Safety Management and Human Factors, Vol. 189, 2025, 65–76
https://doi.org/10.54941/ahfe1006571

Impact of Introducing Sparse Inertial Measurement Units in Computer Vision-Based Motion Capture Systems for Ergonomic Postural Assessment

Hasnaa Ouadoudi Belabzioui[1,2], Pierre Plantard[2], Charles Pontonnier[1], Georges Dumont[1], and Franck Multon[1]

[1]Univ. Rennes, Inria, CNRS, IRISA, 35042 Rennes, France
[2]Moovency, 15 Rue du Chêne Germain, 35510 Cesson-Sévigné, France

ABSTRACT

In ergonomics, worker movement on site is an important factor in assessing the risk of musculoskeletal disorders, among other factors. Several commercial markerless motion capture systems that can be used for this purpose are available, mostly based on monocular or multi RGB (THEIA system)/RGB-D cameras (MS Kinect system). Hybrid systems combining computer vision and Inertial Measurement Units (IMUs) have been introduced, such as the KIMEA (1 RGB-D + 4 IMUs) and the KIMEA Cloud (1 RGB + 4 IMUs) solutions. Although previous works analysed the accuracy of some of these systems, the relevance of coupling computer vision and IMU has not been studied. Hence, we tested the performance of these systems in evaluating bimanual handling tasks, with partial occlusions of the body in the images. The THEIA system exhibits an average of 11.1° error for all the joints, with larger Root Mean Square errors on the wrists and the shoulder (>14° error). KIMEA Cloud with IMU obtained similar global RMS error (10.3° to 10.9° depending on the viewpoint), but with better results for the wrists (3.9° to 4.3°). The impact of coupling RGB-D images and IMU data is even bigger: the RMS error of the Kinect decreased from 17.2° down to 8.9° when adding the IMUs information (KIMEA system). This difference is even bigger for the wrists: 28.3° to 38.5° for the Kinect, and 3.8° to 4° for KIMEA. These results confirm the advantage of introducing a few IMU sensors, especially for the wrists which are badly tracked in the images.

Keywords: Motion capture, Human pose estimation, Evaluation, RULA

INTRODUCTION

In ergonomics, considering a worker's posture and movement is one of the important information to assess the risks of development of musculoskeletal disorders in the workplace. Exposure to biomechanical risk factors, including force, posture, and repetition, along with individual factors affecting the worker, increases the risk of work-related musculoskeletal disorders (WMSDs). Self-assessment, direct measurement, and observational techniques (Li and Buckle, 1999) are common methods to assess this risk. Observational methods, such as the RULA method (McAtamney and Corlett, 1993), involve directly evaluating the performance of the worker at the workstation. The accuracy and validity of the results obtained using these observational methods directly depend on the input information (Fagarasanu and Kumar, 2002). This information can be delivered using two main types of measurement systems: either based on wearable sensors (mainly inertial measurement units IMU nowadays), or based on cameras and computer vision algorithms. On the one hand, some commercial IMU-based systems exhibit an excellent accuracy, comparable to classical optoelectronic motion capture, but are difficult to implement in some real work situations (Li and Buckle, 1999) due to many practical factors, including discomfort, and because they are subject to drift, especially when exposed to magnetic field disturbance (Yunus et al., 2021). On the other hand, camera-based methods benefit from recent advances in computer vision and machine learning, to design markerless systems based on either depth (Microsoft Kinect Azure DK) or RGB cameras (Pavllo et al., 2019). The THEIA system (THEIA) applied this approach to multiple calibrated RGB cameras to enhance the accuracy of markerless human pose estimation. However, camera-based systems can suffer from partial occlusions, and have difficulties to accurately track the movement of small body parts, such as the hands, leading to less accurate human pose estimation.

Several studies have already evaluated these motion capture systems in the field of ergonomics (Menolotto et al., 2020; Humadi et al., 2021). Most of these studies only evaluated systems based on former Kinect systems (Plantard et al., 2017a; Manghisi et al., 2017), no longer commercialized today. More recently, a few studies have attempted to validate the use of monocular (Yuan and Zhou, 2023; Li et al., 2020; Nayak and Kim, 2021) or multiple RGB cameras (Kim et al., 2021) for postural assessment.

Recently, hybrid systems, such as the KIMEA (Moovency) or VIMU (Adjel et al., 2023) systems, proposed to fuse computer vision and sparse IMU data, to benefit from the advantages of both measurements. For example, the KIMEA (monocular RGB-D) and KIMEA Cloud (monocular RGB) systems propose to place IMUs in gloves to enhance accuracy of the wrist joint angle estimation. Indeed, wrists are frequently hidden in manipulation tasks, especially in cluttered environments. Moreover, the size of the wrist in the image is very small, which makes it difficult for computer vision systems to accurately reconstruct wrist joint angles. Hence, when coupling sparse IMUs on the hand and the forearm, the assumption is that it should compensate the inaccurate wrist joint angles of computer vision only systems. Although previous works analyzed the accuracy of some of the sensor-based or camera-based systems, the relevance of coupling computer vision and IMU has not been studied yet. In this paper, we propose to test the performance of these hybrid systems when evaluating bimanual handling tasks, including partial occlusions of the body in the images.

MATERIALS AND METHODS

In this study, we tested several commercial solutions. KIMEA is based on the native Microsoft Kinect pose estimation software, with additional IMUs on the forearms and the hands. Hence, comparing KIMEA and Kinect enabled us to evaluate the impact of using these IMUs. KIMEA cloud is based on a two-steps process: 2D pose estimation (Chen et al., 2017) used as input of a 3D estimation method inspired from (Jena et al., 2021). This is also completed by IMUs placed on the wrists and the forearms. The THEIA multicamera system is supposed to overperform all the monocular systems, as it relies on several points of view and a large training dataset. Comparing THEIA and KIMEA Cloud enabled us to evaluate the impact of adding IMUs.

Experimental Set-Up

The following experimental set up was approved by the Operational Committee for the Evaluation of Legal and Ethical Risks (COERLE) No. 2021-32. 12 participants, 3 women and 9 men (age: 32.6±10 years, height: 1.73±0.079 m, mass: 76±16 kg) participated in this study. Each participant was equipped with the Xsens inertial motion capture system (Roetenberg et al., 2013), considered as the reference system for our experiment (Robert-Lachaine et al., 2017; Kim et al., 2021). An anthropometric measurement and system calibration phase was carried out for each participant, as recommended by the supplier.

Three Orbbec depth cameras (Orbbec) were installed around the participant. The resulting depth images were used to run the Kinect (with Microsoft Kinect Azure DK) and KIMEA systems, with different viewpoints: front, side and behind the participant.

Six RGB cameras were also placed around the subject to provide RGB images required for the THEIA and KIMEA Cloud systems. The THEIA system used data from the 6 camera viewpoints to assess the participant's movement. As KIMEA Cloud is a monocular system, it has been tested with the 6 different viewpoints. Details about the depth and RGB camera placement is given in Figure 1.

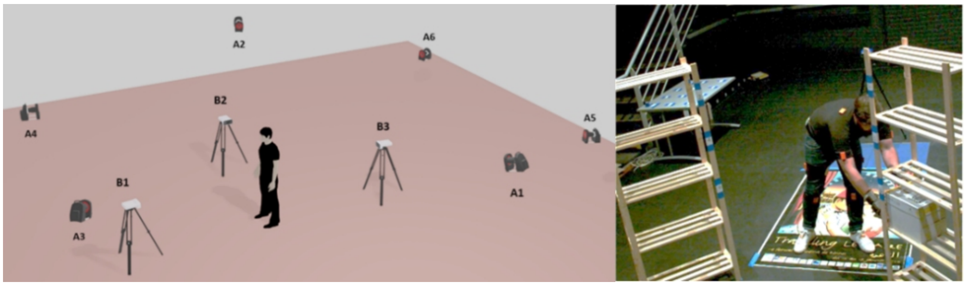

Figure 1: Left) 6 RGB (A1-A6) and 3 RGB-D (B1-B3) camera placements tested in this experiment. Right) picture of the bimanual handling task.

For KIMEA and KIMEA Cloud, four IMUs (integrated into specific gloves) were placed to the midpoint between the styloids, and the dorsal surface of the hand, at level of the third metacarpal bone, for each arm.

The participants simulated bimanual handling tasks: removing an empty cardboard box (size: 39 × 29.5 × 19cm, weight: 250g) from a three-tier shelf and transferring it to another one. This task was repeated five times consecutively with no waiting period in-between. The two shelves were positioned at 45° to the subject, with three different heights: 51 cm, 89 cm and 127 cm from the floor. The average task duration was 25 seconds. The order of shelves was predetermined and remained consistent across subjects. Pick and place order for bimanual handling is illustrated in Figure 1. The execution of the task generated external occlusion (with the box) and self-occlusion according to the different depth and RGB camera viewpoints.

One of the most popular WMSD risk assessment method is RULA (McAtamney and Corlett, 1993), based on joint angles (mainly flexions). We consequently evaluated the impact of joint angle estimation errors on this type of score. To compute the RULA score, each joint angle was assigned a value according to a range of predefined angles. For example, the arm score varied from 1 to 4 if shoulder flexion was between [−20°; 20°], <−20°or between [20°; 45°], between [45°; 90°], or >90° respectively. The same type of threshold was applied to the other joint angles. The scores for each joint were grouped into the A Score, for arms, forearms and wrists, and the B Score, for the neck, trunk and legs. An A score was calculated for the joint of the left upper limb and the right limb respectively. Other elements, such as "Muscle use" (repetitiveness) and "Force score" (external loads) were included in these A and B scores, to give the C Scores (for the left and right upper limbs), and the D Score (for neck, trunk and legs). These additional items were entered manually and set to the same value whatever the motion estimation system was used. Each C Score was combined with the D Score to provide the Final RULA Score for the left and the right parts of the body, ranging from 1 "acceptable" to 7 "immediate changes required". These left and right Final Scores led to a RULA Action Level Score summarized in four levels of intervention (from "acceptable posture" to "workstation requiring immediate changes").

Statistics

We compared the joint angles and corresponding RULA scores (RULA Action Level, left and right Final, left and right C and D scores) estimated with each evaluated system and the Xsens reference one, for each subject. The root mean square error (*RMSE*) was used to quantify this difference (Cao et al., 2017; Plantard et al., 2017b; Kim et al., 2021; Lahkar et al., 2022; Bertram et al., 2023). To compare our results to (Yuan and Zhou, 2023), we also computed the mean absolute error (*MAE*). MAE_j^{eval} is the absolute value of the error between the joint angle θ_j^{eval} of the evaluated system eval for the joint angle j, and the results θ_j^{ref} of the reference system ref for the joint angle j, computed as follows:

$$MAE_j^{eval}\left(\theta_j^{eval},\theta_j^{ref}\right) = \frac{\sum_{i=1}^{n}\left|\theta_j^{eval}(i)-\theta_j^{ref}(i)\right|}{n} \tag{1}$$

We computed the normalized mean absolute error (*nMAE*) to facilitate comparison of the joint angles with different ranges of motion. We normalized the MAE_j^{eval} of each joint j by the range of motion measured by the reference system:

$$nMAE_j^{eval} = \frac{MAE_j^{eval}\left(\theta_j^{eval}, \theta_j^{ref}\right)}{\max\left(\theta_j^{ref}\right) - min\left(\theta_j^{ref}\right)} \tag{2}$$

The correlation between the results of each evaluated system and the reference one was also calculated for the joint angles. A Kolmogorov-Smirnov test was used to verify the normality of the error distribution for these analyses. Since the distributions did not follow a normal distribution for this experiment, Spearman's correlation coefficient (ρ) was used. Finally, we compared the sensitivity of the different systems with the sensitivity of the reference one, by computing the number of times the RULA scores changed during the task. Moreover, we analyzed the Proportion agreement index (Po) of the RULA score (no difference between the RULA score obtained with the reference system, and the one based on the tested systems), for each system and camera placement.

RESULTS

Table 1 shows the *RMSE* in degrees for the 4 tested systems, for depth and RGB cameras placed in front of the subject, as generally recommended. These results show that the *RMSE* of the various calculated joint flexion angles were close to 10°, expect for Kinect (*RMSE*: 17.2°). The *RMSE* of the shoulder and elbow joints were lower for the THEIA system compared to the other evaluated systems.

Table 1: *RMSE* ± standard deviation expressed in degrees [°] for the work task performed during experimentation with frontal camera placement and for the main joint flexion angles required for RULA. Results in bold highlight the smallest errors, per joint flexion angle.

	KIMEA B1	THEIA	Kinect B1	KIMEA Cloud A3	KIMEA Cloud A4
Back	**3.7 ± 1.0**	4.3 ± 1.4	3.7 ± 1.0	5.1 ± 1.7	6.1 ± 1.1
Neck	5.8 ± 2.0	9.7 ± 2.8	5.9 ± 2.0	6.8 ± 1.4	**6.3 ± 1.7**
Left shoulder	15.2 ± 4.8	**14.2 ± 5.4**	15.2 ± 4.8	16.5 ± 3.4	16.6 ± 3.2
Right shoulder	16.0 ± 5.1	**14.2 ± 5.6**	15.9 ± 5.1	18.4 ± 3.0	15.5 ± 2.6
Left elbow	11.9 ± 2.3	**9.7 ± 3.3**	14.7 ± 4.3	15.6 ± 2.7	15.6 ± 4.1
Right elbow	11.2 ± 2.7	**9.4 ± 2.7**	15.7 ± 5.4	16.7 ± 2.7	14.1 ± 1.7
Left wrist	**3.8 ± 2.7**	14.1 ± 3.7	38.5 ±19.2	4.3 ± 3.1	4.3 ± 3.1
Right wrist	4.0 ± 2.4	13.0 ± 3.9	28.3 ± 5.2	4.0 ± 2.2	**3.9 ± 2.2**
Overall	**8.9 ± 2.9**	**11.1 ± 3.6**	**17.2 ± 4.6**	**10.9 ± 2.5**	**10.3 ± 2.5**

Table 2 shows the *MAE* in degrees and *nMAE* in percent for the 4 tested systems, for depth and RGB cameras placed in front of the subject. Kinect exhibited larger errors compared to the other systems, with an average 14.3° (23.8%) error. This error increased up to 32.8° (61.9%) for the wrist. THEIA obtained less accurate wrist joint angles (10.9°, 20.6%) compared to KIMEA

(3.1°) and KIMEA Cloud (3.1°). For these systems, large errors occurring for joints with large movements, such as the shoulders or elbows joints, lead to a lower percentage of error (*nMAE*).

Table 2: *MAE* ± standard deviation expressed in degrees [°] (*nMAE* expressed in percent [%]) when using the frontal camera placement and for the main joint flexion angles required for RULA. Results in bold highlight the smallest errors, per joint flexion angle.

	KIMEA B1	THEIA	Kinect B1	KIMEA Cloud A3	KIMEA Cloud A4
Back	**3.0±0.9 (5.3)**	3.6±1.3 (6.4)	**3.0±0.9 (5.3)**	3.8±1.2 (6.8)	4.8±0.9 (8.5)
Neck	**4.7±1.7 (26.0)**	8.0±2.5 (44.5)	**4.7±1.7 (26.2)**	4.8±0.8 (26.8)	**4.7±1.2 (26.4)**
Left shoulder	13.1±4.7 (12.9)	**12.0±5.2 (11.8)**	13.2±4.7 (12.9)	13.0±2.7 (12.7)	13.0±2.7 (12.7)
Right shoulder	13.4±5.3 (13.1)	**11.7±5.5 (11.6)**	13.3±5.2 (13.2)	14.6±2.5 (14.3)	12.1±2.4 (12.0)
Left elbow	9.6±2.2 (9.1)	**7.9±3.0 (7.5)**	12.1±3.8 (11.6)	12.3±2.1 (11.6)	12.2±4.0 (11.1)
Right elbow	9.3±2.3 (8.7)	**7.8±2.4 (7.3)**	12.9±5.2 (12.1)	12.4±2.1 (11.7)	10.8±1.3 (10.2)
Left wrist	**3.1±2.4 (5.8)**	10.9±3.5 (20.6)	32.8±19.9 (61.9)	3.3±2.4 (6.3)	3.3±2.4 (6.3)
Right wrist	**3.1±1.9 (6.5)**	10.1±3.3 (21.0)	22.6±4.9 (47.1)	**3.1±1.7 (6.5)**	**3.1±1.7 (6.4)**
Overall	**7.4±2.7 (10.9)**	9.0±3.3 (16.3)	14.3±4.5 (23.8)	8.4±1.9 (12.1)	8.0±2.1 (11.8)

Table 3 shows the correlations between estimated and reference joint angles, for the main joints. These results support the hypothesis that wrist angles are difficult to estimate with a computer vision approach. For example, the Kinect B1 exhibited very low correlation for the two wrist angles (0.14 and 0.18 for the left and right wrist flexion respectively). However, hybrid systems with IMUs benefit from additional information for the wrists, leading to higher correlation ($\geq$ 0.88). The neck flexion seemed also difficult to estimate with computer vision methods, especially for KIMEA Cloud system (correlations between 0.34 and 0.39).

Table 3: Spearman's correlation coefficient (ρ) for frontal camera placement and for the main joint flexion angles required for RULA. Results in bold highlight the highest correlations, per joint flexion angle.

	KIMEA B1	THEIA	Kinect B1	KIMEA Cloud A3	KIMEA Cloud A4
Back	0.92 ± 0.03	**0.96 ± 0.02**	0.92 ± 0.03	0.87 ± 0.05	0.77 ± 0.07
Neck	**0.59 ± 0.20**	0.51 ± 0.21	**0.59 ± 0.20**	0.34 ± 0.22	0.39 ± 0.21
Left shoulder	0.95 ± 0.02	**0.96 ± 0.02**	0.95 ± 0.02	0.85 ± 0.05	0.85 ± 0.06
Right shoulder	0.95 ± 0.03	**0.96 ± 0.02**	0.95 ± 0.02	0.81 ± 0.05	0.86 ± 0.03
Left elbow	0.90 ± 0.04	**0.94 ± 0.03**	0.88 ± 0.05	0.86 ± 0.05	0.88 ± 0.06
Right elbow	0.93 ± 0.03	**0.95 ± 0.02**	0.89 ± 0.05	0.85 ± 0.05	0.89 ± 0.05

Continued

Table 3: Continued

	KIMEA B1	THEIA	Kinect B1	KIMEA Cloud A3	KIMEA Cloud A4
Left wrist	**0.88 ± 0.28**	0.55 ± 0.18	0.14 ± 0.14	**0.88 ± 0.23**	**0.88 ± 0.24**
Right wrist	**0.91 ± 0.11**	0.60 ± 0.14	0.18 ± 0.16	0.90 ± 0.14	0.90 ± 0.13
Overall	**0.88 ± 0.09**	0.80 ± 0.08	0.69 ± 0.08	0.80 ± 0.11	0.80 ± 0.11

Table 4 shows the *nMAE* of the joint angles depending on the camera placements. The results show that the camera placement had an important impact on the joint angles estimation. Greater errors occurred for camera placed at the back and the side, for both KIMEA and Kinect (KIMEA back position: 17.4%, side position: 14.0%; Kinect back position: 33.8%, side position: 29.4%). The KIMEA Cloud system seemed less impacted by the camera placement with *nMAE* values ranging between 11.8% and 14.6%.

Table 4: The *nMAE* in precent [%] for the main joint flexion angles required for RULA, for different camera placements. Results in bold highlight the camera placement with the lowest errors, per joint flexion angle.

	KIMEA			THEIA	Kinect			KIMEA Cloud					
	B1	B2	B3		B1	B2	B3	A1	A2	A3	A4	A5	A6
Back	**5**	6	9	6	**5**	6	9	8	10	**7**	9	12	15
Neck	**26**	37	54	44	**26**	39	56	34	39	27	**26**	38	38
Left shoulder	**13**	14	18	12	19	**14**	18	13	18	13	13	**11**	14
Right shoulder	**13**	20	20	12	**13**	20	20	16	12	14	12	**11**	15
Left elbow	**9**	**9**	13	8	**12**	14	17	11	13	12	12	12	**10**
Right elbow	**9**	13	13	7	**12**	16	19	11	**9**	12	10	11	11
Left wrist	**6**	7	6	21	**62**	65	72	**6**	**6**	**6**	**6**	**6**	**6**
Right wrist	**7**	7	6	21	**47**	62	60	7	7	**6**	**6**	**6**	**6**
Overall	**11**	14	17	16	**24**	29	34	13	14	**12**	**12**	14	15

Table 5 shows the proportion agreement index (*Po*) of the global RULA scores obtained with the different tested systems and camera placements. Kinect exhibited the worst performance, with a *RULA Action Level Po* between 71 % to 74%, according to the camera placement. KIMEA obtained highest *Po* values, ranging from 80% to 87% depending on the camera position. KIMEA Cloud also exhibited *Po* scores that depend on the camera placement, between 80 % and 89%, with the highest scores when cameras were placed in front. For KIMEA and KIMEA Cloud, the *C Scores* (upper-limbs) decreased down to 55% and 66% respectively when the camera was placed in the back of the subject. THEIA exhibits high *Po* score for the RULA *Action Level* (86%) and for the *C Scores* (72–73%).

Table 5: Proportion agreement index (*Po*) in percent [%] for RULA scores, for the various motion capture systems and camera placements. Results in bold highlight the camera placement with the higher agreement for each system and for each RULA sub-score.

	KIMEA			THEIA	Kinect			KIMEA Cloud					
	B1	B2	B3		B1	B2	B3	A1	A2	A3	A4	A5	A6
RULA Action Level	**87**	80	81	86	**74**	71	**74**	80	81	87	**89**	85	84
Final RULA Left	**80**	76	71	70	**67**	60	59	72	65	74	**76**	73	69
Final RULA Right	**82**	72	71	71	**72**	59	61	66	72	72	**77**	71	68
Left C score	**72**	69	56	73	**56**	50	42	66	58	70	71	**72**	67
Right C score	**72**	60	55	72	**60**	45	42	60	71	68	**73**	72	66
D score	**74**	63	51	62	**73**	63	51	63	59	69	**70**	57	53

DISCUSSIONS AND CONCLUSION

Joint angles accuracy presented in this paper are consistent with those reported by previous works based on RGB-D cameras (Yuan and Zhou, 2023; Kim et al., 2021; Plantard et al., 2017b). These previous studies evaluated the former Kinect V2 depth camera, but the overall performance using Kinect Azure DK (used in this work), were similar to former versions (Bertram et al., 2023).

(Yuan and Zhou, 2023) reported *RMSE* of 12.9° and a *MAE* of 9.4°, for all the joint angles, when using a monocular RGB system. In our work, *RMSE* for KIMEA Cloud was 10.6°, and *MAE* was 8.2°. (Cao et al., 2017) reported smaller *RMSE* (8.3 °) when using an OpenPose system with 3 cameras, but it was mostly tested on static poses without occlusions. The joint angles errors with THEIA were consistent with those reported for different movements (Lahkar et al., 2022).

Globally, the joint angles obtained with the tested systems exhibited good correlation ($\rho \geq 0.82$) with the reference system for back, shoulder and elbow flexion angles. For neck joint angles, the correlation was weak to moderate (between 0.34 and 0.59). This is partially explained by a small variation of these angles in the studied movements, resulting in a larger normalized error (*nMAE* between 26% to 44.5%). As expected, THEIA and Kinect suffered from more important errors and lower correlation for the wrist joint angles (THEIA: $MAE = 10.5° \pm 3.4$, $\rho \leq 0.60$; Kinect: $MAE = 27.7° \pm 12.4$, $\rho \leq 0.18$), compared to systems embedding IMUs ($MAE \leq 3.3°$, $\rho \geq 0.88$). This result supports the hypothesis that current computer vision algorithms cannot accurately estimate movements of small body parts with large range of motions, and high risk of occlusion, such as the hands. In working tasks, hands are often used, and errors in estimating their motion may lead to unreliable postural assessments. Moreover, WMSDs affecting the forearm, the wrist and the hand are the first to second most common trouble affecting workers whatever the sector of activity (de Costa et al., 2015).

We obtained a proportion agreement (*Po*) for the RULA *Action Level Scores* of 0.83, 0.81, 0.73 and 0.84 for KIMEA, THEIA, Kinect and KIMEA Cloud, respectively. Results showed relatively few variations according to camera placement, with *Po* ranging from 0.80 to 0.87 for KIMEA, 0.71 to 0.74 for Kinect, and 0.81 to 0.89 for KIMEA Cloud. These results are consistent with previous works (Yuan and Zhou, 2023; Kim et al., 2021). Other works reported slightly better RULA estimations using Kinect V2 (Plantard et al., 2017b), but with different methodology and experimental design.

As expected, our results show that the camera placement affected the performance of computer vision methods. We reported lower joint angle errors when the camera was placed in front of the subject (*nMAE* for KIMEA: 10.9%, Kinect: 23.8% and KIMEA Cloud: 12.0%). The impact of the camera placement seemed more limited for the KIMEA Cloud system, with *nMAE* ranging from 11.8% to 14.6% for all camera placements. The most significant joint angle errors were found when the camera was placed on the back for the KIMEA (*nMAE*: 17.4%) and Kinect (*nMAE*: 33.8%). Neck joint error was over 50% when the depth camera was positioned on the back of the subject. The evaluated tasks involved leaning forward movements, which caused head occlusion with this camera placement.

Optoelectronic systems were not used in this work, as the infrared cameras and reflective markers lead to important interferences with depth cameras (Özsoy et al., 2022; Jo et al., 2022). Hence, in this work, we used the Xsens system as a reference system instead, similarly to other previous works (Kim et al., 2021). The measurement error of the Xsens was estimated around 2.8° for handling tasks (Robert-Lachaine et al., 2017), but other authors reported 14.5° errors with different experimental set-ups (Benjaminse et al., 2020). Additionally, these inertial systems may drift over long measurement sequences (Plamondon et al., 2007; Kim and Nussbaum, 2013). Further studies would be necessary to evaluate the impact of using Xsens as a reference system in that case.

Even if we designed the protocol to mimic real work conditions, the experiment was carried out in a laboratory condition. Further works would also be necessary to evaluate these systems on real working conditions. This is a complex task as it is almost impossible to control the test condition (especially the camera placement) and to ensure that a reference system would deliver actual reliable values. This is especially true for tasks involving wrist motions, that might be responsible for disorders, such as carpal tunnel syndrome (Malchaire et al., 1996). It is also important to notice that the various measurement systems rely on custom biomechanical models, which may differ from the ISB recommendations (Wu et al., 2005). The placement and number of anatomical landmarks provided by the different systems for calculating joint angles may also differ. We used the method described by (Kim et al., 2021) to limit the impact of these model differences. However, it would be interesting to add a more accurate calibration phase to approximate anatomical reference points (Robert-Lachaine et al., 2017; Xu et al., 2017) and decrease this potential bias.

To summarize, the different motion capture systems based on computer vision enabled a correct evaluation of the risk of WMSDs, especially when using the RULA assessment method. However, systems based on vision only suffered from high joint angle estimation errors, especially for the wrist joint angles. The results showed that hybrid systems consistently provided more accurate RULA scores, regardless of the camera's placement relative to the worker. Such evaluations contribute to our understanding of the capabilities and limitations of various kinematic data collection systems, thereby informing their practical implementation in ergonomic evaluation practices. The results are relevant to ergonomists who want simple, easy-to-deploy tools for on-site accurate postural assessment with a minimal setup.

ACKNOWLEDGMENT

This work was partially funded by the Cifre convention N° 0936/2021, Moovency company. This work was supported by French government funding managed by the National Research Agency under the Investments for the Future program (PIA) with the grant ANR-21-ESRE-0030 (CONTINUUM project).

REFERENCES

Adjel, M. Sabbah, M. Dumas, R. Mansard, N. Mohammed, S. Watier, B. Bonnet, V. (2023) Multi-modal upper limbs human motion estimation from a reduced set of affordable sensors, proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS).

Benjaminse, A. Bolt, R. Gokeler, A. Otten, B. (2020) A validity study comparing xsens with vicon, proceedings of the ISBS Archive 38.

Bertram, J, Krüger, T. Röhling, H. M. Jelusic, A. Mansow-Model, S. Schniepp, R. Wuehr, M. Otte, K. (2023) Accuracy and repeatability of the microsoft azure kinect for clinical measurement of motor function. Plos one Volume 18.

Cao, Z. Simon, T. Wei, S. E. Sheikh, Y. (2017) Openpose: realtime multi-person 2d pose estimation using part affinity fields, proceedings of the IEEE conference on computer vision and pattern recognition.

Chen, Y. Wang, Z. Peng, Y. Zhang, Z. Yu, G. Sun, J. (2018). Cascaded pyramid network for multi-person pose estimation, proceedings of the IEEE conference on computer vision and pattern recognition.

da Costa, J. Torres, J. Santos Baptista, and Mário Vaz. 2015. Incidence and prevalence of upper-limb work related musculoskeletal disorders: A systematic review. Work 51.4, 635–644.

Fagarasanu, M. Kumar, S. (2002) Measurement instruments and data collection: A consideration of constructs and biases in ergonomics research. International journal of industrial ergonomics Volume 30.

Humadi, A. Nazarahari, M. Ahmad, R. Rouhani, H. 2021. In-field instrumented ergonomic risk assessment: Inertial measurement units versus kinect v2. International Journal of Industrial Ergonomics Volume 84.

Jena, S. (2021) Monocular Human Shape and Pose with Dense Mesh-borne Local Image Features, proceedings of the 16th IEEE International Conference on Automatic Face and Gesture Recognition.

Jo, S. Song, S. Kim, J. Song, C. (2022) Agreement between azure kinect and marker-based motion analysis during functional movements: A feasibility study. Sensors Volume 22.

Kim, S. Nussbaum, M. A. (2013) Performance evaluation of a wearable inertial motion capture system for capturing physical exposures during manual material handling tasks. Ergonomics 56.

Kim, W. Sung, J. Saakes, D. Huang, C. Xiong, S. (2021) Ergonomic postural assessment using a new open-source human pose estimation technology (openpose). International Journal of Industrial Ergonomics Volume 84.

Lahkar, B. K. Muller, A. Dumas, R. Reveret, L. Robert, T. (2022) Accuracy of a markerless motion capture system in estimating upper extremity kinematics during boxing. Frontiers in Sports and Active Living Volume 4.

Li, G. Buckle, P. (1999) Current techniques for assessing physical exposure to work-related musculoskeletal risks, with emphasis on posture-based methods. Ergonomics Volume 42.

Li, L. Martin, T. Xu, X. (2020) A novel vision-based real-time method for evaluating postural risk factors associated with musculoskeletal disorders. Applied Ergonomics Volume 87.

Malchaire, J. B. Cock, N. A. & Robert, A. R. (1996). Prevalence of musculoskeletal disorders at the wrist as a function of angles, forces, repetitiveness and movement velocities. Scandinavian journal of work, environment & health.

Manghisi, V. M. Uva, A. E. Fiorentino, M. Bevilacqua, V. Trotta, G. F. Monno, G. (2017) Real time rula assessment using kinect v2 sensor. Applied Ergonomics Volume 65.

McAtamney, L. Corlett, E. N. (1993) Rula: A survey method for the investigation of work-related upper limb disorders. Applied ergonomics Volume 24.

Menolotto, M. Komaris, D. S. Tedesco, S. O'Flynn, B. Walsh, M. (2020) Motion capture technology in industrial applications: A systematic review. Sensors Volume 20.

Microsoft Kinect Azure DK website: https://azure.microsoft.com/.

Moovency website: https://moovency.com/.

Nayak, G. K. Kim, E. (2021) Development of a fully automated rula assessment system based on computer vision. International Journal of Industrial Ergonomics Volume 86.

Orbbec website: https://www.orbbec.com/microsoft-collaboration/.

Özsoy, U. Yıldırım, Y. Karaşin, S. Şekerci, R. Süzen, L. B. (2022) Reliability and agreement of azure kinect and kinect v2 depth sensors in the shoulder joint range of motion estimation. Journal of Shoulder and Elbow Surgery Volume 31.

Pavllo, D. Feichtenhofer, C. Grangier, D. Auli, M. (2019) 3d human pose estimation in video with temporal convolutions and semi-supervised training, proceedings of the IEEE Conference on Computer Vision and Pattern Recognition CVPR.

Plamondon, A. Delisle, A. Larue, C. Brouillette, D. McFadden, D. Desjardins, P. Larivière, C. (2007) Evaluation of a hybrid system for three-dimensional measurement of trunk posture in motion. Applied Ergonomics Volume 38.

Plantard, P. H. Shum, H. P. Multon, F. (2017a). Filtered pose graph for efficient kinect pose reconstruction. Multimedia Tools and Applications Volume 76.

Plantard, P. Shum, H. P. Le Pierres, A. S. Multon, F. (2017b) Validation of an ergonomic assessment method using kinect data in real workplace conditions. Applied ergonomics Volume 65.

Roetenberg D. Luinge H. Slycke P. (2013) Xsens MVN: Full 6dof human motion tracking using miniature inertial sensors. Xsens Technologies.

Robert-Lachaine, X. Mecheri, H. Larue, C. Plamondon, A. (2017) Validation of inertial measurement units with an optoelectronic system for wholebody motion analysis. Medical & biological engineering & computing Volume 55.

THEIA website: https://www.theiamarkerless.com/.

Wu, G., Van der Helm, F. C., Veeger, H. D., Makhsous, M., Van Roy, P., Anglin, C., Nagels, J., Karduna, A. R., McQuade, K., Wang, X., et al. (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 Volume 38.

Xu, X. Robertson, M. Chen, K. B. hua Lin, J. McGorry, R. W. (2017) Using the microsoft kinect™ to assess 3-D shoulder kinematics during computer use. Applied Ergonomics Volume 65.

Yuan, H. Zhou, Y. (2023) Ergonomic assessment based on monocular RGB camera in elderly care by a new multi-person 3D pose estimation technique (ROMP). International Journal of Industrial Ergonomics Volume 95.

Yunus, M. N. H. Jaafar, M. H. Mohamed, A. S. A. Azraai, N. Z. Hossain, M. S. (2021) Implementation of kinetic and kinematic variables in ergonomic risk assessment using motion capture simulation: A review. International Journal of Environmental Research and Public Health Volume 18.