Gait Analysis Using RGBD Sensors

Ravindu Kumarasiri*, Akila Niroshan†, Zaman Lantra‡, Thanuja Madusanka§,
Chamira U. S. Edussooriya¶, and Ranga Rodrigo‖
Department of Electronic and Telecommunication Engineering, University of Moratuwa,
Moratuwa, Sri Lanka
Email: *mailravindu@gmail.com, †akila.niroshan111@gmail.com, ‡zamanlantra@gmail.com
§pvtmadusanka29@gmail.com,¶chamira@ent.mrt.ac.lk,‖ranga@uom.lk

***Abstract*—Human gait analysis, the study of human locomotion, is possible with low-cost RGBD sensors such as the Kinect sensor. However, due to the inherent depth sensing accuracy limitations of these sensors as the distance from the sensor increases, the distance range of gait analysis too becomes small an inefficient for clinical use. We present a system that uses two independent Kinects in a data fusion framework that increases the distance range of gait analysis from 2.5 m to 4 m with three gait cycles. Our gait parameters are reasonably accurate and comparable with existing systems with 4% error in length measurements and $5°$ error in flexion measurements.**

***Index Terms*—Gait, Kinect, clinical, rehabilitation, diagnostic, vision based, gait parameters, Kalman filter**

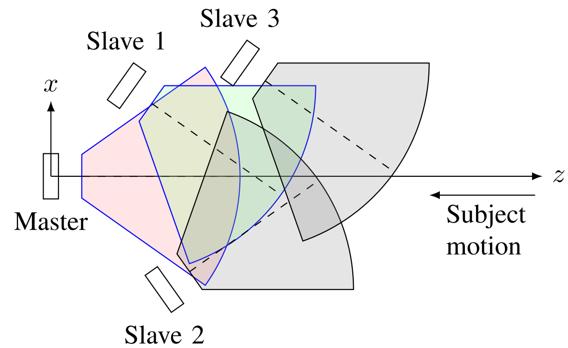


Fig. 1: System setup indicating the fields of view of Kinect sensors. We only use a master and a slave in this paper.

I. Introduction

Gait analysis is the study of human locomotion. The human walking pattern depends on the muscles, the joints and the nervous system [1]. Gait analysis is important in diagnosing patients, who are recovering from accidents, brain strokes, neurological diseases [2], musculoskeletal anomalies and psychiatric disorders [3]. Clinicians prescribe gait analysis test as a standard test to identify and monitor the progress of treatments for aforementioned disorders.

Gait analysing systems can be categorised as vision-based systems and non-vision-based systems. Vision-based systems are further categorised into marker-based systems and markerless systems. Marker-based systems [4] require wearing special cloths with markers, especially-designed high-performance hardware and software to operate. Due to these reasons marker-based systems are expensive. On the other hand, current implementations of markerless systems have a limited operating range which is not sufficient for gait analysis [5] [6]. Even though a treadmill may be employed to overcome this limitation [7] [8], forceful walking on a treadmill changes the natural walking pattern [9]. Besides, patients who are very weak to walk on a treadmill (e.g., patients who walk on crutches) are not be able to use treadmill-based systems.

In this paper, as a solution for the aforementioned problems, we propose an extended-range real-time markerless three dimensional (3-D) gait analysis system using Microsoft Kinect V2 sensors. Extended-range is achieved using multiple Kinect sensors. Optimum placement of Kinect sensors (Fig. 1) and the use of an optimum fusion algorithm improve the accuracy of the system in comparison to similar implementations [5] while improving the operating range. Use of Kinect sensors has reduced the cost of hardware in comparison to marker based systems [4] while reducing the computational complexity.

RGBD sensors, Kinects in particular, have been used for gait analysis. Multiple Kinect sensor fusion for human skeleton tracking using Kalman filtering by Sungphill *et al.* [5] presents a weighted-measurement fusion method that uses Kalman filter for multiple Kinect sensors. They uses five Kinect sensors placed in a semi-circular arch. This limits the system to a fixed operating range. Our solution is scalable for longer operating ranges depending on the requirement. We have achieved a similar level of performance accuracy with a lesser number of Kinect sensors.

The rest of the paper is organized as follows. In section II, we present Kinect camera calibration algorithm, real-time Kalman fusion algorithm and robust gait parameter calculation methodology. Empirical evaluation presented in the section III shows our results are robust and accurate when compared with the ground truth measurements and the readings obtained from a verified inertial-measurement-based gait analysis system. Finally, we present conclusion in section IV.

II. Proposed Multi-Kinect-Based Gait Analysis System

A. Overview

In this section, we discuss the overall architecture of the proposed gait analysis system. Our system, shown in Fig. 2, has two Kinect sensors. The Kinect 1 is the master while the Kinect 2 is the slave. Each Kinect sensor is connected to a

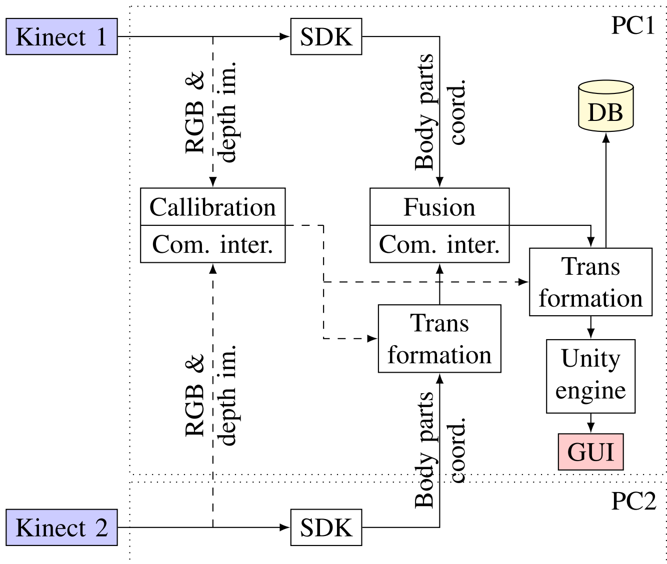


Fig. 2: Architecture of the overall system.

computer that has Microsoft SDK for Kinect to use as the interface.

We have implemented the overall software system using five major modules. Sensor Calibration, the first block, addresses the coordinate system misalignment of the two Kinect sensors. Since the system has two Kinect sensors, we use a data fusion module to combine the sensor outputs. The third is the gait parameter calculation engine. Fused data have x, y and z Cartesian spatial coordinates of 25 human body points. Gait parameter calculation engine uses aforesaid fused data to calculate gait parameters. Last two modules are graphical user interface (GUI) and the database. The database is used to store the calculated gait parameters and patients history for future reference and disease diagnosis process. We discuss sensor calibration, data fusion, gait parameter calculation, the GUI and the database system in the following subsections.

B. Sensor Calibration

Each Kinect sensor works in separate 3-D coordinate systems. Hence, we have to use a calibration mechanism to calculate parameters to transform coordinates in different coordinate systems into a common 3-D coordinate system. Our calibration module, based on Umeyamas algorithm [6] calculates the rotation matrix and the translation vector. In this process we use the master Kinect coordinate system as the common 3-D coordinate system. This works in two steps as following:

1) Detecting 3-D points in a calibration object,
2) Calculating rotation matrix and translation vector.

We use Umeyama's algorithm [10], because it is more accurate, robust, and effective than the singular-value decomposition based algorithm [11]. Using this algorithm, it is possible to calculate the rotation matrix R_{rot}, the translation vector t and the scaling factor C. These calculated parameters are used to transform 3-D coordinates of a human body captured through one Kinect sensor to the other Kinect sensor's coordinate frame. After transformation, the two sets of coordinates are fused as explained in data fusion subsection.

C. Data Fusion

In our system, we use data fusion to the measurements where the vision of both the sensors overlap and also where the sensors operate solely. The Microsoft Kinect V2 body part recognition SDK [12] gives the position coordinates of 25 joints of the human body relative to its coordinate system in meters. After transforming these coordinates to the master Kinect coordinate system (Section II-B) the measurements of each Kinect can be fused.

There is no control input for the tracked coordinate. Microsoft Kinect SDK detects 25 body-parts. The state of the k^{th} body-part coordinate is $\hat{x}_k = \begin{bmatrix} x & \dot{x} & y & \dot{y} & z & \dot{z} \end{bmatrix}^T$. Therefore we have a 150×1 matrix. The matrix A is obtained by linear motion model equations, $\hat{x}_k = \hat{x}_{k-1} + \dot{\hat{x}}_{k-1} \times dt$ and $\dot{\hat{x}}_k = \dot{\hat{x}}_{k-1}$, where dt is the time between two frames ($\frac{1}{30}$ s). P_k are the error covariance matrices and Q is the process noise covariance matrix. The measurement noise covariance matrix R and the observation matrix H of size (75×150) are involved in the correction stage. These depends on the distance from the sensor and hence should be altered [13]. The measurement noise covariance matrix R is with respect to each Kinect. To convert into the coordinate system of master Kinect, R of the slave Kinect is rotated using $R_{rot} \times R \times R_{rot}^T$ where, R_{rot} is the rotation matrix of the slave Kinect with respect to the coordinate system of the master Kinect.

Some coordinates may be occluded by the human body or an object, hence it would be indicated as low reliable coordinates by the Kinect SDK. The SDK gives value 2 for full accuracy, value 1 for 50% accuracy and value 0 for zero accuracy. In our implementation, we have adjusted the H matrix only to get the fully accurate coordinates in the measurement update stage. When there is no reliable measurement at a given time, Kalman prediction is given as the output for that coordinate. When only one Kinect has reliable data, we take only the reliable data of Kinects for fusion.

D. Gait Parameter Calculation

Gait parameters are mainly categorized under angels and distances. Stride length, step length, cadence and spine to knee distance are the distances, and knee, hip, and elbow flexion/extension, lateral and anterior pelvic tilt are angels.

According to the camera setup, the subject walks towards the Kinects along the z-axis (Fig. 3). The stride length and the step length are calculated by identifying the resting position of both legs along the z-axis, corresponding to plateaus of Fig. 3. One-dimensional derivative of the z coordinates will extract the plateau points. Distance between two consecutive plateaus of the same leg produces the stride length. Distance between two consecutive plateaus for the left leg and the right leg produces the step length.

Cadence is the number of steps per minute. Cadence is calculated after obtaining the number of steps over the time

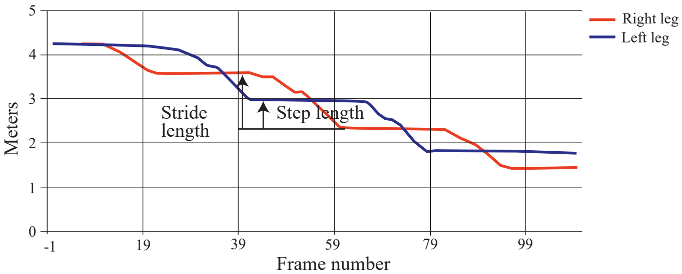


Fig. 3: The stride length and the step length.

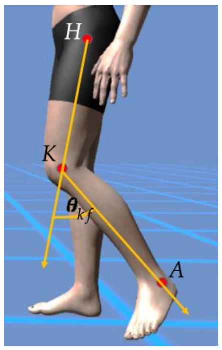


Fig. 4: Hip-knee ($\overrightarrow{HK}$) vector, knee-ankle ($\overrightarrow{KA}$) vector and knee flexion/extension (θ_{kf}).

difference. Knee flexion/extension is the movement around knee joint. This is represented by the angel θ_{kf} (Fig. 4) [1], [14] and calculated as

$$\theta_{kf} = \cos^{-1}\left(\frac{\overrightarrow{HK}\cdot\overrightarrow{KA}}{\left\|\overrightarrow{HK}\right\|\left\|\overrightarrow{KA}\right\|}\right), \tag{1}$$

where $\overrightarrow{HK}$ and $\overrightarrow{KA}$ are the hip-knee vector and the knee-ankle vector, respectively. Hip flexion/extension is represent by the angle between a leg and the vertical coronal plane of the human body. Hip flexion/extension is defined with respect to both the legs. It is calculated by getting the angle between vertical vector downwards ($\hat{v}$) through the coronal plane and the vector representing each leg. With respect to the vector used, the left and the right hip flexion θ_{hf}, lateral pelvic tilt θ_{lat} and anterior pelvic tilt θ_{ant} are calculated using (2), (3) and (4), respectively, where B and M are the coordinates of the spine base and the spine mid, respectively, and H is the coordinate of the hip.

$$\theta_{hf} = \cos^{-1}\left(\frac{\overrightarrow{KH}\cdot\hat{v}}{\left\|\overrightarrow{KH}\right\|\|\hat{v}\|}\right) \tag{2}$$

$$\theta_{lat} = 90^\circ - \cos^{-1}\left(\frac{\overrightarrow{BH}\cdot\overrightarrow{BM}}{\left\|\overrightarrow{BH}\right\|\left\|\overrightarrow{BM}\right\|}\right) \tag{3}$$

$$\theta_{ant} = \cos^{-1}\left(\frac{\overrightarrow{BM}\cdot\hat{u}}{\left\|\overrightarrow{BM}\right\|\|\hat{u}\|}\right) \tag{4}$$

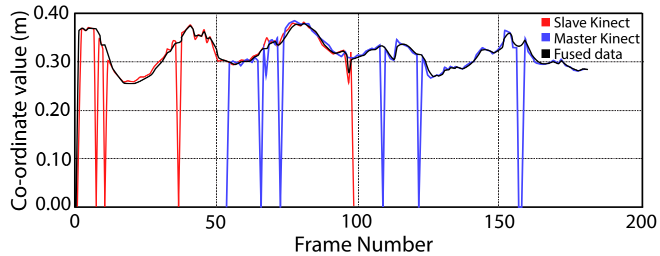


Fig. 5: Gait data of a random coordinate.

Due to injuries or diseases, patients tend to show waddling gait (legs move in a circular path when walking), circumduction (conical movement of limb). Observation of spine mid to knee distance will indicates those abnormalities. The system can calculate the lateral distances between the knees using the joint 3-D coordinates on the frontal plane. The x, y and z coordinates are passed to the gait parameter calculator with standard errors.

III. Results and Verification

In this section, we first describe the results pertaining to the individual blocks of our system: system calibration, data fusion, and gait parameter calculation. Then we present the results that verify the performance of the system along with comparisons.

A. System Calibration

We used 49 distinct corners of a vertically positioned 7×7 chess board to find the rotation R_{rot}, translation t and scaling factor C using the Umeyama's algorithm. We computed the re-projection error for 24 points out of the 49 using the estimated R, C, and t. The root mean square (RMS) re-projection error was 10^{-6} m and the maximum re-projection error was 5×10^{-3} m. Furthermore, rotation matrices were orthonormal and the scaling factor was 0.99 which should ideally be 1. RMS re-projection error for randomly selected samples of data turned out to be 10^{-2} m. This indicate that the calibration is reasonably accurate.

B. Data Fusion

Fig. 5 shows the gait data using only the individual Kinects and the fused gait data for two random coordinates. Spurious null values are due to frame losses. It is clear that the fused data follows very smoothly over the data from both the Kinects, though there are noise peaks in the data set. In addition, we can see that the calibration is successful, since data form both Kinects substantially overlap on each other after transformation.

C. Gait Parameter Calculation

Table I shows the gait parameter output from our system and from clinical gait analysis [15] for an average data set. This comparison is to investigate whether our system gives meaningful gait parameters. We observe in the average data set that the stride length is almost as twice as the step length

TABLE I: Parameter comparison table

Parameter	Our system	Averaged system
Stride length	1.1 m	0.76 m – 0.84 m
Step length	0.5 m	0.41 m – 0.45 m
Cadence	84.6 steps/min	82 – 88 steps/min
Speed	0.98 m/s	0.54 m/s – 0.6 m/s

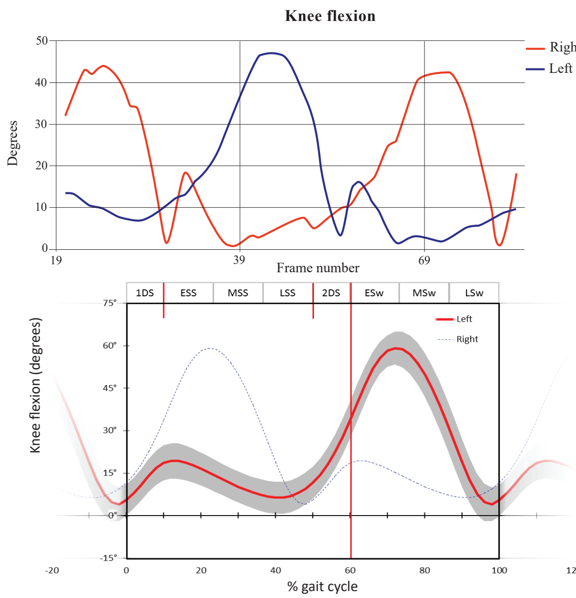


Fig. 6: Knee flexion/extension: our system vs. standard (source: wwrichard.net)

in the averaged system. In our system too those values have similar association.

Fig. 6 shows the knee flexion graph obtain by our system (top plot) and the standard plot (bottom plot) [16]. It is possible to interpret our systems output by comparing with this reference qualitatively. Fig. 7 and 8 show the hip flexion and spine-mid to knee distance of a healthy person respectively along with the standard curves. The shape and periodicity of the hip flexion graphs (Fig. 7) match with the standard curve. We do not report the lateral and anterior pelvic tilts to keep the presentation concise. Spine-mid to knee distance is a new parameter which is calculated by our system. A periodically changing pattern is the expectation from this graph and the validation can be directly done by measuring the distance manually.

Fig. 9 is a set of gait parameter graphs of a patient suffering from Cerebral Palsy having a waddling, circumduction and shuffling in the walking. Medical doctors and physiotherapists indicated that these graphs qualitatively match with the patients walking pattern.

D. Verification Experiments

1) Depth measurement of static objects: Initial experiment was to check the depth measurement accuracy of the Kinect against the ground truth. We kept static objects in front of the Kinect sensor and measured the depth manually. Then using the *coordinate mapper* functions of Kinect SDK, we obtain the depth measurements for interested points from the Kinect

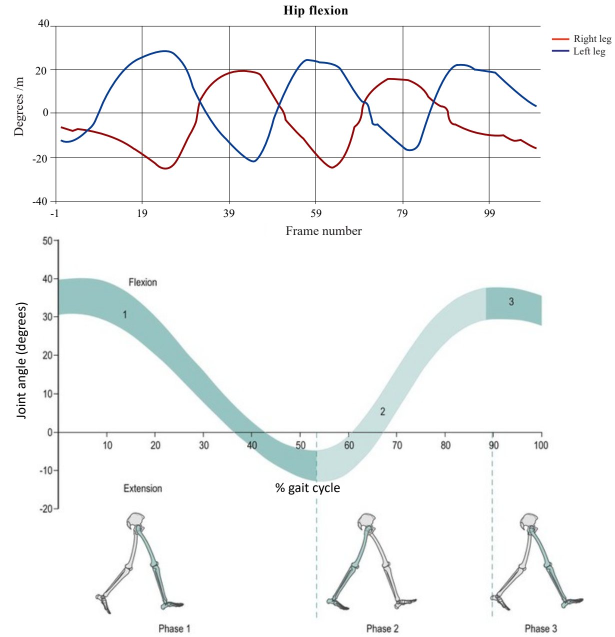


Fig. 7: Hip flexion/extension: our system vs. standard (source: musculoskeletalkey.com)

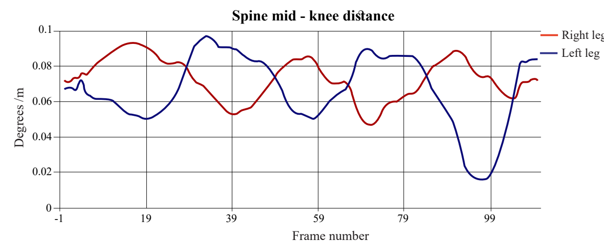


Fig. 8: Spine-mid to knee distance (our system)

sensor. We took a sample of 504 data points over a range of 280 cm. Depth error of the Kinect increases with the increase of distance from the Kinect and the overall average depth error was be 4 mm.

2) Human body part length and human body joint angle measurements of a stationary person: This experiment was conducted taking Kinect measurements while a person was standing motionless in front of the Kinect. The body part length detection accuracy was found by comparing the body part lengths obtained from the Kinect body part recognition, depth measurements obtained from the *coordinate mapper* function and the actual physical length measurements between joints using a meter ruler. Distances were calculated for the upper arm, forearm, shin and the thigh. The body part angle detection error was found by comparing the angles between body parts calculated using Kinect body part coordinates and actual body angles measured using a Goniometer. The right-knee, left-knee and the elbow flexion/extension were calculated in degrees.

Measurements were made from 1m to 3.2m from the Kinect with 20 cm step. The human body part length detection error of the Kinect system averages to 35 mm, and averaged angle detection error was $4°$ (see Figs. 10, 11). The idea of data

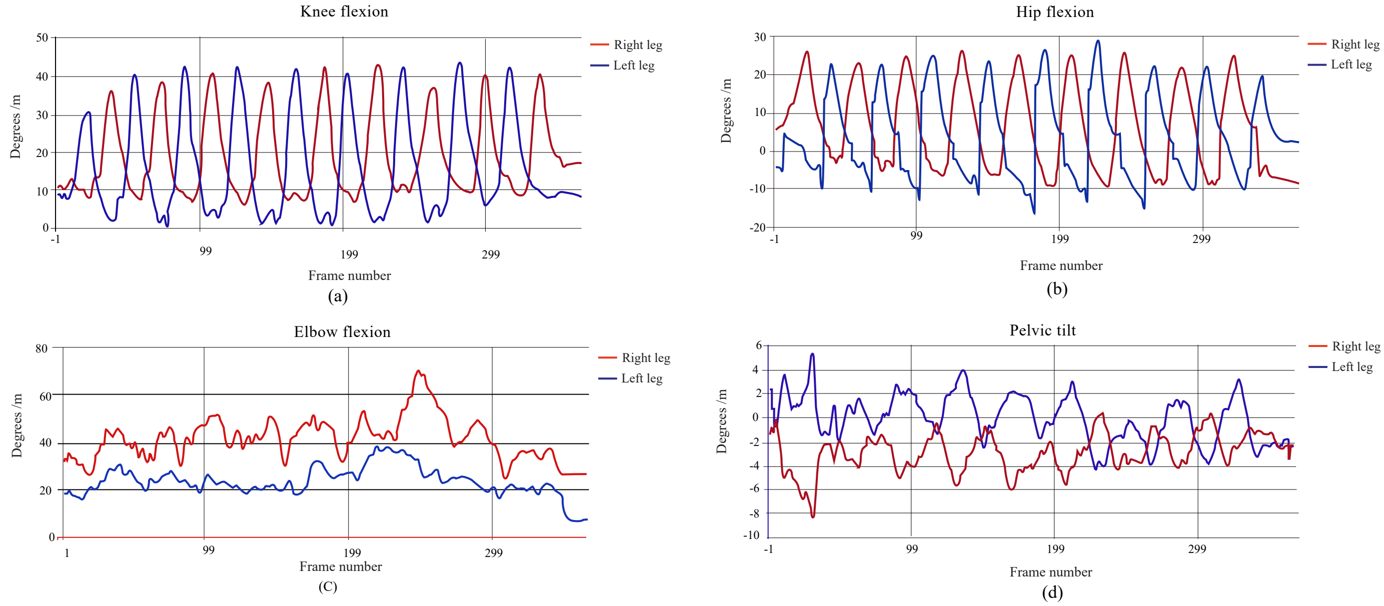


Fig. 9: (a) Knee flexion, (b) hip flexion, (c) elbow flexion, (d) lateral pelvic tilt (for patient data).

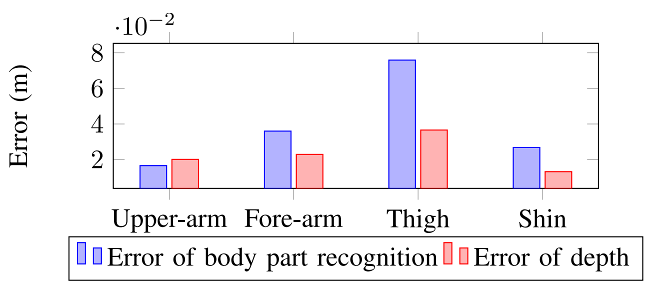


Fig. 10: Average body part length detection error.

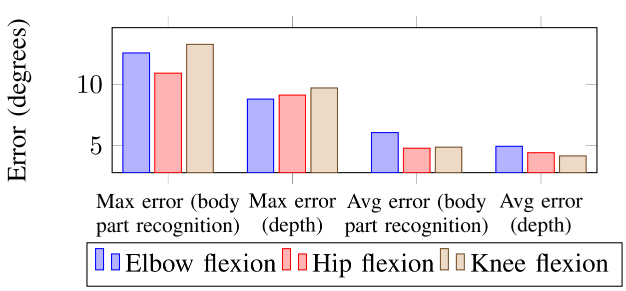
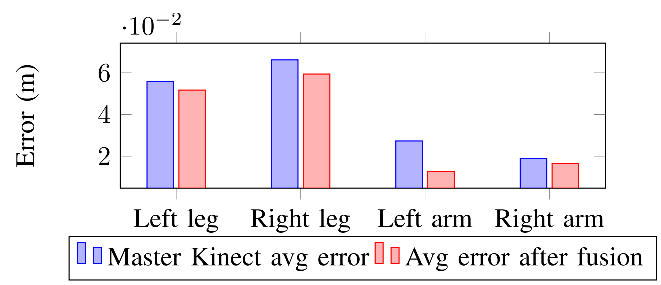


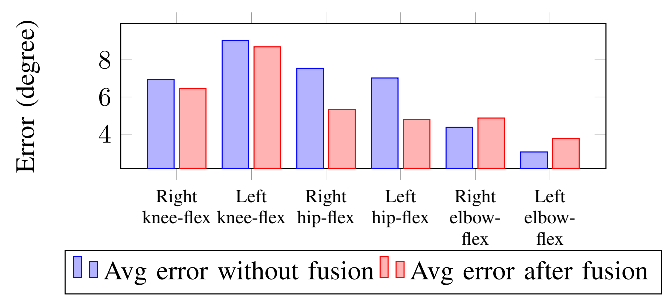
Fig. 11: Maximum and Average human joint angle detection error.

fusion is to increase the range from 2.5 m to 4.0 m. Therefore, there will be only a limited range of 1.5 m in which both the sensors gets data simultaneously. Fig. 12 shows that data fusion has increased the accuracy of the system, 5% in body length detection and 7% in joint angle detection.

3) Human body joint angle measurement of a moving character: Dynamic joint angles were measured against the readings from Kairos Sensing, which is a system built based on inertial measurement unit sensors. It has been verified against a robotic arm and has a maximum error of 3°. As shown in Fig. 13, the deviation of knee flexion and the hip flexion is 5°, when compared with the Kairos Sensing system.



(a) Body part length detection average error.



(b) Joint angle detection average error.

Fig. 12: Length and angle detection error before and after data fusion.

TABLE II: Average difference of stride lengths.

Parameter	Error
Left stride	5.0 cm
Right stride	5.8 cm
Step length	4.8 cm

4) Step and stride length detection of a moving character: Step and stride length are the two main gait parameters to identify any abnormalities in the walking pattern. In the verification process, step and stride length were captured by a

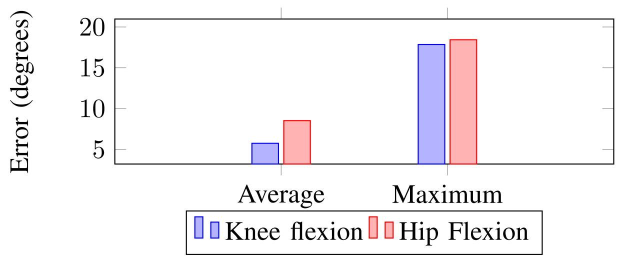


Fig. 13: Average and maximum deviation of joint angles obtained by Kairos Sensing.

TABLE III: Results comparison

	Stride & error		FE
	ME	EP	
Multiple Kinect sensor fusion [5]	0.069 m	6%	-
Biomechanical validation [6]	-	-	7°
2-D markerless Gait analysis [17]	0.04 m	3%	4.53°
Our system	0.05 m	4%	5°

video footage and compared with the results obtained by our system. We found that the overall average error as 5 cm. The smallest measurement of the used measuring tape is 1 cm.

E. Results Comparison

In Table III, the mean error (ME) in meters, error percentage (EP) and the flexion error (FE) in degrees are compared. Moon *et al.* [5] is the most closely related implementation to ours. Due to our optimal camera placement, we have achieved a similar, slightly superior performance with less number of cameras. Bio mechanical validation [6] by Ferandez-Baena *et al.* is a validation done for single Kinect system. Use of multiple Kinects and running the Kalman filter on the data have improved our accuracy over a single Kinect system. Two-dimensional (2-D) markerless gait analysis [17] by Castelli *et al.* is a 2-D vision based system. Their error performance is more superior to our system. However, due to very high computational demand required for their point cloud projections their system is not real-time like our system.

IV. Conclusion

As we have shown in our work, use of multiple Kinects and fusing the data using the Kalman filter has increased the operating range of 2.5 m of a single Kinect to 4 m while increasing the accuracy. This is achieved using only two Kinects. Average calibration ME is 10^{-6} m while the back projection error is 5×10^{-3} m which are well under the acceptable level of accuracy. Verification for a static body reveals an error of 35 mm in length measurements and 4° in angle measurements. Results for a dynamic body in comparison with Kairos sensing gives a variation of 5°. The maximum reported propagation error of our system is 7°. These errors and the variations are within the acceptable level for gait analysis applications. Capture range of this system could be easily improved using a cascade of sensors.

Acknowledgment

The authors would like to thank the Senate Research Committee of the University of Moratuwa for financial support through the grant SRC/LT/2016/04; and Pujitha Silva, and Peshala G. Jayasekara for technical contributions.

References

[1] J. Perry and J. M. Burnfield, *Gait Analysis: Normal and Pathological Function*, 2nd ed. New Jersey: Slack Incorporated, 2010.

[2] J. M. Hausdorff, A. Lertratanakul, M. E. Cudkowicz, A. L. Peterson, D. Kaliton, and A. L. Goldberger, "Dynamic markers of altered gait rhythm in amyotrophic lateral sclerosis," *J. of Appl. physiology*, vol. 88, no. 6, pp. 2045–2053, June 2000.

[3] D. Hodgins, "The importance of measuring human gait," *Medical Device Technol.*, vol. 19, no. 5, pp. 42–44, September 2008.

[4] "Motion capture systems," http://www.vicon.com/, accessed 19- Jul-2016.

[5] S. Moon, Y. Park, D. W. Ko, and I. H. Suh, "Multiple Kinect sensor fusion for human skeleton tracking using Kalman filtering," *Int. J. of Adv. Robotic Syst.*, vol. 13, no. 2, p. 65, 2016.

[6] A. Fern'ndez-Baena, A. Susín, and X. Lligadas, "Biomechanical validation of upper-body and lower-body joint movements of Kinect motion capture data for rehabilitation treatments," in *Int. Conf. on Intell. Netw. and Collaborative Syst.*, Bucharest, Romania, September 2012, pp. 656–661.

[7] "Run3 - 3D gait analysis," http://3dgaitanalysis.com, accessed 19-Jul-2016.

[8] A. Pfister, A. M. West, S. Bronner, and J. A. Noah, "Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis," *J. of Med. Eng. & Technol.*, vol. 38, no. 5, pp. 274–280, July 2014.

[9] S. Lee and J. Hidler, "Biomechanics of overground vs. treadmill walking in healthy individuals," *J. of Appl. physiology*, vol. 104, no. 3, pp. 747–755, March 2008.

[10] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, pp. 376–380, April 1991.

[11] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-9, pp. 698–700, September 1987.

[12] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, "Real-time human pose recognition in parts from single depth images," in *Proc. IEEE Conf. Comp. Vis. Patt. Recogn.*, June 2011, pp. 1297–1304.

[13] M. A. Livingston, J. Sebastian, Z. Ai, and J. W. Decker, "Performance measurements for the microsoft Kinect skeleton," in *Proc. IEEE Virtual Relity Workshops*, Costa Mesa, CA, March 2012, pp. 119–120.

[14] P. K. Levangie and C. C. Norkin, *Joint Structure and Function: A Comprehensive Analysis*. FA Davis, 2011.

[15] C. Kirtley, *Clinical Gait Analysis: Theory and Practice*. Elsevier Health Sciences, 2006.

[16] R. Baker, "Temporal and spatial data, the gait cycle and gait graphs," University of Salford, Tech. Rep., 2012.

[17] A. Castelli, G. Paolini, A. Cereatti, and U. Della Croce, "A 2D markerless gait analysis methodology: Validation on healthy subjects," *Comput. and Math. Methods in Med.*, vol. 2015, p. 11, 2015.