Medical Engineering and Physics 96 (2021) 22–28

Contents lists available at ScienceDirect

Medical Engineering and Physics

journal homepage: www.elsevier.com/locate/medengphy

Original Article

Concurrent validity of a custom computer vision algorithm for measuring lumbar spine motion from RGB-D camera depth data

Wantuir C. Ramos Jr. [a], Kristen H.E. Beange [b,c], Ryan B. Graham [a,c,*]

[a] School of Human Kinetics, Faculty of Health Sciences, University of Ottawa, 200 Lees Avenue, Ottawa, ON K1N 6N5, Canada
[b] Department of Systems and Computer Engineering, Faculty of Engineering and Design, Carleton University, 1125 Colonel By Drive, Ottawa, ON K1S 5B6, Canada
[c] Ottawa-Carleton Institute for Biomedical Engineering, Ottawa, ON, Canada

ARTICLE INFO

Keywords:
RGB-D cameras
Computer vision
Depth camera
Low back pain
Movement quality

ABSTRACT

Using RGB-D cameras as an alternative motion capture device can be advantageous for biomechanical spine motion assessments of movement quality and dysfunction due to their lower cost and complexity. In this study, we evaluated RGB-D camera performance relative to gold-standard optoelectronic motion capture equipment. Twelve healthy young adults (6M, 6F) were recruited to perform repetitive spine flexion-extension, while wearing infrared reflective marker clusters placed over their T_{10}-T_{12} spinous processes and sacrum, and motion capture data were recorded simultaneously by both systems. Custom computer vision algorithms were developed to extract spine angles from depth data. Root mean square error *(RMSE)* was calculated for continuous Euler angles, and intraclass correlation coefficients *($ICC_{2,1}$)* were calculated between minimum and maximum angles and range of motion in all movement planes. *RMSE* was low *($RMSE \leq 2.05°$)* and reliability was good to excellent *($0.849 \leq ICC_{2,1} \leq 0.979$)* across all movement planes. In conclusion, the proposed algorithm for tracking 3D lumbar spine motion during a sagittal movement task from one RGB-D camera is reliable in comparison to gold-standard motion tracking equipment. Future research will investigate accuracy and validity in a wider variety of movements, and will also investigate the development of novel methods to measure spine motion without using infrared reflective markers.

1. Introduction

Low back pain is one of the most prevalent and costly musculoskeletal disorders in the world, affecting up to 80% of people at some point in their lifetime. While a small number of cases have a definitive diagnosis and associated treatment, the vast majority of cases are classified as non-specific, meaning there is no known pathoanatomical cause [1]. For this group of patients, researchers and clinicians typically assess spine movement quality during a set of movement dysfunction tests with the goal of classifying normal and abnormal (i.e., dysfunctional or harmful) movements [2-3]. Clinical assessments often involve subjective visual appraisal of movement quality performed by a trained clinician; however, there is poor inter- and intra-rater reliability [3-6]. Conversely, researchers in laboratory settings typically utilize optoelectronic motion capture to obtain objective measurements of spine movement quality; however, these systems are costly and technologically challenging, which limits their feasibility for use in clinics.

To overcome these challenges and improve clinical spine movement assessment, researchers have investigated the feasibility of alternative devices to objectively measure spine motion in clinics. For example, wearable inertial measurement units (IMUs) have been validated for tracking trunk motion and assessing spine neuromuscular control during repetitive trunk flexion-extension (FE) tasks [7], as well as for assessments of common clinical movement dysfunction tests [8]. It was established in these studies that spine range of motion (ROM) measurements were within or slightly above the clinical benchmark for accuracy of human movement tracking (i.e., $< 2°$ error, or between $2°$ and $5°$ with additional subjective interpretation) [9], and that they were reliable for measuring typical movement dysfunction parameters [7-8]. While IMU performance is regarded as clinically acceptable for motion tracking, the sensors are subject to magnetic disturbance, and sensor location is also difficult to obtain (i.e., an IMU is only able to track *orientation* relative to itself, or another IMU). Many researchers attempt to track IMU location in global space (e.g., [10-11]); however, these procedures are computationally expensive and advanced, and are dependent on careful IMU-to-segment placement and calibration

* Corresponding author at: University of Ottawa, Faculty of Health Sciences, Ottawa, ON K1N 6N5, Canada
E-mail address: rgraham@uottawa.ca (R.B. Graham).

https://doi.org/10.1016/j.medengphy.2021.08.005
Received 11 February 2021; Received in revised form 10 August 2021; Accepted 12 August 2021
Available online 14 August 2021

techniques. Some commercially available devices are capable of providing estimates of lumbar spine ROM and movement quality in clinical settings (e.g., DorsaVi Professional Suite; DorsaVi, Victoria, AUS); however, these systems come with proprietary motion-tracking algorithms, which are prohibitive for conducting custom assessments of spine movement quality and dysfunction. As an alternative, researchers are exploring the use of depth sensors for tracking people's movement, as these sensors are able to provide position and orientation data, and it is possible to customize methods of movement assessment from raw depth data.

Red, green, blue, depth (RGB-D) cameras are gaining popularity as an alternative method for motion tracking in clinical and research settings. Many of these cameras use their depth data stream and native machine learning algorithms to locate pre-defined joint centres for tracking full-body motion (aka the built-in "skeletal model"). RGB-D cameras are less expensive and require less time than traditional optoelectronic motion capture equipment for tracking human motion, and they can be easily installed within clinical settings. The performance and implementation of these cameras have been studied for the assessment of various experimental and clinical outcome measures relative to gold-standard motion tracking equipment, such as: 1) clinical measurements of motor function and postural control [12-14]; 2) spatiotemporal gait parameters during overground and treadmill walking [14-15]; and 3) spatial accuracy (i.e., location and orientation) of landmarks during various movement tasks [12,16-17].

The majority of studies that utilize RGB-D camera systems for tracking human motion achieve adequate results using the integrated skeletal model [12,15,18]. Specifically, Otte et al. [12] found excellent reliability while evaluating the spatial accuracy of RGB-D cameras for clinical assessments of whole body motion, obtaining nearly perfect correlation ($r > 0.99$) with optoelectronic motion capture when tracking "spine base" (approximately L_5/S_1), "spine mid" (approximately T_{12}/L_1), and "spine shoulder" (approximately T_4) landmark locations (as defined by the native skeletal model) in the sagittal plane, strong correlation in the frontal plane ($r \geq 0.85$), and moderate correlation in the transverse plane ($0.46 \leq r \leq 0.70$). Clark et al. [18] compared measurements of spinal posture between the skeletal data relative to optoelectronic motion capture equipment; pairwise comparison of *ICCs* showed no significant differences between devices during single leg balance and both lateral and forward reach, and the correlation between systems was very strong across all movement axes ($0.93 \leq r \leq 0.99$). Lastly, Dolatabadi and colleagues [15] evaluated the performance of the native skeletal model in estimating spatiotemporal gait parameters (e.g., stance time, step time, step length, and velocity) relative to optoelectronic motion capture, and achieved moderate to excellent reliability between systems ($0.73 \leq ICC_{2,1} \leq 0.98$). Researchers in these studies recorded data using a frontal view of the participants; however, for evaluating lumbar spine movement tasks, the camera must be positioned in a posterior-anterior view (i.e., behind the participant) to capture the anatomical landmarks of interest, which restricts the ability to use the built-in skeletal model. Thus, development of custom motion tracking algorithms from raw depth data are necessary for clinical assessment of lumbar spine motion. Prior to clinical integration, these algorithms must be validated relative to gold-standard motion tracking equipment to meet the clinical benchmark for human motion tracking.

In a meta-analysis conducted by Cuesta-Vargas and colleagues [19] concerning the use of IMUs as an alternative for human motion tracking, it was concluded that IMUs should be considered only as a tool, and that their validity for tracking human movement is task- and site-specific, and also heavily dependent on the specific pre-/post-processing algorithms used to assess motion from raw sensor data [19]. Applying the same premise to RGB-D cameras: they are tools for tracking motion, and, therefore, it is necessary to evaluate their performance according to the same standard.

Some researchers have developed custom methods of tracking human motion from raw RGB-D camera depth data and achieved adequate results. Macpherson and colleagues [16] positioned the RGB-D camera behind participants and recorded depth data in order to measure pelvis and trunk kinematics during treadmill locomotion, finding weak to very strong correlations ($0.3 \leq r \leq 0.9$) during walking and running. Similarly, Cippitelli and colleagues [17] developed a method using the RGB-D camera's depth sensor to measure changes in full-body positioning using sagittal plane views, and achieved superior performance of their method over the built-in skeletal model when compared to gold-standard optoelectronic motion capture [17]. Lastly, Auvinet et al. [14] developed a custom framework to extract lower extremity kinematics from the RGB-D camera's depth data (using both anterior and posterior view points) to measure intersegmental coordination and gait asymmetry during treadmill walking. Results for this study showed excellent correlation between the RGB-D camera's depth data and gold-standard optoelectronic motion capture equipment when measuring gait asymmetry from both view points; however, intersegmental coordination computed with the RGB-D camera's skeletal data was not reliable [14].

Of the aforementioned studies, two looked at trunk kinematics using the built-in skeletal model (i.e., [12-13]), and one investigated lower back kinematics using custom motion tracking algorithms (i.e., [16]). To the best of the authors' knowledge, there are no studies validating RGB-D camera depth data against an optoelectronic motion capture system when comparing lumbar spine motion during a dynamic trunk FE task. Therefore, the goal of this study was to develop and validate a computer vision method to track lumbar spine motion from an RGB-D camera's depth data using infrared (IR) reflective markers during a repetitive FE task relative to gold-standard optoelectronic motion capture equipment. Based on results from previous studies, as well as previous work that evaluated the performance of wearable IMUs for spine motion tracking using similar methods, it was expected that the RGB-D camera would achieve a similar level of performance (i.e., low root mean square error (*RMSE*) for continuous angles and excellent reliability when measuring minimum (min) and maximum (max) angles and ROM). *RMSE* values below 2° are considered low and clinically acceptable, and between 2-5° are clinically acceptable with additional subjective interpretation [9]; $ICC_{2,1}$ values of below 0.5, 0.5-0.75, 0.75-0.9, and above 0.9 correspond to poor, moderate, good, and excellent reliability, respectively [20].

2. Methods

2.1. Participants

Twelve healthy adults (6M/6F) were recruited from the general university population via poster-advertising and word of mouth. Participant demographics are reported in Table 1. Participants who were unable to complete the trunk movement tasks without experiencing pain or fatigue, and participants who had suffered from a low back injury within six months prior to data collection were excluded from this study. Informed consent was obtained from all participants prior to experimentation.

2.2. Equipment and experimental setup

Passive IR reflective marker clusters (12.7 mm, B & L Engineering, USA) were placed superficial to the T_{10}-T_{12} vertebrae (i.e., "trunk") and

Table 1
Mean (SD) Participant Demographics.

	Height (cm)	Weight (kg)	Age (years)	BMI (kg/m^2)
Male	176.2 (5.2)	77.4 (2.6)	26.7 (6.0)	25.0 (1.6)
Female	163.4 (6.1)	55.4 (11.1)	21.3 (2.5)	20.6 (2.9)
All	169.8 (8.6)	66.4 (13.9)	24.0 (5.2)	22.8 (3.2)

SD = standard deviation; BMI = body mass index.

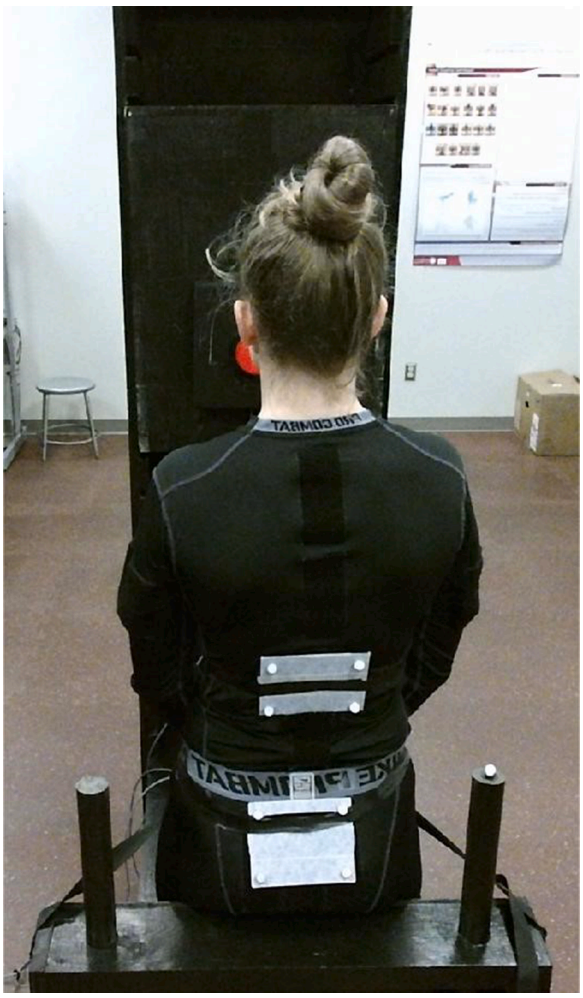

Fig. 1. Marker placement and experimental setup.

the sacrum (i.e., "pelvis"; Fig. 1). During the data collection session, kinematic data were collected simultaneously from one RGB-D camera (Kinect v2, Microsoft Corporation, Redmond, WA, USA) at 30 Hz and a 10-camera motion capture system (Vantage V5, Vicon, UK) at 120 Hz. The RGB-D camera was positioned 1.0 m behind the participant at a height of 1.70 m, with the camera pointing 30° downward from the horizontal towards the participant's back (Fig. 2). This camera position allowed adequate visualization of the participant's back during execution of the movement task, avoiding camera occlusions and ensuring the detection of their full ROM. Prior to collection, the RGB-D camera was pre-heated for over 25 minutes in order to avoid temperature-induced fluctuations in depth accuracy [21]. Depth images were calibrated as per the intrinsic camera parameters provided by the manufacturer. The RGB-D camera's depth data were captured using the "Kinect Image Acquisition Toolbox" from MATLAB (R2018b, The MathWorks Inc., USA). To synchronize both the RGB-D camera and the optoelectronic motion capture systems, a common temporal event was created by tapping a passive IR reflective marker placed behind the participant causing it to accelerate and signal start time. The starting frame for each system was determined by selecting the first frame where movement was detected (i.e., when the IR reflective marker leaves the stationary position; indicated by a sudden, drastic change in the position-time graphs).

2.3. Movement protocol

Participants performed 35 cycles of trunk FE at a rate of 30 beats/minute controlled by a metronome (i.e., 15 cycles/minute) while constrained at the hip; this was done to ensure the measured ROM represents trunk FE movement exclusively, rather than combined ROM from the lower limbs, pelvis, and trunk. To control for ROM, participants touched targets placed at specific locations with arms outstretched in front of them; one target was located 50 cm anteriorly at knee height, oriented parallel to the floor, and the other at shoulder height and at arm's length away, oriented perpendicular to the floor [7,22]. One cycle was defined as the participant moving from standing position (touching shoulder height target), into forward flexion (touching knee height target), and back to standing (touching shoulder height target; Fig. 2). All aspects of this protocol were approved by the University of Ottawa Research Ethics Board (H02-17-11; H08-17-26).

2.4. Data processing and analysis

To compare spine kinematics between systems, each system underwent identical post-processing for direct comparison (as outlined in the following subsections). For the RGB-D camera's depth data, an additional step was required that involved a series of image processing algorithms to track the 3D location of IR reflective markers from depth data. Additionally, optoelectronic motion capture system data were down-sampled to ensure the trials for each system were synchronous and equal-length. Lastly, data from both systems were filtered using a 2nd order dual-pass low-pass Butterworth filter, with a cut-off frequency of 2 Hz.

2.4.1. Depth image computer vision algorithm

Several steps were involved in the custom-developed computer vision algorithm; these included: 1) manually selecting a region of interest from the depth images where IR reflective markers would be likely

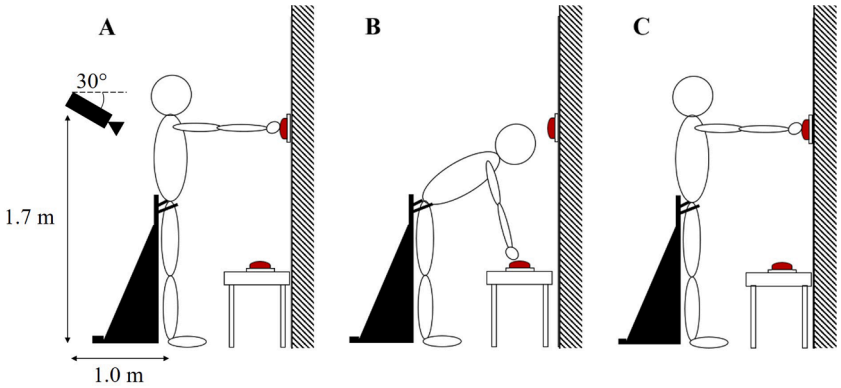

Fig. 2. Experimental setup and task protocol. (A) Participants begin in upright standing position with arms outstretched in front of them; then, (B) participants move into full flexion; and finally, (C) participants return to upright standing position for the completion of one full cycle.

to appear throughout the video; 2) applying a logical operator on each image to find pixels that had value equal to zero (indicating reflection of IR light and therefore the presence of markers); 3) filtering out pixel regions that do not match the IR reflective markers' size and shape (i.e., 12.7 mm diameter sphere) and are therefore not IR reflective markers, based on their morphological properties (i.e., major axis length and area; this was done using the *"regionprops"* function in MATLAB); 4) locating the centroid position *(i,j)* of each retained pixel region on the image plane and defining as *(x,y)* global coordinates for each IR reflective marker; and finally 5) calculating marker *(z)* global coordinates based on the average of the pixel values in the area surrounding each pixel region (i.e., a 1-pixel-thick perimeter around each IR reflective marker; this indirect method was required as the pixel value for *x* and *y* coordinates was zero). This process produced 3D global coordinates *(x,y,z)* for each IR reflective marker.

2.4.2. Kinematics

Both systems underwent identical calculation of lumbar spine kinematic variables for direct comparison. First, individual local coordinate systems were defined using trunk and pelvis marker clusters [23]; individual markers were labelled: T_{Rupper}, T_{Rlower}, T_{Lupper}, T_{Llower}, P_{Rupper}, P_{Rlower}, P_{Lupper}, and P_{Llower} (Fig. 3; where T and P refer to trunk and pelvis marker clusters, respectively, and subscripts R and L represent right and left markers, respectively, when facing the participants' back). The origin of the trunk cluster was defined at T_{Rlower}. To create the mediolateral x-component of the trunk, a unit vector $\hat{\imath}$ was defined using Eq. 1. Next, an auxiliary vector $\hat{v}$ was created as a temporary y-component unit vector (i.e., inferosuperior direction; Eq. 2). This vector was used to create the z-component unit vector $\hat{k}$ (i.e., anteroposterior direction), which was defined as the cross product between $\hat{\imath}$ and $\hat{v}$ (Eq. 3). Finally, the y-component unit vector $\hat{\jmath}$, was adjusted by taking the cross product between $\hat{k}$ and $\hat{\imath}$ (Eq. 4).

$$\hat{i} = \frac{T_{Llower} - T_{Rlower}}{||T_{Llower} - T_{Rlower}||} \tag{1}$$

$$\hat{v} = \frac{T_{Rupper} - T_{Rlower}}{||T_{Rupper} - T_{Rlower}||} \tag{2}$$

$$\hat{k} = \hat{i} \times \hat{v} \tag{3}$$

$$\hat{j} = \hat{k} \times \hat{i} \tag{4}$$

The resulting right-handed rotation matrix (Eq. 5) describes the orientation of the rigid body in 3D global space:

$$R_{Trunk} = \begin{bmatrix} \hat{i}_x & \hat{i}_y & \hat{i}_z \\ \hat{j}_x & \hat{j}_y & \hat{j}_z \\ \hat{k}_x & \hat{k}_y & \hat{k}_z \end{bmatrix} \tag{5}$$

This process was repeated for the pelvis cluster (i.e., the origin was defined as P_{Rlower}, and the rotation matrix R_{Pelvis} was obtained identically). A rotation matrix describing relative motion between trunk and pelvis marker clusters was calculated, and Euler angles were extracted using an X-Z-Y rotation sequence (corresponding to FE-lateral bend (LB)-axial twist (AT) axes).

2.4.3. Range of motion

To obtain lumbar spine ROM (i.e., trunk relative to pelvis), the following steps were taken: 1) the FE, LB, and AT time-series were divided into individual movement cycles by automatically identifying peak angles (corresponding to upright standing) with a mean peak prominence of $> 15°$ in the FE plane (so as not to capture local maxima that occur as a result of neuromuscular tremor) [7]; then, 2) each individual cycle was normalized to 101 data points (i.e., 0-100% of the cycle); next, 3) the mean ensemble curve (representing the average movement across all cycles) was obtained by calculating the mean angle across all cycles, at each percentage of the movement; and finally 4) FE, LB, and AT ROM were calculated according to Eq. 6 - 8.

$$ROM_{FE} = \left|\theta^{FE}_{max} - \theta^{FE}_{min}\right| \tag{6}$$

$$ROM_{LB} = \left|\theta^{LB}_{max} - \theta^{LB}_{min}\right| \tag{7}$$

$$ROM_{AT} = \left|\theta^{AT}_{max} - \theta^{AT}_{min}\right| \tag{8}$$

2.5. Statistical analysis

To quantify the magnitude of error of the RGB-D camera relative to optoelectronic motion capture system, *RMSE* was calculated across the entire trial; two-way random intraclass correlation coefficient's ($ICC_{2,1}$) were used to assess reliability between the min and max angles and ROM, and Bland-Altman plots were constructed to assess absolute agreement between ROM measurements obtained by each system. All statistical analyses were performed using SPSS statistical software for windows (SPSS 23, IBM Corporation, USA) and custom MATLAB scripts.

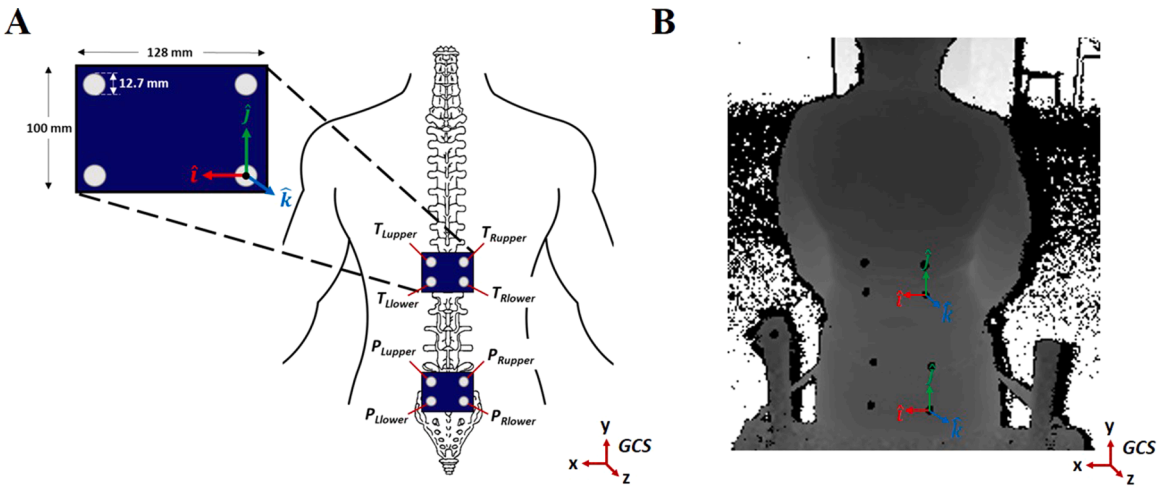

Fig. 3. Infrared reflective marker setup. (A) Trunk and pelvic marker clusters with individual marker labels. The trunk marker cluster with the local coordinate system is shown in the zoomed-in image. The global coordinate system is shown in the bottom right. (B) Depth image with local coordinate systems defined over each marker cluster. The global coordinate system is shown in the bottom right.

3. Results

RGB-D camera and optoelectronic motion capture system relative lumbar spine angles were extracted and compared. *RMSE* of continuous planar angles, and $ICC_{2,1}$*'s* of min, max, and ROM angular measurements were calculated to quantify performance of the RGB-D camera depth data and proposed image processing algorithms relative to the optoelectronic motion capture system. Overall, our method of extracting angles from the RGB-D camera's depth data produced very low *RMSE* in comparison with the optoelectronic motion capture system across all participants ($RMSE_{FE} = 2.05^\circ \pm 0.97^\circ$; $RMSE_{LB} = 0.65^\circ \pm 0.37^\circ$; $RMSE_{AT} = 0.84^\circ \pm 0.38^\circ$). Mean ensemble curves (i.e., representing the average motion across all cycles and across all participants) and associated standard deviations (SDs) were plotted for visual appraisal (Fig. 4).

Cycle-to-cycle min and max angles and ROM and their respective SDs from both systems were calculated and directly compared; they are organized and presented in Table 2. For analysis of reliability between systems, $ICC_{2,1}$'s for relative min and max angles and ROM of cycle-to-cycle data were compared across the FE, LB, and AT movement axes, and exhibited good to excellent reliability ($0.849 \leq ICC_{2,1} \leq 0.979$) in comparison with gold-standard motion capture equipment; specific results are listed in Table 3. Bland-Altman plots were also constructed to identify any instrument bias when measuring planar ROM (Fig. 5); these plots show that the RGB-D camera systematically underestimated ROM in the FE plane, and overestimated ROM in the LB and AT planes.

4. Discussion

In this study, we developed a custom image-processing method to track lumbar spine motion from RGB-D camera depth data during repetitive spine FE, and we were able to show that our method performs well when compared to a gold-standard optoelectronic motion capture system. *RMSE* of continuous motion, and absolute agreement and reliability of lumbar spine angular measurements were assessed across all movement axes between systems. Overall, our method of tracking lumbar spine motion from RGB-D camera's depth data produced low *RMSE* ($RMSE \leq 2.05^\circ \pm 0.97^\circ$) with optoelectronic motion capture system data. Although *RMSE* was highest overall in the FE plane, the error as a percentage of total ROM was lowest (i.e., ~ 4%); percentage error in the LB and AT planes was relatively high (i.e., ≥ 26%), but a large proportion of this can likely be attributed to random measurement noise, as the out-of-plane movement during FE is minimal [7]. Additionally, the values obtained for min, max, and ROM in the current study are comparable to similarly collected data from previous research [7]. $ICC_{2,1}$*'s* were good to excellent between systematic measurements of min and max angles, and ROM it was excellent in the FE plane (i.e., along the main axis of movement; $ICC_{2,1,FE} \geq 0.930$), and good to excellent in LB and AT planes ($0.865 \leq ICC_{2,1,LB} \leq 0.914$; $0.849 \leq ICC_{2,1,AT} \leq 0.921$). Lastly, Bland-Altman plots revealed that measurement agreement between systems was excellent [24], but there was minor systematic bias when measuring ROM; nevertheless, this difference was within the acceptable limits for clinical motion tracking (i.e., less than 2°) [9]. These results combined show that using RGB-D cameras is a

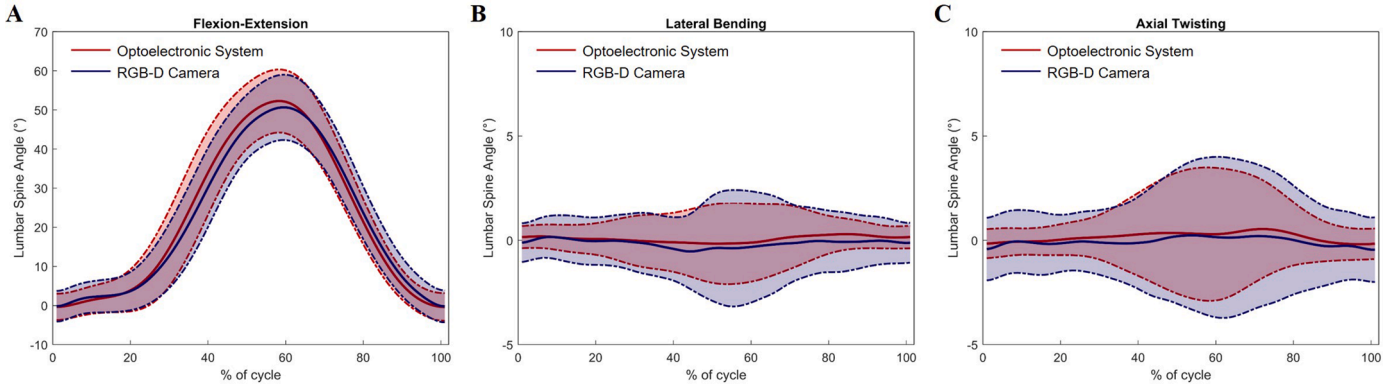

Fig. 4. RGB-D camera (blue) vs optoelectronic motion capture system (red) averages (solid) and standard deviations (dashed) of the lumbar spine joint angle for (A) flexion-extension, (B) lateral bending, and (C) axial twisting movement axes. Y-axis range for (B) and (C) are different than (A) for better visualization.

Table 2
Mean (SD) minimum and maximum angles and range of motion.

	Flexion-Extension		Lateral Bending		Axial Twisting	
	Vicon	**Kinect**	**Vicon**	**Kinect**	**Vicon**	**Kinect**
ROM (°)	53.4 (7.7)	51.5 (7.6)	1.9 (1.0)	2.5 (1.4)	2.7 (1.8)	3.0 (1.4)
max (°)	52.9 (8.6)	51.2 (8.5)	0.9 (1.2)	1.0 (2.0)	1.5 (1.6)	1.4 (2.4)
min (°)	-0.4 (3.5)	-0.3 (4.1)	-1.0 (1.1)	-1.5 (1.6)	-1.2 (2.3)	-1.6 (2.7)

ROM = range of motion; max = maximum; min = minimum; SD = standard deviation.

Table 3
Reliability analysis ($ICC_{2,1}$; *95% CI bounds*) between optoelectronic and RGB-D lumbar spine angles.

		Flexion-Extension	Lateral Bending	Axial Twisting
$ICC_{2,1}$	***ROM***	0.979 (0.453-0.996)	0.889 (0.290-0.973)	0.898 (0.654-0.970)
	max	0.973 (0.851-0.993)	0.914 (0.700-0.975)	0.849 (0.460-0.957)
	min	0.930 (0.754-0.980)	0.865 (0.521-0.961)	0.921 (0.736-0.977)

ROM = range of motion; max = maximum; min = minimum; $ICC_{2,1}$ = intraclass correlation coefficient; CI = confidence interval.

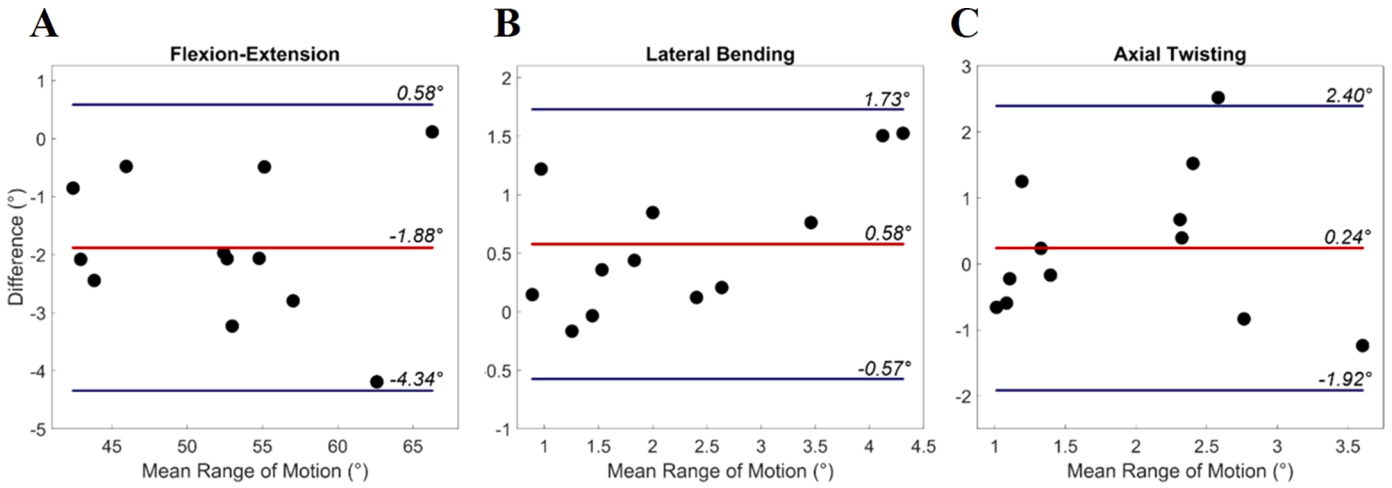

Fig. 5. Bland-Altman plots for lumbar spine (A) flexion-extension, (B) lateral bending, and (C) axial twisting range of motion measurements between optoelectronic and RGB-D systems. The red lines represent the mean measurement difference between systems, and the blue lines represent mean ± 1.96SD (standard deviation).

realistic and feasible option for clinical motion capture of the spine.

Similar to the current study, Macpherson and colleagues [16] tracked an RGB-D camera's depth data with a custom developed algorithm that identified the centroid of individual IR reflective markers. They compared linear and angular displacements of the trunk and pelvis between the RGB-D camera and the optoelectronic motion capture system during walking and running on a treadmill. While correlations between linear displacements were very strong to nearly perfect, correlations of angular displacements ranged from moderate to strong. It is likely that their methods to obtain angular displacement from linear marker trajectories lacked the sophistication needed to accurately obtain these angles; that is, Macpherson and colleagues [16] utilized only 4 IR reflective markers (placed over the right and left posterior superior iliac spines and bilaterally over the 10^{th} rib) to track motion of the trunk and pelvis, and for each marker, a 42-point rectangular cloud was superimposed onto the image surrounding the marker. The mean marker depth was used as the camera-marker distance (i.e., anterior-posterior direction), and the medial-lateral and superior-inferior locations of the marker were calculated using trigonometry and manufacturer information regarding the field of view. To calculate angular positions, vectors joining left-sided and right-sided markers were created (representing the FE plane axis), and the projected angles of these vectors on the LB and AT planes were recorded for the trunk and pelvis rotations. This process was quite different than the method proposed in our study, with the main difference being the utilization of synthetic superimposed marker data to track angular positions. The use of synthetic data was necessary in their study as their marker setup involved only 2 markers per anatomical region (their marker tracking algorithm was unable to discern more closely positioned markers). We believe that the method used in the current study performed better as it did not involve the use of synthetic markers (it is recommended that 4 markers – or at least 3 – are sufficient for position/orientation tracking of marker clusters), and the process for obtaining rotations from rigid-body marker clusters placed over the trunk and pelvis has been used extensively in literature [25–27]. Additionally, Macpherson and colleagues [16] used a different RGB-D camera (i.e., the Kinect v1), which was a previous version of the system; our study used the Kinect v2 which has been proven to have better accuracy and precision when tracking motion, [21,28]. Lastly, Macpherson and colleagues' [16] movement protocol for the involved fast-paced treadmill running, which can cause excess skin-motion artifact noise at heel contact, and can affect motion tracking; we believe that our movement protocol is less strenuous and therefore likely to avoid excessive noise caused by skin-motion.

While the results from the current study are promising, there are some limitations to consider. First, the spatial resolution of the selected RGB-D camera decays with distance [29]; as such, there may be increased error (and poorer agreement and reliability) as a direct result of the nature of the given task (i.e., as the participant moves farther away into a fully flexed position, spatial resolution decays). To address this in future studies, the utilization and synchronization of two or more RGB-D cameras that are capable of recording the full capture volume with high-quality spatial resolution may be a feasible option. This may also help to reliably capture more complex movements in future studies (e.g., repeated LB, multidirectional movements, etc.). In addition, the method to obtain z global coordinates relies on a heuristic approximation, which has potential to influence accuracy. For example, if a larger pixel vicinity is selected, it is possible that some data points that do not represent the pixel cluster region could be included (e.g., participants' garment), introducing uncertainty into the measured shape and positioning of these pixel values. Our method accounts for this uncertainty by: 1) selecting a pixel vicinity region that is 1 pixel thick (i.e., not too large so that it includes the participants' clothing or other objects outside of the marker cluster), and 2) optimizing camera position and orientation so that the flat rigid base of the vicinity of the IR reflective marker clusters is visible throughout the entire movement protocol (i.e., to ensure smaller pixel regions within the vicinity of the IR marker cluster are always visible, without capturing clothing or other possible artifacts). Therefore, it is not expected for these factors to have a significant affect on accuracy. Furthermore, although the selected RGB-D camera has been discontinued, our methods and algorithms are directly usable and transferable to other RGB-D time-of-flight cameras, including the Azure Kinect (Microsoft Corporation, Redmond, WA, USA) which contains an RGB-D camera with advanced AI sensors for building computer vision models, and can be connected to Azure cloud databases for future big-data analyses. Lastly, this study aimed to validate the use of RGB-D cameras as quantitative tools for measuring spine kinematics during a repetitive trunk FE task on healthy participants; in order to be considered feasible for clinical motion tracking of the spine in people with low back pain, it must first be validated on this group of patients. Overall, this study generates evidence of low levels of error, and good to excellent agreement and reliability between systems, making it a feasible option for lumbar spine motion tracking in clinical settings.

5. Conclusion

The proposed method performed well (i.e., low error and good to excellent agreement and reliability) for tracking lumbar spine motion using a single RGB-D camera's depth data in a standardized setting compared to optoelectronic motion capture equipment; thus, the

proposed method can be considered as a valid method for quantifying lumbar spine FE movement using one RGB-D camera, and can be extended to various applications beyond spine motion assessment (e.g., other anatomical segments which require custom camera setup and post-processing that deviates from the standard full-body skeletal model). Further work is being conducted to develop deep learning methods to automatically segment and classify trunk/spine anatomical landmarks, reducing the amount of time and specific expertise required to process data, and allowing similar kinematic analysis to be done without the use of IR reflective markers. Future studies should investigate the validity of using RGB-D cameras to track human movement in a wider variety of settings (i.e., alternate camera positions and angles), during additional movement tasks (e.g., repeated LB and AT, and activities of daily living), and on clinical populations.

Ethical approval

This study was approved by the University of Ottawa Research Ethics Board (H02-17-11; H08-17-26).

Declaration of Competing Interest

The authors have no conflicts of interest to declare.

Acknowledgments

This study was funded by the Natural Sciences and Engineering Research Council of Canada (RGPIN-2014-05560 [Ryan Graham]) and an Ontario Early Researcher Award (Ryan Graham).

References

[1] Balagué F, Mannion AF, Pellisé F, Cedraschi C. Non-specific low back pain. Lancet 2012;379:482–91. https://doi.org/10.1016/S0140-6736(11)60610-7.
[2] Silfies SP, Bhattacharya A, Biely S, Smith SS, Giszter S. Trunk control during standing reach: A dynamical system analysis of movement strategies in patients with mechanical low back pain. Gait Posture 2009;29(3):370–6. https://doi.org/10.1016/j.gaitpost.2008.10.053.
[3] Biely SA, Silfies SP, Smith SS, Hicks GE. Clinical observation of standing trunk movements: What do the aberrant movement patterns tell us? J Orthop Sports Phys Ther 2014;44(4):262–72. https://doi.org/10.2519/jospt.2014.4988.
[4] Fritz JM, Cleland JA, Childs JD. Subgrouping patients with low back pain: evolution of a classification approach to physical therapy. J Orthop Sports Phys Ther 2007;37(6):290–302. https://doi.org/10.2519/jospt.2007.2498.
[5] Hicks GE, Fritz JM, Delitto A, Mishock J. Interrater reliability of clinical examination measures for identification of lumbar segmental instability. Arch Phys Med Rehabil 2003;84(12):1858–64. https://doi.org/10.1016/S0003-9993(03)00365-4.
[6] Stanton TR, Fritz JM, Hancock MJ, Latimer J, Maher CG, Wand BM, et al. Evaluation of a treatment-based classification algorithm for low back pain: A cross-sectional study. Phys Ther 2011;91(4):496–509. https://doi.org/10.2522/ptj.20100272.
[7] Beange KHE, Chan ADC, Beaudette SM, Graham RB. Concurrent validity of a wearable IMU for objective assessments of functional movement quality and control of the lumbar spine. J Biomech 2019;97(3):109356. https://doi.org/10.1016/j.jbiomech.2019.109356.
[8] Bauer CM, Rast FM, Ernst MJ, Kool J, Oetiker S, Rissanen SM, et al. Concurrent validity and reliability of a novel wireless inertial measurement system to assess trunk movement. J Electromyogr Kinesiol 2015;25(5):782–90. https://doi.org/10.1016/j.jelekin.2015.06.001.
[9] McGinley JL, Baker R, Wolfe R, Morris ME. The reliability of three-dimensional kinematic gait measurements: A systematic review. Gait Posture 2009;29(3):360–9. https://doi.org/10.1016/j.gaitpost.2008.09.003.
[10] Ricci L, Taffoni F, Formica D. On the orientation error of IMU: Investigating static and dynamic accuracy targeting human motion. PLoS One 2016;11(9):e0161940. https://doi.org/10.1371/journal.pone.0161940.
[11] Zimmermann T, Taetz B, Bleser G. IMU-to-segment assignment and orientation alignment for the lower body using deep learning. Sensors (Switzerland) 2018;18:E302. https://doi.org/10.3390/s18010302.
[12] Otte K, Kayser B, Mansow-Model S, Verrel J, Paul F, Brandt AU, et al. Accuracy and reliability of the kinect version 2 for clinical measurement of motor function. PLoS One 2016;11(11):e0166532. https://doi.org/10.1371/journal.pone.0166532.
[13] Clark RA, Pua Y-H, Fortin K, Ritchie C, Webster KE, Denehy L, et al. Validity of the Microsoft Kinect for assessment of postural control. Gait Posture 2012;36(3):372–7. https://doi.org/10.1016/j.gaitpost.2012.03.033.
[14] Auvinet E, Multon F, Manning V, Meunier J, Cobb JP. Validity and sensitivity of the longitudinal asymmetry index to detect gait asymmetry using Microsoft Kinect data. Gait Posture 2017;51:162–8. https://doi.org/10.1016/j.gaitpost.2016.08.022.
[15] Dolatabadi E, Taati B, Mihailidis A. Concurrent validity of the Microsoft Kinect for Windows v2 for measuring spatiotemporal gait parameters. Med Eng Phys 2016;38(9):952–8. https://doi.org/10.1016/j.medengphy.2016.06.015.
[16] Macpherson TW, Taylor J, McBain T, et al. Real-time measurement of pelvis and trunk kinematics during treadmill locomotion using a low-cost depth-sensing camera: A concurrent validity study. J Biomech 2016;49(3):474–8. https://doi.org/10.1016/j.jbiomech.2015.12.008.
[17] Cippitelli E, Gasparrini S, Spinsante S, Gambi E. Kinect as a tool for gait analysis: Validation of a real-time joint extraction algorithm working in side view. Sensors (Switzerland) 2015;15:1417–34. https://doi.org/10.3390/s150101417.
[18] Clark RA, Pua YH, Bryant AL, Hunt MA. Validity of the Microsoft Kinect for providing lateral trunk lean feedback during gait retraining. Gait Posture 2013;38(4):1064–6. https://doi.org/10.1016/j.gaitpost.2013.03.029.
[19] Cuesta-Vargas AI, Galán-Mercant A, Williams JM. The use of inertial sensors system for human motion analysis. Phys Ther Rev 2010;15(6):462–73. https://doi.org/10.1179/1743288X11Y.0000000006.
[20] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15(2):155–63. https://doi.org/10.1016/j.jcm.2016.02.012.
[21] Wasenmüller O, Stricker D. Comparison of Kinect v1 and v2 depth images in terms of accuracy and precision. In: Chen C-S, Lu J, Ma K-K, editors. Comput. Vis. – ACCV 2016 Workshops. 1st. Cham: Springer International Publishing; 2016. p. 34–45.. https://doi.org/10.1007/978-3-319-54427-4_3.
[22] Granata KP, England SA. Stability of dynamic trunk movement. Spine 2006;31(10):E271–6. https://doi.org/10.1097/01.brs.0000216445.28943.d1.
[23] Cappozzo A, Cappello A, Croce UD, Pensalfini F. Surface-marker cluster design criteria for 3-D bone movement reconstruction. IEEE Trans Biomed Eng 1997;44(12):1165–74. https://doi.org/10.1109/10.649988.
[24] Giavarina D. Understanding Bland Altman analysis. Biochem Medica 2015;25(2):141–51. https://doi.org/10.11613/BM.2015.015.
[25] Beange K, Chan ADC, Graham RB. Wearable sensor performance for motion tracking of the lumbar spine. CMBES Proc 2019;42(May). https://proceedings.cmbes.ca/index.php/proceedings/article/view/859/852.
[26] Mavor MP, Graham RB. Exploring the relationship between local and global dynamic trunk stabilities during repetitive lifting tasks. J Biomech 2015;48(14):3955–60. https://doi.org/10.1016/j.jbiomech.2015.09.026.
[27] Ross GB, Mavor MP, Brown SHM, Graham RB. The Effects of Experimentally Induced Low Back Pain on Spine Rotational Stiffness and Local Dynamic Stability. Ann Biomed Eng 2015;43(9):2120–30. https://doi.org/10.1007/s10439-015-1268-9.
[28] Samir M, Golkar E, Rahni AAA. Comparison between the Kinect™ V1 and Kinect™ V2 for respiratory motion tracking. In: IEEE 2015 Int Conf Signal Image Process Appl ICSIPA 2015 - Proc. IEEE; 2016. p. 150–5. https://doi.org/10.1109/ICSIPA.2015.7412180. https://doi.org/10.1109/ICSIPA.2015.7412180.
[29] Yang L, Zhang L, Dong H, et al. Evaluating and improving the depth accuracy of Kinect for Windows v2. IEEE Sens J 2015;15(8):4275–85. https://doi.org/10.1109/JSEN.2015.2416651.