Heliyon 9 (2023) e21606

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Azure Kinect performance evaluation for human motion and upper limb biomechanical analysis

Cristina Brambilla [a], Roberto Marani [a], Laura Romeo [a,b], Matteo Lavit Nicora [a,c], Fabio A. Storm [d], Gianluigi Reni [e], Matteo Malosio [a], Tiziana D'Orazio [a], Alessandro Scano [a,*]

[a] Institute of Intelligent Industrial Systems and Technologies for Advanced Manufacturing (STIIMA), Italian Council of National Research (CNR), Italy
[b] Department of Electrical and Information Engineering (DEI), Polytechnic of Bari, Bari, Italy
[c] Industrial Engineering Department, University of Bologna, Bologna, Italy
[d] Bioengineering Laboratory, Scientific Institute, IRCCS "Eugenio Medea", 23842 Bosisio Parini, Lecco, Italy
[e] Informatics Department, Autonomous Province of Bolzano, Bolzano, Italy

ARTICLE INFO

Keywords:
Azure Kinect
Biomechanics
Human tracking
Kinematics
Upper limb
Vicon

ABSTRACT

Human motion tracking is a valuable task for many medical applications. Marker-based opto-electronic systems are considered the gold standard in human motion tracking. However, their use is not always feasible in clinics and industrial environments. On the other hand, marker-less sensors became valuable tools, as they are inexpensive, noninvasive and easy to use. However, their accuracy can depend on many factors including sensor positioning, light conditions and body occlusions. In this study, following previous works on the feasibility of marker-less systems for human motion monitoring, we investigate the performance of the Microsoft Azure Kinect sensor in computing kinematic and dynamic measurements of static postures and dynamic movements. According to our knowledge, it is the first time that this sensor is compared with a Vicon marker-based system to assess the best camera positioning while observing the upper body part movements of people performing several tasks. Twenty-five healthy volunteers were monitored to evaluate the effects of the several testing conditions, including the Azure Kinect positions, the light conditions, and lower limbs occlusions, on the tracking accuracy of kinematic, dynamic, and motor control parameters. From the statistical analysis of the performed measurements, the camera in the frontal position was the most reliable, the lighting conditions had almost no effects on the tracking accuracy, while the lower limbs occlusion worsened the accuracy of the upper limbs. The assessment of human static postures and dynamic movements based on experimental data proves the feasibility of applying the Azure Kinect to the biomechanical monitoring of human motion in several fields.

* Corresponding author. Institute of Intelligent Industrial Systems and Technologies for Advanced Manufacturing (STIIMA), Italian Council of National Research (CNR), Via A.Corti 12, 20133 Milan, Italy.
E-mail address: alessandro.scano@stiima.cnr.it (A. Scano).

https://doi.org/10.1016/j.heliyon.2023.e21606
Received 31 May 2023; Received in revised form 21 September 2023; Accepted 24 October 2023
Available online 4 November 2023

1. Introduction

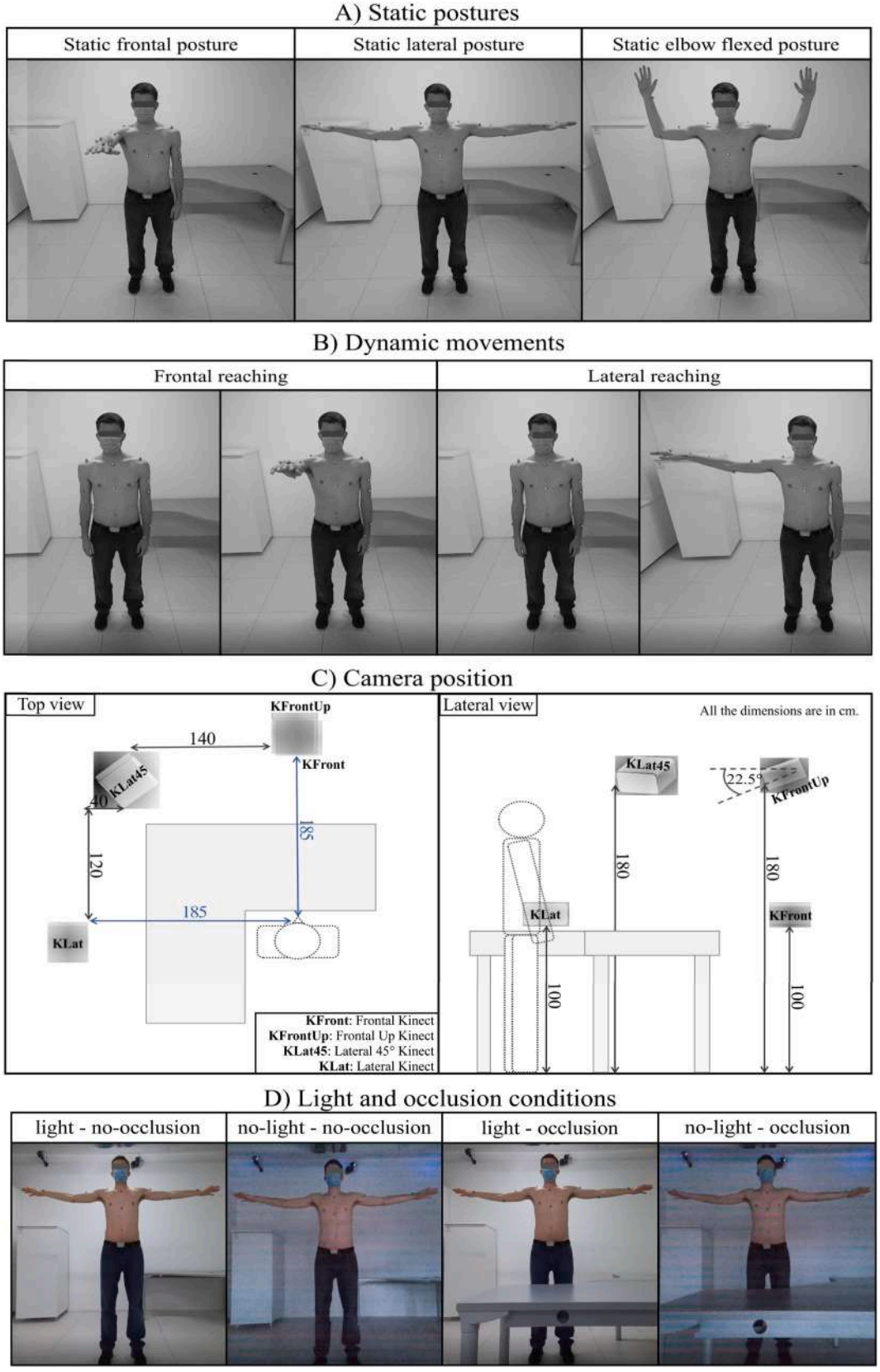
Human motion tracking is a valuable instrument, typically employed in laboratory or clinical environments [1], for studying the kinematics of movements and detecting motor and neurological impairments. The most used techniques to track human motion are marker-based optoelectronic systems, based on a set of infrared cameras that record the position of retro-reflective markers placed on the anatomical landmarks. The 3D positions of the markers are reconstructed and are used to derive kinematic variables, including human joint angular motion. Due to high accuracy, marker-based systems are considered the gold standard for motion tracking. However, these camera systems are expensive; they require dedicated areas, a long preparation time and large set-ups in motion tracking laboratories [2]. Therefore, their use is usually not feasible in small clinics or home rehabilitation, as they require specific setups and procedures [3]. Moreover, the use of markers placed on the body of the operator is generally not feasible, when monitoring movement in real world scenarios. Despite these limitations, the use of tracking devices has been fostered also in the industrial environment to predict human intention in collaborative tasks [4], to evaluate the ergonomics [5] and to assess fatigue of the workers [6]. In fact, with the introduction of collaborative robots (cobots) in manufacturing processes, the use of human motion tracking systems in industrial environments is increasing for evaluating the level of collaboration, interaction with the device, fatigue and psychological state of the worker. With this aim, several recent research projects were launched that monitor kinematic parameters [7–11]. Due to their high precision, marker-based systems are also employed as gold standard methods for benchmarking sensors [12]. In the last decade, inertial sensors and marker-less motion capture systems have been considered promising alternatives for recording human motion tracking [13]. These devices are less expensive, less invasive, have a faster set-up, and have higher flexibility in the acquisition environment. In this framework, motion capture systems based on RGB and depth sensors demonstrated their feasibility for kinematic analysis in clinical and rehabilitation assessments [14]. New devices are continuously launched on the market with more improvements in terms of accuracy. One of the most used marker-less systems is the Azure Kinect sensor (from hereon, also referred only as “Kinect”), developed by Microsoft for entertainment purposes. It became really popular also in clinical contexts [15] with a variety of applications. Thanks to a combination of RGB and depth sensors, it can track human motions and reconstruct body skeletons. A set of tracking libraries is also provided with the Kinect sensors, but the details of employed libraries are not fully specified. For this reason, the accuracy and reliability of the previous generations of Kinect (V1 and V2) have been studied in the literature and results have been applied to different fields [16]. Postural control and balance were investigated by Clark and colleagues, finding excellent performance with both Kinect V1 [17] and V2 [18]. In rehabilitation contexts, the tracking accuracy of Kinect V2 showed better results for the joints of the upper body than the lower body [19] and the spatiotemporal gait parameters showed excellent validity with respect to a maker-based system [20]. Kinect V2 was used for the evaluation of the symmetry of gait and could discriminate symmetrical and asymmetrical gait [21]. Moreover, Kinect V2 position data were employed as input for automated classification system to assess exercise movement quality [22]. However, in other studies the kinematic validity of lower limb [23,24] and upper limb joints [25] was poor and high errors were found in tracking fine movements [26] and hand motion [27]. While the first two generations of Kinect were deeply assessed in literature, few studies compared the tracking performance of the new Azure Kinect to optoelectronic systems or the previous versions of the Kinect camera. The Azure Kinect showed high reliability for joint angles in gait [28] and sit-to-stand tasks [29]. Moreover, excellent agreement was found for the spatiotemporal gait parameters between Azure Kinect and a marker-based system [30], finding relevant improvements in the accuracy of spatial parameters of gait with respect to the previous version V2 [31]. However, the tracking accuracy of the Kinect cameras may be influenced by multiple factors. Many studies [32–34] found that the Kinect camera cannot accurately estimate the joint position when some body parts are hidden. Furthermore, the accuracy of the Kinect system is in principle sensitive to light conditions [35,36]. Finally, the Kinect accuracy also depends on the distance and the orientation of the camera with respect to the subject [37,38].

In industrial contexts, Kinect sensors have been used to preserve human safety and to improve the efficiency of human-robot collaboration. Recently, they were employed for the ergonomic evaluation of workplaces, substituting the observational risk assessment, in order to improve working postures and decrease the risk of developing musculoskeletal disorders (MSD) [39]. Furthermore, the Kinect sensor was used to test the psychophysiological responses of operators interacting with a cobot in a virtual reality environment [40]. In a very recent work [41], a large set of multi-modal data were recorded with the Azure Kinect during the execution of industrial tasks and the data were spatio-temporally assessed for movement segmentation and recognition purposes.

In all the considered contexts, several issues must be faced when dealing with a marker-less system such as the Azure Kinect:

- The technologies adopted for the Kinect sensor and the algorithms used for identifying 3D positions of the joints are different from those employed in gold standard marker-based systems. Kinect accuracy should be quantified in the target scenarios.
- The positioning of the sensor with respect to the subject must adapt to the set-up, since the environment might not be flexible and the light conditions may vary depending on the place where the set-up is installed and cannot be controlled.
- Typical real scenarios – including telerehabilitation, robotic rehabilitation or working activities with industrial cobots – may be part of complex set-ups where occlusions may occur and, especially, often lower limbs are obstructed, for example by tables.
- Very few studies expanded the analysis from pure kinematic assessments to dynamic and motor control variables involving relevant data connected to people’s behavior, such as fatigue and effort.

Following this rationale, the main objective of this study is the evaluation of the tracking performance of the Azure Kinect compared to the Vicon optoelectronic system, used as a reference gold standard. To the best of our knowledge, this is one of first studies to analyze the reliability of the Azure Kinect in the specific context of human monitoring in real applications, taking into consideration movements that involve mainly the upper limbs and derived biomechanical parameters.



(caption on next page)

Fig. 1. Motor tasks and testing conditions. In panel A, static postures used to assess the tracking capability of the devices are shown: (i) static frontal posture, (ii) static lateral posture and (iii) static elbow flexed postures. In panel B, the resting position and the maximum extension of the arm are depicted for both dynamic movements (frontal reaching on the left, lateral reaching on the right). In panel C, a schematic representation of the nominal positioning of the Azure Kinect cameras from the top and lateral views are shown with measures in cm. In panel D, the four combinations of light and occlusion conditions are illustrated.

The main contributions of the work are:

- A novel experimental campaign was conducted involving a cohort of healthy subjects. The extensive data collected by the opto-electronic systems and the Azure Kinect cameras are available to the scientific community.
- A complete experimental protocol of multiple static poses and dynamic tasks was defined to evaluate the influence of four camera positions/orientations, different light conditions, and the effects of lower joint occlusion on the tracking accuracy.
- A statistical analysis evaluated the minimum target number of subjects that provide significant statistical power. ANOVA tests were used to identify the significant differences between the settings and between conditions.
- The effects of joint tracking errors on low-level parameters (body segment lengths), on medium-level parameters (joint angles) and high-level parameters (motor control and dynamics, obtained with the biomechanical model) were analyzed with respect to a marker-based system for each of the Kinect camera configuration separately.

The remaining part of this work is organized as follows. Section 2 deals with materials and methods, detailing the participants, the equipment, and the experimental protocol. The data analysis with the derived measures evaluated, and the statistical analysis performed are described in Section 3. In Section 4, extensive results obtained in static and dynamic experiments are detailed, while a deep discussion is reported in Section 5.

2. Materials and methods

2.1. Participants

Participants recruited for the study were 25 young volunteers without neurological or musculoskeletal impairments (age: 28.5 $\pm$ 4.9; height: 175.6 $\pm$ 9.6 cm; weight: 69.4 $\pm$ 10.0 kg; 18 M and 7F). Ethical approval was granted by the referring ethical committee (Approval Prot. N. 19/20 – CE, April 20th, 2020), and the experimental trial was conducted in compliance with the Declaration of Helsinki [42] and its updates. Every participant provided written informed consent to take part in this study and to allow publication of any data included in this article. The participants were asked to wear very tight vests or bare-chested at the time of trial, to facilitate placement of the markers on the anatomical landmarks.

2.2. Equipment

Movements were acquired in the motion acquisition laboratory of the Italian Council of National Research (CNR) in Lecco, Italy. The acquisitions were made with two different systems, a marker-based and a marker-less tracking system. The laboratory had the following equipment:

- A Vicon Vero system (v2.2, Vicon Motion Systems Ltd., Oxford, UK), composed of ten infrared cameras. A set of 25 retro-reflective markers was attached to each subject to track the movement by the cameras.
- Four Azure Kinect cameras recorded the movements from four points of view. Acquisitions were time-synchronized by a trigger signal dispatched through an AUX cable connecting the cameras in a daisy-chain configuration.
- An L-shape table. It was used for testing the occlusion condition and simulating a working surface.
- Tripods for supporting Kinect cameras.
- A lux meter for measuring the light intensity.

2.3. Experimental protocol

2.3.1. Motor task

The tracking capability of the devices was assessed in static postures and dynamic movements:

a) Static postures: frontal arm, lateral arms and lateral arms with flexed elbows

In static trials, the subject held the fixed position for 10 s. We selected three static postures, shown in Fig. 1: (i) with the right arm raised frontally, (ii) with both arms raised laterally and (iii) with both arms raised laterally and the elbows flexed. In posture (i), the subject stood with the right arm raised frontally at 90° with the palm facing downwards. In posture (ii), the subject maintained the arms raised laterally at 90° with the elbows extended and the palms facing downwards. In posture (iii), the subject held the arms raised laterally at 90° with the elbows flexed at 90° and the palms facing forward.

b) Dynamic movements: frontal reaching and lateral reaching

Frontal reaching is a simple movement frequently used in human motion analysis. We assessed both frontal and lateral reaching as a motion primitive of several upper limb tasks. The subject started from the resting position, defined as the standing position with the arm relaxed by the side, with extended elbow and the shoulder not elevated. For the frontal reaching, the subject was asked to raise the right arm frontally at 90°, with the palms facing downwards and turn back to the resting position. For the lateral reaching, the subject was asked to raise the right arm laterally at 90° and return to the resting position. Each movement was repeated ten times. A schematic illustration of the movements is presented in Fig. 1.

2.3.2. Testing conditions

We tested the accuracy of the Azure Kinect sensor under three conditions: acquisition from different camera positions, in absence and presence of occlusions, and under two light conditions.

a) Camera position:

Since the camera point of view can affect the joint position estimation of the Azure Kinect, we assessed how the positioning of the camera influences the tracking accuracy. Four Kinect cameras were used, investigating four positions and orientations with respect to the subject in simultaneous acquisitions. In order to cover a wide field of view, the four chosen positions of the camera were: (a) in front of the subject (frontal view) at 1-m height from the ground (frontal Kinect – KFront); (b) in front of the subject, from above at 1.80-m height from the ground with an inclination of −22.5° (pitch) pointing downwards (frontal up Kinect – KFrontUp); (c) on the lateral side at 45° (yaw), from above at 1.80-m height from the ground with an inclination −22.5° (pitch) pointing downwards (lateral 45° Kinect – KLat45); (d) on the lateral side (lateral view) at 90° (yaw) at 1-m height from the ground (lateral Kinect – KLat). The lateral side is opposite to the side on which the movement is executed, i.e. the left side. A schematic representation of the camera positioning is shown in Fig. 1 from top view and lateral view. Data from each camera were analyzed independently from the others in order to identify the camera position with the highest accuracy.

b) Occlusion:

The tasks were performed with and without joint occlusions to quantify the inaccuracies due to partial occlusions. In particular, we were interested in quantifying whether the occlusion of the lower limbs affects performance with respect to non-obstructed conditions and to the gold standard marker-based system, on both the static postures and the dynamic movements.

c) Light condition:

Light exposure influences the tracking accuracy of the Kinect and also the performance of the gold standard device. In order to standardize light conditions, tests were performed using artificial light, with no-light from the outside. Two light conditions were assessed: illuminance was 268.76 ± 23.3 Lux in the 'light' condition and 18.2 ± 6.1 Lux in the 'no-light' condition. The four combinations of occlusion and light intensity are shown in Fig. 1.

2.3.3. Acquisition

Markers were attached to the subject on anatomical landmarks corresponding to the Vicon Upper limb model requirements (7th cervical vertebra, Right back, 10th thoracic vertebra, Clavicle, Sternum, Left shoulder, Left upper arm marker A, Left upper arm marker B, Left upper arm marker C, Left elbow, Left medial epicondyle, Left forearm, Left wrist marker A, Left wrist marker B, Left finger, Right shoulder, Right upper arm marker A, Right upper arm marker B, Right upper arm marker C, Right elbow, Right medial epicondyle, Right forearm, Right wrist marker A, Right wrist marker B, Right finger). For details, please refer to the Vicon guide [43]. The acquisitions were recorded with the Vicon and Azure Kinect systems simultaneously: the Vicon system started recording before the Azure Kinect; instead, at the end of the acquisition, the Azure Kinect stopped recording before Vicon. The static posture was held for 10 s while dynamic tasks were repeated ten times with a 1-s pause between repetitions. Moreover, each acquisition was performed twice to have backups where needed.

The data acquired by the Vicon were elaborated in the Vicon Nexus software to track and label the markers: the output obtained was the 3D coordinates of markers and joint center positions sampled at 100 Hz. The acquisition from the Azure Kinect cameras was elaborated with the Microsoft Azure Body Tracking SDK (v1.1.1): the output was the 3D positions of 32 joints and quaternions indicating segment orientation sampled at 30 Hz.

2.4. Temporal synchronization between systems

Since the Vicon system and the Azure Kinect cameras were not temporally synchronized, movement onset and offset were identified in the data from each system, and the corresponding phases were aligned with a post-processing procedure [29,44] in MATLAB® R2021a (MathWorks, Natick, USA). To identify the movement phases of the dynamic movements, we chose the articular angle that was the most representative of the movement (i.e., the one that had a wider range of motion and clear onset/offset points). For the frontal reaching, the shoulder flexion angle was used; the shoulder abduction angle was employed for the lateral reaching. To identify

movement phases, the time series of the angle were filtered with a 4th-order Butterworth low pass filter at a cut-off frequency of 3 Hz and the first derivative (angular velocity) was computed with the aim of identifying the phases of the movement. Therefore, the onset and offset of each movement were identified as the start and end points where the shoulder angular velocity was over a threshold equal to 1 % of its absolute maximum value.

3. Data analysis

For data analysis, the raw data obtained from both systems were filtered with a 4th-order Butterworth low pass filter at a cut-off frequency = 5 Hz in order to remove noise artifacts. Therefore, since the acquisition systems have different sampling rates (100 Hz for the Vicon system and 30 Hz for the Kinect), the Azure Kinect data were up-sampled to 100 Hz with a shape-preserving piecewise cubic interpolation [45] to allow data comparison from the two systems with a procedure already adopted in previous studies to compare marker-based and Kinect system [31,38]. Afterwards, the data were elaborated with a biomechanical model [46,47], which allowed the computation of kinematic and dynamic variables and motor control parameters. The biomechanical model takes as input the 3D position of markers and joint center for the Vicon system and the 3D coordinates of joints and the quaternions for the Kinect. Then, it reconstructs the 3D coordinate system of each segment and joint and computes joint angles. Joint moment and forces are computed with Newton-Euler equations. Power exerted at joint level and expended energy are computed with equations reported in Appendix A. For both Vicon and Kinect, the mass properties (mass, inertia matrix and center of mass of each segment) are the same for each subject and they are estimated with anthropometric tables [48], using the height and the weight of each participant.

3.1. Outcome measures

The parameters computed for the analysis were divided into three categories based on the level of detail they provide. *Basic* parameters were not associated with a biomechanical model and included:

- Normalized body segment lengths (arm and forearm), as the limb length divided by the subject's height. Normalization is needed for reliable inter-subject comparisons.
- Execution time, as the time needed to execute each movement.

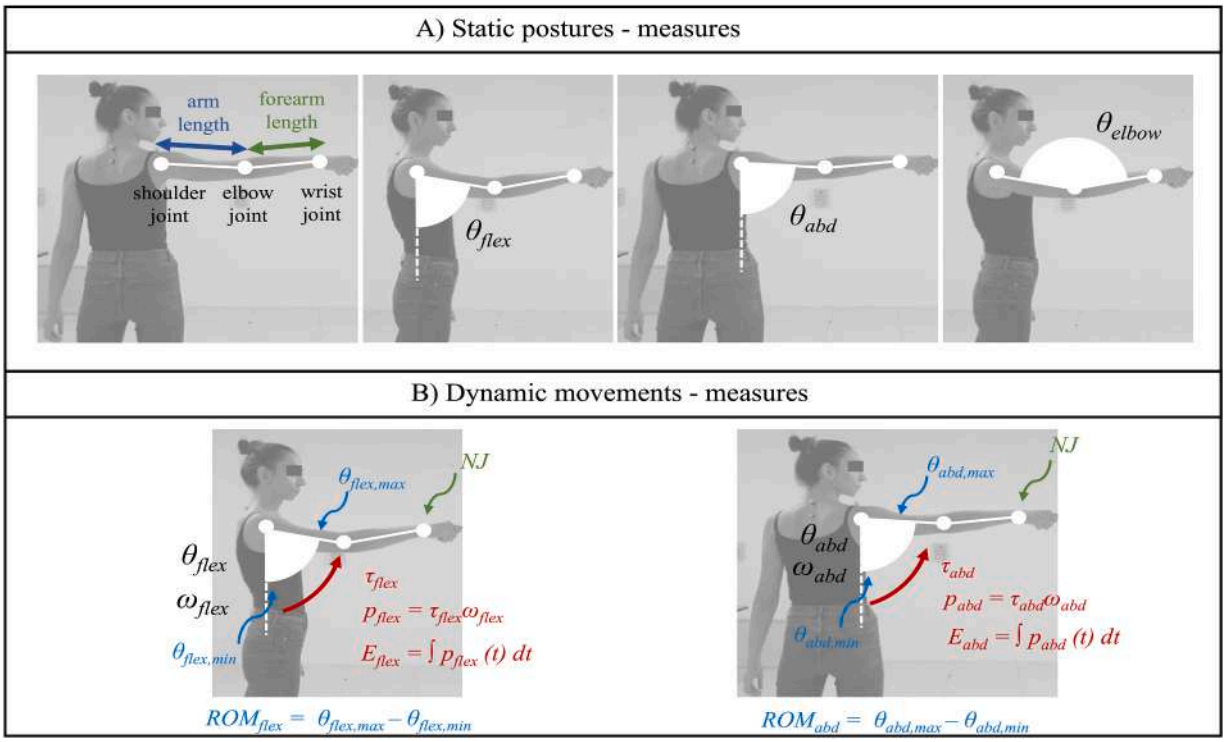


Fig. 2. Schematic representation of the parameters assessed for the Kinect-Vicon comparison. Panel A shows the parameters of the assessment computed for the static postures: limb lengths and articular angles θ are reported in the upper panel. (flex = flexion, abd = abduction). Panel B shows the parameters of the assessment computed for the dynamic movements: shoulder angle (flexion angle θ_{flex} for the frontal reaching and abduction angle θ_{abd} for the lateral reaching), minimum (min) and maximum (max) angles, range of motion ROM, angular velocity ω, torque τ, power p, energy E and normalized jerk NJ are reported in the lower panel.

The limb lengths were computed as the Euclidean 3D distance between wrist and elbow for estimating the forearm length and between elbow and shoulder for the arm length. For the Vicon system, the joint centers of the shoulder, elbow and wrist estimated with the Vicon Nexus software were considered for the computation. In contrast, the skeleton's shoulder, elbow and wrist joints were used for the Azure Kinect. Direct comparisons of Kinect joints estimation and markers or joint centers position computed with Nexus software were done in previous similar studies [24,26,27]. The procedure was done for both the left and the right arm.

The *kinematic* parameters were computed with the biomechanical model:

- joint angle (θ)
  - shoulder flexion angle (θ_{flex})
  - shoulder abduction angle (θ_{abd})
  - elbow flexion angle (θ_{elbow})
- minimum joint angle ($\theta_{\min}$)
  - shoulder flexion angle (θflex,min)
  - shoulder abduction angle ($\theta_{abd,min}$)
- maximum joint angle ($\theta_{\max}$)
  - shoulder flexion angle ($\theta_{flex,max}$)
  - shoulder abduction angle ($\theta_{abd,max}$)
- range of motion (*ROM*), as the difference between the maximum and the minimum angle reached
  - shoulder flexion angle (ROM_{flex})
  - shoulder abduction angle (ROM_{abd})
- angular velocity (ω), as the derivative of the joint angle
  - shoulder flexion angle (ω_{flex})
  - shoulder abduction angle (ω_{abd})

Finally, the *dynamic* parameters regarded motor control and dynamics, obtained with the biomechanical model through the Newton-Euler equations:

- joint torque (τ)
- peak power ($p_{\max}$), as the maximum power exerted at the joint level during the movement. The computation is shown in (1) in Appendix A.
- expended energy (E), as the integral of the power within movement phases. The computation is shown in (2) in Appendix A.
- normalized jerk (*NJ*), as a measure of the movement smoothness [49], computed as in (3) in Appendix A.

The presented parameters are summarized in Fig. 2.

The joint torque, power and energy were normalized, dividing each parameter by the arm's length and mass to allow inter-subject comparison, as done in previous studies [50,51]. All the dynamics parameters (e.g. inertial tensors) were taken from anthropometric tables [48]. Since the static postures do not involve movement, only limb length and joint angles were considered. For the dynamic movement, all kinematic and dynamic measures were computed. The comparison between the settings was performed on the mean values of the parameters that were averaged across repetitions for each subject. The use of the mean value of the parameters allows to preserve the information about their magnitude and to be more informative than the mean errors, as already done in similar studies [24,32]. Finally, Pearson's correlation coefficients were used to assess the similarity of the time series of the shoulder angle, torque, power and energy between each Kinect setting and the Vicon system in the dynamic tasks.

3.2. Statistical analyses

For the definition of the sample size, a priori power computation was performed in GPower 3.1.9.7 software (Heinrich Heine University, Dusseldorf, Germany) [52]. Considering a statistical power of 0.80 and a significance level equal to 0.05, the minimum target number of subjects to be enrolled is 13 when we hypothesize a large effect size (0.45). In contrast, 25 subjects are needed to achieve statistical power higher than 0.80 with a medium effect size (0.32).

All data distributions were tested for normality through the Kolmogorov-Smirnov test. 1-Way ANOVA tests were used to assess differences between each Azure Kinect setting and the Vicon. The ANOVA tests were performed to evaluate the effect of light, comparing each parameter of the "light" condition with the corresponding one of the "no-light" condition of each setting. The same tests were performed to evaluate the effect of occlusion, comparing each parameter of the "no-occlusion" condition with the corresponding one of the "occlusion" condition of each system. The alpha level of significance was set at 0.05 in all the tests. In the results section, the results of the tests between each Kinect and Vicon are identified with * (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$); the tests comparing the occlusion condition are indicated with ‡ ($^{\ddagger}p < 0.05$, $^{\ddagger\ddagger}p < 0.01$, $^{\ddagger\ddagger\ddagger}p < 0.001$); the tests comparing the light condition are indicated with $\hat{}$ ($\hat{p} < 0.05$, $\hat{\hat{}}p < 0.01$, $\hat{\hat{}}\hat{}p < 0.001$).

4. Results

The repeatability of all the parameters of the Vicon data was tested among the conditions to determine whether it could be

considered as a reference measure, without statistical differences among the trials (i.e., in all conditions, the performed movements were the same). Since this condition was met, the parameters computed from Kinect data could be compared to the Vicon.

4.1. Static measures

4.1.1. Static frontal posture

This section reports biomechanical results from the subjects performing a static frontal posture. First, the results on the normalized limb lengths and, then, the results on the articular angles are shown:

a) Limb lengths

The normalized lengths of the arm and forearm of both the right and left sides are shown (mean and standard deviation) for all the acquisition settings and in each testing condition in Fig. 3. The Kinect sensors significantly underestimated the limb lengths with respect to the Vicon system in all the conditions ($p < 0.001$). Moreover, the limb lengths of the KLat were significantly lower than the other Kinects ($p < 0.001$) in all the conditions, while the limb lengths of KFront were higher than KFrontUp and KLat45 in the "occlusion" conditions. The light condition did not significantly affect the results ($p > 0.37$), while the occlusion condition affected the results of all the Kinect settings.

b) Articular angles

Fig. 3 presents the means and standard deviations of the shoulder and elbow flexion angles in each testing condition. The Kinect cameras significantly underestimated the shoulder flexion angle ($p < 0.001$) and overestimated the elbow flexion angle ($p < 0.001$) with respect to the Vicon system in all the conditions. KFront gave the best results in both shoulder and elbow angles, showing significantly higher abduction angles and lower elbow flexion angles with respect to KFrontUp, KLat45 and KLat. The light condition did not significantly affect the results ($p > 0.28$), while the occlusion condition affected only KLat45 ($p < 0.05$) in both shoulder and elbow flexion angles.

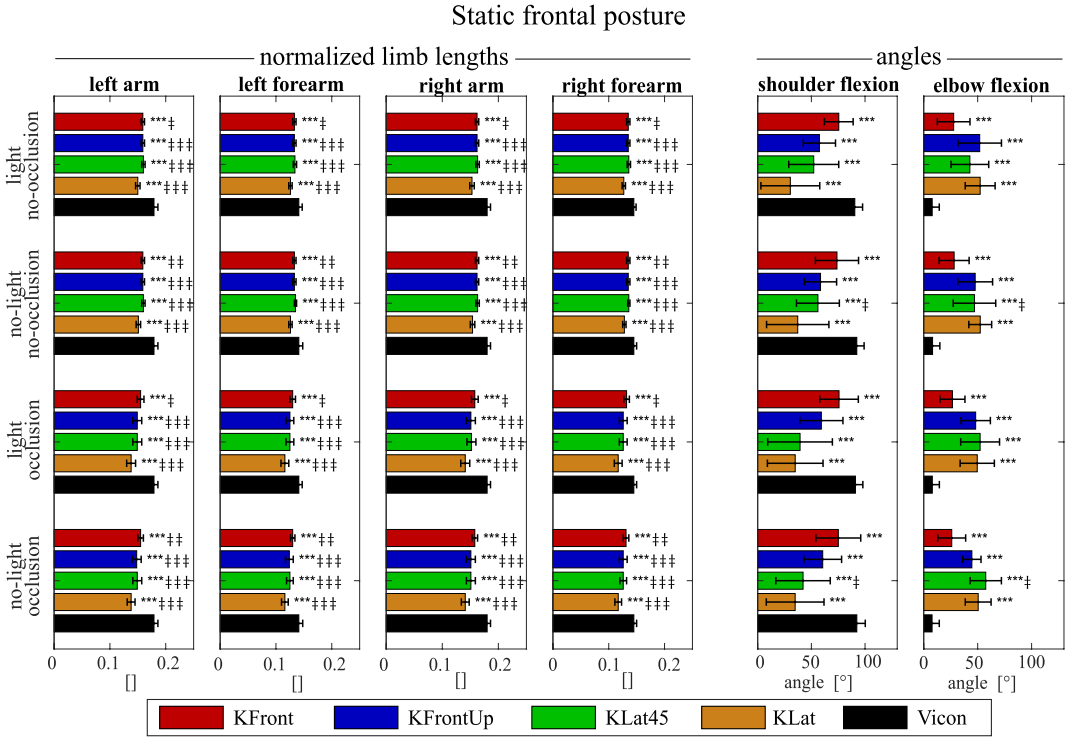


Fig. 3. Static frontal posture – Normalized limb lengths and angles. Means and standard deviations of the normalized lengths of the left arm (first column), left forearm (second column), right arm (third column) and right forearm (fourth column) are reported in the left panel. Means and standard deviations of the shoulder flexion angle and elbow flexion angle are reported in the right panel. The conditions (light/no-light, occlusion/no-occlusion) are reported in the rows. Settings are represented in different colors: KFront in red, KFrontUp in blue, KLat45 in green, KLat in orange and Vicon in black. Results of statistical analysis are reported with * for setting, ‡ for occlusion and ˆ for light. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

4.1.2. Static lateral posture

In this section, the results of the static lateral posture are reported. First, the results on the normalized limb lengths are presented and, then, the results on the articular angles are shown:

a) Limb lengths

Fig. 4 shows the means and standard deviations of the right and left sides' arm and forearm normalized lengths for all the settings and in all the testing conditions. The Kinect cameras significantly underestimated the normalized limb lengths with respect to the Vicon system in all the conditions ($p < 0.001$). Moreover, the normalized limb lengths of the KLat were significantly lower than the other Kinect settings ($p < 0.001$) in the "no-occlusion" conditions. The light condition did not significantly affect the results ($p > 0.26$), while the occlusion condition affected the results of all the Kinects.

b) Articular angles

Fig. 4 also presents the means and standard deviations of the shoulder abduction angle and elbow flexion angle in each testing condition. The Kinect cameras significantly underestimated the shoulder abduction angle and overestimated the elbow flexion angle with respect to the Vicon system in all conditions. KFront and KFrontUp gave the best results in both shoulder and elbow angles, showing significantly higher abduction angles and lower elbow flexion angles with respect to KLat45 and KLat. The light condition did not significantly affect the results ($p > 0.13$), while the occlusion condition affected only KFrontUp ($p < 0.001$) in both shoulder abduction and elbow flexion angles.

4.1.3. Static elbow flexed posture

This section presents the analysis of static elbow flexed, reporting the results on the normalized limb lengths and, then, the results on the articular angles:

a) Limb lengths

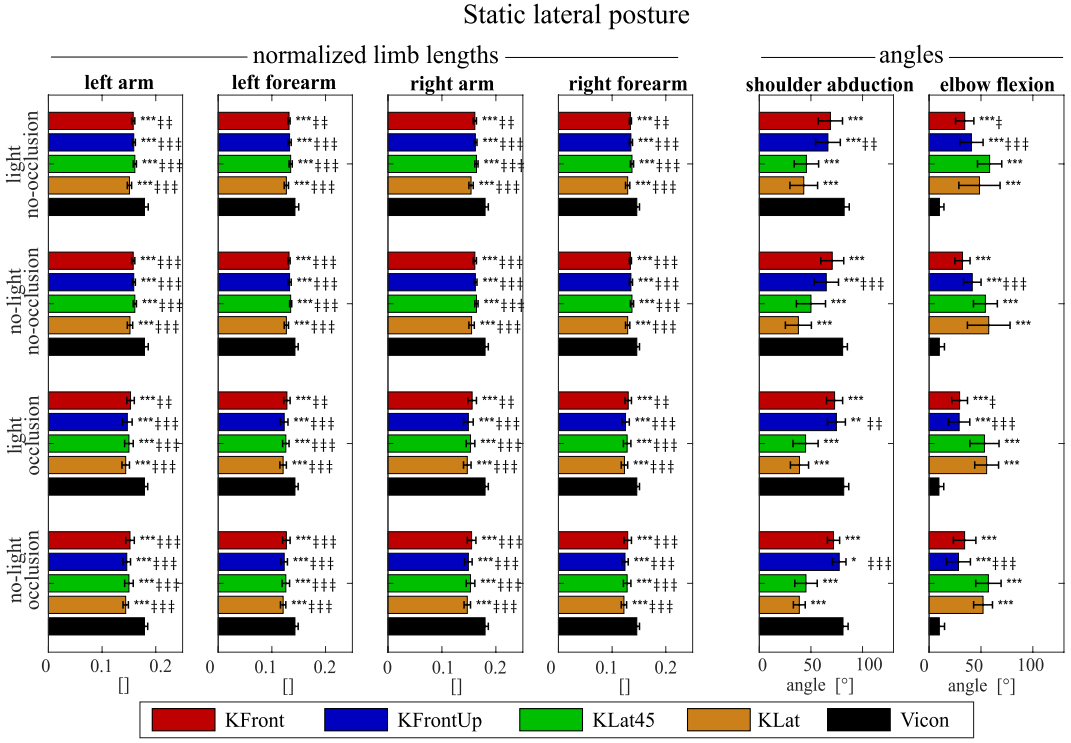


Fig. 4. Static lateral posture – Normalized limb lengths and angles. Means and standard deviations of the normalized lengths of the left arm (first column), left forearm (second column), right arm (third column) and right forearm (fourth column) are reported in the left panel. Means and standard deviations of the shoulder abduction angle and elbow flexion angle are reported in the right panel. The conditions (light/no-light, occlusion/no-occlusion) are reported in the rows. Settings are represented in different colors: KFront in red, KFrontUp in blue, KLat45 in green, KLat in orange and Vicon in black. Results of statistical analysis are reported with * for setting, ‡ for occlusion and ˆ for light. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

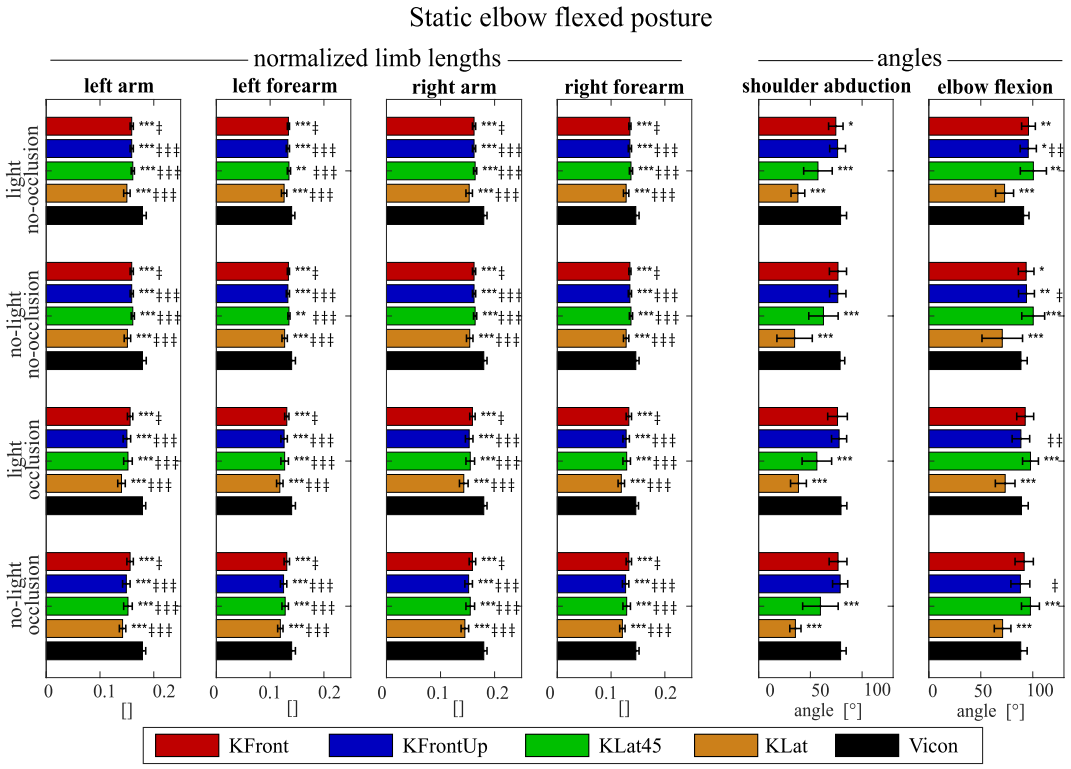


Fig. 5. Static elbow flexed posture – Normalized limb lengths and angles. Means and standard deviations of the normalized lengths of the left arm (first column), left forearm (second column), right arm (third column) and right forearm (fourth column) are reported in the left panel. Means and standard deviations of the shoulder abduction angle and elbow flexion angle are reported in the right panel. The conditions (light/no-light, occlusion/no-occlusion) are reported in the rows. Settings are represented in different colors: KFront in red, KFrontUp in blue, KLat45 in green, KLat in orange and Vicon in black. Results of statistical analysis are reported with * for setting, ‡ for occlusion and ˆ for light. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Fig. 5 shows the means and standard deviations of the normalized lengths of the arm and forearm of both the right and left sides for all the settings and in all testing conditions. The Kinect cameras significantly underestimated the normalized limb lengths with respect to the Vicon system in all the conditions ($p < 0.001$). The normalized limb lengths of the KLat were significantly lower than the other Kinects ($p < 0.001$) in all the conditions. The normalized limb lengths of KFront were higher than KFrontUp in the "light occlusion" condition ($p < 0.05$) and in the "no-light, occlusion" condition ($p < 0.01$). The light condition did not significantly affect the results ($p > 0.50$), while the occlusion condition affected the results of KFrontUp, KLat45 and KLat ($p < 0.001$) and KFront ($p < 0.01$).

b) Articular angles

Fig. 5 also presents the means and standard deviations of the shoulder abduction angle and elbow flexion angle for each testing condition. KFront and KFrontUp gave the best results with respect to the Vicon system, even if KFront significantly underestimated the shoulder abduction in the "light, no-occlusion" condition and both KFront and KFrontUp overestimated the elbow flexion in the "no-occlusion" conditions. KLat45 and KLat underestimated the shoulder abduction angle and overestimated the elbow flexion angle with respect to V, KFront and KFrontUp. The light condition did not significantly affect the results ($p > 0.11$), while the occlusion condition affected the KFrontUp in both shoulder abduction and elbow flexion angles under both "light" ($p < 0.01$) and "no-light" ($p < 0.05$) conditions.

4.2. Dynamic measures

4.2.1. Frontal reaching

In this section, the results of the frontal reaching movement are reported. First, we reported the synthetic indexes and, then, the results on the time series.

a) Synthetic indexes

Fig. 6 shows the means and the standard deviations of the parameters computed from all the settings in each testing condition. In

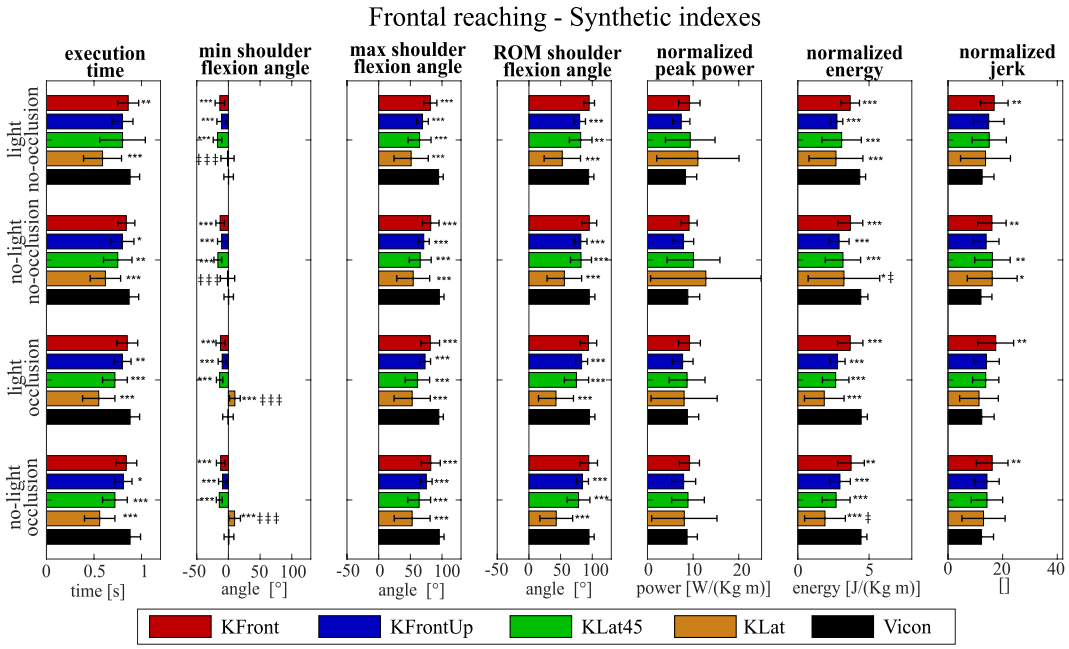


Fig. 6. Frontal reaching – Synthetic indexes. The means and the standard deviations of all the parameters are reported in columns (execution time, minimum shoulder flexion angle, maximum shoulder flexion angle, shoulder flexion range of motion, normalized shoulder peak power, normalized shoulder energy and normalized jerk). The conditions (light/no-light, occlusion/no-occlusion) are reported in the rows. Settings are represented in different colors: KFront in red, KFrontUp in blue, KLat45 in green, KLat in orange and Vicon in black. Results of statistical analysis are reported with * for setting, ‡ for occlusion and ˆ for light. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

general, the execution time of the Kinects was lower than the Vicon system and the KFront gave the best results, with no significant differences with respect to V. The KLat, instead, gave the shortest execution time. KFront, KFrontUp and KLat45 significantly underestimated both the minimum and the maximum shoulder flexion angle, but only KFrontUp and KLat45 showed a significantly lower ROM. Instead, the minimum angle of KLat was higher in the "occlusion" conditions with respect to Vicon system and in all the conditions with respect to the other Kinects; the maximum angle and the ROM were significantly reduced with respect to V and the other Kinects. The peak power showed no significant differences between the settings in all the conditions. The shoulder energy of the Kinects, instead, was significantly lower than the Vicon system in all the conditions. KLat showed the worst results with respect to KFront, KFrontUp and KLat in the "occlusion" conditions. The normalized jerk of the Vicon system was significantly lower than KFront ($p < 0.01$) in all the conditions and KLat45 ($p < 0.01$) and KLat ($p < 0.05$) in the "no-light, no-occlusion" condition. The occlusion condition had significant effects only on the maximum shoulder angle in KLat ($p < 0.001$), both with light and no-light, while the light condition did not affect the results.

b) Time series indexes

Figs. 7 and 8 show the time series of shoulder flexion angle, normalized torque, normalized power and normalized energy, averaged across subjects in all four testing conditions. The Kinect underestimated the shoulder flexion angle, torque, power and energy in every setting with respect to the Vicon system. Moreover, the time series showed higher standard deviations, while the time series of the Vicon system had lower variability between subjects. The frontal Kinect presented the best results compared to the Vicon, while the lateral Kinect showed the worst results.

The correlation coefficients between signals, comparing each Kinect to the gold standard were computed for shoulder flexion angle, torque, power and energy. The shoulder angle and energy showed the highest correlation coefficients, while the shoulder power showed the lowest ones. KLat showed lower correlation coefficients of angle, torque, power and energy with respect to the other Kinects in all the testing conditions. KFront and KFup showed the best correlation coefficients with correlation >0.80 for angle, torque and energy and >0.50 for power. The correlation coefficients of KLat45 were >0.65 for angle, torque and energy and >0.40 for power and only the correlation of the shoulder torque was lower than KFront and KFup in all the conditions.

4.2.2. Lateral reaching

In this section, the results of the lateral reaching movement are presented. First, we reported the synthetic indexes and, then, the results on the time series.

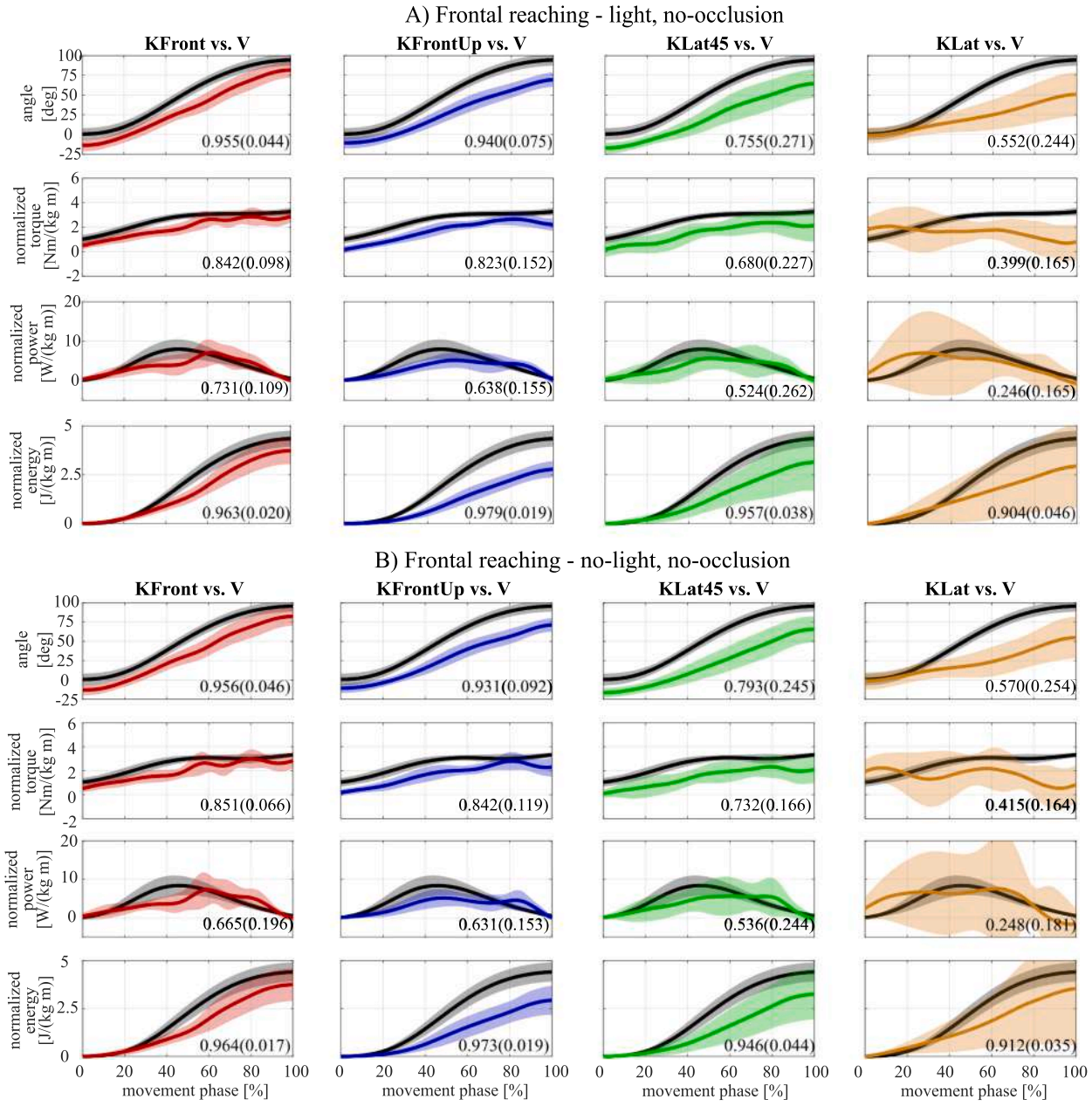


Fig. 7. Frontal reaching – Time series, no-occlusion. Panel A shows the light, no-occlusion condition. Panel B shows the no-light, no-occlusion condition. Shoulder flexion angle, normalized torque, normalized power and normalized energy are reported averaged across subjects. Solid lines are the means, while the shaded areas are the standard deviations. Signals are from the Vicon system (black), frontal Kinect (red), frontal up Kinect (blue), lateral 45 Kinect (green) and lateral Kinect (orange). Correlation coefficients (mean and standard deviations) are reported in each graph. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

a) Synthetic indexes

Fig. 9 shows the means and the standard deviations of the parameters computed from all the settings in each testing condition. Since the KLat could not track the movement well, data could not be reliably segmented. Thus, we used the phase segmentation performed on the videos of the KLat45, which was the nearest Kinect, since a reference trigger signal guarantees time synchronization between KLat and KLat45.

The KFront and KFrontUp obtained similar results in all the testing conditions, with no significant differences between the parameters. The execution time of KLat45 and KLat was significantly lower than V, KFront and KFrontUp in all the conditions, except in the "no-light, no-occlusion" condition. The minimum abduction angle of all the Kinects was higher with respect to Vicon in all the

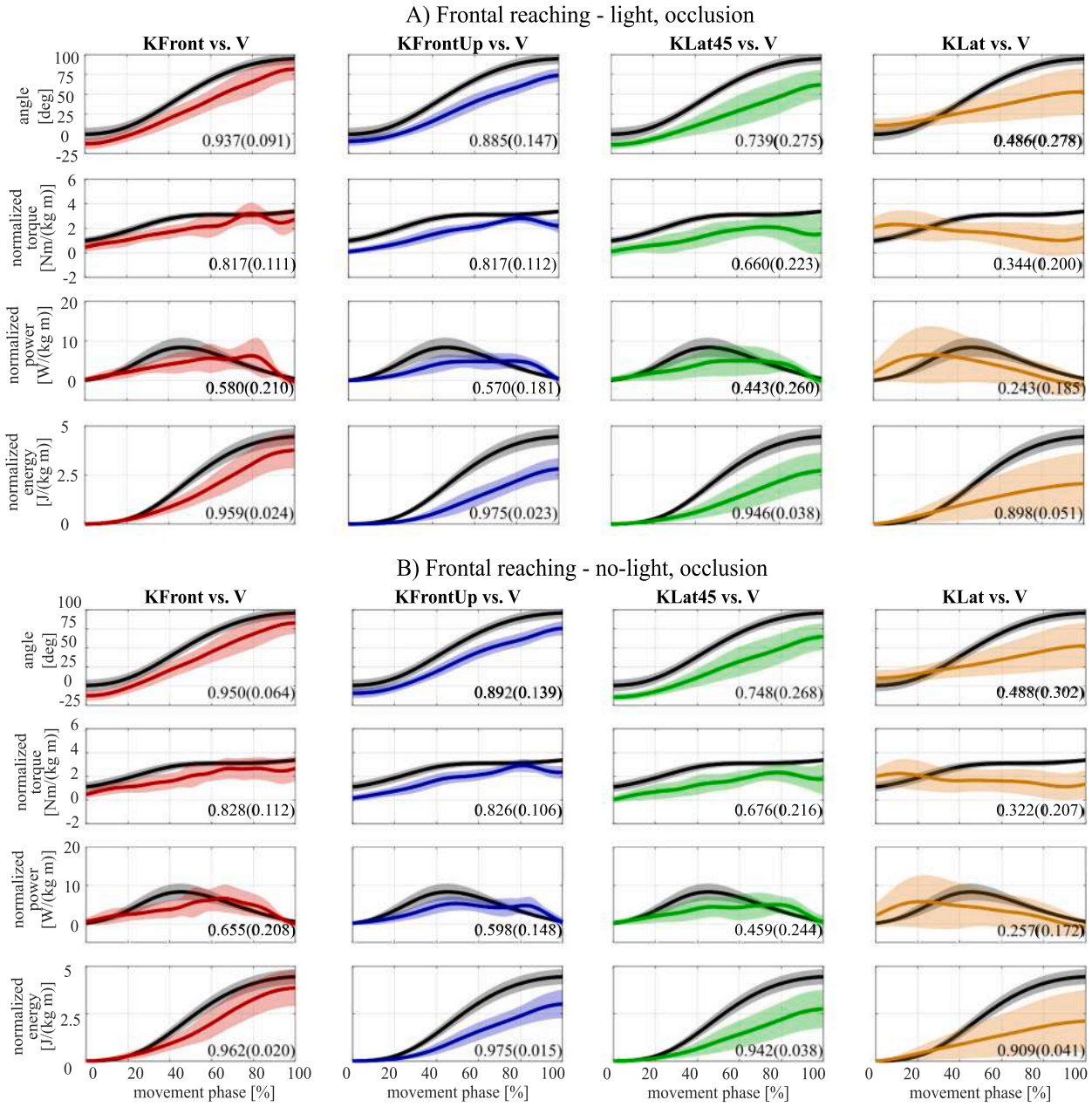


Fig. 8. Frontal reaching – Time series, occlusion. Panel A shows the light, occlusion condition. Panel B shows the no-light, occlusion condition. Time series of shoulder flexion angle, normalized torque, normalized power and normalized energy are reported averaged across subjects. Solid lines are the means, while the shaded areas are the standard deviations. Signals are from the Vicon system (black), frontal Kinect (red), frontal up Kinect (blue), lateral 45 Kinect (green) and lateral Kinect (orange). Correlation coefficients (mean and standard deviations) are reported in each graph. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

conditions and the minimum angle of KLat was higher than the other Kinect settings, except for the "no-light, no-occlusion" condition. The maximum abduction angle of KLat45 and KLat was significantly lower V, KFront and KFrontUp in all the conditions. The ROM abduction angle of the Kinects was always lower than Vicon, except for KFront in the "no-light, no-occlusion" condition and the ROM of KLat45 and KLat was lower than KFront and KFrontUp. The peak power of KLat was significantly lower than all the other settings, while the peak power of V was higher than KFront in the "light" conditions, KFrontUp in the "light, no-occlusion" condition. The shoulder energy of KLat was lower than the other settings, while the energy of KFront, KFrontUp and KLat45 was lower than V in the "light" conditions. The normalized jerk of the KFront and KFrontUp was higher than V in all the conditions except in the "no-light, no-occlusion" condition; the normalized jerk of KLat45 was higher than V in the "no-light" conditions, while the normalized jerk of KLat was higher than V in all the conditions, except in the "light, occlusion" condition. The occlusion condition had significant effects only

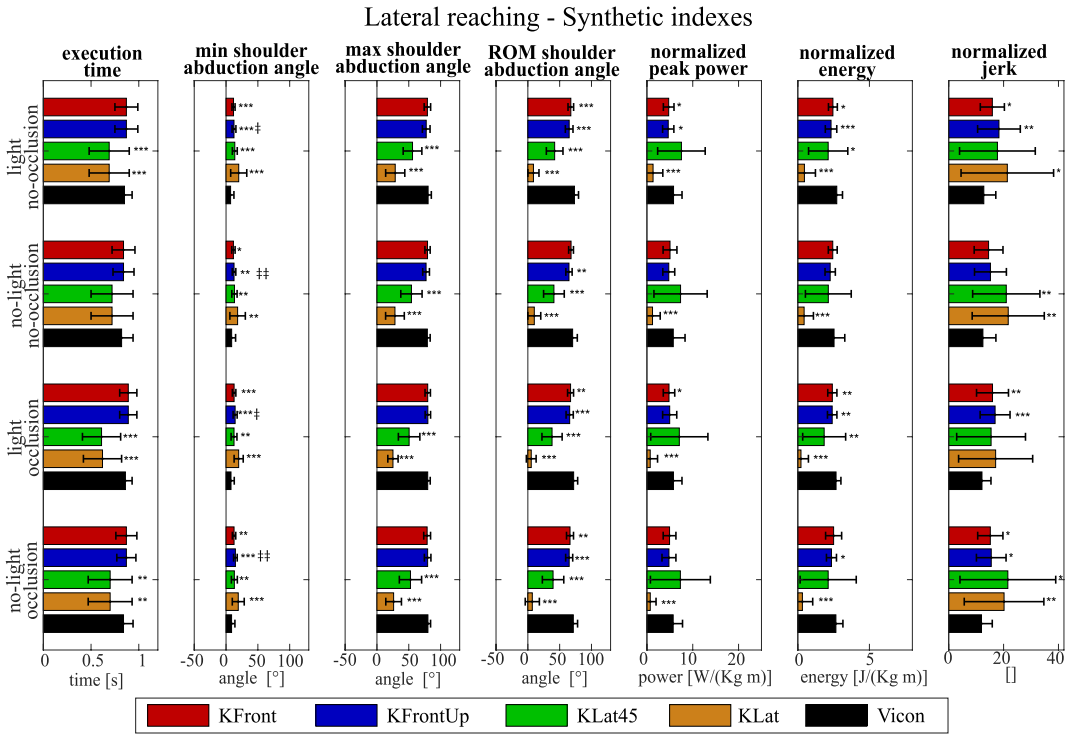


Fig. 9. Lateral reaching – Synthetic indexes. The means and the standard deviations of all the parameters are reported in columns (execution time, minimum shoulder abduction angle, maximum abduction flexion angle, shoulder abduction range of motion, normalized shoulder peak power, normalized shoulder energy and normalized jerk). The conditions (light/no-light, occlusion/no-occlusion) are reported in the rows. Settings are represented in different colors: KFront in red, KFrontUp in blue, KLat45 in green, KLat in orange and Vicon in black. Results of statistical analysis are reported with * for setting, ‡ for occlusion and ˆ for light. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

on the minimum shoulder angle in KFrontUp ($p < 0.05$), both with light and no-light, while the light condition did not affect the results.

b) Time series indexes

Figs. 10 and 11 show the time series of shoulder abduction angle, normalized torque, normalized power and normalized energy over the movement phase, averaged across subjects.

The KFront and KFrontUp tracked similarly to the Vicon system, while the KLat45 showed high variability between subjects and high levels of noise. Instead, the KLat gave the worst results, probably because it could not track the skeleton with good accuracy. The correlation coefficients between signals, comparing each Kinect to the gold standard were computed for shoulder abduction angle, torque, power and energy were considered. The shoulder angle and energy showed the highest correlation coefficients, while the shoulder torque showed the lowest ones. The KFront presented the highest correlation coefficients: angle and energy correlations were >0.90, power correlations were >0.75 and torque correlations were >0.50. The correlation coefficients of KFrontUp were similar to KFront, except for the power in the "no-occlusion" conditions. KFront and KFrontUp showed higher correlations than KLat45 and KLat in all the conditions, while KLat45 showed higher correlations than KLat in all the conditions except for the energy correlations. The occlusion condition significantly affected only the correlation of the shoulder angle of KLat45 in the "light" condition ($p < 0.01$). In contrast, the light condition had significant effects on the shoulder torque of KLat in the "no-occlusion" condition ($p < 0.05$).

5. Discussion

This study aimed to support the use of human-centered approaches based on motion tracking to evaluate and monitor people during real activities by means of a marker-less motion capture system. A novel experimental campaign was conducted involving a cohort of healthy subjects to evaluate the tracking accuracy in multiple static poses and dynamic tasks and the effects of joint tracking errors on kinematic, dynamic and motor control measures, depending on (i) the camera point of view, (ii) the light conditions and (iii) lower joint occlusions. Biomechanical parameters were computed with a biomechanical model and the analysis was conducted considering three static postures and two dynamic movements. Results are compared with those of the Vicon optoelectronic system, used as the gold standard.

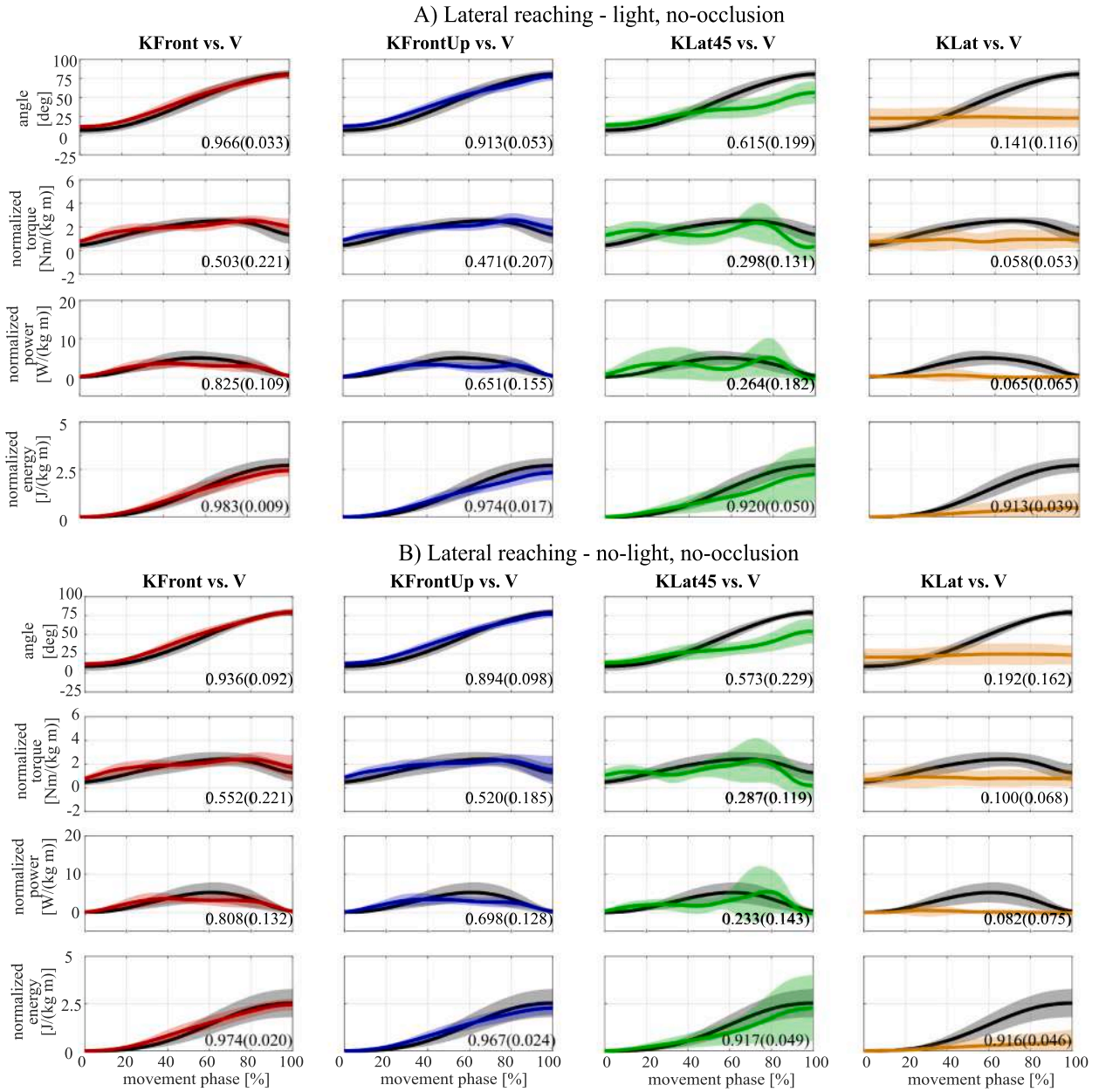


Fig. 10. Lateral reaching – Time series, no-occlusion. Panel A shows the light, no-occlusion condition. Panel B shows the no-light, no-occlusion condition. Time series of shoulder abduction angle, normalized torque, normalized power and normalized energy are reported averaged across subjects. Solid lines are the means, while the shaded areas are the standard deviations. Signals are from the Vicon system (black), frontal Kinect (red), frontal up Kinect (blue), lateral 45 Kinect (green) and lateral Kinect (orange). Correlation coefficients (mean and standard deviations) are reported in each graph. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

A first general consideration concerns all the measures provided by the Azure Kinect sensor. After the normalization of the limb lengths obtained dividing the measures by the height of each subject, the Azure Kinect underestimates the lengths. When considering articular angles, the Azure Kinect systematically underestimates the shoulder angles, and overestimates the elbow flexion angles. These results are in accordance with those obtained by previous works. The evaluation of the Kinect performance was investigated in the literature mainly on lower limbs and for estimating the spatiotemporal parameters of gait [53,54] and few studies regarding the upper limbs considered the old versions of the camera. Differences of few centimeters were found for both Kinect V1 and Kinect V2 [26]. Albert et al. [31] found that the estimation of joints provided by the Azure Kinect differed of nearly 2 cm for shoulder and elbow joints and 3 cm for the wrist with respect to marker-based tracking systems. Comparable errors in shoulder joint angles were found in upper limb reaching movements with the Kinect V1 [55] and Kinect V2 on both healthy subjects [34] and patients [56]. Moreover,

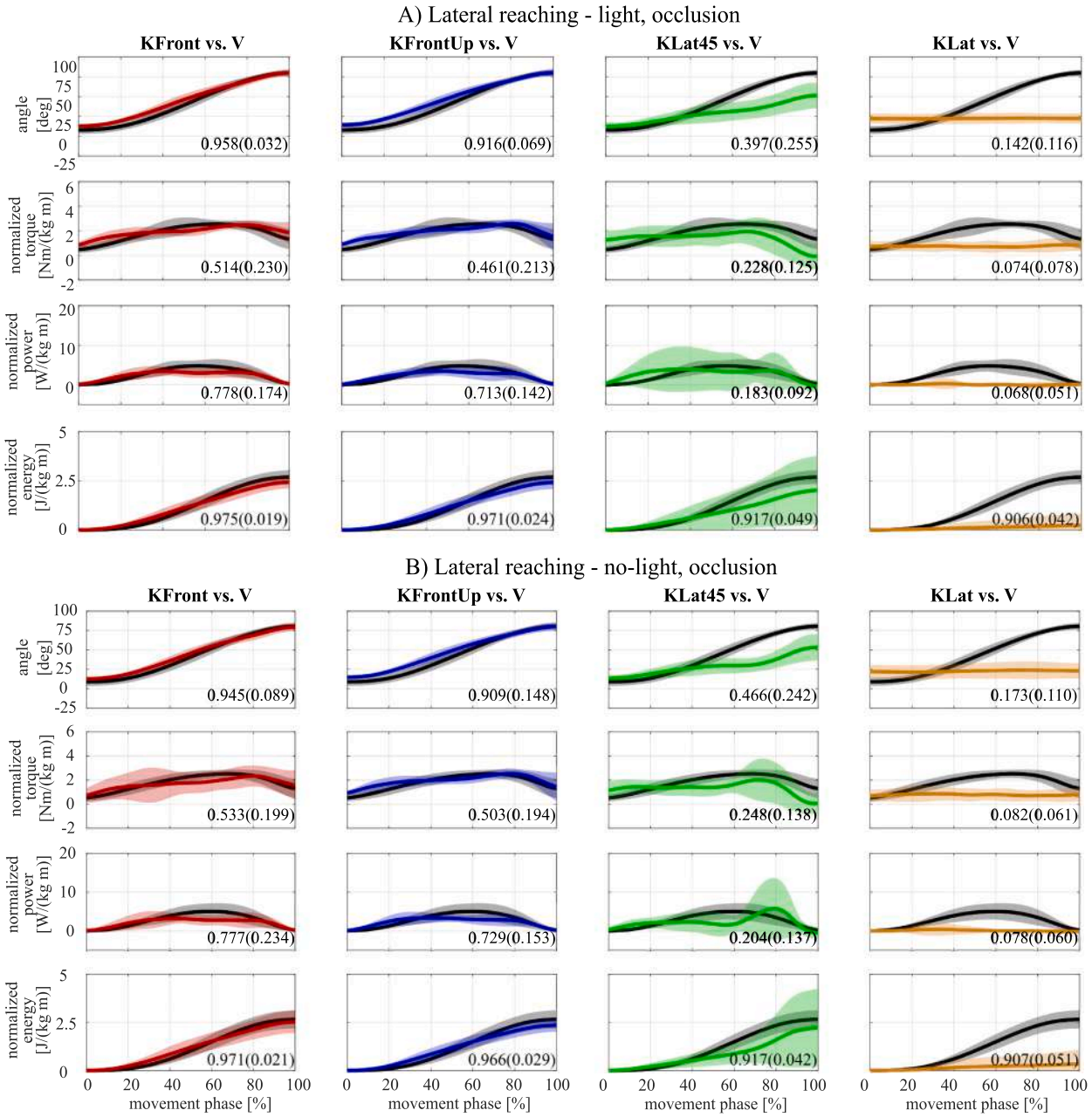


Fig. 11. Lateral reaching – Time series, occlusion. Panel A shows the light, occlusion condition. Panel B shows the no-light, occlusion condition. Time series of shoulder abduction angle, normalized torque, normalized power and normalized energy are reported averaged across. Solid lines are the means, while the shaded areas are the standard deviations. Signals are from the Vicon system (black), frontal Kinect (red), frontal up Kinect (blue), lateral 45 Kinect (green) and lateral Kinect (orange). Correlation coefficients (mean and standard deviations) are reported in each graph. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

other studies analyzed simple movements with the Kinect V1, finding an overestimation of the elbow angle [57], as in our results. Similar errors on shoulder and elbow angles were also found by Antico et al. [58] with the Azure Kinect in both frontal and lateral reaching movements.

The main reason for these different estimations is the way in which the systems provide the 3D coordinates of the articular centers. Vicon system calculates the joint centers with internal kinematic models provided by the Nexus software; on the contrary, the Kinects estimate the joints based on the RGB and depth streams, processed by the SDK body tracking. It is actually difficult to establish the exact correspondence of the detected joint positions provided by the compared sensors.

In the following parts of this section we comment and discuss the specific results obtained in the different testing conditions.

5.1. Camera positioning

In the static posture with the elbow flexed, the KFront and KFrontUp provided the most precise tracking when evaluating the limb lengths and the articular angles. In frontal reaching movements, the frontal Kinect always performed better than the other Kinects, even if it underestimated the minimum and the maximum shoulder angle, the expended energy and the normalized jerk. Both KFront and KFrontUp showed high correlation coefficients with Vicon data. The KLat had the lowest correlation with the Vicon system. In the lateral reaching movement, the KFront and KFrontUp showed the best results with low errors and high correlation coefficients.

The lateral Kinect provided the worst estimation of biomechanical parameters and the higher variability between subjects, potentially because the lateral Kinect was in an unfavorable position. For the lateral posture and the lateral reaching movement in which the body occluded the arm, the lateral Kinect was unsuitable for motion tracking. Since the camera was positioned on the contralateral side with respect to the movement execution, the arm and especially the shoulder were partially occluded also during frontal reaching task and static frontal posture [37]. In fact, the estimation of the shoulder was challenging and may have affected the results. The frontal Kinect, instead, had the best point of view for all the static postures and dynamic movements, showing the best results. Therefore, the most suitable position of the camera is in front of the subject. In this way, the camera can better capture various movements in different directions in space.

5.2. Effects of lower limb occlusion

The presence of lower limb occlusions influenced negatively the accuracy of the limb lengths of all the Kinects and the articular angles estimation only of KLat45 in the frontal posture and KFrontUp in the lateral posture and in the elbow flexed posture. In dynamic movements, the occlusion condition affected the minimum shoulder angle in KLat45 (frontal reaching) and in KFrontUp (lateral reaching) and the correlation coefficients of the shoulder angle of KFront (frontal reaching) and KLat45 (lateral reaching). Therefore, the best configuration is the one with no occlusions. In case human tracking is needed in which lower limbs are obstructed due to environmental encumbrances, occlusions might be tolerated or not depending on the applications, and their quantitative effects are provided in detail in this study. Furthermore, other camera angles can be considered to reduce occlusions and better capture movement features. For many assessed conditions, the frontal up Kinect gave similar results to the frontal camera. These results could be exploited if the set-up does not allow the camera to be placed frontally to the user. Other studies found that partial occlusion of body parts can increase the instability of the estimation of joints that are far from the occlusion with Kinect V1 [33], Kinect V2 [32,34] and Azure Kinect [35].

5.3. Effects of light condition

Interestingly, the light condition did not affect the results, except for just one case in very unfavorable tracking conditions (that showed low correlation of shoulder torque of KLat in the lateral reaching). The Kinect cameras are suitable for tracking movement in settings where the light intensity cannot be easily controlled (as long as the Kinect is not exposed to direct sunlight), as in rehabilitation at home, in clinical scenarios or industrial contexts, without altering the biomechanical assessment. In our configuration, in which the distance of the subject from the sensor about 2 m, no significant influence of the light condition was found, in accordance with Romeo et al. [35], who found that light exposure could affect the tracking accuracy especially when the distance from the Kinect camera is higher than 2 m.

5.4. Feasibility of Azure Kinects in real life scenarios

Our study included a multi-factorial assessment that considered not only a standard kinematic analysis but also a biomechanical assessment with dynamic and motor control parameters. In the literature, very few studies investigated the effect of the tracking accuracy of the V1 and V2 Kinect cameras on biomechanical parameters, such as joint torques [59] and normalized jerk [55]. These analyses provided a complete characterization of the subject, and includes parameters relevant for medium-term monitoring of subjects, including identifying fatigue and monitoring mental health conditions during repetitive activities and working cycles. Moreover, the computation of dynamic parameters allows the evaluation of the risks of musculoskeletal injuries. Furthermore, the wide spectrum of analyzed parameters makes the employed protocol useable for other research domains and on-field applications. Indeed, biomechanical and motor control assessments are employed in clinical environments to evaluate neurological patients' clinical course of pathology and the effects of rehabilitative therapy [60–62].

The results found in our study demonstrate that the Azure Kinect could be acceptable in industrial applications for the biomechanical assessment of the worker or in clinical environments for the evaluation of motor performance, since the biomechanical parameters were well correlated with the gold standard measures. Azure Kinect cameras could also be employed for obstacle avoidance in human-robot collaboration, but the differences in the estimation of limb lengths and articular angles with respect to the Vicon reference should be considered while implementing the algorithms.

5.5. Future works

Our study showed that marker-less sensors are already applicable in the industrial environment for the biomechanical evaluation of workers [10]. Moreover, both static postures and dynamic movements were analyzed, with lower body joint occlusions and different

lighting conditions. Our protocol includes all realistic environmental conditions that can usually be found in real contexts, where the occlusion regards principally the lower limbs and the lighting depends on artificial light. Due to the numerous combinations of different conditions (multiple Kinects, light, occlusion), we had to limit the number of tasks to propose a reasonable protocol in terms of time and number of tasks. Future applications may involve a wide variety of tasks and more extreme light conditions.

Further analyses are necessary to assess the performance of the Azure Kinects for the segmentation and recognition of more complex movements. In this study, the experiments demonstrated that the Azure Kinects could be suitable for tracking movements with a large range of motion and with clear distinct phases. Further investigations are required for fine and small movements, as they were not correctly identified with previous Kinect versions [26].

Finally, starting from the results of this paper on the camera positioning, future works could include the data fusion between multiple cameras monitoring the same volume, as in Refs. [63,64], to improve the overall tracking accuracy [65], or new similar devices, such as Orbbec Gemini 2, could be analyzed to evaluate the improvements in tracking accuracy.

6. Conclusions

This study provided a comprehensive analysis of the performance of the Azure Kinect camera for upper limb tracking, comparing a large set of kinematic and dynamic parameters on both static postures and dynamic movements. Moreover, we analyzed different testing conditions that could influence the accuracy of the Kinect camera, such as camera point of view, light intensity and presence of occlusions. Our study demonstrated that the frontal position is the most reliable; finds almost no difference when comparing artificial lightning with the no-light condition; quantifies the worsening of the performance in tracking the upper limbs when lower limbs are occluded. In our opinion, the main result of the proposed study is the demonstration, by statistical analyses and comparisons of results, that a low-cost and non-invasive sensor can be used in many medical and industrial applications. Measures provided in static and dynamic tasks and in derived biomechanical measures are stable and reliable, entirely in line with those provided by expensive and invasive optoelectronic systems.

Author contribution statement

Cristina Brambilla: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Roberto Marani: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Laura Romeo: Contributed reagents, materials, analysis tools or data; Wrote the paper.

Matteo Lavit Nicora: Contributed reagents, materials, analysis tools or data; Wrote the paper.

Fabio A. Storm: Conceived and designed the experiments; Wrote the paper.

Gianluigi Reni: Conceived and designed the experiments; Wrote the paper.

Matteo Malosio: Conceived and designed the experiments; Wrote the paper.

Tiziana D'Orazio: Analyzed and interpreted the data; Conceived and designed the experiments; Wrote the paper.

Alessandro Scano: Conceived and designed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Data availability statement

Data will be made available on request.

Ethics statement

Ethical approval was granted by the referring ethical committee (Approval Prot. N. 19/20 – CE, April 20th, 2020).

Funding

This research received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 847926.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Computation of dynamic and motor control parameters

The power $P_i(t)$ exerted at the joint level was calculated as in (1):

$$P_i(t) = \overline{\tau}_i(t)\omega_i(t) \tag{1}$$

where $\bar{\tau}_i(t)$ is the torque normalized to the length of the subject's arm (La) and weight (Ma) and $\omega_i(t)$ the angular velocity of the joint. The peak power was taken as the maximum power P_{MAX}.

The expended energy E_i is the integral of the power exerted at the joint level during the execution of the task, computed as in (2):

$$E_i = \int_{t_0}^{t_{end}} P_i(t)\, dt \tag{2}$$

Where t_0 and t_{end} are the initial and the ending time of the task, and $P_i(t)$ is the power time course. The dynamic quantities were normalized in order to allow inter-subject analysis.

The normalized jerk NJ [49] was used to evaluate the smoothness of the movement and was computed as in (3):

$$NJ = \sqrt{\frac{1}{2}\frac{t_{tot}^5}{L^2}\int j^2\, dt} \tag{3}$$

where t_{tot} is the task execution time, j is the third derivative of the 3D position of the wrist and L is the length of the wrist trajectory during the execution of the task.

References

[1] H. Zhou, H. Hu, Human motion tracking for rehabilitation-A survey, Biomed. Signal Process Control 3 (2008) 1–18, https://doi.org/10.1016/j.bspc.2007.09.001.

[2] S.R. Simon, Quantification of human motion: gait analysis - benefits and limitations to its application to clinical problems, J. Biomech. 37 (2004) 1869–1880, https://doi.org/10.1016/j.jbiomech.2004.02.047.

[3] M.H. Khan, M. Zöller, M.S. Farid, M. Grzegorzek, Marker-based movement analysis of human body parts in therapeutic procedure, Sensors 20 (2020), https://doi.org/10.3390/s20113312, 1–19.

[4] M.L. Lee, W. Liu, S. Behdad, X. Liang, M. Zheng, Robot-Assisted disassembly sequence planning with real-time human motion prediction, IEEE Trans. Syst. Man, Cybern. Syst. (2022), https://doi.org/10.1109/TSMC.2022.3185889.

[5] M.N.H. Yunus, M.H. Jaafar, A.S.A. Mohamed, N.Z. Azraai, M.S. Hossain, Implementation of kinetic and kinematic variables in ergonomic risk assessment using motion capture simulation: a review, Int. J. Environ. Res. Publ. Health 18 (2021), https://doi.org/10.3390/ijerph18168342.

[6] D. Panariello, S. Grazioso, T. Caporaso, A. Palomba, G. Di Gironimo, A. Lanzotti, Biomechanical analysis of the upper body during overhead industrial tasks using electromyography and motion capture integrated with digital human models, Int. J. Interact. Des. Manuf. 16 (2022) 733–752, https://doi.org/10.1007/s12008-022-00862-9.

[7] E.M. Argyle, A. Marinescu, M.L. Wilson, G. Lawson, S. Sharples, Physiological indicators of task demand, fatigue, and cognition in future digital manufacturing environments, Int. J. Hum. Comput. Stud. 145 (2021), 102522, https://doi.org/10.1016/j.ijhcs.2020.102522.

[8] L. Gualtieri, I. Palomba, F.A. Merati, E. Rauch, R. Vidoni, Design of human-centered collaborative assembly workstations for the improvement of operators' physical ergonomics and production efficiency: a case study, Sustain. Times 12 (2020) 3606, https://doi.org/10.3390/su12093606.

[9] M. Lagomarsino, M. Lorenzini, P. Balatti, E. De Momi, A. Ajoudani, Pick the right Co-worker: online assessment of cognitive ergonomics in human-robot collaborative assembly, IEEE Trans. Cogn. Dev. Syst. (2022), https://doi.org/10.1109/TCDS.2022.3182811.

[10] M. Lavit Nicora, E. Andre, D. Berkmans, C. Carissoli, T. D'Orazio, A.D. Fave, P. Gebhard, R. Marani, R.M. Mira, L. Negri, F. Nunnari, A.P. Fernandez, A. Scano, G. Reni, M. Malosio, A human-driven control architecture for promoting good mental health in collaborative robot scenarios, in: 2021 30th IEEE Int. Conf. Robot Hum. Interact. Commun. RO-MAN 2021, Institute of Electrical and Electronics Engineers (IEEE), 2021, pp. 285–291, https://doi.org/10.1109/RO-MAN50785.2021.9515315.

[11] A. Ranavolo, G. Chini, F. Draicchio, A. Silvetti, T. Varrecchia, L. Fiori, A. Tatarelli, P.H. Rosen, S. Wischniewski, P. Albrecht, L. Vogt, M. Bianchi, G. Averta, A. Cherubini, L. Fritzsche, M. Sartori, B. Vanderborght, R. Govaerts, A. Ajoudani, Human-Robot Collaboration (HRC) Technologies for Reducing Work-Related Musculoskeletal Diseases in Industry 4.0, Lect, 223 LNNS, Notes Networks Syst, 2022, pp. 335–342, https://doi.org/10.1007/978-3-030-74614-8_40/COVER.

[12] A.I. Cuesta-Vargas, A. Galán-Mercant, J.M. Williams, The use of inertial sensors system for human motion analysis, Phys. Ther. Rev. 15 (2010) 462–473, https://doi.org/10.1179/1743288X11Y.0000000006.

[13] S.L. Colyer, M. Evans, D.P. Cosker, A.I.T. Salo, A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system, Sport. Med. - Open. 4 (2018) 1–15, https://doi.org/10.1186/s40798-018-0139-y.

[14] N.K. Mangal, A.K. Tiwari, A review of the evolution of scientific literature on technology-assisted approaches using RGB-D sensors for musculoskeletal health monitoring, Comput. Biol. Med. 132 (2021), 104316, https://doi.org/10.1016/j.compbiomed.2021.104316.

[15] W.W.T. Lam, Y.M. Tang, K.N.K. Fong, A systematic review of the applications of markerless motion capture (MMC) technology for clinical measurement in rehabilitation, J. NeuroEng. Rehabil. 20 (2023), https://doi.org/10.1186/s12984-023-01186-9.

[16] B. Milosevic, A. Leardini, E. Farella, Kinect and wearable inertial sensors for motor rehabilitation programs at home: state of the art and an experimental comparison, Biomed. Eng. Online 19 (2020) 1–26, https://doi.org/10.1186/s12938-020-00762-7.

[17] R.A. Clark, Y.H. Pua, K. Fortin, C. Ritchie, K.E. Webster, L. Denehy, A.L. Bryant, Validity of the Microsoft Kinect for assessment of postural control, Gait Posture 36 (2012) 372–377, https://doi.org/10.1016/j.gaitpost.2012.03.033.

[18] R.A. Clark, Y.H. Pua, C.C. Oliveira, K.J. Bower, S. Thilarajah, R. McGaw, K. Hasanki, B.F. Mentiplay, Reliability and concurrent validity of the Microsoft Xbox One Kinect for assessment of standing balance and postural control, Gait Posture 42 (2015) 210–213, https://doi.org/10.1016/j.gaitpost.2015.03.005.

[19] M. Capecci, M.G. Ceravolo, F.F. Ferracuti, S. Iarlori, S. Longhi, L. Romeo, S.N. Russi, F. Verdini, Accuracy evaluation of the Kinect v2 sensor during dynamic movements in a rehabilitation scenario, in: Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, Annu Int Conf IEEE Eng Med Biol Soc, 2016, pp. 5409–5412, https://doi.org/10.1109/EMBC.2016.7591950.

[20] M. Eltoukhy, J. Oh, C. Kuenze, J. Signorile, Improved kinect-based spatiotemporal and kinematic treadmill gait assessment, Gait Posture 51 (2017) 77–83, https://doi.org/10.1016/j.gaitpost.2016.10.001.

[21] T.N. Nguyen, H.H. Huynh, J. Meunier, Human gait symmetry assessment using a depth camera and mirrors, Comput. Biol. Med. 101 (2018) 174–183, https://doi.org/10.1016/j.compbiomed.2018.08.021.

[22] P.F. Dajime, H. Smith, Y. Zhang, Automated classification of movement quality using the Microsoft Kinect V2 sensor, Comput. Biol. Med. 125 (2020), 104021, https://doi.org/10.1016/j.compbiomed.2020.104021.

[23] B.F. Mentiplay, L.G. Perraton, K.J. Bower, Y.H. Pua, R. McGaw, S. Heywood, R.A. Clark, Gait assessment using the Microsoft Xbox One Kinect: concurrent validity and inter-day reliability of spatiotemporal and kinematic variables, J. Biomech. 48 (2015) 2166–2170, https://doi.org/10.1016/j.jbiomech.2015.05.021.

[24] M. Wochatz, N. Tilgner, S. Mueller, S. Rabe, S. Eichler, M. John, H. Völler, F. Mayer, Reliability and validity of the Kinect V2 for the assessment of lower extremity rehabilitation exercises, Gait Posture 70 (2019) 330–335, https://doi.org/10.1016/j.gaitpost.2019.03.020.

[25] G. Faity, D. Mottet, J. Froger, Validity and reliability of kinect v2 for quantifying upper body kinematics during seated reaching, Sensors 22 (2022), https://doi.org/10.3390/s22072735.

[26] K. Otte, B. Kayser, S. Mansow-Model, J. Verrel, F. Paul, A.U. Brandt, T. Schmitz-Hübsch, Accuracy and reliability of the kinect version 2 for clinical measurement of motor function, PLoS One 11 (2016), https://doi.org/10.1371/journal.pone.0166532.

[27] B. Galna, G. Barry, D. Jackson, D. Mhiripiri, P. Olivier, L. Rochester, Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease, Gait Posture 39 (2014) 1062–1068, https://doi.org/10.1016/j.gaitpost.2014.01.008.

[28] Y. Ma, B. Sheng, R. Hart, Y. Zhang, The validity of a dual Azure Kinect-based motion capture system for gait analysis: a preliminary study, in: 2020 Asia-Pacific Signal Inf. Process, Assoc. Annu. Summit Conf. APSIPA ASC 2020 - Proc., 2020, pp. 1201–1206.

[29] J. Thomas, J.B. Hall, R. Bliss, T.M. Guess, Comparison of Azure Kinect and optical retroreflective motion capture for kinematic and spatiotemporal evaluation of the sit-to-stand test, Gait Posture 94 (2022) 153–159, https://doi.org/10.1016/j.gaitpost.2022.03.011.

[30] T.M. Guess, R. Bliss, J.B. Hall, A.M. Kiselica, Comparison of Azure Kinect overground gait spatiotemporal parameters to marker based optical motion capture, Gait Posture 96 (2022) 130–136, https://doi.org/10.1016/j.gaitpost.2022.05.021.

[31] J.A. Albert, V. Owolabi, A. Gebel, C.M. Brahms, U. Granacher, B. Arnrich, Evaluation of the pose tracking performance of the azure kinect and kinect v2 for gait analysis in comparison with a gold standard: a pilot study, Sensors 20 (2020) 1–22, https://doi.org/10.3390/s20185104.

[32] L. Cai, Y. Ma, S. Xiong, Y. Zhang, Validity and reliability of upper limb functional assessment using the Microsoft kinect V2 sensor, Appl. Bionics Biomech. 2019 (2019), https://doi.org/10.1155/2019/7175240.

[33] R.P. Kuster, B. Heinlein, C.M. Bauer, E.S. Graf, Accuracy of KinectOne to quantify kinematics of the upper body, Gait Posture 47 (2016) 80–85, https://doi.org/10.1016/j.gaitpost.2016.04.004.

[34] A. Scano, R.M. Mira, P. Cerveri, L.M. Tosatti, M. Sacco, Analysis of upper-limb and trunk kinematic variability: accuracy and reliability of an RGB-D sensor, Multimodal Technol. Interact. 4 (2020) 14, https://doi.org/10.3390/mti4020014.

[35] L. Romeo, R. Marani, M. Malosio, A.G. Perri, T. D'Orazio, Performance Analysis of Body Tracking with the Microsoft Azure Kinect, Institute of Electrical and Electronics Engineers (IEEE), 2021, pp. 572–577, https://doi.org/10.1109/med51440.2021.9480177.

[36] H. Sarbolandi, D. Lefloch, A. Kolb, Kinect range sensing: structured-light versus time-of-flight kinect, Comput. Vis. Image Underst. 139 (2015) 1–20, https://doi.org/10.1016/j.cviu.2015.05.006.

[37] L. Cai, D. Liu, Y. Ma, Placement recommendations for single kinect-based motion capture system in unilateral dynamic motion analysis, Healthc 9 (2021), https://doi.org/10.3390/healthcare9081076.

[38] L.F. Yeung, Z. Yang, K.C.C. Cheng, D. Du, R.K.Y. Tong, Effects of camera viewing angles on tracking kinematic gait patterns using Azure Kinect, Kinect v2 and Orbbec Astra Pro v2, Gait Posture 87 (2021) 19–26, https://doi.org/10.1016/j.gaitpost.2021.04.005.

[39] A. Lunin, C.H. Glock, Systematic review of Kinect-based solutions for physical risk assessment in manual materials handling in industrial and laboratory environments, Comput. Ind. Eng. 162 (2021), 107660, https://doi.org/10.1016/j.cie.2021.107660.

[40] R. Etzi, S. Huang, G.W. Scurati, S. Lyu, F. Ferrise, A. Gallace, A. Gaggioli, A. Chirico, M. Carulli, M. Bordegoni, Using virtual reality to test human-robot interaction during a collaborative task, in: Proc. ASME Des. Eng. Tech. Conf., American Society of Mechanical Engineers (ASME), 2019, https://doi.org/10.1115/DETC2019-97415.

[41] G. Cicirelli, R. Marani, L. Romeo, M.G. Domínguez, J. Heras, A.G. Perri, T. D'Orazio, The HA4M dataset: multi-modal monitoring of an assembly task for human action recognition in manufacturing, 2022 91, Sci. Data 9 (2022) 1–12, https://doi.org/10.1038/s41597-022-01843-z.

[42] P.P. Rickham, Human experimentation code of ethics of the world medical association, Br. Med. J. 2 (1964) 177, https://doi.org/10.1136/bmj.2.5402.177.

[43] Nexus 2.12 Documentation - Vicon Documentation, Vicon Motion Syst. Ltd., 2016. https://docs.vicon.com/display/Nexus214?preview=/83296552/83296566/Model_UpperLimb_ProductGuide_Rev1.0_2007Jul.pdf. (Accessed 13 December 2022).

[44] E. Ceseracciu, Z. Sawacha, C. Cobelli, Comparison of markerless and marker-based motion capture technologies through simultaneous data collection during gait: proof of concept, PLoS One 9 (2014), e87640, https://doi.org/10.1371/JOURNAL.PONE.0087640.

[45] F.N. Fritsch, R.E. Carlson, Monotone piecewise cubic interpolation, SIAM J. Numer. Anal. 17 (1980) 238–246, https://doi.org/10.1137/0717021.

[46] C. Brambilla, M. Malosio, G. Reni, A. Scano, Optimal biomechanical performance in upper-limb gestures depends on velocity and carried load, Biology 11 (2022), https://doi.org/10.3390/BIOLOGY11030391.

[47] R. Dumas, L. Chèze, 3D inverse dynamics in non-orthonormal segment coordinate system, Med. Biol. Eng. Comput. 45 (2007) 315–322, https://doi.org/10.1007/s11517-006-0156-8.

[48] J.M. Christensen, J.W. McBarron, J.T. McConville, W.R. Pogue, R.C. Williges, W.E. Woodson, Man-Systems Integration Standards NASA-STD-3000, 1995.

[49] H.L. Teulings, J.L. Contreras-Vidal, G.E. Stelmach, C.H. Adler, Parkinsonism reduces coordination of fingers, wrist, and arm in fine motor control, Exp. Neurol. 146 (1997) 159–170, https://doi.org/10.1006/exnr.1997.6507.

[50] M.K.Y. Mak, O. Levin, J. Mizrahi, C.W.Y. Hui-Chan, Joint torques during sit-to-stand in healthy subjects and people with Parkinson's disease, Clin. Biomech. 18 (2003) 197–206, https://doi.org/10.1016/S0268-0033(02)00191-2.

[51] N. Sylla, V. Bonnet, F. Colledani, P. Fraisse, Ergonomic contribution of ABLE exoskeleton in automotive industry, Int. J. Ind. Ergon. 44 (2014) 475–481, https://doi.org/10.1016/j.ergon.2014.03.008.

[52] F. Faul, E. Erdfelder, A.G. Lang, A. Buchner, G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences, in: Behav. Res. Methods, Behav Res Methods, 2007, pp. 175–191, https://doi.org/10.3758/BF03193146.

[53] J. Latorre, R. Llorens, C. Colomer, M. Alcañiz, Reliability and comparison of Kinect-based methods for estimating spatiotemporal gait parameters of healthy and post-stroke individuals, J. Biomech. 72 (2018) 268–273, https://doi.org/10.1016/j.jbiomech.2018.03.008.

[54] A. Pfister, A.M. West, S. Bronner, J.A. Noah, Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis, J. Med. Eng. Technol. 38 (2014) 274–280, https://doi.org/10.3109/03091902.2014.909540.

[55] A. Scano, M. Caimmi, M. Malosio, L.M. Tosatti, Using Kinect for upper-limb functional evaluation in home rehabilitation: a comparison with a 3D stereoscopic passive marker system, in: Proc. IEEE RAS EMBS Int. Conf. Biomed. Robot. Biomechatronics, IEEE Computer Society, 2014, pp. 561–566, https://doi.org/10.1109/biorob.2014.6913837.

[56] A. Scano, A. Chiavenna, M. Malosio, L. Molinari Tosatti, F. Molteni, Kinect V2 implementation and testing of the reaching performance scale for motor evaluation of patients with neurological impairment, Med. Eng. Phys. 56 (2018) 54–58, https://doi.org/10.1016/j.medengphy.2018.04.005.

[57] B. Bonnechère, B. Jansen, P. Salvia, H. Bouzahouene, L. Omelina, F. Moiseev, V. Sholukha, J. Cornelis, M. Rooze, S. Van Sint Jan, Validity and reliability of the Kinect within functional assessment activities: comparison with standard stereophotogrammetry, Gait Posture 39 (2014) 593–598, https://doi.org/10.1016/j.gaitpost.2013.09.018.

[58] M. Antico, N. Balletti, G. Laudato, A. Lazich, M. Notarantonio, R. Oliveto, S. Ricciardi, S. Scalabrino, J. Simeone, Postural control assessment via Microsoft azure kinect DK: an evaluation study, comput, Methods Programs Biomed 209 (2021), https://doi.org/10.1016/j.cmpb.2021.106324.

[59] P. Plantard, A. Muller, C. Pontonnier, G. Dumont, H.P.H. Shum, F. Multon, Inverse dynamics based on occlusion-resistant Kinect data: is it useable for ergonomics? Int. J. Ind. Ergon. 61 (2017) 71–80, https://doi.org/10.1016/j.ergon.2017.05.010.

[60] A. Scano, E. Guanziroli, R.M. Mira, C. Brambilla, L. Molinari Tosatti, F. Molteni, Biomechanical assessment of the ipsilesional upper limb in post-stroke patients during multi-joint reaching tasks: a quantitative study, Front. Rehabil. Sci. 3 (2022), https://doi.org/10.3389/fresc.2022.943397.

[61] M.L. Schiefelbein, A.P. Salazar, R.R. Marchese, K.D. Rech, G.P. Schifino, C.S. Figueiredo, V. Cimolin, A.S. Pagnussat, Upper-limb movement smoothness after stroke and its relationship with measures of body function/structure and activity – a cross-sectional study, J. Neurol. Sci. 401 (2019) 75–78, https://doi.org/10.1016/j.jns.2019.04.017.
[62] J. Van Kordelaar, E. Van Wegen, G. Kwakkel, Impact of time on quality of motor control of the paretic upper limb after stroke, Arch. Phys. Med. Rehabil. 95 (2014) 338–344, https://doi.org/10.1016/j.apmr.2013.10.006.
[63] H. Du, Y. Zhao, J. Han, Z. Wang, G. Song, Data fusion of multiple kinect sensors for a rehabilitation system, in: Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, Annu Int Conf IEEE Eng Med Biol Soc, 2016, pp. 4869–4872, https://doi.org/10.1109/EMBC.2016.7591818.
[64] S. Hazra, A.A. Pratap, D. Tripathy, A. Nandy, Novel data fusion strategy for human gait analysis using multiple kinect sensors, Biomed. Signal Process Control 67 (2021), 102512, https://doi.org/10.1016/j.bspc.2021.102512.
[65] J.C. Núñez, R. Cabido, A.S. Montemayor, J.J. Pantrigo, Real-time human body tracking based on data fusion from multiple RGB-D sensors, Multimed. Tools Appl. 76 (2017) 4249–4271, https://doi.org/10.1007/s11042-016-3759-6.