



**HAL**  
open science

## Usability of corrected Kinect measurement for ergonomic evaluation in constrained environment

Pierre Plantard, Hubert Shum, Franck Multon

► **To cite this version:**

Pierre Plantard, Hubert Shum, Franck Multon. Usability of corrected Kinect measurement for ergonomic evaluation in constrained environment. *International Journal of Human Factors Modelling and Simulation*, 2017, 5 (4), pp.338 - 353. 10.1504/IJHFMS.2017.087018 . hal-01612939

**HAL Id: hal-01612939**

**<https://inria.hal.science/hal-01612939v1>**

Submitted on 9 Oct 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

---

## **Usability of corrected Kinect measurement for ergonomic evaluation in constrained environment**

---

**Pierre Plantard\***

M2S Laboratory,  
University Rennes 2,  
INRIA Rennes  
Rennes, France  
Email: pierre.plantard88@gmail.com  
\*Corresponding author

**Hubert P.H. Shum**

Faculty of Engineering and Environment,  
Northumbria University,  
Newcastle upon Tyne, UK  
Email: hubert.shum@northumbria.ac.uk

**Franck Multon**

M2S Laboratory,  
University Rennes 2,  
INRIA Rennes  
Rennes, France  
Email: franck.multon@irisa.fr

**Abstract:** Evaluation of potential risks of musculoskeletal disorders in real workstations is challenging as the environment is cluttered, which makes it difficult to correctly assess the pose of a worker. Being marker-free and calibration-free, Microsoft Kinect is a promising device to assess these poses, but it can deliver unreliable poses especially when occlusions occur. To overcome this problem, we propose to detect and correct badly recognised body parts thanks to a database of example poses. We applied the proposed method to compute rapid upper limb assessment (RULA) score in a realistic environment that involved sub-optimal Kinect placement and several types of occlusions. Results showed that when occlusions occur, the inaccurate raw Kinect data could be significantly improved using our correction method, leading to acceptable joint angles and RULA scores. Our method opens new perspectives to define new fatigue or solicitation indexes based on continuous measurement contrary to classical static images used in ergonomics.

**Keywords:** Kinect; pose correction; occlusions; ergonomics assessment; RULA score; joint angles.

**Reference** to this paper should be made as follows: Plantard, P., Shum, H.P.H. and Multon, F. (201X) ‘Usability of corrected Kinect measurement for ergonomic evaluation in constrained environment’, *Int. J. Human Factors Modelling and Simulation*, Vol. x, No. x, pp.xxx–xxx.

**Biographical notes:** Pierre Plantard received his PhD in Biomechanics from the Rennes 2 University in 2016 and his MSc in Computer Science from the Rennes 1 University in 2012. He is currently a Research Engineer in the M2S Laboratory at the Rennes 2 University. He is also member of the INRIA research team named MimeTIC. His research interests include ergonomics, biomechanics, human motion analysis and markerless motion capture.

Hubert P.H. Shum is a Senior Lecturer (Associate Professor) at the Northumbria University. Before joining the university, he worked as a Lecturer in the University of Worcester, a Post-doctoral Researcher in the RIKEN Japan, as well as a Research Assistant in the City University of Hong Kong. He received his PhD from the School of Informatics in the University of Edinburgh. His research interests include character animations, machine learning, human motion analysis and computer vision.

Franck Multon is a Professor in the University Rennes 2 in France. He is performing his research in Biomechanics (M2S Lab) and in character simulation (MimeTIC/Inria group). His research interests are biomechanics and character simulation, as well as interaction between real and virtual humans. He defended his PhD in 1998 in INRIA Rennes on motion control of virtual humans. Since 1999, he was Assistant Professor in the University Rennes 2, has defended his ‘authorisation to supervise research’ in 2006, and has been hired as Full Professor in 2008.

This paper is a revised and expanded version of a paper entitled ‘Ergonomics measurements using Kinect with a pose correction framework’ presented at Digital Human Modeling (DHM), Montreal, Canada, 15–17 June 2016.

---

## 1 Introduction

Microsoft Kinect is used nowadays to measure user performance in many application domains. Initially designed for video games, the low-cost and easy-to-use motion capture device has been applied in clinical gait analysis (Auvinet et al., 2012, 2014; Galna et al., 2014), human-computer interactions (Wang et al., 2013), sign-language analysis (Gameiro et al., 2014; Pedersoli et al., 2014), sport training (Cassola et al., 2014) and ergonomics (Diego-Mas and Alcaide-Marzal, 2014; Vignais et al., 2013). In ergonomics, the posture and movement of the worker provides important information for determining the risk of musculoskeletal injury in the workplace (Vieira and Kumar, 2004). Consequently, several works have proposed assessment grids based on body posture, such as the famous rapid upper limb assessment (RULA) grid (McAtamney and Corlett, 1993). This ergonomics tools rate the pose of the worker based on an estimation of the joint angles. They then provide a global risk score that represent the risk of injury

based on the performed postures. Recent works in ergonomics (Vignais et al., 2013) have demonstrated that real-time ergonomic feedback through head mounted display positively influences the motion of workers, decreasing locally hazardous RULA values. However, the method was based on inertial sensors and feedback devices that can change the way people perform the motion. Optical motion capture systems require positioning sensors or markers on the user's body, calibrating the system or measuring user dimensions, which is not always possible in real work conditions. Sensors may be incompatible with security constraints and can also be disrupted by electromagnetic interference. These motion capture problems are also encountered in other application domains such as sports, training or rehabilitation.

Recent papers evaluated the accuracy of the Kinect skeleton data mostly for simple motions and in accordance with the Kinect recommendation (sensor placed in front of the subject) (Clark et al., 2012; Kurillo et al., 2013; Bonnechere et al., 2014). It has been shown that this errors rapidly increase for complex motions with auto-occlusions and when the sensor is not placed in the recommended position (Plantard et al., 2015).

Several methods have been proposed to correct badly reconstructed poses provided by the Kinect. Since human motion is highly nonlinear, learning statistical dynamic models (based on database of examples) as a motion prior can produce higher quality movements (Chai and Hodgins, 2007). Applying these methods to reconstruct Kinect poses has a major drawback as each body joint position is assumed to be accurately reconstructed whereas Kinect data delivers noisy or even incorrect information. To overcome this limitation, recent works have proposed to take the reliability of the Kinect data into account in the correction process. Reliability can then be integrated into a lazy learning framework (Shum et al., 2013) or a Gaussian model (Zhou et al., 2014; Liu et al., 2016) to reconstruct a more reliable pose.

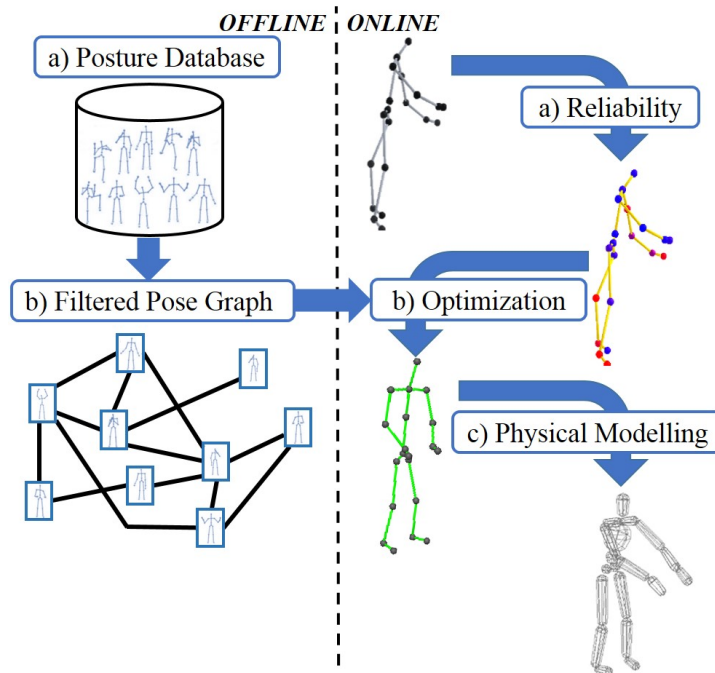
Recently, an example-based correction approach has been proposed to deal with situations where when occlusions occur (Plantard et al., 2017a). A new data structure named filtered pose graph guarantees efficient preselection of a relevant subset of poses before correction, ensuring continuity and maximising reliability even when important occlusions occur. This enhances both computation speed and reconstruction quality of the system. In real working situations, the resulting ergonomic assessments demonstrated similar scores than those provided by ergonomic experts (Plantard et al., 2017b), but no accurate comparison to reference posture measurements have been carried-out. Hence, correction methods have not been strictly evaluated in constrained conditions, with many occlusions and uncontrolled sensor placement. The main contribution of this paper is to evaluate the actual usability of this method in constrained environments, in an ergonomic perspective. To this end, we propose to compare ergonomic scores computed with both corrected Kinect data and reference motion capture data.

The paper is organised as follows. The method used in this paper to improve the interpretation of Kinect data is recalled in Section 2. The computation of the RULA grid and the protocol used for the reliability evaluation of our method are given in Section 3. Results of the joint angles evaluation and the RULA score estimation in constrained environments are given in Section 4. We discuss and conclude the work in Sections 5 and 6.

## 2 Correction framework

In this section, we give an overview of the correction framework presented in Plantard et al. (2017a). The readers are referred to Plantard et al. (2017a) for more details. The correction method improves the interpretation of Kinect data thanks to an example-based approach. The correction framework is composed of an offline and an online process as shown in Figure 1.

**Figure 1** Overview of the pose correction method (see online version for colours)



Notes: Offline database preprocessing: a) posture database to b) filtered graph representation. Online pose correction: a) reliability estimation, b) pose optimisation and c) physical filtering.

Source: Adapted from Plantard et al. (2017a)

The offline process organises the database of poses extracted from motion capture clips to produce a so-called filtered pose graph (a and b in the offline part in Figure 1). The filtered pose graph organises a set of captured postures in a graph structure that enables efficient pose correction. In such a graph, the nodes are individual human pose, and the edges are potential links between the two poses if they could be connected without discontinuities (i.e., the distance between poses is below a given threshold). The graph is filtered both to eliminate redundant poses and avoid creating an overly dense graph with numerous edges and nodes. The resulting graph enables us to rapidly select poses that are close to a given current pose, which could be considered as potential next poses in the studied motion. Hence, in the online correction phase, the idea is to rapidly gather the set of pose examples that could help to correct badly reconstructed body parts.

During the online correction process, the system first estimates the reliability of each joint center reconstructed by the Kinect [Figure 1(a)]. Then, based on the reliable information delivered by the Kinect, it selects the potential nodes (i.e., pose examples) in the filtered pose graph that can help to correct unreliable joint positions. The resulting pose examples are combined using an optimisation process to correct unreliable information using prior knowledge while preserving continuity and similarity to the reliable information [Figure 1(b)]. Finally, a physical model is used to filter the resultant pose and avoid jerky motion [Figure 1(c)].

### 3 Material and methods

This section describes the method and experimental protocol used to evaluate the method introduced above in an ergonomic context.

#### 3.1 Computation of the RULA grid

##### 3.1.1 Introduction to RULA grid

In ergonomics, one of the most popular observational methods is the RULA (McAtamney and Corlett, 1993). This tool requires to rate the pose of the worker based on an estimation of the upper-body joint angles. Each joint angle is associated with a joint score according to a predefined range of angles. One has to notice that additional conditions can increase the local body part scores, such as when the shoulder is raised or the upper arm is abducted. These scores are combined to provide a global risk score for the left and right body parts, ranging from one (posture is acceptable) to seven (workstation requires investigation and changes immediately).

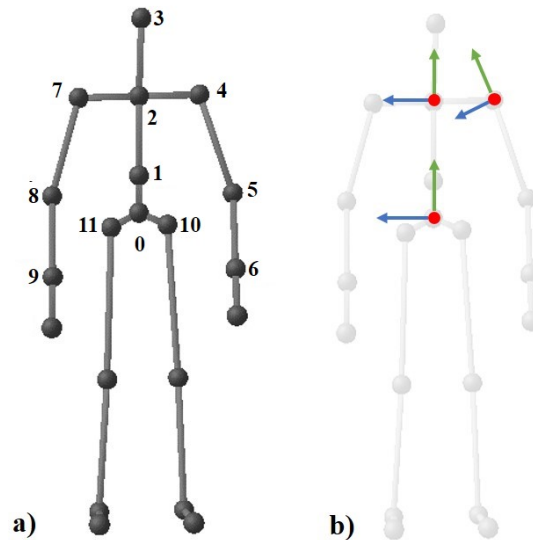
##### 3.1.2 Computation of joint angles using the Kinect data

To use the RULA method, relevant joint angles have to be computed based on the Kinect skeleton data (see Figure 2). A Kinect pose is defined as  $p = \{x_j, y_j, z_j\}_{j=1..N}$  where  $N$  stands for the number of joints in the pose, and  $x_j, y_j, z_j$  stand for the 3D Cartesian coordinates of the  $j^{th}$  joint. According to the estimated joint positions, joint angles should be computed using the ISB recommendation (Wu et al., 2002, 2005), as it is common and recognised manner to estimate bone motion thanks to skin markers. However, the Kinect skeleton is not fully compatible with this recommendation as it does not provide all the required anatomical landmarks.

We consequently slightly adapted the joint angle definition to take the available Kinect joints [named with numbers in Figure 2(a)] into account.

The pelvis coordinate is define relative to the recommendation (Wu et al., 2002). The Y-axis is along the trunk axes represented by the vector from the hip center (0) to the spine (1). The X-axis is defined as the normal of the plane formed by the Y-axis, the left (10) and the right (11) hips. Finally the Z-axis is computed as the normal of the X and Y axes.

**Figure 2** (a) Skeleton model provided by the Kinect and the correction method (b) Body part coordinates (pelvic, trunk and shoulder) (see online version for colours)



Notes: (0) hip centre, (1) spine, (2) shoulder centre, (3) head, (4) left and (7) right shoulders, (5) left and (8) right elbows, (6) left and (9) right wrists, (10) left and (11) right hips. X-axis in red pointing forward, Y-axis in green pointing upward and Z-axis in blue pointing to the right.

Source: Adapted from Plantard et al. (2017b)

For the trunk coordinate system, the Y-axis is represented by the vector from the spine joint (1) to the shoulder center joint (2). The X-axis is defined as the normal of the plan formed by the Y-axis, the left (4) and the right (7) shoulders. Finally the Z-axis is computed as the normal of the X and Y axes.

The shoulder coordinate system is defined according to the ISB recommendation. The Y-axis is given by the vector from elbow joint (5 or 8) to shoulder joint (4 or 7). The Z-axis is the normal of the plane formed by the Y-axis and the lower arm defined from wrist joint (6 or 9) to elbow joint (5 or 8). The X-axis is the normal of the plane formed by the two previous axes.

These three coordinate systems were placed to the hip centre (0), shoulder centre (2) and shoulder joints (4 or 7) respectively, as depicted in Figure 2(b).

The trunk and shoulder joint angles were then computed according to the ISB recommendation. We changed the matrix decomposition sequences of the shoulder joint angle computation from YXY to ZXY, to obtain abduction values and to limit gimbal lock problems as suggested in Senk and Cheze (2006).

The Kinect skeleton does not provide enough points to compute the neck and elbow local coordinate systems. We computed the elbow joint angles using the vector convention detailed by Bonnechere et al. (2014).

The neck joint angles were computed by planar projection of the neck vector (2 to 3) expressed into the local trunk coordinate system.

As there is not enough information to compute the wrist angles, the wrist, and wrist twist RULA scores are set manually by the user. It consists in manually filling-it filed

in the user interface. Finally, all the threshold values that are not provided by the RULA method are set to  $20^\circ$ , such as Aptel et al. (2000) for the shoulder joint abduction.

### 3.2 Experimental setup

In this section, we present the experimental protocol used to evaluate the relevance of the proposed method in constrained conditions, such as work conditions. To this end, we carried-out an experimental protocol with 12 male participants (age:  $30.1 \pm 7.0$  years, height:  $1.75 \pm 0.046$  m, mass:  $62.2 \pm 7$  kg). They were equipped with 47 reflective markers positioned at standardised anatomical landmarks, as suggested in Wu et al. (2005). The motion of the participants was recorded by both Microsoft Kinect two sensor and a 15 camera Vicon optical motion capture system.

The subject had to perform getting and putting motions. More precisely, the subject had to carry a 40 cm by 30 cm by 17 cm box with two hands, place it in front of their abdomen, wait a few seconds and place it in the original position. The box (attached to a magnet) had two target placements, in order to generate two different motions. In the first placement named F (i.e., front) the target was located in front of the subject, at 1.70 m high, 0.35 m left and 0.50 m in front. In the second placement S (i.e., side) the target was located on the left of the subject, aligned with the two shoulders at the same height and 0.55 m left.

To simulate workplace environmental constraints, three experimental setups were defined, including manipulation box (to add occlusions during the manipulation task) and various Kinect placements:

- [NB – no box condition]: the manipulation of the box was simulated by the subject without using actually a box to avoid occlusions. The Kinect was placed in front of the subjects, as recommended by Microsoft. It enabled us to test the robustness of the Kinect sensor under favourable conditions. In this condition, the subject simply reached to the position of the attachment where the box would usually be located.
- [B – box]: the manipulation is actually performed with the box to create occlusions of parts of the body, as in real working situations. The Kinect was again placed in front of the subject, as recommended by Microsoft.
- [B45 – box and  $45^\circ$  sensor placement]: as in the B condition the subject actually manipulated the box but the Kinect was placed  $45^\circ$  left forward of the subject, as in real cluttered environments. In this condition, occlusion was more important.

The subject repeated each gesture five times: getting, and putting, for each conditions and box placement ( $F_{NB}$ ,  $F_{B45}$ ,  $F_B$ ,  $S_{NB}$ ,  $S_{B45}$  and  $S_B$ ):  $5 \times 3 \times 2 = 30$  motions were recorded for each subject.

In this experimentation, the correction was performed using a filtered motion graph made-up with 130 professional example gestures leading to 532,624 poses. These poses were captured using a traditional optical motion capture system to ensure motion naturalness and involve a range of object handling movement. The readers are referred to (Plantard et al., 2017c) for more details. The poses were then filtered into 2,048 nodes with an average of almost 7.81 links per node. The filtration intensity was chosen relative to the optimal condition used in Shum et al. (2013).



In real workstations the occlusion and camera placement may be much more important. Indeed, the sensor placement is highly constrained by the environment and many occlusions may occur, due to the objects which are being manipulated. Consequently, as a proof of concept, we evaluate the system in two simulated workstation scenarios, involving the manipulation of a real car seat, as depicted in the Figure 3.

**Figure 3** The two simulated workstations scenarios, (a) lightly occluded scenario (b) heavily occluded scenario (see online version for colours)



(a)



(b)

*Source:* Courtesy of Plantard et al. (2017b)

### 3.3 Comparison between corrected Kinect and Vicon data

Based on the joint angles computed with the two systems, we evaluated the accuracy of the joint angles measured with raw Kinect data or corrected ones, as these angles were used in RULA. To this end, we computed the root mean square error (RMSE) between the joint angles measured with the reference Vicon motion capture system, and those computed with the raw and corrected Kinect data, as described in Section 3.1.2. This RMSE has been applied to the eight main angles used in the RULA score:  $\alpha_T$  and  $\gamma_T$  for flexion and side bending of the torso respectively;  $\alpha_{LS/RS}$  and  $\beta_{LS/RS}$  for flexion and abduction of the the left/right shoulders respectively;  $\alpha_{LE/RE}$  for the left/right elbow flexion.

Some angles in the body posture exhibit low ranges, such as the trunk flexion  $\alpha_T$ , whereas the others exhibit large ranges, such as the shoulder flexion  $\alpha_{LS}$ . Consequently, displaying RMSE in a unique figure for all the joint angles may be difficult to analyse. Thus, to have a synthetic view of all the results in a unique figure, we defined a normalised RMSE value to minimise the effect of the range of angles:

$$nRMSE(\theta_i) = \frac{RMSE(\theta_i)}{\max(\theta_i) - \min(\theta_i)} \quad (1)$$

Kolmogorov-Smirnov test was used to check the normality of the distribution of the nRMSE for this analysis. The distributions did not follow a normal law. A Wilcoxon signed rank test was used for detect significant differences between Kinect error and corrected error for all the subject in each condition.

Secondly, based on the joint angles computed with the raw and corrected Kinect data, we computed the corresponding RULA score, as described in Section 3.1. In the same way, we computed the RMSE between the RULA score computed using the reference Vicon data and the two Kinect ones. Let us recall here that the RULA score ranges from one to seven only for each body side.

## 4 Results

Table 1 reports the joint angles ranges for all the joint angles and all the conditions, obtained with reference Vicon data.

**Table 1** Joint angle ranges for the six scenarios in degree

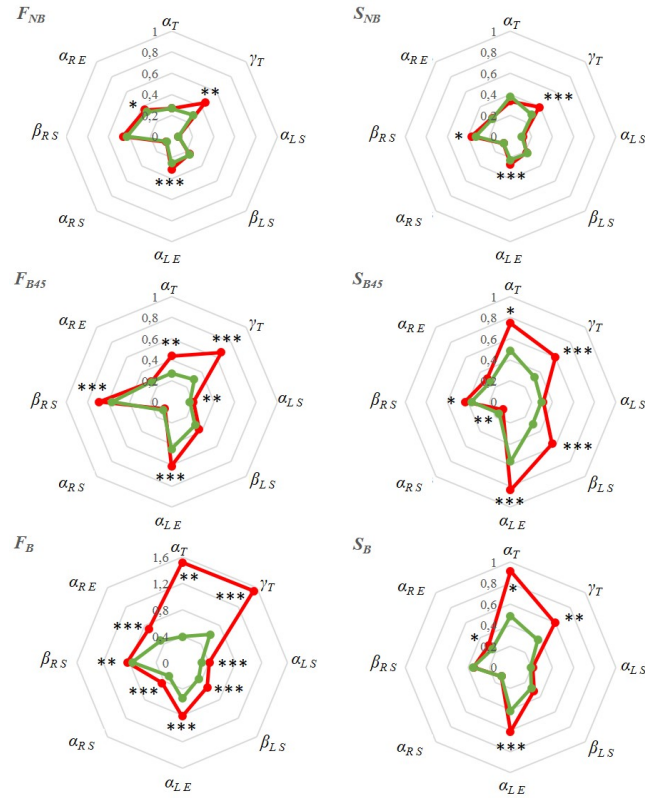
	$F_{NB}$	$F_{B45}$	$F_B$	$S_{NB}$	$S_{B45}$	$S_B$
$\alpha_T$	11.4	12.6	13.2	7.4	5.8	6.1
$\gamma_T$	19.2	19.0	18.2	33.0	29.0	29.4
$\alpha_{LS}$	130.8	104.6	108.6	111.5	93.5	90.9
$\beta_{LS}$	47.1	37.7	38.7	57.0	56.8	55.0
$\alpha_{LE}$	76.8	67.9	67.4	85.9	71.9	74.5
$\alpha_{RS}$	125.4	99.3	99.5	111.9	83.8	85.8
$\beta_{RS}$	30.7	25.1	24.3	38.7	36.8	35.8
$\alpha_{RE}$	73.9	69.0	68.4	83.1	73.9	75.5

Figure 4 shows the normalised RMSE for the eight main angles (one star diagram per type of trial) using equation (1). The results exhibit the nRMSE between zero (no error) to one (error corresponding to the range of motion). It is displayed for the six studied conditions.

In this figure, no-occlusion scenarios ( $F_{NB}$  and  $S_{NB}$ ) exhibit lower errors compared to those involving partial occlusion ( $F_{B45}$ ,  $S_{B45}$ ,  $F_B$  and  $S_B$ ). In no-occlusion scenarios, correction of the Kinect data did not significantly decrease this error. On the other hand, when occlusions occurred, corrected Kinect data leads to significantly ( $p < 0.001$ ) better estimation of joint angles in comparison to reference Vicon data.

For scenario  $F_B$  (displacing a box), nRMSE reached higher values than one for two torso angles:  $\alpha_T$  and  $\gamma_T$ . This is mainly due to the fact that the joint angles varied within a small range while occlusions due to the box led to high errors when using a Kinect placed in front.

**Figure 4** Normalised RMSE between reference angles using the Vicon data and using both raw Kinect (in red) and corrected Kinect (in green), for the six situations (see onlineversion for colours)



Note: Significance between the two performances is given by \*\*\* for  $p < 0.001$ .

Errors in RULA scores for corrected and raw Kinect data are reported in Table 2. Significant differences between the RMSE are noted with \*, \*\*, and \*\*\* for  $p < 0.05$ ,  $p < 0.01$  and  $p < 0.001$  respectively.

**Table 2** RMSE between the reference RULA score computed with Vicon data compared to using direct or corrected Kinect measurements

<i>Motion</i>	<i>RULA left</i>			<i>RULA right</i>		
	<i>Kinect</i>	<i>Correct</i>	<i>p</i>	<i>Kinect</i>	<i>Correct</i>	<i>p</i>
$F_{NB}$	0.49	0.50	NS	0.45	0.41	*
$F_{B45}$	0.66	0.65	NS	0.66	0.55	***
$F_B$	1.30	0.63	***	1.40	0.49	***
$S_{NB}$	0.55	0.63	*	0.65	0.62	**
$S_{B45}$	0.59	0.60	NS	0.62	0.45	***
$S_B$	0.51	0.51	NS	0.42	0.36	**

Note: Significance between the two performances is given by \*\*\* for  $p < 0.001$ .

In this table, one can see that the average error is below one for most of the scenarios, except  $F_B$  where occlusions with the box occurred. For scenarios with occlusions, frequently observed in real workstations, the correction method provides significantly better angles and RULA scores. Scores differ between left and right as it corresponds to different occlusion conditions. Kinect was placed  $45^\circ$  on the left of the subject, leading to more occlusions for right part of the body. These right-part occlusions increased when the user placed the box on his right ( $S$  conditions).

**Table 3** Percentage of correctly computed RULA score for the left and right body parts, using the direct Kinect measurement or the corrected one

<i>Motion</i>	<i>RULA left</i>			<i>RULA right</i>		
	<i>Kinect</i>	<i>Correct</i>	<i>p</i>	<i>Kinect</i>	<i>Correct</i>	<i>p</i>
$F_{NB}$	77	75	NS	80	82	*
$F_{B45}$	71	69	NS	62	75	***
$F_B$	52	69	***	55	79	***
$S_{NB}$	76	71	NS	76	78	**
$S_{B45}$	74	71	NS	64	80	***
$S_B$	78	77	NS	82	86	**

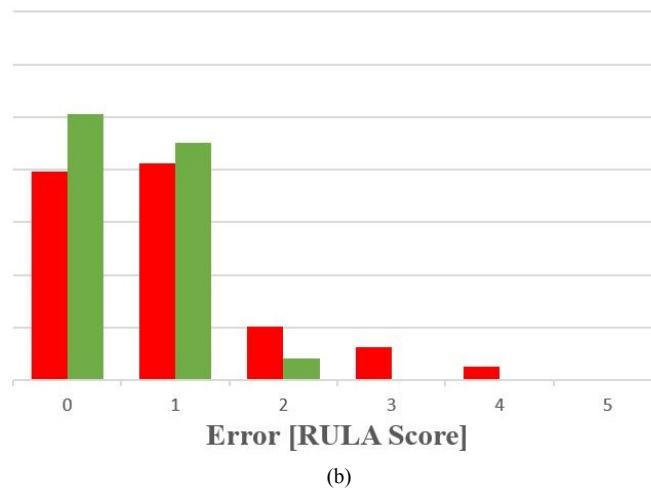
Note: Significance between the two performances is given by \*\*\* for  $p < 0.001$ .

RMSE is based on averaged errors and it would be interesting to also analyse the performance of using the Kinect correction to correctly compute the RULA score. To this end, Table 3 reports the percentage of correctly computed RULA scores (zero difference between the Vicon-based score and the Kinect-based scores) in all the conditions. For RULA scores based on raw Kinect data, this percentage is between 51% (most occluded condition) and 82% (few occlusions condition). In the worst case, with many occlusions, this percentage significantly raised from 52% (resp. 55%) to 69% (resp. 79%) for the left (resp. right) upper-limb. The above analyses were carried-out with the six controlled laboratory conditions.

As shown in Figure 3, the camera was not placed in the position recommended by Microsoft. The first scenario involved light occlusion, whereas heavy occlusion occurred in the second one. The Vicon system was also placed in the environment to measure reference data, as for the previous controlled conditions. For this two simulated workstations scenarios, Figure 5 depicts the histogram of the RULA score errors when

using raw and corrected Kinect data: the percentage of images where the error was equal to 0, 1, 2, 3, 4 and 5. These results show that almost all the errors greater than one disappeared when using the corrected data instead of raw Kinect data. As expected, the second scenario with many occlusions [Figure 5(b)] exhibited a lower occurrence of zero-error images in comparison to the first scenario [Figure 5(a)]. However, correction of the Kinect data enabled us to eliminate almost all the errors greater than one, while it corresponded to almost 20% of the cases without correction.

**Figure 5** Histogram showing the percentage of frames for which the RULA score error is 0, 1...5 when using raw (red) and corrected (green) Kinect data (see online version for colours)



Note: Two realistic scenarios are studied: a) simple one with few occlusions, b) complex one with many occlusions.

## 5 Discussion

The results were in accordance with those presented in Manghisi et al. (2017) without occlusions. However, it was difficult to compare other of our results because this recent study was limited to 15 static postures performed by one volunteer and without external occlusions. Our results showed that correction of Kinect data allowed us to significantly improve the joint angle accuracy, particularly when the body was partially occluded. RULA score was more reliable with corrected data, although angular thresholds tended to minimise the effect of noise when using Kinect raw data. Our results confirmed previous studies in real conditions (Plantard et al., 2017b) but added comparison to reference measurements in controlled situations (especially the occurrence of occlusions).

As RULA is based on angular thresholds, it tends to minimise the effect of noise when the angle is far from the thresholds. Hence, it leads to more acceptable errors than simply looking at the joint angles. However, in a multimedia interface delivering real-time feedback to a user about his RULA performance, this could be a problem. Indeed, if the user can visualise an avatar with a badly reconstructed pose, different from his actual performance, he may not be able to understand and adapt his performance to decrease the RULA score. In this type of real-time feedback system (Vignais et al., 2013), the coherence between the user's motion, the avatar pose and the RULA score is very important. In most of no-occlusion scenarios we could expect to have acceptable results. This statement could be confirmed by carrying-out perceptual studies.

The skeleton delivered by the Kinect did not contain all the required information to compute all the joint angles as accurately as using Vicon data with ISB recommendations. Particularly, Kinect delivered very noisy and unreliable information about the hand. Hand configuration is a key point in ergonomics, as reported in the RULA assessment scores. As it is not correctly measured by the Kinect most of the time, further research would be necessary to address this particular point. Hence, motion involving dexterous manipulation and fine motion of the wrist cannot be studied with such a system. The method used for correction involves that a minimum set of reliable information is delivered by the Kinect, which is not guaranteed for the wrist in Kinect v1 and v2.

Another limitation of the method is the use of a database that may not correspond to the actual use of the system. In this paper, we used a database trained with working motions, similar to those performed by the subject. For other type of motions, involving poses that have never been recorded before, especially for larger ranges of motions, the performance of the correction method would not be so good.

The method is also based on a set of parameters, such as the number of candidates used to run the optimisation, or the thresholds applied to prune the database and eliminate redundant information. It would be interesting to evaluate the actual impact of these parameters on the performance of the correction method.

Despite the reported limitations, the results of the current study are promising for the ergonomic evaluation of workstations. Kinect has already been considered as a promising tool to evaluate ergonomics on-site (Diego-Mas and Alcaide-Marzal, 2014; Patrizi et al., 2015), but only with simple and inaccurate posture representation. This study shows the applicability of our framework for a wider use and global evaluation tool. Using such automatic system enables to deliver a score at each frame (30 Hz with a Kinect), which is an improvement compared to traditional methods based on few key

frames. Indeed it could provide the amount of time spent above a given score as an additional information for the ergonomist.

**Figure 6** Example of ergonomics application based on our Kinect data correction framework (see online version for colours)



*Source:* Courtesy of Plantard et al. (2017b).

## 6 Conclusions

This paper presents an evaluation of the quality of angular and RULA score values when using a Kinect with software correction. The results showed that both corrected and uncorrected Kinect data enabled to compute acceptable to reliable angular and RULA score data in occlusion-free conditions. However, in more challenging environments with occlusions, kinematic data provided by the Kinect was more noisy, leading to inaccurate estimation of the joint angles. The proposed correction framework enables us to consider encumbered capture area (e.g., production chain) that leads to such occlusions or bad sensor placements. Uncorrected Kinect data exhibited much higher errors than corrected ones, which may lead to difficulties when using the system in real working environments. However, the correction method (Plantard et al., 2017a) shows a practical capacity to correctly provide ergonomic evaluation for working tasks with a cheap and easy-to-use system. Figure 6 depicts an example of potential application based on our framework, where joint angles and resulting RULA score are given to the ergonomists. The user can visualise the video, the 3D character, joint angles, and RULA scores for each frame of the recording, at 30 Hz. It provides supplementary temporal information, such as the time spent above a given RULA score.

The correction method used in this experiment makes it possible to obtain reliable information, necessary to estimate internal joint torques using inverse dynamics, even if occlusions occur (Plantard et al., 2017c). This opens new perspectives to define new fatigue or solicitation indexes based on continuous measurement contrary to classical static images generally used in ergonomics.

Finally, one has to notice that correction runs in real-time and allows the possibility to implement real-time user feedback, with potential application in training or virtual prototyping, as suggested by Vignais et al. (2013).

### Acknowledgements

This work was partially funded by the Cifre convention N1222/2012 and Faurecia Company. It was also partially supported by the Engineering and Physical Sciences Research Council (EPSRC) (Ref: EP/M002632/1) and the Royal Society (Ref: IE160609).

### References

- Aptel, M., Lahaye, S. and Gerling, A. (2000) 'Un outil d'évaluation des facteurs biomecaniques: OREGE (Outil de Reperage et Evaluation des Gestes)', *INRS (eds), Documents pour le medecin du travail*, Vol. 83, pp.217–223.
- Auvinet, E., Multon, F., Aubin, C.E., Meunier, J. and Raison, M. (2014) 'Detection of gait cycles in treadmill walking using a Kinect', *Gait and Posture*, Vol. 41, No. 2, pp.722–725.
- Auvinet, E., Multon, F. and Meunier, J. (2012) 'Lower limb movement asymmetry measurement with a depth camera', *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp.6793–6796.
- Bonnechere, B., Jansen, B., Salvia, P., Bouzahouene, H., Omelina, L., Moiseev, F., Sholukha, V., Cornelis, J., Rooze, M. and Jan, S.V.S. (2014) 'Validity and reliability of the Kinect within functional assessment activities: comparison with standard stereo-photogrammetry', *Gait and Posture*, Vol. 39, No. 1, pp.593–598.
- Cassola, F., Morgado, L., de Carvalho, F., Paredes, H., Fonseca, B. and Martins, P. (2014) 'Online-gym: a 3D virtual gymnasium using Kinect interaction', *Procedia Technology*, Vol. 13, pp.130–138.
- Chai, J. and Hodgins, J.K. (2007) 'Constraint-based motion optimization using a statistical dynamic model', *ACM Trans. Graph*, Vol. 26, No. 3, p.8.
- Clark, R.A., Pua, Y.H., Fortin, K., Ritchie, C., Webster, K.E., Denehy, L. and Bryant, A.L. (2012) 'Validity of the Microsoft Kinect for assessment of postural control', *Gait and Posture*, Vol. 36, No. 3, pp.372–377.
- Diego-Mas, J.A. and Alcaide-Marzal, J. (2014) 'Using Kinect sensor in observational methods for assessing postures at work', *Applied Ergonomics*, Vol. 45, No. 4, pp.976–985.
- Galna, B., Barry, G., Jackson, D., Mhiripiri, D., Olivier, P. and Rochester, L. (2014) 'Accuracy of the Microsoft Kinect sensor for measuring movement in people with parkinson's disease', *Gait and posture*, Vol. 39, No. 4, pp.1062–1068.
- Gameiro, J., Cardoso, T. and Rybarczyk, Y. (2014) 'Kinect-sign, teaching sign language to 'listener' through agame', *Procedia Technology*, Vol. 17, pp.384–391.
- Kurillo, G., Chen, A., Bajcsy, R. and Han, J.J. (2013) 'Evaluation of upper extremity reachable workspace using Kinect camera', *Technol. Healthcare*, Vol. 21, No. 6, pp.641–656.
- Liu, Z., Zhou, L., Leung, H. and Shum, H.P.H. (2016) 'Kinect posture reconstruction based on a local mixture of Gaussian process models', *IEEE Transactions on Visualization and Computer Graphics*, Vol. 22, No. 11, pp.2437–2450.
- Manghisi, V.M., Uva, A.E., Fiorentino, M., Bevilacqua, V., Trotta, G.F. and Monno, G. (2017) 'Real time RULA assessment using Kinect v2 sensor', *Applied Ergonomics*, in press.



- McAtamney, L. and Corlett, E.N. (1993) 'RULA: a survey method for the investigation of work-related upper limb disorders', *Applied Ergonomics*, Vol. 24, No. 2, pp.91–99.
- Pedersoli, F., Benini, S., Adami, N. and Leonardi, R. (2014) 'Xkin: an open source framework for hand pose and gesture recognition using Kinect', *Vis Comput*, Vol. 30, No. 10, pp.1107–1122.
- Patrizi, A., Pennestri, E. and Valentini, P.P. (2015) 'Comparison between low-cost marker-less and high-end marker-based motion capture systems for the computer-aided assessment of working ergonomics', *Ergonomics*, Vol. 59, No. 1, pp.155–162.
- Plantard, P., Auvinet, E., Le Pierres, A-S. and Multon, F. (2015) 'Pose estimation with a Kinect for ergonomic studies: evaluation of the accuracy using a virtual mannequin', *Sensors*, Vol. 15, No. 1, pp.1785–1803.
- Plantard, P., Shum, H.P.H. and Multon, F. (2017a) 'Filtered pose graph for efficient Kinect pose reconstruction', *Multimedia Tools and Applications*, Vol. 76, No. 3, pp.4291–4312.
- Plantard, P., Shum, H.P.H., Le Pierres, A-S. and Multon, F. (2017b) 'Validation of an ergonomic assessment method using Kinect data in real workplace conditions', *Applied Ergonomics*, Vol. 65, pp.562–569.
- Plantard, P., Muller, A., Pontonnier, C., Dumont, G., Shum, H.P.H. and Multon, F. (2017c) 'Inverse dynamics based on occlusion-resistant Kinect data: is it usable for ergonomics?', *International Journal of Industrial Ergonomics*, Vol. 61, pp.71–80.
- Senk, M. and Cheze, L. (2006) 'Rotation sequence as an important factor in shoulder kinematics', *Clinical Biomechanics*, Vol. 21, No. 1, pp.3–8.
- Shum, H.P.H., Ho, E.S., Jiang, Y. and Takagi, S. (2013) 'Real-time posture reconstruction for Microsoft Kinect', *IEEE Transactions on Cybernetics*, Vol. 43, No. 5, pp.1357–1369.
- Vieira, E.R. and Kumar, S. (2004) 'Working postures: a literature review', *Journal of Occupational Rehabilitation*, Vol. 14, No. 2, pp.143–159.
- Vignais, N., Miezal, M., Bleser, G., Mura, K., Gorecky, D. and Marin, F. (2013) 'Innovative system for real-time ergonomic feedback in industrial manufacturing', *Applied Ergonomics*, Vol. 44, No. 4, pp.566–574.
- Wang, Y., Dubey, R., Magnenat-Thalmann, N. and Thalmann, D. (2013) 'An immersive multi-agent system for interactive applications', *Vis Comput*, Vol. 29, No. 5, pp.323–332.
- Wu, G., Siegler, S., Allard, P., Kirtley, C., Leardini, A., Rosenbaum, D. and Others. (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', *Journal of Biomechanics*, Vol. 35, No. 4, pp.543–548.
- Wu, G., van der Helm, F.C., Veeger, H.D., Makhsous, M., Roy, P.V., Anglin, C., Nagels, J., Karduna, A.R., McQuade, K., Wang, X., Werner, F.W. and Buchholz, B. (2005) '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', *Journal of Biomechanics*, Vol. 38, No. 5, pp.981–992.
- Zhou, L., Liu, Z., Leung, H. and Shum, H.P.H. (2014) 'Posture reconstruction using Kinect with a probabilistic model', *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, pp.117–125