Inverse dynamics based on occlusion-resistant Kinect data: Is it usable for ergonomics?

Pierre Plantard [a, b, f, *], Antoine Muller [b, c], Charles Pontonnier [b, c, d], Georges Dumont [b, c], Hubert P.H. Shum [e], Franck Multon [b, f]

[a] FAURECIA Automotive Seating, Etampes, France
[b] IRISA/M2S/INRIA MimeTIC, Rennes, France
[c] ENS Rennes, Bruz, France
[d] Ecoles de Saint Cyr Coëtquidan, Guer, France
[e] Northumbria University, Newcastle upon Tyne, UK
[f] Université de Rennes 2, Rennes, France

ARTICLE INFO

ABSTRACT

Joint torques and forces are relevant quantities to estimate the biomechanical constraints of working tasks in ergonomics. However, inverse dynamics requires accurate motion capture data, which are generally not available in real manufacturing plants. Markerless and calibrationless measurement systems based on depth cameras, such as the Microsoft Kinect, are promising means to measure 3D poses in real time. Recent works have proposed methods to obtain reliable continuous skeleton data in cluttered environments, with occlusions and inappropriate sensor placement. In this paper, we evaluate the reliability of an inverse dynamics method based on this corrected skeleton data and its potential use to estimate joint torques and forces in such cluttered environments. To this end, we compared the calculated joint torques with those obtained with a reference inverse dynamics method based on an opto-electronic motion capture system. Results show that the Kinect skeleton data enabled the inverse dynamics process to deliver reliable joint torques in occlusion-free (r = 0.99 for the left shoulder elevation) and occluded (r = 0.91 for the left shoulder elevation) environments. However, differences remain between joint torques estimations. Such reliable joint torques open appealing perspectives for the use of new fatigue or solicitation indexes based on internal efforts measured on site.

Relevance to industry: The study demonstrates that corrected Kinect data could be used to estimate internal joint torques, using an adapted inverse dynamics method. The method could be applied on-site because it can handle some cases with occlusions. The resulting Kinect-based method is easy-to-use, real-time and could assist ergonomists in risk evaluation on site.

© 2017 Published by Elsevier B.V.

1. Introduction

Posture and movement of workers are important information for determining the risk of musculoskeletal injury in the workplace (Vieira and Kumar 2004). Based on accurate kinematic data and external forces, inverse dynamics provides ergonomists with internal efforts, such as joint forces and torques (De Looze et al., 2000), or even muscle tensions (Rasmussen et al., 2003; Pontonnier et al., 2014) that are useful to better understand the risk of musculoskeletal injury. Inverse dynamics can be performed by isolating each body segment and using the Newton-Euler methods to retrieve the joint forces and torques (Featherstone, 2014). Another approach is to drive a dynamic model into the kinematic measurements using optimization (Damsgaard et al., 2006). In both approaches, inaccuracies in the kinematic data would strongly influence the resulting joint torques and forces (Riemer et al., 2008).

As a result, accurate motion capture systems, such as the optoelectronic systems with complex setup and calibration, are generally required. Such optoelectronic systems require placing multiple infrared cameras in the environment, positioning skin markers/sensors over standardized anatomical landmarks, calibrating the

setup, and post-processing the data. On-site, in real work conditions, this motion capture process is not possible and could interfere with the current task the subject is performing. Recent development of cheap, markerless and calibrationless sensors, such as the Microsoft Kinect, provides an alternative to these motion capture systems in various application domains, such as clinical gait analysis (Clark et al., 2013; Springer and Seligmann, 2016; Auvinet et al., 2014, 2015), fall-risk assessment (Stone and Skubic, 2015), evaluation of the upper-extremity reachable workspace (Kurillo et al., 2013) and computer graphics (Wei et al., 2012). In ergonomics, previous works have evaluated the ability of the Kinect to measure reliable 3D positions (Dutta, 2012), individual morphologies (Bonnechère et al., 2014; Bonnet and Venture, 2015), assess postures at work (Diego-Mas and Alcaide-Marzal, 2013; Spector et al., 2014; Plantard et al., 2016), and provide real-time feedback to the workers (Martin et al., 2012).

However, recent works have shown that joint angles could be badly estimated in some situations, especially those with occlusions or with inappropriate Kinect placement (Plantard et al., 2015, 2017). These constraints generally occur in real manufacturing plants due to cluttered workstations. The resulting inaccuracies could consequently strongly impact posture assessment and further processes such as inverse dynamics.

Several methods have been proposed to enhance the quality of Kinect skeleton data delivered by the associated software (Shotton et al., 2011). Recently, several authors have proposed to reconstruct badly estimated skeleton data by more plausible one, using a database of accurately captured examples (Shum et al., 2013; Shen et al., 2014). To ensure continuity of the resulting posture sequence, recent works have proposed to organize the database of examples as a graph connecting two postures without discontinuity (Plantard et al., 2017).

The relevance of such corrected postures in an inverse dynamics process has not been tested in previous studies. The aim of this paper is to evaluate if inverse dynamics based on these corrected postures could provide accurate joint torques for an ergonomic purpose, even in bad measurement conditions: occlusions and unsuitable sensor placement. To address this question, we compared results obtained with this method with those computed with a reference method based on classical motion capture data measured with an optoelectronic system. The first part of the paper deals with the materials and methods used to develop the experimental protocol, the dynamic calculation from the two type of data and statistics. The second part of the paper presents and discuss the results of the experiments.

2. Materials and methods

The aim of this study is to evaluate the possibility to correctly estimate joint torques from Kinect data. To this end, we have carried out an experiment comparing joint torques computed with a reference method from accurate Vicon data (assumed to be the ground truth) and those computed from Kinect data. We first detail the experimental protocol used to achieve this comparison. Then, we explain how to compute the joint torques from data provided by the two systems.

2.1. Participants

12 male participants (age: 30.1 ± 7.0 years, height: 1.75 ± 0.046 m, mass: 62 ± 2.7 kg) were voluntary to participate in this study. The study was approved by the Research Ethical Committee of the M2S Laboratory from the University of Rennes 2.

2.2. Protocol

In real work conditions, one of the main constraints is the occlusion occurrences, induced by manipulation of external objects or tools. To reproduce this situation in laboratory conditions, the subjects had to perform *Getting* and *Putting* tasks, with a $40 \times 30 \times 17$ cm empty cardboard box, as depicted in the right part of Fig. 1. In this protocol, we have chosen an empty cardboard box to have a minimum weight manipulated by the subject (200 g), leading to negligible external forces but introducing occlusions. The *Getting* task consisted of a carrying box motion from initial position to the front of the hips. The *Putting* task involved replacing the box to the starting position. The box was attached in the air using a wire and a magnet with low resistance so that external forces were negligible all along the motion. The initial position of the box was set at two possible locations, in order to generate motion variability. In placement P1 the target was located on the left of the subject, aligned with the two shoulders at 1.70 m high and 0.55 m left. In placement P2 the target was located at the same height, but 0.35 m left 0.50 m in front of the subject, as depicted in the left part of Fig. 1.

The manipulated box is supposed to generate more or less occlusions according to its placement in relation to the position of the Kinect. We tested different scenarios with and without the box, and various positions of the Kinect, in order to analyse the impact of different types of occlusions:

- NB: without box condition. The subject had to mimic the manipulating motion without actually using a box, leading to a situation without occlusion. Under this condition, subjects were simply asked to reach the position with their hands where the box would normally be. The Kinect was placed in front of the subjects, as recommended by Microsoft. This scenario allowed us to test the robustness of the Kinect in optimal conditions.
- B: with box. The manipulation was realized with the box, leading to occlusions of body parts, as in real work conditions. The Kinect camera was again placed in front of the subject, as recommended by Microsoft.
- B45: with box and camera placement 45° to the right. The only difference with condition B was that the Kinect was placed 45° on the right of the subject. This type of non-recommended Kinect placement generally occurs in cluttered environments. Under this condition, the risk of occlusions was greater than in all previous conditions.

The above conditions (NB, B and B45) have therefore been combined with two target placements of the box (P1 and P2) for a total of six experimental conditions (P1-NB, P1-B, P1-B45, P2-NB, P2-B and P2-B45). These experimental conditions are summarized in Table 1. The subject repeated each task (*Getting* and *Putting*) 5 times in a unique trial, in each experimental condition.

In order to ensure that the inverse dynamic method based on optoelectronic data delivers actually accurate data, we compared recorded ground reaction forces to those predicted by inverse dynamics using Vicon data. Therefore the subjects were placed on two force plates AMTI 120 by 60 cm (frequency of 1000 Hz), calibrated regularly throughout the experiment. The subject positioned each foot on a different platform to measure ground reaction forces under each foot. The weight of each subject was measured using the two force plates. Anatomical landmarks used for marker placements were defined by the International Society of Biomechanics ISB (Wu et al., 2002, 2005).

2.3. Dynamics estimation method

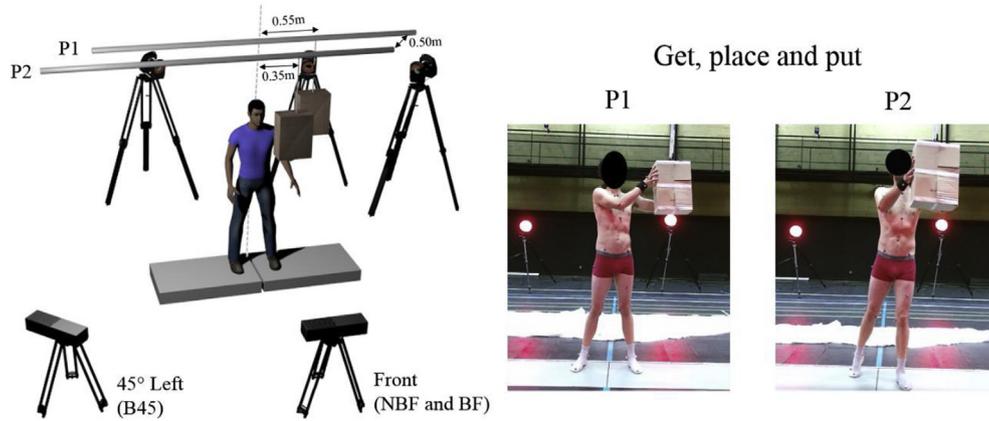In this experiment, we proposed to compare joint torques

Fig 1. Experimental set-up with two placements (P1: object on the left or P2: front-left), with two Kinect positions (NBF and BF: in front or B45: 45° left).

Table 1
Experimental conditions tested in this study.

Kinect Placement	Box/No Box	Box target placement	
		P1	P2
Front	No Box	P1-NB	P2-NB
	Box	P1-B	P2-B
45° Right	Box	P1-B45	P2-B45

computed from Kinect data with those computed from reference motion capture data. We focused this study on the left shoulder and elbow, because these body parts were heavily occluded when the box was manipulated, especially when the Kinect was not positioned optimally (Conditions B and B45). The overall process is depicted in Fig. 2.

Output data from the two motion capture systems did not deliver the same kinematic information. Indeed, the Vicon system measured the 3D positions of each external marker i, named χ_i^{ref}, whereas Kinect provided an estimation of 3D positions for 20 main internal joints i, denoted χ_i^{kin}. Therefore, it required to set up two calculation pipelines, as shown in blue and green in Fig. 2, for the Vicon and Kinect systems respectively.

The torque estimation pipeline is divided into three steps: 1) Firstly, raw kinematic data were corrected to deal with occlusions, using dedicated methods for both Vicon and Kinect data. 2). Then, we used inverse kinematics in order to determine joint angles θ_i^{ref} and θ_i^{kin} along the axis i. Inverse kinematics used either 3D positions of external markers delivered by the Vicon system p_i^{ref}, or by joint centers estimated by the Kinect p_i^{kin}. 3) Finally, the joint torque τ_i^{ref} and τ_i^{kin} along the axis i, is computed thanks to "top-down" inverse dynamics method using joint angles θ_i^{ref} and θ_i^{kin}, respectively. Let us detail now each part of this pipeline.

2.3.1. Vicon pipeline

Vicon data was processed thanks to the Nexus™ software.[1] Occluded marker trajectories were reconstructed under the hypothesis of rigid bodies.

Fig. 3 shows the whole body biomechanical model used in the computation pipeline based on Vicon data. It includes the kinematic skeleton model associated with inertial parameters of each body segment. The markers used in the experiment were also modelled as relative positions expressed in the local coordinates

system of each body segment.

The geometrical parameters, mainly the lengths of the body segments, were initialized thanks to a scaling process, based on the subject size. Then, a geometrical calibration was performed to adapt the segment lengths and marker positions of the model to those of each subject. This calibration was formulated into an optimization problem trying to minimize the difference between the body segment lengths of the model and those of the subject, obtained from Vicon data. Concurrently we also minimized the distance between the marker positions of the model in the local body segment reference frame and those of the subject, obtained from Vicon data (Muller et al., 2015a,b). Body segment inertial parameters (BSIP) were estimated with the regression method proposed by (Dumas et al., 2007).

The joint angles were estimated thanks to an inverse kinematics step consisting in a global optimization at each frame that minimized the distance between the experimental and the model markers global position (Lu and O'connor, 1999) (Equation (1)).

$$\min_{\theta_i^{ref}} \sum_i \left\| p_i^{ref} - p_i^{model}\left(\theta_i^{ref}\right) \right\|^2 \tag{1}$$

where θ_i^{ref} is the vector of generalized coordinates, $p_i^{model}(\theta_i^{ref})$ is a function (named forward kinematics function) delivering the co-ordinates of marker i according to θ_i^{ref}, p_i^{ref} the experimental co-ordinates of marker i.

Joint angles were low-pass filtered (5 Hz) thanks to a 4th order Butterworth filter with no phase shift. Joint velocities and accelerations were calculated with finite difference method.

The joint torques were obtained from joint positions, velocities and accelerations using a recursive Newton-Euler algorithm (Featherstone, 2014). The process is applied from the body extremities to the root of the kinematic chain by linking the forces acting on each body segment i to its motion (see Fig. 4 and Equation (2)):

$$f_i = f_i^B - f_i^x + \sum_{j \in \mu(i)} f_i \tag{2}$$

Where f_i is to the forces applied to the body segment i by its parent. f_i^B is the net force acting on body segment i related to its acceleration, f_i^x is the external forces applied on body segment i, corresponding in this case to the gravitational acceleration, and $\mu(i)$ corresponds to children of body segment i.

The joint torques applied by its parent is the consequence of the force applied along the joint axis.
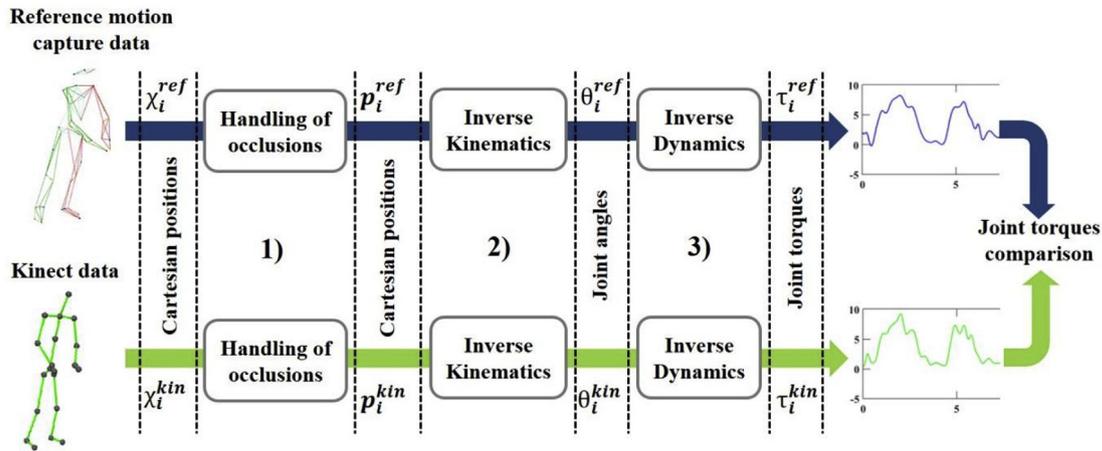
Fig 2. Overview of the two pipelines allowing the joint torque comparisons, using both Kinect data (in green) and reference Vicon data (in blue). Joint torque estimation was divided in three steps: 1) Handling of occlusions; 2) inverse kinematic computation and 3) inverse dynamics computation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
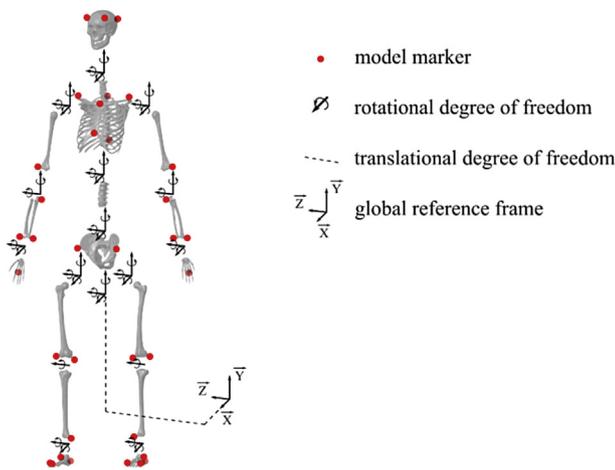

Fig 3. Biomechanical model and markers position for the reference inverse dynamics pipeline (bones geometry were extracted from the AnyBody Managed Model Repository[21]). A virtual 6 degrees of freedom (DoF) joint connects the pelvis to the global reference frame to convert a floating-based system into an equivalent fixed based system.
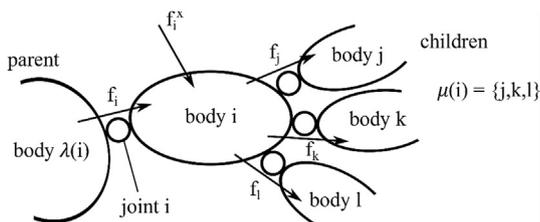

Fig 4. Forces acting on body i (Featherstone, 2014). f_i corresponds to the forces applied on the body segment i by its parent $\lambda(i)$. f_i^x is the external forces applied on body segment i corresponding in this case to the gravitational acceleration. $\mu(i)$ corresponds to children of body segment i, in this case, body i has three children: j, k and l.

2.3.2. Kinect pipeline

As shown in previous studies, the raw Kinect data lead to inaccurate kinematic results when occlusions occur (Plantard et al., 2015; Wang et al., 2015). Indeed, the skeleton provided by the Kinect cannot keep the segment length constant over time in such conditions (Obdrzalek et al., 2012).

Recent work has shown that correction of the Kinect data allows to correctly perform an ergonomic assessment in such occluded environment whereas the raw Kinect data lead to large error (Plantard et al., 2017). We, therefore, propose to use such a correction method in order to limit the impact of occlusions on the Kinect measurement accuracy. This correction method proposed to replace the joint position badly estimated by other, more plausible, using an example-based approach. This method is fully automatic and correct Kinect data in real time. The output data provide Kinect-like skeleton composed of 20 joint 3D positions. Readers are referred to (Plantard et al., 2017) for more details about the pose correction method.

Fig. 5 shows the biomechanical model used for the Kinect calculation pipeline. The position and the orientation of the thorax are assumed to be the basis of the floating-based system. Since the hand position obtained with the Kinect was not accurate enough, the wrist was considered as locked. Moreover, information provided by the Kinect does not provide the pronation/supination movement of the forearm. Therefore, we have chosen to model the elbow joint as a revolute joint. The method used to correct Kinect data is based on a physical model that ensures to maintain constant the distance between two adjacent joints. The segment lengths of the model were then set to these values. BSIP were estimated with the regression method proposed by (Dumas et al., 2007).

International Society of Biomechanics ISB (Wu et al., 2005) recommends a specific method to compute joint angles from anatomical landmark positions, using local coordinate system of each segment. However, Kinect data are not fully compatible with these recommendations because they do not provide all the necessary anatomical landmarks. Therefore, we slightly modified the calculation of the coordinate system associated with each
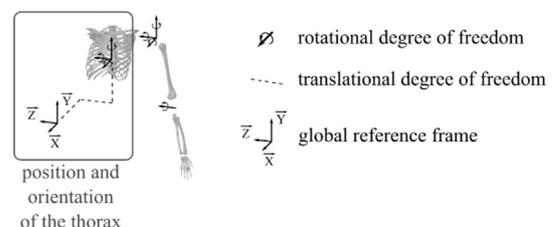

Fig 5. The biomechanical model of the Kinect inverse dynamic pipeline (bones geometry was extracted from the AnyBody Managed Model Repository).

segment to take into account the joint centers returned by the Kinect. The coordinate system associated with the thorax and the shoulder segments was chosen as suggested by Plantard et al. (2016). The joint angles of the shoulders were then obtained following decomposition sequence recommended by the ISB, namely YXY. The first rotation along the Y axis (Y1) defines the elevation plane, the rotation along the axis X corresponds to the elevation and the second rotation along the Y-axis (Y2) represents the internal rotation/external. To limit the impact of geometrical singularities, the external/internal rotation of the arm was fixed (i.e. equal to the previous value) when a singularity (gimbal lock) was detected at a given frame. Finally, the posture given by the Kinect does not provide all the required anatomical landmarks to calculate the local coordinate system of the forearm, as recommended by the ISB. We then computed the elbow flexion according to the method detailed in Bonnechère et al. (2014). The elbow flexion is here defined as the rotation along the Z axis. Joint angles were low-pass filtered (3 Hz) thanks to a 4th-order Butterworth filter with no phase shift. Joint velocities and accelerations were calculated with finite difference method.

According to this simplified biomechanical model, the inverse dynamic step was computed similarly to the reference inverse dynamics pipeline (see Equation (2) and Fig. 4).

2.4. Statistics

In order to ensure that the experimental conditions introduce different levels of occlusion, we computed the mean reliability of all the Kinect joints used for angle computations (trunk, left shoulder, left elbow and left wrist). Reliability of the Kinect joints is computed during the first step of the pose correction method used in this experiment. It consists in a real number between 0 (bad reliability) and 1 (good reliability). Hence reliability is proportional to the influence of the occlusions on the Kinect joint reconstruction errors. Readers are referred to (Plantard et al., 2017) for more details.

Before comparing the results from the Kinect and Vicon pipelines, we evaluated the validity of the joint torques returned by the Vicon pipeline, expected to be the reference one. If the inverse dynamics calculation for the complete body provides accurate results, the resulting 6 degrees of freedom (6-dof) forces and torques for the ground-to-pelvis joint should be close to zero. These values are defined as residual forces and torques. The 6-dof joint between the pelvis and the ground is an artificial joint enabling floating base dynamics. In an ideal situation where the motion is perfectly captured and the external forces are perfectly measured and applied to the model, the forces arising in this 6-dof joint are zero since this joint does not exhibit any actuation. However, in most full body dynamics simulations, these forces are non-negligible since the 6-dof joint compensates kinematics and model errors. A good simulation should exhibit low dynamics residuals (low 6-dof forces) as a proof of its accuracy (Muller et al., 2017). To achieve this validation, we performed the inverse dynamics calculation with Vicon data for the full-body model, while applying the external forces provided by the force plates. We then analysed the forces and torques applied to the 6-dof ground-to-pelvis joint. For each experimental condition, the RMS of these residual forces and torques were calculated for each axis, expressed as mean values and standard deviations for all the subjects. Residual forces were normalized by the body weight of the subject (BW) and the residual torques by the body weight of the subject multiplied by his size (BW × H).

We then evaluated the results obtained using the corrected Kinect data along YXY shoulder axis, and the Z elbow axis, according to the ISB recommendations (Wu et al., 2005).

Firstly, we compared the angles (θ_i) and joint torques (τ_i) along axis i, obtained using the Kinect data with those calculated from Vicon data (assumed to be the reference values). The cross-correlation coefficient (r) and the time lag (τ_{lag}) were calculated for each sequence of motions (including 5 repetitions for each condition). These results were expressed as mean values for each condition. Cross-correlation aimed at measuring the similarity between two signals. Additionally, an intra-class correlation coefficient (ICC) was performed to evaluate the consistency of the joint torque results between those computed from Kinect data and those computed from Vicon data, for each sequence of motions. The maximal joint torques of each sequence were used as a criterion, in order to get discriminative values from one task to another.

In a second time, the recorded motions were segmented in order to focus the analysis on the dynamic *Getting* and *Putting* tasks. Thus, static poses at the beginning and the end of each motion had been eliminated. We then calculated the absolute and relative error (*RMSE* and *nRMSE*) between the joint torques (τ_i) along axis i, estimated from Vicon pipeline and Kinect pipeline. We normalized the RMSE by the amplitude of the reference values (τ_i^{ref}), along the i axis of rotation, as follows:

$$nRMSE\left(\tau_i^{ref}, \tau_i^{kin}\right) = \frac{RMSE\left(\tau_i^{ref}, \tau_i^{kin}\right)}{\max\left(\tau_i^{ref}\right) - \min\left(\tau_i^{ref}\right)}$$

The results of *RMSE* and *nRMSE* are expressed in N.m and % respectively, and are reported as mean values and standard deviations.

Finally, a repeated measures analysis of variance (ANOVA) with a turkey's HSD post-hoc test ($p < 0.05$) was performed in order to assess possible interactions between the Kinect placement and the occlusion condition.

3. Results

In order to quantify the severity of occlusion simulated in this experiment, we computed the joint reliability for each experimental conditions. The mean reliability scores for P1 and P2 placements are 0.89 and 0.91 for NB, 0.78 and 0.71 for B and 0.70 and 0.74 for B45, respectively.

The first part of the results aims at evaluating the accuracy of the joint torques calculated with reference Vicon data. This validation step ensures that the resulting data is suitable to be considered as reference values. To this end, the normalized values of residual forces and torques (at the theoretical ground-to-pelvis joint) were calculated for all conditions. Table 2 shows the average RMS (and standard deviation) of these residual forces and torques.

The average residual forces were below 3.5%, and the standard deviation was below 1%. The largest residual forces were obtained along the vertical axis. Moreover, the results were relatively unaffected by the experimental conditions (NB, B, B45).

Table 3 reports the correlation (r) and time lag (τ_{lag}) values between the joint angles computed from Kinect and reference data, for the YXY shoulder axis and the Z elbow axis.

Joint angles obtained from the Kinect and the reference motion capture system were correlated in all the conditions in and along all rotation axis. However, we can notice that the lower correlation values are found for the first and second rotation around the Y axis of the left shoulder (Y$_1$: r = 0.70 and Y$_2$: r = 0.65).

Table 4 provides the same results for joint torque values.

Joint torques were correlated (r > 0.77) in all conditions and

---

Table 2
Mean RMS ±SD (%), of the residual forces and torques applied to the 6 dof ground-to-pelvis joint for each condition. Let us recall that the z-axis was placed along the vertical axis.

		RMS [%]					
		Fx/BW	Fy/BW	Fz/BW	Mx/(BW*H)	My/(BW*H)	Mz/(BW*H)
P1	NB	2.5 ± 0.9	1.6 ± 0.2	2.5 ± 0.4	0.8 ± 0.2	0.7 ± 0.2	2.8 ± 0.5
	B	2.0 ± 0.5	1.5 ± 0.1	2.7 ± 0.5	0.9 ± 0.2	0.7 ± 0.3	2.6 ± 0.5
	B45	1.9 ± 0.3	1.4 ± 0.2	2.6 ± 0.5	0.8 ± 0.2	0.7 ± 0.3	2.6 ± 0.6
P2	NB	1.8 ± 0.4	1.4 ± 0.3	2.9 ± 0.7	0.4 ± 0.2	0.8 ± 0.3	2.5 ± 0.3
	B	1.6 ± 0.6	1.5 ± 0.6	3.3 ± 0.8	0.4 ± 0.1	0.8 ± 0.3	2.4 ± 0.4
	B45	1.6 ± 0.6	1.4 ± 0.7	3.4 ± 0.9	0.3 ± 0.2	0.8 ± 0.3	2.4 ± 0.4

Table 3
Mean cross-correlation coefficient (r) and mean time lag (τ_{lag}) expressed in seconds, of the joint angles along the YXY shoulder axis and along the Z elbow axis, in all the conditions.

		r				τ_{lag}			
		Shoulder			Elbow	Shoulder			Elbow
		Y_1	X	Y_2	Z	Y_1	X	Y_2	Z
P1	NB	0.89	0.99	0.92	0.98	0.01	0.03	0.04	0.0
	B	0.86	0.96	0.84	0.95	0.02	0.03	0.04	0.01
	B45	0.70	0.92	0.72	0.94	0.37	0.02	0.20	0.15
P2	NB	0.92	1.0	0.91	0.98	0.04	0.03	0.06	0.01
	B	0.80	0.90	0.65	0.92	0.04	0.03	0.66	0.13
	B45	0.88	0.98	0.81	0.95	0.05	0.03	0.08	0.0

along all the joint axis, except for the first rotation along the Y shoulder axis, which corresponds to the orientation of the left shoulder elevation plane. For this axis, the results are poorly correlated (r ranging from 0.26 to 0.50). This low correlation value explains the important values of temporal lag along the same rotation axis. In order to evaluate the consistency of the resulting joint torques, an ICC was performed. The results showed high correlations between the two measurement systems. ICC results were 0.98, 0.98, 0.99, 0.99 along the Y_1, X and Y_2 shoulder axis, and along the Z elbow axis respectively.

Table 5 reports the RMSE and normalized RMSE values of the joint torques along YXY shoulder axis and the Z elbow axis, for the *Getting* and *Putting* tasks, in all conditions.

The most significant absolute error (RMSE = 2.82 N m, nRMSE = 29.5%) was found along the X left shoulder axis, corresponding to the left shoulder joint elevation for the *Getting* task, in P1-B45 condition. Relative error reached a maximum of 36.2% along the left shoulder Y2-axis also in the P1-B45 condition. Interaction hypothesis for experimental conditions was tested and the results showed significant interaction between the occlusion condition and the Kinect placement for most joint torques investigated.

In this experiment, computation time for joint torques based on Vicon and Kinect data were measured. Recall that reconstruction of occluded trajectories was performed manually for the Vicon data,

while the Kinect data were automatically corrected in real time. Mean computation times for inverse kinematics were 250 ms and 0.09 ms for the Vicon-based and Kinect-based calculations respectively. Inverse dynamics has requested a mean computation time of 10 ms and 0.67 ms when using the Vicon and Kinect data respectively. These values are consistent with the sampling frequency of the Kinect (30 Hz), opening perspectives for real-time calculation.

4. Discussion

The current study aimed at evaluating the ability to obtain correct joint torques estimation from Kinect data, in an ergonomic context. To this end, we conducted an experiment consisting in *Getting* and *Putting* tasks under different experimental conditions, e.g. occlusion level (with-box/no-box) and Kinect orientation (Front/45°). The resulting joint torques were compared to those calculated from Vicon data.

Residual forces and torques in the 6 dof ground-to-pelvis joint were below 3.5%, and the standard deviation of these values was below 1%. These values computed from Vicon data were similar to previously published ones (Hansen et al., 2014). Moreover, ranges of shoulder flexion and elbow flexion torques were in accordance with those reported for reaching tasks in the literature (Hollerbach and Flash, 1982). Such results allowed us to consider the joint

Table 4
Mean cross-correlation coefficient (r) and time lag (τ_{lag}) in seconds of the joint torques along the YXY shoulder axis and the Z elbow axis, in all the conditions.

		r				τ_{lag}			
		Shoulder			Elbow	Shoulder			Elbow
		Y_1	X	Y_2	Z	Y_1	X	Y_2	Z
P1	NB	0.50	0.98	0.82	0.97	0.18	0.03	0.17	0.03
	B	0.28	0.96	0.82	0.92	2.65	0.03	0.20	0.04
	B45	0.31	0.91	0.77	0.91	1.82	0.03	0.46	0.09
P2	NB	0.35	0.99	0.92	0.97	0.60	0.03	0.04	0.03
	B	0.28	0.87	0.81	0.90	4.0	0.02	0.21	0.06
	B45	0.26	0.97	0.87	0.95	0.30	0.03	0.14	0.04

Table 5
Mean RMSE ± SD expressed in N.m and nRMSE expressed in (%) of the joint torques along YXY shoulder axis and the Z elbow axis, for *Getting* and *Putting* motion, in all conditions.

		RMSE [N.m] (nRMSE [%])			
		Shoulder			Elbow
		Y_1	X	Y_2	Z
Getting					
P1	NB	0.83 ± 0.38 (20.2)	1.22 ± 0.50 (12.9)	0.66 ± 0.32 (24.4)	0.60 ± 0.21 (17.7)
	B	1.05 ± 0.55 (20.8)	1.74 ± 0.94 (17.3)	0.79 ± 0.34 (29.7)	0.81 ± 0.34 (21.9)
	B45	1.25 ± 0.65 (28.4)	2.82 ± 1.45 (29.5)	0.90 ± 0.56 (36.2)	0.91 ± 0.41 (25.1)
P2	NB	0.66 ± 0.26 (13.8)	1.14 ± 0.56 (10.6)	0.58 ± 0.18 (22.7)	0.46 ± 0.19 (12.9)
	B	1.17 ± 0.54 (27.1)	2.54 ± 1.34 (24.6)	0.74 ± 0.35 (28.8)	0.94 ± 0.41 (25.0)
	B45	0.98 ± 0.39 (20.1)	1.87 ± 0.93 (18.1)	0.66 ± 0.27 (24.8)	0.63 ± 0.26 (17.3)
Putting					
P1	NB	0.68 ± 0.29 (16.5)	1.17 ± 0.49 (12.3)	0.70 ± 0.31 (25.7)	0.60 ± 0.24 (17.3)
	B	1.27 ± 0.91 (26.3)	1.51 ± 0.88 (15.0)	0.64 ± 0.29 (23.5)	0.99 ± 0.47 (26.4)
	B45	1.18 ± 0.65 (28.4)	2.12 ± 1.05 (22.2)	0.73 ± 0.35 (29.8)	0.91 ± 0.44 (24.7)
P2	NB	0.94 ± 0.45 (19.5)	1.13 ± 0.53 (10.7)	0.50 ± 0.20 (18.9)	0.57 ± 0.27 (16.3)
	B	1.07 ± 0.84 (23.5)	2.12 ± 1.09 (20.2)	0.72 ± 0.26 (27.8)	0.93 ± 0.47 (24.8)
	B45	1.01 ± 0.72 (21.3)	1.56 ± 0.98 (15.2)	0.64 ± 0.24 (23.3)	0.74 ± 0.32 (20.5)

torques obtained with the Vicon data as a reference, and could consequently be compared to the Kinect-based method. Remaining error is mostly explained by the scaling of the biomechanical model, BSIP estimation that is relatively rough, and the kinematic error generated by inverse kinematics.

Kinect-based and reference-based estimated joint torques were compared qualitatively and quantitatively thanks to two criteria: cross correlation and RMSE respectively.

Firstly, qualitative results show a high similarity of torque shapes among trials, excepted for the orientation of the shoulder elevation plane (Y_1). These poor results along Y_1-axis can be explained by the small variation of the computed joint torques for the studied situations. Indeed, variations in joint torques values along this axis are so small that the signal/noise ratio becomes very important, leading to poor statistical results, especially when trying to correlate noisy signals together. This type of result is also found for the Y_2-axis, even if the correlation values remained higher than for the Y_1-axis. Indeed, Fig. 6 shows the joint torques calculated from the Kinect data (in red) and reference data (in blue) along the YXY shoulder axis and the Z elbow axis. We noticed significant variations in joint torques values along X-axis, unlike Y_1-axis and Y_2-axis. In an ergonomic context, the elevation of the shoulder is an extremely important rotation as it has been reported to be linked to many well-known diseases. This rotation is also the best computed in inverse dynamic process based on Kinect data for this type of motion.

Secondly, quantitative results showed differences between the two methods that can become relatively high. For the X axis, the maximum RMSE was 2.82 ± 1.45 N m and 2.54 ± 1.34 N m for the Side-B45 and Front-B conditions, respectively. These absolute errors should be related to the high range of joint torque values calculated for this rotation axis (10.1 ± 1.6 N m). Thus, RMSE obtained from the Kinect data represents an average value of 17.4% of the joint torque range for this rotation axis. On the opposite, for the Y_2-axis, although the absolute RMSE were lower (less than 1 N m), relative to the joint torques range for this axis (2.7 ± 0.6 Nm), this represents an average relative error of 26.3%. Again, one may consider the signal/noise ratio for such axis. Significant interaction between occlusion conditions and Kinect placements show the importance of the point of view according to the task performed. Indeed, the B45 occlusion led to significantly better results in P2 placement compared to B, whereas B led to significantly better results for P1. This observation means that even if occlusion was more important for any placement in B45 than in B, the most

occluded joints were not necessary the ones evaluated in the study. In conclusion, one has to carefully choose the Kinect placement with regard to the task to investigate and the environment in order to obtain the most reliable results. ICC results highlighted a very important feature that such a method has to exhibit. Indeed, the Kinect-based maximum torques were consistent with those obtained with the Vicon data for most of the trials. This leads to the conclusion that the method is able to discriminate properly two different work situations in terms of torque level. Thus, it could be used to discriminate discomfort between work situations. nRMSE reported in Table 5 may show that the system estimated approximately the absolute values of the joint torques. However, ICC proves that such a method may be used to assess the torque of a given task and compare it to another task. These results are promising, but it could be useful to evaluate the robustness of the proposed method according to different work task conditions (with various external forces and velocities).

The correction method of Kinect data is based on parameters which may affect its performance, especially, such as the filtering step of the corrected data. Indeed, the filter used in the correction can introduce a delay that could lead to a time lag in the final forces and torques curves. However, the results presented in Table 4 show that the observed time lag between the results obtained using Kinect and Vicon data ranged from 0.02 s to 0.03s for the shoulder elevation, and from 0.03 s to 0.09 s for the elbow flexion. This time lag induced by the filter seems to consequently have no significant impact on the joint torques estimation, as shown in Figure 6.

The results found in this experiment support the potential usability of the proposed method to correctly estimate dynamic quantities based on corrected Kinect data. However, several limitations can be identified.

The sampling frequency of the Kinect can impact its usability for estimating joint torques, especially for fast motions. Indeed, the Kinect remains a low-frequency acquisition system (30 Hz) whereas the Vicon was sampled at 100 Hz. This low-frequency acquisition could have an impact on the processing of fast movements. Indeed, the low frequency can introduce derivative errors, resulting in a poor estimation of the speed and acceleration, and consequently increasing the residual forces and torques of the theoretical 6-dof pelvis-to-ground joint. However, in a work context, execution speed is generally limited and could be compatible with this frequency for most of the tasks. However, it would be important to assess the impact of speed on joint torques estimation. To this end, further studies on a larger number of tasks
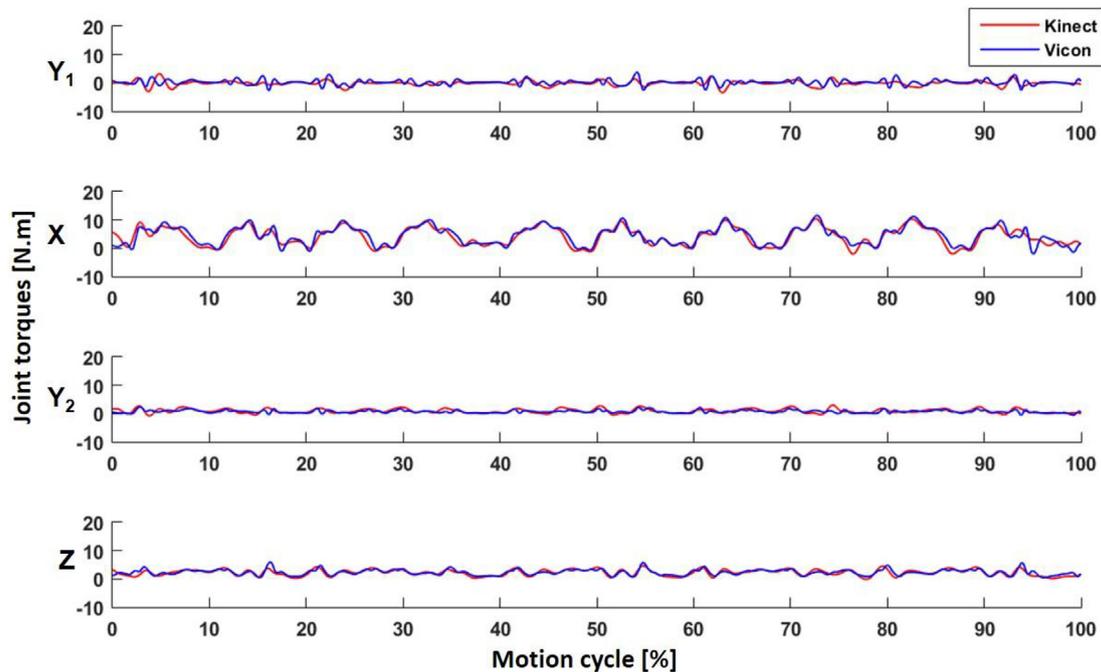
Fig 6. Examples of estimated joint torques (in N.m), along the YXY shoulder axis and the Z elbow axis, computed from the Kinect data (red) and reference data (blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

are necessary, in order to evaluate the relevance of the approach for a wider range of real work conditions.

The data provided by the Kinect does not contain all the information required to accurately compute all of the joint angles, as recommended by the ISB. A possible solution to tackle this problem is to develop a more complex kinematic model, as proposed by Bonnechère et al. (2014). However, the estimation of new anatomical landmarks from available joints should be based on precisely reconstructed information. Recent articles showed that the joints provided by the Kinect could introduce significant errors in specific postures when self-occlusions occurred (Plantard et al., 2015). Because of these large potential errors, it seems difficult to reliably estimate new landmarks and more research would be needed to improve data quality first. Indeed, although results reported in this study are promising, some errors remain for conditions where occlusions persist over a long period.

Another limitation is linked to the ability of the various measurement systems to capture small details, such as the hand motion. Indeed, the Kinect data was not accurate enough to capture all the dof estimated by the Vicon system. Particularly, wrist motion was not taken into account. This involves limiting the use of this type of system for tasks where these motions are not relevant. For motion involving large angular ranges of the shoulder and elbow, the system seems to be promising in working condition. However, we should consider how this error changes according to the external efforts involved by the task. It would be interesting to conduct a sensitivity study to evaluate how correctly external forces (such as changing the mass of the manipulated box) would be taken into account with this method. We could assume that the error decreases with larger external forces, as the signal-noise ratio would decrease accordingly.

Another consideration is that any inverse dynamics calculation involves precise knowledge of external forces. In the current task, no external load was taken into account. This is due to the fact that the 200 g box was empty. With a task involving external forces, we would have to consider these values to get proper joint torques. It

seems possible to model these external forces by an estimated mass of the object for simple work tasks. However, for more complex tasks, these values need to be measured, which is difficult on-site without disturbing the worker. However, if such measures are available, they can be directly used by the inverse dynamics method, such as input data of the Newton-Euler recursive algorithm (Featherstone 2014).

5. Conclusion

This study presents an evaluation of joint forces and torques estimation based on corrected Kinect data. The results show that these data are accurate enough to compute reliable joint torques values in challenging experimental conditions (when occlusions occur or sensor placement is not optimal).

Despite the reported usability limitations, the results of the current study are promising for the ergonomic evaluation of workstations and assessment of physical risks. Kinect has already been considered as a promising tool to evaluate ergonomics on-site (Dutta, 2012; Diego-Mas and Alcaide-Marzal, 2013; Patrizi et al., 2015), but only at a postural level. The current work shows a practical capacity to estimate dynamics in the same experimental conditions. Moreover, the proposed correction method allows performing such estimations in challenging environments, e.g. cluttered production chains introducing occlusions. Particularly, the method fairly estimated absolute joint torques values, but also properly followed the joint torque trends during the trials and rated with consistency the tasks with regard to the VICON results. Internal forces, such as joint torques and muscle forces could be very useful for ergonomists to better understand the risk of musculo-skeletal injury and compare work situation. This result opens appealing perspectives for the definition and the use of new fatigue or solicitation indexes based on joint forces and torques estimated on-site (Ma et al., 2009; Pontonnier et al. 2014).

Finally, the joint torques estimation proposed in this study achieved real-time performances (0.09 ms and 0.67 ms for inverse

kinematics and inverse dynamics, respectively). It would be interesting to test the benefit of producing a real-time feedback to the worker based on mobilized internal forces, as realized for postural risk (Vignais et al., 2012). Indeed, these previous studies suggested that an ergonomic feedback in real-time, based on a motion capture system, influences how workers perform their tasks, reducing the values of MSD risk scores.

Acknowledgement

References

Auvinet, E., Multon, F., Aubin, C.E., Meunier, J., Raison, M., 2014. Detection of gait cycles in treadmill walking using a Kinect. Gait Posture 41 (2), 722–725.

Auvinet, E., Multon, F., Meunier, J., 2015. New lower-limb gait asymmetry indices based on a depth camera. Sensors 15 (3), 4605–4623.

Bonnechère, B., Jansen, B., Salvia, P., Bouzahouene, H., Omelina, L., Moiseev, F., Sholukha, V., Cornelis, J., Rooze, M., S. V. S, Jan. 2014. Determination of the precision and accuracy of morphological measurements using the Kinect™ sensor: comparison with standard stereophotogrammetry. Ergonomics 57 (4), 622–631.

Bonnet, V., Venture, G., 2015. Fast determination of the planar body segment inertial parameters using affordable sensors. IEEE Trans. Neural Syst. Rehabil. Eng. 23 (4), 628–635.

Clark, R.A., Bower, K.J., Mentiplay, B.F., Paterson, K., Pua, Y., 2013. Concurrent validity of the Microsoft Kinect for assessment of spatiotemporal gait variables. J. Biomech. 46 (15), 2722–2725.

Damsgaard, M., Rasmussen, J., Christensen, S.T., Surma, E., de Zee, M., 2006. Analysis of musculoskeletal systems in the AnyBody modeling system. Simul. Model. Pract. Theory 14 (8), 1100–1111.

Diego-Mas, J.A., Alcaide-Marzal, J., 2013. Using Kinect™ sensor in observational methods for assessing postures at work. Appl. Ergon. 45 (4), 976–685.

Dumas, R., Cheze, L., Verriest, J.P., 2007. Adjustments to McConville et al. and Young, et al. body segment inertial parameters. J. Biomech. 40 (3), 543–553.

Dutta, T., 2012. Evaluation of the Kinect sensor for 3-D kinematic measurement in the workplace. Appl. Ergon. 43 (4), 645–649.

Featherstone, R., 2014. Rigid Body Dynamics Algorithms. Springer.

Hansen, C., Venture, G., Rezzoug, N., Gorce, P., Isableu, B., 2014. An individual and dynamic Body Segment Inertial Parameter validation method using ground reaction forces. J. Biomech. 47 (7), 1577–1581.

Hollerbach, J.M., Flash, T., 1982. Dynamic interactions between limb segments during planar arm movement. Biol. Cybern. 44 (1), 67–77.

Kurillo, G., Chen, A., Bajcsy, R., Han, J.J., 2013. Evaluation of upper extremity reachable workspace using Kinect camera. Technol. Health Care 21 (6), 641–656.

De Looze, M.P., Van Greuningen, K., Rebel, J., Kingma, I., Kuijer, P.P.F.M., 2000. Force direction and physical load in dynamic pushing and pulling. Ergonomics 43 (3), 377–390.

Lu, T.-W., O'connor, J., 1999. Bone position estimation from skin marker coordinates using global optimisation with joint constraints. J. Biomech. 32 (2), 129–134.

Ma, L., Chablat, D., Bennis, F., Zhang, W., 2009. A new simple dynamic muscle fatigue model and its validation. Int. J. Ind. Ergon. 39 (1), 211–220.

Martin, C.C., Burkert, D.C., Choi, K.R., Wieczorek, N.B., McGregor, P.M., Herrmann, R.A., Beling, et.P.A., 2012. A real-time ergonomic monitoring system using the Microsoft Kinect. In: In IEEE Systems and Information Design Symposium (SIEDS).

Muller, A., Germain, C., Pontonnier, C., Dumont, G., 2015a. A simple method to calibrate kinematical invariants: application to overhead throwing. In: 33rd International Conference on Biomechanics in Sports (ISBS). Poitier, France, June 29-July 3.

Muller, A., Pontonnier, C., Germain, C., Dumont, G., 2015b. Dealing with modularity of multibody models. Comput. methods Biomech. Biomed. Eng. 18 (1), 2008–2009.

Muller, A., Pontonnier, C., Dumont, G., 2017. Uncertainty propagation in multibody human model dynamics. Multibody Syst. Dyn. 1–16.

Obdrzalek, S., Kurillo, G., Oi, F., Bajcsy, R., Seto, E., Jimison, H., Pavel, M., 2012. Accuracy and robustness of Kinect pose estimation in the context of coaching of elderly population. In: In Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).

Patrizi, A., Pennestrì, E., Valentini, P.P., 2015. Comparison between low-cost markerless and high-end marker-based motion capture systems for the computer-aided assessment of working ergonomics. Ergonomics 59 (1), 155–162.

Plantard, P., Shum, H.P.H., Multon, et F., 2017. Filtered pose graph for efficient kinect pose reconstruction. Multimedia Tools Appl. 76 (3), 4291–4312.

Plantard, P., Auvinet, E., Le Pierres, A.S., Multon, F., 2015. Pose estimation with a kinect for ergonomic studies: evaluation of the accuracy using a virtual mannequin. Sensors 15 (1), 1785–1803.

Plantard, P., Shum, H.P.H., Le Pierres, A.S., Multon, F., 2016. Validation of an Ergonomic Assessment Method Using Kinect Data in Real Workplace Conditions. Applied ergonomics, in Press.

Pontonnier, C., de Zee, M., Samani, A., Dumont, G., Madeleine, P., 2014. Strengths and limitations of a musculoskeletal model for an analysis of simulated meat cutting tasks. Appl. Ergon. 45 (3), 592–600.

Rasmussen, J., Damsgaard, M., Surma, E., Christensen, S.T., de Zee, M., Vondraket, V., 2003. Anybody-a software system for ergonomic optimization. In: Fifth World Congress on Structural and Multidisciplinary Optimization. Venice, Italy, May 19–23.

Riemer, R., Hsiao-Wecksler, E.T., Zhang, X., 2008. Uncertainties in inverse dynamics solutions: a comprehensive analysis and an application to gait. Gait posture 27 (4), 578–588.

Shen, W., Deng, K., Bai, X., Leyvand, T., Guo, B., Tu, Z., 2014. Exemplar based human action pose correction. IEEE Trans. Cybern. 44 (7), 1053–1066.

Shotton, J., Fitzgibbon, A., Cook, M., Sharp, T., Finocchio, M., Moore, R., Kipman, A., Blake, A., 2011. Real-time human pose recognition in parts from single depth images. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Providence, RI, USA, June 20–25.

Shum, H.P.H., Ho, E.S.L., Jiang, Y., Takagi, S., 2013. Real-time posture reconstruction for microsoft kinect. IEEE Trans. Cybern. 43 (5), 1357–1369.

Spector, J.T., Lieblich, M., Bao, S., McQuade, K., Hughes, M., 2014. "Automation of workplace lifting hazard assessment for musculoskeletal injury prevention. Ann. Occup. Environ. Med. 26 (1), 1–8.

Springer, S., Seligmann, G.Y., 2016. Validity of the kinect for gait assessment: a focused review. Sensors 16 (2), 194.

Stone, E., Skubic, M., 2015. Fall detection in homes of older adults using the microsoft kinect. IEEE J. Biomed. Health Inf. 19 (1), 290–301.

Vieira, E.R., Kumar, S., 2004. Working postures: a literature review. J. Occup. Rehabil. 14 (1), 143–159.

Vignais, N., Miezal, M., Bleser, G., Mura, K., Gorecky, D., Marin, F., 2012. Innovative system for real-time ergonomic feedback in industrial manufacturing. Appl. Ergon. 44 (4), 566–574.

Wang, Q., Kurillo, G., Oi, F., Bajcsy, R., 2015. Evaluation of pose tracking accuracy in the first and second generations of Microsoft Kinect. In: In IEEE International Conference on Healthcare Informatics (ICHI).

Wei, X., Zhang, P., Chai, J., 2012. Accurate realtime full-body motion capture using a single depth camera. ACM Trans. Graph. 31 (6), 1–12.

Wu, G., Siegler, S., Allard, P., Kirtley, C., Leardini, A., Rosenbaum, D., Whittle, M., et al., 2002. ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion–part I: ankle, hip, and spine. International Society of Biomechanics. J. Biomech. 35 (4), 543–548.

Wu, G., van der Helm, F.C., Veeger, H.E., Makhsous, M., Van Roy, P., Anglin, C., Nagels, J., et al., 2005. International Society of Biomechanics, ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion – Part II: shoulder, elbow, wrist and hand. J. Biomech. 38 (5), 981–992.

Pierre Plantard received his M.Sc. in Biomechanics from Rennes 2 University in 2011 and his M.Sc. in Computer Science from Rennes 1 University in 2012. He is currently a Ph.D. Student in the M2S laboratory at Rennes 2 University in collaboration with Faurecia Company. He is also member of the INRIA research team named MimeTIC. His research interests include ergonomics, health prevention, human motion analysis and markerless motion capture.

Antoine Muller is an assistant professor at École Normale Supérieure de Rennes in mechanical science. He is a Ph.D student in INRIA/IRISA MimeTIC research team since 2014. His research focuses on the development of subject-specific models and methods for determining forces related to human motion. The goal is to analyse the motion and to compute the muscle forces of the human activity for ergonomics or for performance applications.

Charles Pontonnier received his PhD degree in mechanics in 2010 from Rennes 1 University. He is associate professor at École Normale Supérieure de Rennes on secondment at military academy of Saint-Cyr Coëtquidan since 2012. His current research focuses on muscle-based motion analysis-synthesis for ergonomics and virtual reality tools for ergonomics.

Georges Dumont is a Professor in Mechanical Sciences at École Normale Supérieure de Rennes in Rennes, France. He received his PhD in computer science from Rennes 1 University in 1990 and his habilitation in mechanical science in 2005. His main research interests include physical simulation, mechanics, biomechanics, haptic rendering, interactive collaboration and virtual reality.

Hubert P.H. Shum is an Associate Professor at Northumbria University. Before joining the university, he worked as a Lecturer at the University of Worcester, a post-doctoral researcher in RIKEN Japan, as well as a research assistant at the City University of Hong

Kong. He received his PhD degree from the School of Informatics at the University of Edinburgh. His research interests include character animations, machine learning, human motion analysis and computer vision.

Franck Multon is Professor in the M2S Lab. (Univ. Rennes2, France) and is leading the MimeTIC Inria team. He obtained his PhD degree in computer science in 1998, and his habilitation in computer sciences in 2006, from University Rennes 1. His research interests are motion analysis, virtual humans, biomechanics, ergonomics, and virtual reality in sports.