, King Abdulaziz University, P.O. Box 80204, Jeddah 21589, Saudi Arabia.

[3]3Physical Therapy Department, Faculty of Applied Medical Sciences, King Abdulaziz University, PO Box 80324 Jeddah 21589.

[4]Faculty of Physical Therapy, Cairo University, Cairo 21589, Egypt.

gestures for a user interface, which they did by using a single time-of-flight cam-era. The time-of-flight camera provided 3D image data, giving the depth as well as the color information at each point. This provided more reliable depth information, but only used recognized specific gestures with marker-less tracking.

Marker-less motion tracking offers a novel solution to over-come the problem faced by markers based systems. The marker-less systems have their complexity of acquiring accurate 3D Kinematics. Estimating the free motion of human body is using multiple camera is under-constrained without the spatial and temporal correspondence that marker based system provide [11].

Problem arising in Marker-less systems are addressed by using model-based approaches. Complexity of model can be reduced by using a priori model of subject and can be used to decrease the total number of degree of freedom another method is to increase the number of cameras providing more measured data to resolve the issue of degree of freedom resulting in the robustness to Marker-less approach. This strategy was implemented by using human model to identify the motion of subjects, while several other model based approaches had been proposed by modeling the human body or parts with rigid an non rigid segments[12,13].

Model based approaches had accuracy issues for 3D kinematics of body segments and limited the number of body segments that can be studied [12]. To overcome this approach of mathematical formulation for model joints exponential maps provided the necessary tools to simplify the estimation of pose leading to identification of poses of kinematics. In order to choose an approach for Marker-less motion capture system is the formulation of cost function used to match the representation of subject (2D silhouette, 3D visual hull, 2D features etc. to model) [11].

Quantitative assessment for Kinect X box 360 with a Marker-based system for monitoring respiratory motion of the anterior surface was found to have small deviation for the motion analysis results between the two systems and proved that Kinect Xbox could be used as a potent tool for motion analysis[12]. Feasibility study of marker-less system for gait analysis used Kinect to track joint movements for right shoulder, right elbow, right hip, right knee and right ankle from sitting to standing task. Kinect camera was used to track the markers placed over the subject and Kinect depth feature was used to track the marker-less joints. The correlation coefficient be-tween Kinect camera's Marker-based and Marker-less systems was found to be 90.16% [1].

The objective of this study is to evaluate the accuracy of Kinect system as low cost, non-invasive system that can track and analyze a wide range of motion. The angles formed by the motion of certain joins (Knee, Hip, and Ankle) will be calculated using a marker-less Kinect system and a marker-based Myomotion system that is widely used by rehabilitators in motion analysis. Results will help identify whether a Kinect system can be used for this motion tracking.

The objective of this study is to evaluate the accuracy of Kinect system as low cost, non-invasive system that can track and analyze a wide range of motion. The angles formed by the motion of certain joins (Knee, Hip, and Ankle) will be calculated using a marker-less Kinect system and a marker-based Myomotion system that is widely used by rehabilitators in motion analysis. Results will help identify whether a Kinect system can be used for this motion tracking.

2. METHODOLOGY

MATERIALS REQUIRED:

- Myomotion Inertial Sensor and its proprietary software
- Two Kinect Camera and 2 tripod
- Kinect SDK software
- Visual studio C 2010
- Developed algorithm to calculate joint angles
- One chair without hand support
- Subjects to record the data

2.1. MYOMOTION:

Myomotion Inertial sensor system is a marker-based and a 3D portable motion capture system designed by Noraxon to study joint kinematics. The system has a static accuracy of +/- 0.40 and dynamic accuracy of +/- 1.20[14]. In this study it was used as a reference to measure the angles formed by the motion of six joints (left knee, right knee, left ankle, right an-kle, and the hip joint).

The Myomotion uses wireless sensors that can track 3D degrees of freedom (DOF) motion in a virtual reality environment. The system is portable and the sensors don't have a no-line of sight restrictions. The sensor includes a small amplifier with a wifi radio module that transmits Myomotion (MM) data in real time up to 20 meters.

Figure 1: Subject with Inertial sensor over his body.

The MM transmission sensors are strapped on to the subject body parts based on standard application of Noraxon system. The transmission probe transmits the data to the MM receiver which in turn is connected to PC using USB cable. Figure 1 shows an example of how the sensors are strapped on the body.Myomotion provides full body modelling using 16 sensors tracking the motion of 15 joints through calculating 22 anatomical angles as shown in Figure 2.

Myomotion provides full body modeling using 16 sensors tracking the motion of 15 joints through calculating 22 anatomical angles as shown in Figure 3.

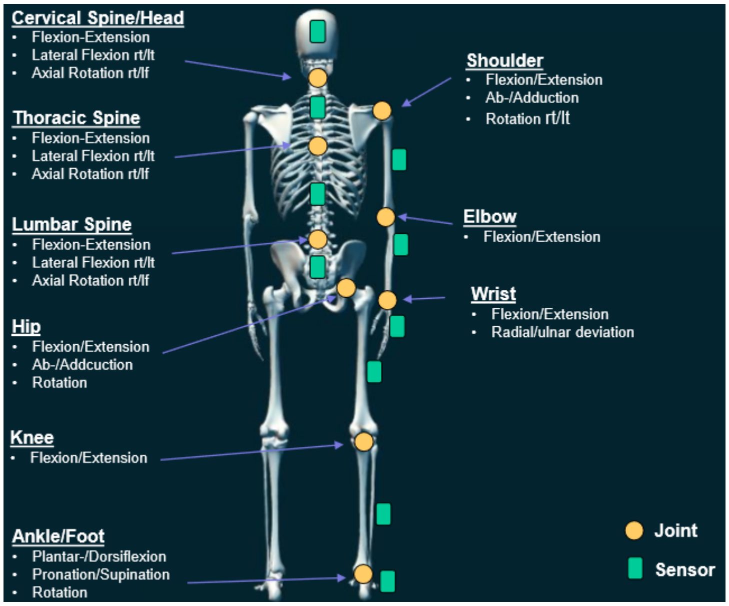


Figure 2: Myomotion provides full body modelling using 16 sensors tracking the motion of 15 joints through calculating 22 anatomical angles.

2.2. KINECT DESCRIPTION::

Microsoft Kinect is a motion sensing device developed for popular gaming console XBOX. User control and gesture interaction being the key feature which eliminates the need for game controller [15]. Kinect uses a range camera technology that recognizes specific gesture making it hands free control by using infrared projector and camera, it also contains a special microchip to track the movement of objects and individuals in 3D [16].

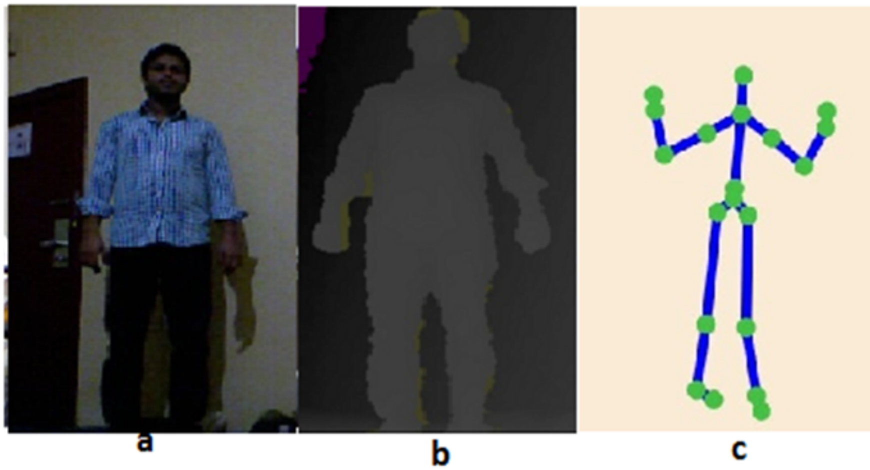


Figure 3: Subject with Inertial sensor over his body.

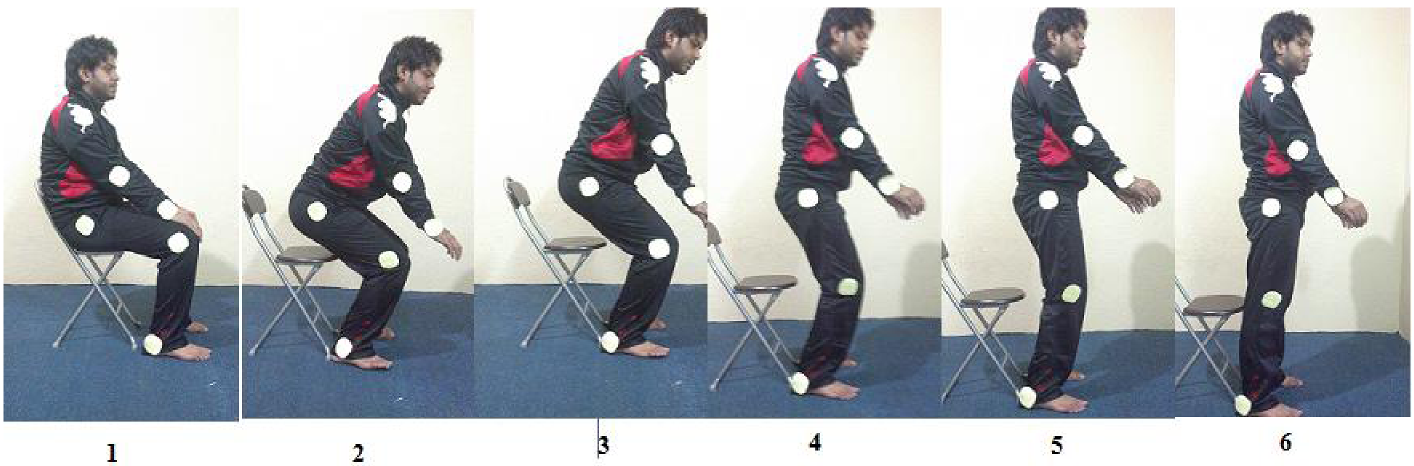


Figure 4: Subject with Inertial sensor over his body.

The system includes a sensor consisting of an RGB camera, a 3D depth sensor in the front of the Kinect, and multiple microphone arrays at the sides. The sensor also has a motorized tilt that allows users to change the viewing angle plus or mi-nus 27 degrees. The depth sensor consists of two parts, an infrared projector, which is located on the left side of the RGB camera, and a receiver which is located on the right side of the camera.

The infrared projector of Kinect projects a fixed pattern of infrared spots onto the object, then the receiver captures a shifted grid of these spots, the processor then calculates the offset of each of the spots to generate a depth map. The Kinect sensor can measure the distance of an object 2 meters away within 1 cm accuracy [17]. Kinect RGB uses 8 bit VGA resolution (640 x 480) with Bayer color filter and capable of handling higher resolution up to 1024 x 1024 but at a lower frame rate [18].

The Kinect depth sensor consists of an infrared laser projector with monochrome CMOS sensor, which enable Kinect camera to capture the 3D depth information in any ambient light condition as shown in Figure 2.(b). Kinect provides adjustable depth sensor and it is able to calibrate the sensor based on physical environment [19]. Kinect skeletal tracking identifies subject in the field of view, the location of the user is also reported in the depth map. Kinect skeleton tracks human motion with 20 joints defined by Microsoft proprietary software. For every pixel in the depth map, the three lowest-order bits contain subject index information. Kinect can recognize a maximum of 6 subjects at a time. This pattern of light is used to calculate the depth of the people in the field of view allowing the recognition of different people and different body parts [20]. Kinect provides three kinds of output RGB, Depth and Skeleton shown in Figure.5 a, b, c show RGB, Depth and Skeleton output. These features of Kinect had led researchers to use it for clinical monitoring and gait analysis [1,18].

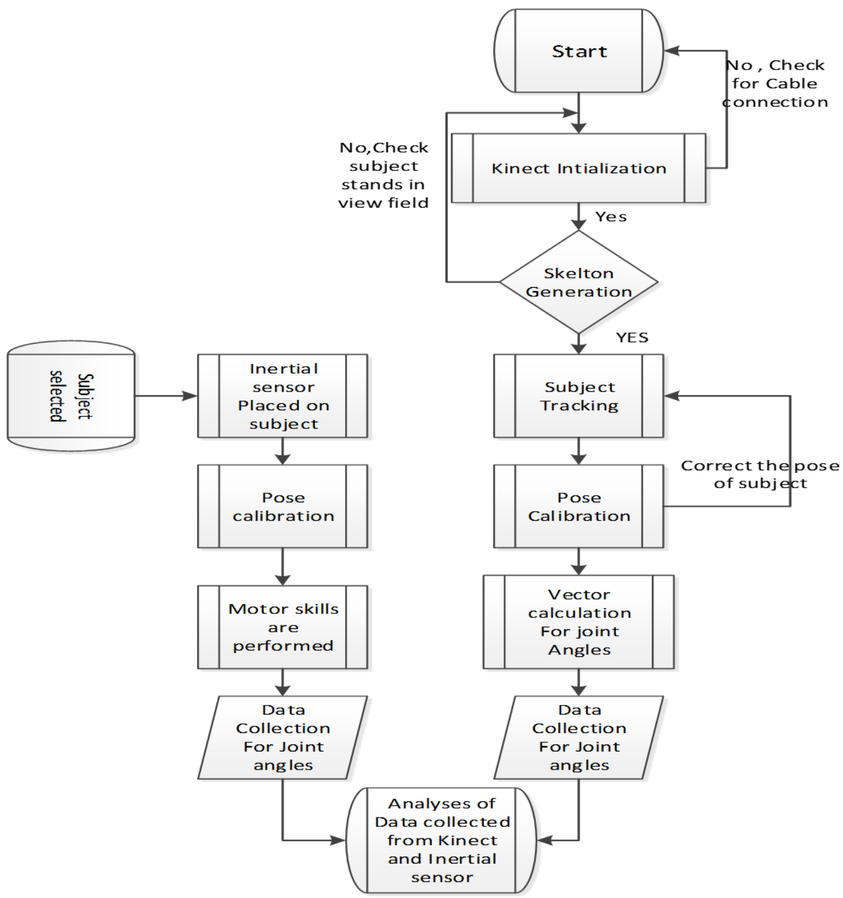


Figure 5: Subject with Inertial sensor over his body.

2.3. SYSTEM DESCRIPTION

In order to track the joints, Kinect Skeletal was created to track the joint angles by an algorithm developed using Kinect SDK (Software Development Kit), C# programming language and coding4fun toolkit. Kinect SDK is available from Kinect website, C# programming language is widely available over internet, Coding4fun website provides toolkit to develop code for Kinect [1].

ROOM SETTING: Room is a bright, quiet, contrast on back ground with no moving object or place objects in viewing field of Kinect camera. Subjects: Normal subject varying between the age group 18-30 yrs. were used in trials. The motor capabilities of subjects are of normal people.

SELECTED TASK: A gross motor skill of Sit to Stand (STS) was selected for the trials on the recommendation of physiotherapist as shown in Figure 6.

Figure 7 shows the work flow of this study. The inertia sensors are calibrated by having the subject to sit in a stable position. The Inertial sensors are attached to the subject on middle of the Posterior Superior Iliac Spine, Lateral upper third aspect of both thighs, lateral upper third aspect of both legs and lateral aspect of both ankles. The subject then is asked to stand while his arms beside his body for pose calibration. The next step is to calibrate the Kinect. The subject is asked to stand with his arms away from body facing the Kinect so that Kinect is able to detect the subject. The subject is then asked to sit down and asked to perform the task. The data for the task was collected simultaneously for Kinect and Myomotion . Later data was exported from both systems in an excel sheets

2.4. JOINT ANGLE CALCULATION:

Vector calculus was used to calculate the joints between an-gles using Kinect as shown in Figure 8. if a and b are assumed to be two vectors between joints and intersect at a point. The angle ϕ can be calculated as following:

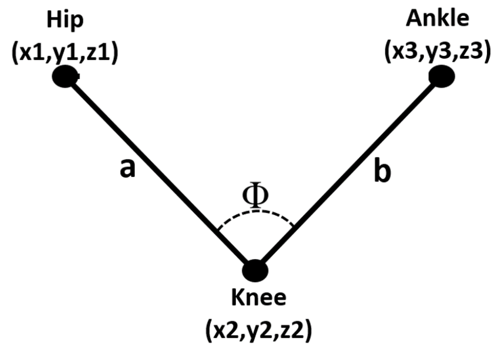


Figure 6: Vector calculus was used to calculate the joints between an-gles using Kinect as shown in the Figure .

$$unit(a) = \frac{a}{|a|} \tag{1}$$

$$unit(b) = \frac{b}{|b|} \tag{2}$$

$$a\ .\ b = |a|\ |b|\ cos(\phi) \tag{3}$$

$$unit(\phi) = cos^{-1}\ (\frac{a.b}{|a||b|}) \tag{4}$$

3. Data Processing

The data collected from Kinect in terms of Joint coordinates was used to find the angle between joints as explained in Joint angle calculation. The joint angles calculated in terms of degree.

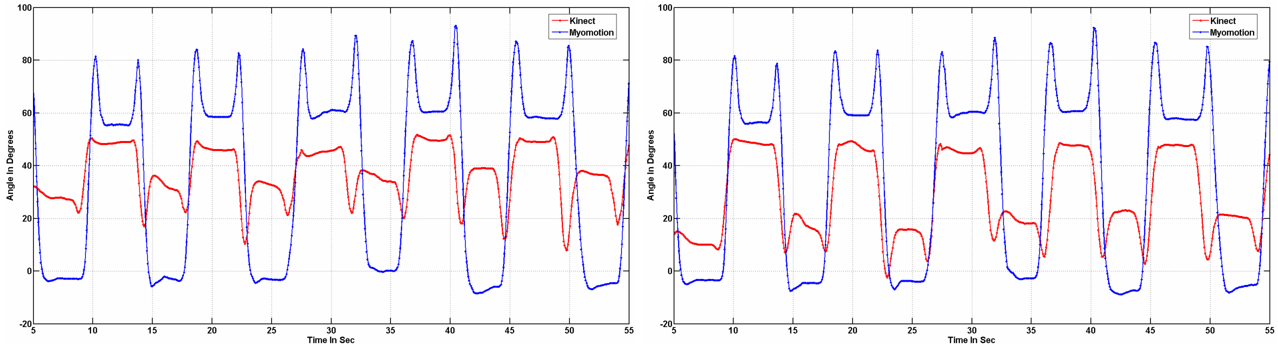
Inertial sensor collects data at 100 frames per second and Kinect collects data at 30 frames per second. The disparity in frames rates of data collections, required a solution to achieve the desired results. Matlab was used to process the data from Kinect and Inertial sensor. Kinect data was scaled to the time frame of inertial sensor so that length of Kinect and Inertial sensor vectors be-come same. After making the vectors of same length interpolation was used to construct new data points for Kinect with-in the range of a discrete set of known data points of Inertial sensor. The algorithm 1 is a code snippet used for time scaling and interpolation of data.

Algorithm 1 Time scaling and data interpolation.

if max(ta0) < max(tb0) **then**
    $dt = ta0(end) - ta0(end-1)$
    $n = fix((max(tb0) - max(ta0))/dt)$
    $a0 = [a0; zeros(n, 1)]$
    $ta0 = [ta0; max(ta0) + (1 : n + 1)' * dt]$
else
    if max(tb0)<max(ta0), **then**
        $n = fix((max(ta0) - max(tb0))/dt)$
        $b0 = [b0; zeros(n, 1)]$
        $tb0 = [tb0; max(tb0) + (1 : n)' * dt]$
    end if
end if
$b0 = interp1(tb0, b0, ta0,' pchip')$

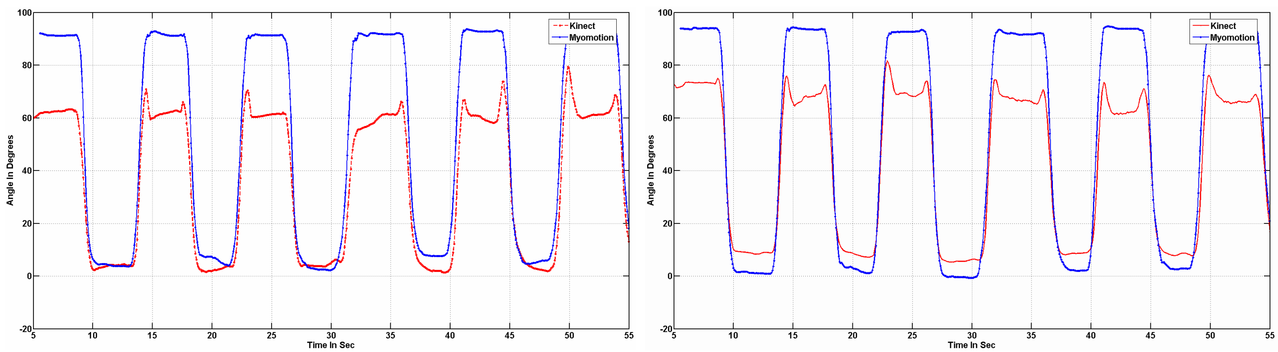
4. RESULTS:

The data from Kinect and inertial sensors were recorded in terms of coordinates for x, y, z axis for the mentioned six joints. Joint angle calculation was done using the same algorithm as mentioned in the methodology section. The



(a) Left hip joints Kinect and Myomotion results. (b) Right hip joints Kinect and Myomotion results.

Figure 7: An example Figure of the hip joints Kinect and Myomotion results where 9(a) illustrates the left hip and 9(b) shows the right hip data.



(a) Left Knee joints Kinect and Myomotion results. (b) Right Knee joints Kinect and Myomotion results.

Figure 8: An example Figure of the Knee joints Kinect and Myomotion results where 9(a) illustrates the left knee and 9(b) shows the right knee data.

data from Kinect and Myomotion Inertial sensors were collected in a synchronized manner. The data from Kinect and Inertial was used to generate the RMSE (Root mean square error) for the selected task of sit to stand. Root mean square (RMSE) was derived over normalized vector Inertial sensor data and Kinect data using Equation 5 where M and K are the Inertial sensor and Kinect data respectively and $i = 1, 2, ...N$ is the number of tracked position.

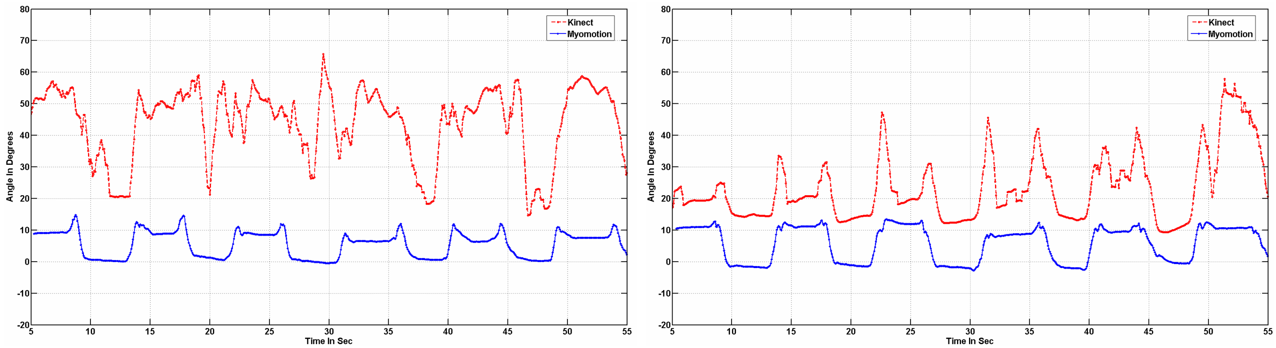
$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(M_i - K_i)^2} \tag{5}$$

The Correlation $\rho_{M,K}$ between the Kinect data and inertial sensor was calculated using Equation 6. where σ_M and σ_K are the standard deviations, E is the expected value operator, cov means covariance, and, corr a widely used alternative notation for the correlation coefficient.

$$\rho_{M,K} = Corr(M, K) = \frac{cov(M, K)}{\sigma_M \sigma_K} = \frac{(M - \bar{M})(K - \bar{K})}{\sigma_M \sigma_K} \tag{6}$$

Figure 7 shows the hip joint angles collected during the task of sit to stand. The Blue color indicates the data from Inertial sensors and Red indicates the Kinect data. The graph shows change in angle in degrees over time in sec for Kinect data and Myomotion data and they follow each other. The RMSE (Root mean square error) is 25.8 and correlation value is 0.46 for Left Hip and RMSE value is 23.5 and correlation value is 0.56 for Right Hip.

Figure 8 shows the knee joint angles collected during the task of a sit to stand movement. The graph shows change in the angle in degrees over time (sec) for Kinect data and Myomotion data. The RMSE (root mean square



(a) Left Ankle joints Kinect and Myomotion results. (b) Right Ankle joints Kinect and Myomotion results.

Figure 9: An example Figure of the Ankle joints Kinect and Myomotion results where 9(a) illustrates the left Ankle and 9(b) shows the right Ankle data.

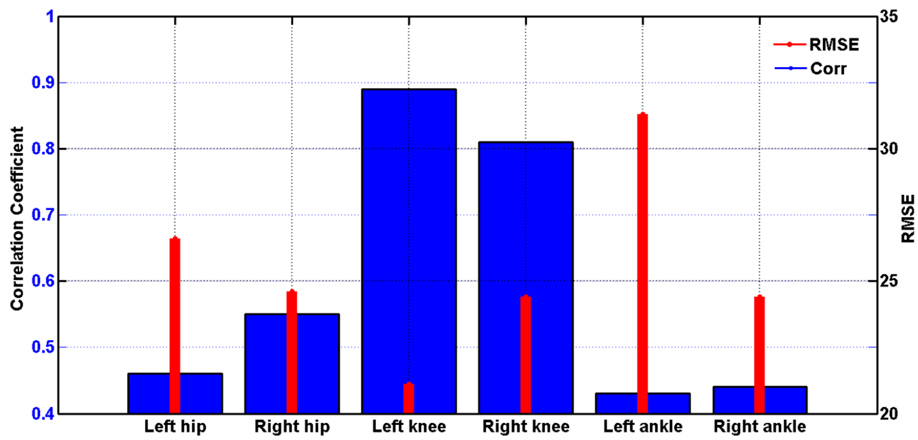


Figure 10: The figure shows the RMSE and Correlation for joints between Kinect and Inertial sensor.

error) is 21.6 and correlation value is 0.97 for Left Knee and RMSE value is 26.2 and correlation value is 0.97 for Right Knee. Where Figure 9 shows the hip joint angles collected during the task of sit to stand. The graph shows change in angle in degrees over time in sec for Kinect data and Myomotion data and they follow each other. The RMSE (root mean square error) is 23.5 and correlation value is 0.36 for Left ankle and RMSE value is 19.7 and correlation value is 0.51 for Right ankle.

Table 1 shows the RMSE (Root means square error) and Correlation between Kinect Data and Myomotion Inertial sensor for the defined task the data was collected simultaneously. The data for each subject was collected individually and average was calculated for the joints for RMSE and Correlation.

Table 1: RMSE and Correlation between Kinect and My-omotion (Avg. of 4 Subjects).

Kinect Position	Left hip	Right Hip	Left Knee	Right Knee	Left ankle	Right ankle
RMSE	26.6	24.6	21.1	24.4	31.3	24.4
Correlation	0.46	0.55	0.89	0.81	0.43	0.44

Figure 10 shows the RMSE VS Correlation for joints between Kinect and Inertial sensor. The Graph shows high correlation and low error for Knee joints. There is a low correlation be-tween ankle and hip with higher error rates.

5. Discussion

This study focused on accuracy of marker based system with the standard marker-based clinical system motion analysis system. The joint tracking with Kinect had measured reason-able well for some joints and for some joints the results were did not measure well with Inertial sensors. Kinect has shown high correlation values and low error rates for Knee Joints enhancing the accuracy of Knee joint tracking for both the legs when compared with Myomotion. The accuracy is around 85% for some subjects and accuracy for some subjects had reached to 99%. The variation in the accuracy for subjects is due to fact every subject performs the same Sit to Stand task uniquely which would bring some errors and tracking the subject without markers brought variable results. Kinect was able to track the Knee with much accuracy then ankle and hip as it is the angle made between two long bones which is very much visible with 1 degree of freedom but the ankle and hip are complex joints and with more degree of freedom.

The other reason being difference in approach for joint angle calculation between Kinect and Myomotion. Kinect uses depth data to generate a skeleton by segmenting the depth image into a dense probabilistic body part labeling using randomized decision tree, with the parts defined to be spatially localized near skeletal joints of interest. Re-projecting the inferred parts into world space, the localized spatial modes of each part distribution and thus generate (possibly several) confidence-weighted proposals for the 3D locations of each skeletal joint [21].

But Myomotion uses sensors specifically placed over region of interest. The sensors provide Quaternion, Linear acceleration and magnetic distortion data. These data's are the processed using the Myomotion proprietory software to generate angles for joints and output is exported in terms of excel files [13].

The Kinect failed to live up to the expectation of tracking the ankle and hip joints with high accuracy as shown in Table 1. The error rates were high and correlation values were low ankle and hip joints.

Kinect provides a portable and viable option as compared to expensive marker-based system but it has its own limitation. The body joints defined by Kinect are not precise when compared with a marker-based system. As Kinect works on the principle of depth range using infrared pattern it is not very effective in outdoor and bright lights. Kinect detects human presence by recognize the body posture and face but sometimes it might create skeleton on some non-human objects like table or chair especially when object parts my look like extremities. Kinect requires its user to stand at a distance of 6-8 feet and has viewing range of 430 vertical and 570 vertical so the possibility of using it in a small room is ruled out.

6. Conclusion

In this paper, a new approach for Marker-less human motion capture from a single Kinect camera and comparing the joint angles with a clinical marker-based inertial sensor system. The aim of this study was to extract the data from Kinect to generate the joint angles and calculate them during the task and compare the results with inertial sensor and check for the ac-curacy. The above experiments conclude that Kinect is able to track the joints of lower limbs with Knee tracking being most-ly in sync with inertial sensor and achieving high correlation value but the ankle and hip have low correlation.

With above mentioned benefits and limitation Kinect shows promising results and beholds a future for marker-less tracking. The most important feature of Kinect is that provides a cheaper 3D system, depth data, and skeleton data and RGB data. It also comes as a user friendly, portable device, can work in dark without any light source.

The developed software was compatible with windows 7 and windows 8, no restrictive modeling was used as the approach was generic. Future development in depth range technology and its availability over Kinect could bring better results.

7. Bibliography styles

References

[1] R. Feynman, F. Vernon Jr., The theory of a general quantum system interacting with a linear dissipative system, Annals of Physics 24 (1963) 118–173. `doi:10.1016/0003-4916(63)90068-X`.

[2] P. Dirac, The lorentz transformation and absolute time, Physica 19 (1-12) (1953) 888–896. `doi:10.1016/S0031-8914(53)80099-6`.