The validity of a dual Azure Kinect-based motion capture system for gait analysis: a preliminary study

Yunru Ma[*], Bo Sheng[*], Rylea Hart[*] and Yanxin Zhang[*]
[*]The University of Auckland, Auckland, New Zealand
E-mail: yanxin.zhang@auckland.ac.nz

***Abstract*— The Microsoft Kinect can track human motions in various motor tasks. The recently released Azure Kinect is reported to have an improved image sensing technology. However, the validity of this newest sensor for gait analysis is still unknown. In this study, a dual Azure Kinect-based motion capture system was developed. Gait analysis was conducted with five healthy adults. Joint angles calculated by this system were compared with that acquired by the Vicon motion capture system. The coefficient of multiple correlations (CMC) and root mean square errors (RMSE) were computed. The dual Azure Kinect system could provide accurate knee angles (CMC=0.87±0.06, RMSE=11.9°±3.4°). Hip sagittal angles demonstrated moderate agreement with the reference (CMC=0.60±0.34, RMSE=15.1°±6.5°). The hip frontal, transversal, and ankle angles demonstrated poor validity. Although levels of accuracy for each joint varied, the dual Azure Kinect system demonstrated an overall improved validity than the Kinect V2. Future studies should involve more participants and patient populations, and compare different versions of sensors in the same experimental setup simultaneously to warrant the findings derived from this study. Furthermore, it is also necessary to standardize the experimental setup and involve more sensors to provide adequate depth images for analysis.**

***Keywords*—Azure Kinect, Gait Analysis, Validity, dual Kinect**

I. INTRODUCTION

Three-dimensional (3D) clinical gait analysis provides quantitative information that assists in treatment decision making and outcome assessment. However, its high financial expenditure and technical requirements limit its clinical application. Due to its marker-based tracking strategies, participants may feel observed and demonstrate over-performance rather that their natural daily gait [1]. The Kinect sensor and its SDK developed by Microsoft is capable of capturing human motion and providing corresponding 3D joint coordinates in real-time, without using any markers or handheld controllers. It presents the potential to be utilized as a cost-effective portable markerless motion capture tool for in-clinic and home-based gait observation.

Prior studies and two recently published reviews [2, 3] reported that the tracking accuracy of the older versions of Kinect (Kinect for Windows V1&2) was poor for some gait kinematic variables. Recently, the newest version of Azure Kinect DK has been released that promises improved image sensing technologies. It captures the human motion with two optional field-of-view modes, a higher resolution, a more compact appearance, and nearly half the weight of the previous version. The Azure Kinect Body Tracking SDK allows the sensor to provide position and orientation information of 32 joints that flow from the centre of the body to the extremities. Therefore, these improved features may improve its joint tracking accuracy and portability, potentially enhance its clinical feasibility for investigating and screening gait. It is logical that evaluate the validity of the Kinect Azure sensor before it could be utilized in clinical or home-based scenarios.

In a single Kinect motion capture system, occlusions could happen when the body segment obstruct each other. For patients with restricted mobility, walking aids or other assistive devices may be needed to help maintain stability, which could probably interfere with the tracking accuracy. Therefore, two or more Kinect sensors can be integrated to track human motion simultaneously to solve the full-body tracking problem [4-6].

Our previous work showed that a single Kinect V2 demonstrated an overall poor validity to investigate overground gait kinematics for children with cerebral palsy [7]. Therefore, a dual Azure Kinect motion capture system was developed in this study to provide adequate depth images for skeletal tracking. Currently, there is no data on the validity of the newest version of the Kinect. The aim of this pilot study was to evaluate the validity of this updated markerless motion capture system and discuss its feasibility in clinical application. Lower limb kinematic parameters of comfortable paced overground gait were assessed by this system and compared with the standard optoelectric Vicon motion capture system.

II. METHODOLOGY

A. Dual Azure Kinect Motion Capture System

This system consisted of two components: two Azure Kinect sensors that connected to one computer, and the iPi Mocap Studio software (iPi Soft, LLC, Moscow, Russia). The two sensors were positioned 4m from each other oppositely, elevated 1.1m from the ground, resulting in a distance of approximately 0.9m from the sensors to the edge of the proposed measurement volume (Fig. 1).

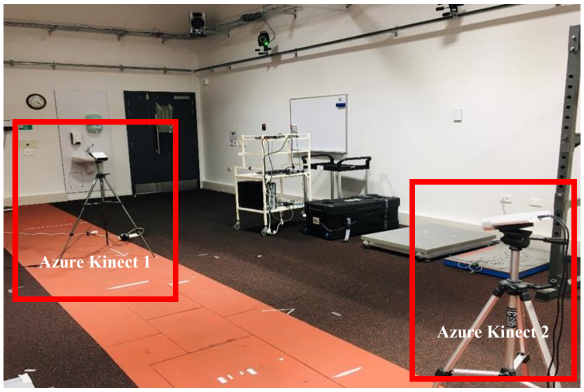

Fig. 1 Experimental setup for the dual Azure Kinect motion capture system

B. Calibration

Before the data collection, the background (without any moving objects) was evaluated. A dynamic calibration was conducted to determine the accurate positions and orientations of the two sensors. An operator used a flashlight whose head was unscrewed and slowly waved it throughout the entire measurement volume. This process started from the top, and the flashlight was moved in a descending spiral motion. It was accessible that the calibration quality was assessed as "Perfect" in the iPi Mocap Studio software (Fig. 2).

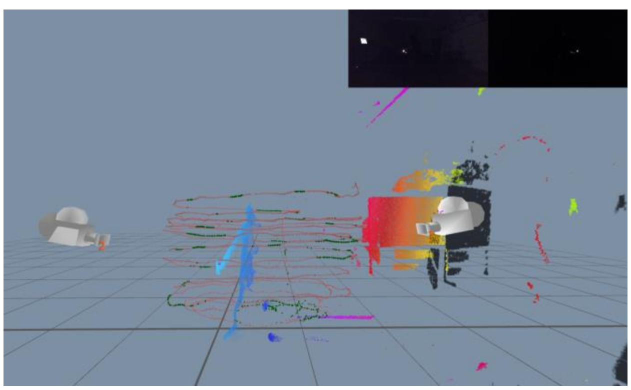

Fig. 2 The calibration for the dual Azure Kinect motion capture system.

C. Gait Analysis

Five injury-free adults (age: 29.8±5.8 years, height: 169.6±10.1 cm, mass: 72.9±15.4 kg, male: 2, female: 3) volunteered to participate.

Twenty-seven reflective markers were placed on each participant's sacrum, as well as the left and right anterior superior iliac spine (ASIS), thigh (cluster markers), lateral and medial condyles of the knee, shank (cluster markers), lateral and medial malleoli of the ankle, each calcaneus, and the second metatarsal head of both feet, according to a modified Cleveland marker set [8, 9]. Each participant performed a "T-pose" before the gait trial to fit the generic skeleton (Fig. 3). Then the participant started walking from one sensor to another with a comfortable speed. The eight-camera Vicon motion capture system (Oxford Metrics Group, Oxford, UK) captured reflective marker trajectories at a sampling rate of 100 Hz. The Azure Kinect sensors captured depth image data at a sampling rate of 30Hz and extracted the subjects' movement data from the background depth image information by using a 22-joint anthropometric model (Fig. 4 (a)). Three gait trials were acquired for each participant.

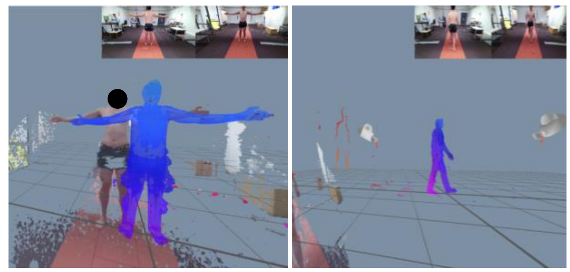

Fig. 3 "T" pose for skeleton fitting (left). Gait analysis using the dual Azure Kinect motion capture system (right).

D. Data analysis

For the Vicon motion capture system, reflective marker position data were filtered through a fourth-order Butterworth low-pass digital filter with a cut-off frequency of 6 Hz. The definition of the lower limb coordinate system and was established according to the ISB recommendations [10]. The hip, knee, and ankle angles were calculated as Euler angles that following the rotation sequence of flexion/extension, adduction/abduction, and internal/external rotation. All the data processing was conducted via Visual 3D Version 6 (C-Motion Inc., Germantown, MA, USA).

Anatomical landmarks collected by the dual Kinect Azure motion capture system were filtered through a fourth-order Butterworth low-pass filter with a cut-off frequency of 6 Hz. Pelvis and femur coordinate systems were established in a similar way as the ISB recommendations [7].

The formula below showed the standard for technical frame (TF) coordinate systems (Equation 1). The origin of the pelvis and femur TF was firstly determined. Line 1 was the vector represented the y-axis of each segment. The cross product of line 1 and line 2 decided the second axis, and then the third axis was defined as the cross product of the y-axis and the second axis according to the definition order. Marker names and the definition of the pelvis and femur coordinate system were presented in Table 1 and Fig. 4.

$$TF = [Segment\ origin, Line\ 1, Line\ 2, Definition\ order] \quad (1)$$

Equation 2 and 3 showed the ZXY order rotation matrix $R_{ZXY}(\alpha, \beta, \gamma)$ that rotated the TF_{pelvis} to the TF_{femur}.

$$TF_{femur} = R_{ZXY}(\alpha, \beta, \gamma) \times TF_{pelvis} \quad (2)$$

$$R_{ZXY}(\alpha,\beta,\gamma) = \begin{bmatrix} -s\alpha s\beta s\gamma + c\alpha c\gamma & -s\alpha c\beta & s\alpha s\beta c\gamma + c\alpha c\gamma \\ c\alpha s\beta s\gamma + s\alpha c\gamma & c\alpha c\beta & -c\alpha s\beta c\gamma + s\alpha s\gamma \\ -c\beta s\gamma & s\beta & c\beta c\gamma \end{bmatrix} \quad (3)$$

The α,β,γ angles could be computed through Equation 4 to 7, where α,β,γ represented the hip flexion/extension, adduction/abduction, internal/external rotation angle, respectively.

$$R_{ZXY}(\alpha,\beta,\gamma) = \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \quad (4)$$

$$\beta = \tan^{-1}\left(e_{32}/\sqrt{e_{12}{}^2 + e_{22}{}^2}\right) \quad (5)$$

$$\alpha = \tan^{-1}\left(-\frac{e_{12}}{e_{22}}\right) \quad (6)$$

$$\gamma = \tan^{-1}\left(-\frac{e_{31}}{e_{33}}\right) \quad (7)$$

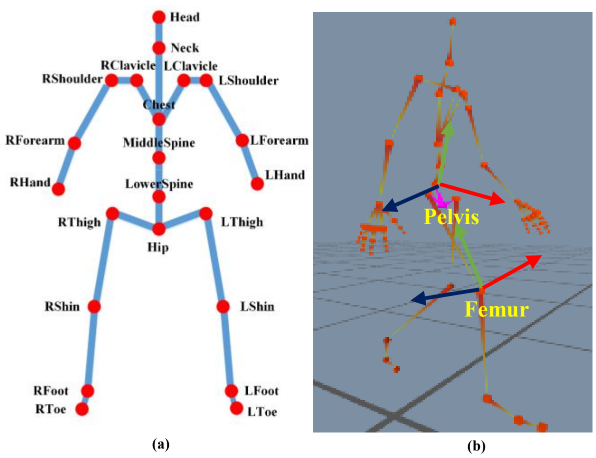

Fig. 4 Stick figure model obtained from the Azure Kinect in the iPi software (a). The definition of the pelvis and femur coordinate system (b).

The knee and ankle angle was defined as the angle between two segmental vectors; the formula was as follows:

$$\theta_{knee} = \cos^{-1}\left(\frac{\overrightarrow{v_{thigh}} \cdot \overrightarrow{v_{shank}}}{|\overrightarrow{v_{thigh}}||\overrightarrow{v_{shank}}|}\right) \quad (8)$$

$$\theta_{ankle} = \cos^{-1}\left(\frac{\overrightarrow{v_{foot}} \cdot \overrightarrow{v_{shank}}}{|\overrightarrow{v_{foot}}||\overrightarrow{v_{shank}}|}\right) - 90° \quad (9)$$

where θ_{knee} and θ_{ankle} was knee and ankle joint angle, $\overrightarrow{v_{thigh}}$ was the vector pointing from "RShin" to "RThigh", $\overrightarrow{v_{shank}}$ was the vector pointing from "RFoot" to "RShin", and $\overrightarrow{v_{foot}}$ was the vector pointing from "RFoot" to "RToe".

All the data were processed by a customized program in Matlab R2018b (MathWorks Inc., Natick, MA).

E. Statistics

Joint angles for the right limb of each gait cycle were normalized to 101-time steps. The coefficient of multiple correlations (CMC) was computed for each subject following Kadaba's approach [11]. The CMC values could be explained as: excellent similarity (0.95-1); very good similarity (0.85-0.94); good similarity (0.75-0.84); moderate similarity (0.6-0.74) and poor similarity (0-0.59) [12]. The root mean square error (RMSE) was calculated to compare the differences between the two devices over a gait cycle. All the statistical analysis was processed by a customized program in Matlab R2018b (MathWorks Inc., Natick, MA).

III. Results

For the integrated joint angle profiles, ensemble curve analyses for the hip, knee, and ankle angles during a gait cycle is shown in Fig. 5 and Table 2. As a comparison, the results extracted from our previous single Kinect V2 study were presented in Table 2. The results showed that only the knee flexion/extension angles between two systems showed very good similarity (CMC=0.87±0.06, RMSE=11.9°±3.4°). The agreement of hip moderate (CMC=0.60±0.34, RMSE=7.2°±4.7°). The hip frontal angle and ankle dorsi/plantar flexion angle showed poor similarity between two systems (CMC=0.48±0.45 and 0.55±0.09, RMSE=7.2°±4.7° and 11.6°±2.4°). The hip internal/external rotation angle calculated by the dual Azure Kinect motion capture system showed the worst validity with CMC <0.001 and RMSE of 32.3°±22.2°.

Table 1 Definition of pelvis and femur coordinate system.

Segment	Segment origin	Line 1	Line 2	Definition order
Pelvis	"LowerSpine"	The vector from "LowerSpine" to "MiddleSpine".	The vector from "RThigh" to "LThigh".	YXZ
Right Femur	"RShin/LShin"	The vector from "RShin/LShin" to "RThigh/LThigh".	The vector from "RFoot/LFoot" to "RShin/LShin".	YZX

Table 2 CMC and RMSE (±SD) between the joint angle profiles using the dual Azure Kinect and Vicon (n=5). Comparative results between the single Kinect V2 and standard optoelectric motion capture system are extracted from previous work [7].

Joint angle	CMC		RMSE (°)	
	Dual Azure Kinect	Kinect V2	Dual Azure Kinect	Kinect V2
Hip flexion/extension	0.60±0.34	0.45±0.36	15.1±6.5	20.7±8.8
Hip adduction/abduction	0.48±0.45	<0.001	7.2±4.7	12.5±3.4
Hip internal/external rotation	<0.001	<0.001	32.3±22.2	40.2±22.6
Knee flexion/extension	0.87±0.06	0.70±0.12	11.9±3.4	16.7±4.2
Ankle dorsi/plantar flexion	0.55±0.09	<0.001	11.6±2.4	23.0±5.0

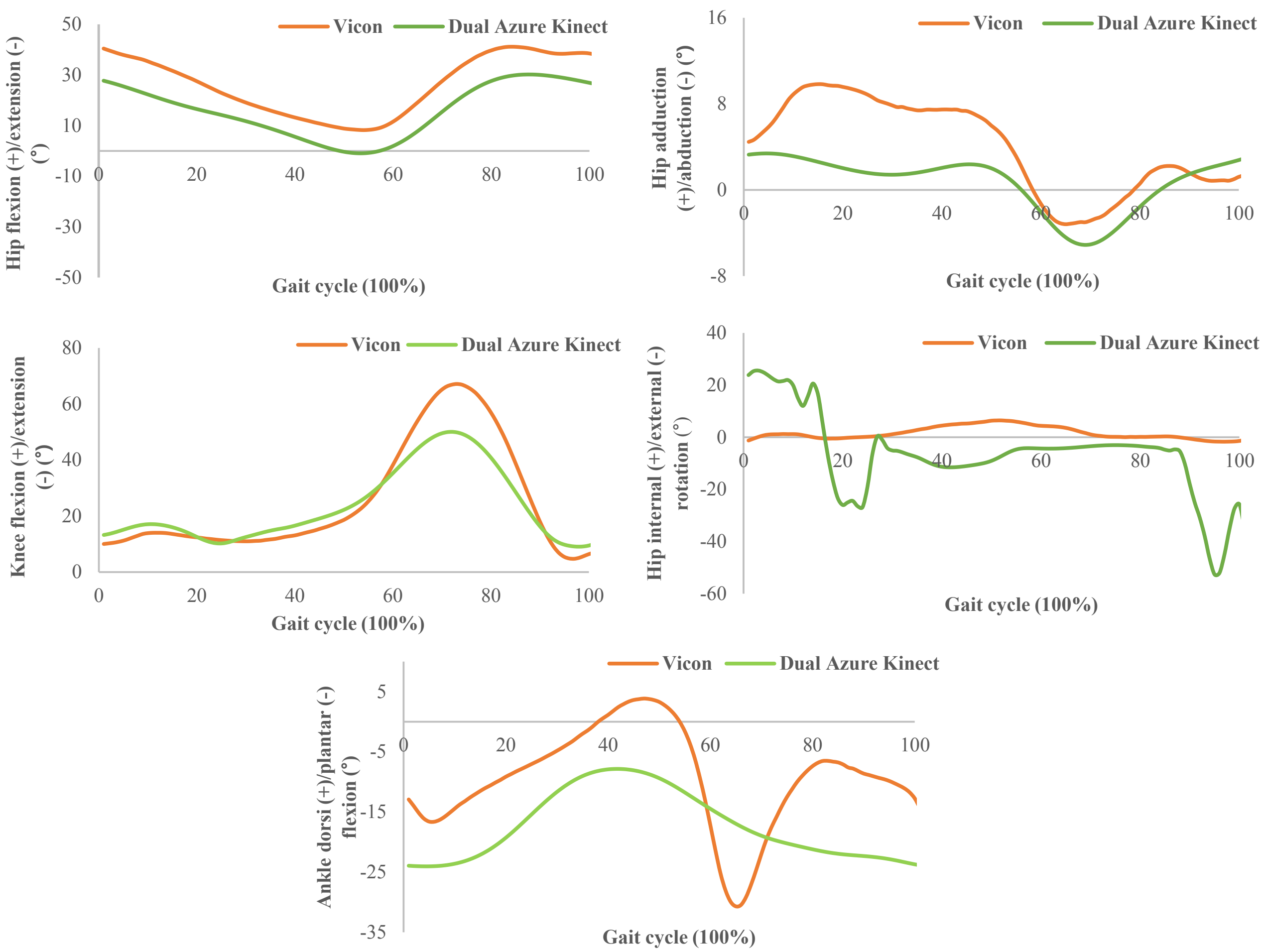

Fig. 5 The average hip, knee, and ankle angle profiles over a gait cycle across all gait trials (n=5).

IV. Discussions

The results showed that the validity of a dual Kinect Azure system could provide relatively accurate knee angles in an overground gait when compared with a standard 3D gait analysis system. Various accuracy levels for different lower limb kinematic parameters were found. Moreover, measurement errors at each degree of freedom were smaller than the single Kinect V2 system.

A. *Hip flexion/extension angles*

The definition for the pelvis coordinate system was different from what we previously developed in a Kinect V2 system [7] due to the different pelvis skeleton structures between the two versions of the Kinect SDK. The “Hip” landmark in Kinect Azure was located at the anterior inferior of the two thigh landmarks. Additionally, with a “LowerSpine” that situated posterior superior of the two thigh landmarks, the pelvis defined by those landmarks showed a posterior tilt. If the pelvis y-axis is defined by the vector pointing from the “Hip” to the “LowerSpine”, the hip angle calculated by the Kinect Azure would be hyperextended. Although the measurement errors exceed the clinically acceptable range [13], it could be observed from Fig. 5 that hip angles calculated from Kinect followed the trend of its referential counterpart, which was consistent with previous findings [3, 14, 15]. The measurement errors could be reduced by performing parallel translation [14] or linear regression [7]. Therefore, the dual Azure Kinect demonstrates the potential of providing promising hip sagittal kinematic assessment.

B. Hip adduction/abduction angles

The validity of hip adduction/abduction angles varied among subjects in this study. Moreover, the measurement errors seemed to be less than that observed in a single Kinect V2 system. In a dual Kinect motion capture system, subjects' skeletons were registered to fit a skeleton predefined by the iPi software. If the skeleton were not well-matched, further tracking errors would happen. The dimensions of the segment also differed among individuals, which increased the risk of inaccurate landmark trajectories. The markerset utilized in this study contained “T-shape” marker clusters that attached to the lateral of lower limbs, which could also cause interference for the configuration of the depth images.

C. Hip internal/external angles

The hip internal/external rotation angle was sensitive to the relative position of anatomical markers, even in a marker-based motion capture system [13]. From Figure 5, it was found that the abnormal spikes appeared in the early stance and late swing phase, which might mainly result in the low agreement and measurement errors. Except for the inaccurate pelvic structure mentioned before, the small knee angles happed in the two gait phases implied a nearly collinear relationship among the three markers that defined the femur coordinate system. Due to the insufficient markers provided in the Kinect skeleton, it was a concession that added the foot marker to the femur coordinate system, which may help explain these measurement errors.

D. Knee flexion/extension angles

It is generally found that the Kinect is an effective tool to assess sagittal knee kinematics [14, 16-18]. Therefore, it was not surprising to find that sagittal knee angles always showed a good agreement with the referential counterpart in various motor tasks. From Figure 5, knee joint angles generated by two motion capture systems demonstrate a good consistency, especially in the stance phase. Moreover, the dual Azure Kinect system demonstrated a better agreement with the referential optoelectric system and smaller measurement errors when compared with the Kinect V2, indicating an improved accuracy in assessing knee kinematics. Ankle dorsi/plantarflexion angles

Foot tracking accuracy is a common problem among studies [16, 19, 20]. Although the agreement between two motion capture systems was poor in this study, the RMSE was less than that reported in our previous single Kinect V2-based study. It could be found that the Kinect was limited to track the ankle dorsiflexion, which was consistent with the previous study [7, 16]. In this study, the participants performed barefoot walking on a carnation track, and the low contrast between human skin and the track could make it difficult to distinguish the foot contact.

V. LIMITATIONS

Except for the limited sample size and age group, several limitations should be addressed. First, marker clusters were attached to participants' lower limbs in this study. It is found that the reflective markers may have a disruptive effect on the tracking algorithm of the Kinect. Therefore, a more streamlined markerset could be applied to reduce misreading [21]. Secondly, this study did not compare the Azure Kinect and Kinect V2 simultaneously. The data about Kinect V2 were extracted from our early work. One study has compared the performance of Kinect V1 and V2 in the same experimental setup [22]. Future studies should compare the newest sensor with previous versions to study its accuracy in a possible clinical or home-based scenario.

VI. CONCLUSIONS

The dual Kinect Azure motion capture system showed an overall improved accuracy. It should be noted that the updated skeleton landmarks might cause misleading when interpreted the hip sagittal kinematics. Two recent studies applied skeleton information captured by the Kinect to create a musculoskeletal modelling workflow for further biomechanical simulation [4, 23], which showed the potential of the multiple Kinect motion capture system to be utilized as a cost-effective option for movement analysis in future. Therefore, future studies should standardize the experimental setup, and more sensors could be included to provide adequate depth images for analysis. More participants should be recruited to warrant the findings derived from this study.

REFERENCES

[1] J. C. van den Noort, A. Ferrari, A. G. Cutti, J. G. Becher, and J. Harlaar, "Gait analysis in children with cerebral palsy via inertial and magnetic sensors," *Medical & biological engineering & computing,* vol. 51, no. 4, pp. 377-386, 2013.

[2] R. A. Clark, B. F. Mentiplay, E. Hough, and Y. H. Pua, "Three-dimensional cameras and skeleton pose tracking for physical function assessment: a review of uses, validity, current developments and Kinect alternatives," *Gait & posture,* vol. 68, pp. 193-200, 2019.

[3] S. Springer and G. Yogev Seligmann, "Validity of the kinect for gait assessment: A focused review," *Sensors,* vol. 16, no. 2, p. 194, 2016.

[4] S. Skals, K. P. Rasmussen, K. M. Bendtsen, J. Yang, and M. S. Andersen, "A musculoskeletal model driven by dual Microsoft Kinect Sensor data," *Multibody System Dynamics,* vol. 41, no. 4, pp. 297-316, 2017.

[5] A. Kotsifaki, R. Whiteley, and C. Hansen, "Dual Kinect v2 system can capture lower limb kinematics reasonably well in a

clinical setting: concurrent validity of a dual camera markerless motion capture system in professional football players," *BMJ open sport & exercise medicine,* vol. 4, no. 1, p. e000441, 2018.

[6] D. J. Geerse, B. H. Coolen, and M. Roerdink, "Kinematic validation of a multi-Kinect v2 instrumented 10-meter walkway for quantitative gait assessments," *PloS one,* vol. 10, no. 10, p. e0139913, 2015.

[7] Y. Ma, K. Mithraratne, N. C. Wilson, X. Wang, Y. Ma, and Y. Zhang, "The Validity and Reliability of a Kinect v2-Based Gait Analysis System for Children with Cerebral Palsy," *Sensors,* vol. 19, no. 7, p. 1660, 2019.

[8] D. H. Sutherland, "The evolution of clinical gait analysis: Part II Kinematics," *Gait & posture,* vol. 16, no. 2, pp. 159-179, 2002.

[9] A. H. Mackey, S. E. Walt, G. A. Lobb, and N. S. Stott, "Reliability of upper and lower limb three-dimensional kinematics in children with hemiplegia," *Gait & posture,* vol. 22, no. 1, pp. 1-9, 2005.

[10] G. Wu *et al.*, "ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion—part I: ankle, hip, and spine," *Journal of biomechanics,* vol. 35, no. 4, pp. 543-548, 2002.

[11] M. Kadaba, H. Ramakrishnan, M. Wootten, J. Gainey, G. Gorton, and G. Cochran, "Repeatability of kinematic, kinetic, and electromyographic data in normal adult gait," *Journal of Orthopaedic Research,* vol. 7, no. 6, pp. 849-860, 1989.

[12] P. Garofalo *et al.*, "Inter-operator reliability and prediction bands of a novel protocol to measure the coordinated movements of shoulder-girdle and humerus in clinical settings," *Medical & biological engineering & computing,* vol. 47, no. 5, pp. 475-486, 2009.

[13] J. L. McGinley, R. Baker, R. Wolfe, and M. E. Morris, "The reliability of three-dimensional kinematic gait measurements: a systematic review," *Gait & posture,* vol. 29, no. 3, pp. 360-369, 2009.

[14] R. Tanaka, H. Takimoto, T. Yamasaki, and A. Higashi, "Validity of time series kinematical data as measured by a markerless motion capture system on a flatland for gait assessment," *Journal of biomechanics,* vol. 71, pp. 281-285, 2018.

[15] X. Xu, R. W. McGorry, L.-S. Chou, J.-h. Lin, and C.-c. Chang, "Accuracy of the Microsoft Kinect™ for measuring gait parameters during treadmill walking," *Gait & posture,* vol. 42, no. 2, pp. 145-151, 2015.

[16] M. Eltoukhy, J. Oh, C. Kuenze, and J. Signorile, "Improved kinect-based spatiotemporal and kinematic treadmill gait assessment," *Gait & posture,* vol. 51, pp. 77-83, 2017.

[17] M. Wochatz *et al.*, "Reliability and validity of the Kinect V2 for the assessment of lower extremity rehabilitation exercises," *Gait & Posture,* vol. 70, pp. 330-335, 2019.

[18] T. M. Guess, S. Razu, A. Jahandar, M. Skubic, and Z. Huo, "Comparison of 3D joint angles measured with the kinect 2.0 skeletal tracker versus a marker-based motion capture system," *Journal of applied biomechanics,* vol. 33, no. 2, pp. 176-181, 2017.

[19] B. F. Mentiplay *et al.*, "Gait assessment using the Microsoft Xbox One Kinect: Concurrent validity and inter-day reliability of spatiotemporal and kinematic variables," *Journal of biomechanics,* vol. 48, no. 10, pp. 2166-2170, 2015.

[20] J. Oh, C. Kuenze, M. Jacopetti, J. F. Signorile, and M. Eltoukhy, "Validity of the Microsoft Kinect™ in assessing spatiotemporal and lower extremity kinematics during stair ascent and descent in healthy young individuals," *Medical engineering & physics,* vol. 60, pp. 70-76, 2018.

[21] M. Naeemabadi, B. Dinesen, O. K. Andersen, and J. Hansen, "Influence of a marker-based motion capture system on the performance of Microsoft Kinect v2 skeleton algorithm," *IEEE Sensors Journal,* vol. 19, no. 1, pp. 171-179, 2018.

[22] M. do Carmo Vilas-Boas, H. M. P. Choupina, A. P. Rocha, J. M. Fernandes, and J. P. S. Cunha, "Full-body motion assessment: Concurrent validation of two body tracking depth sensors versus a gold standard system during gait," *Journal of biomechanics,* vol. 87, pp. 189-196, 2019.

[23] J. Oh, M. Eltoukhy, C. Kuenze, M. S. Andersen, and J. F. Signorile, "Comparison of predicted kinetic variables between Parkinson’s disease patients and healthy age-matched control using a depth sensor-driven full-body musculoskeletal model," *Gait & Posture,* vol. 76, pp. 151-156, 2020.