International Journal of Industrial Ergonomics 80 (2020) 103025

Contents lists available at ScienceDirect

International Journal of Industrial Ergonomics

journal homepage: http://www.elsevier.com/locate/ergon

Inter-rater reliability of an inertial measurement unit sensor-based posture-matching method: A pilot study

Wonil Lee *, Jia-Hua Lin, Stephen Bao

Safety & Health Assessment & Research for Prevention (SHARP) Program, Washington State Department of Labor and Industries, Olympia, WA, 98504, USA

ARTICLE INFO

Keywords:
Work-related musculoskeletal disorders
Posture matching
Inertial measurement unit
Inter-rater reliability

ABSTRACT

Posture quantification is important for analyzing trunk and upper extremity loading and estimating the risk of work-related musculoskeletal disorders. This study introduced an inertial measurement unit sensor-based posture-matching (ISPM) method. The ISPM method is the concept that involves observers wearing sensors and simulating postures in photos or videos obtained at job sites to assess the postural risks for the trunk and upper extremity. We tested the inter-rater reliability of the method in 4 novice observers. For the trunk, shoulder/elbow, and wrist joint angle category, weighted kappa scores were 0.89, 0.52, and 0.26, respectively. The results of the ISPM method indicated moderate or high inter-rater reliability for trunk flexion, shoulder flexion/extension, shoulder abduction/adduction, and elbow flexion. The ISPM method was more reliable than conventional observation-based posture assessment tools used for trunk posture analyses. However, the assessment of upper extremity angles indicated that the reliability of ISPM was lower than that of conventional observation-based posture assessment tools. This was a proof-of-concept study conducted using a few samples. Therefore, further testing is necessary to support the findings.

Relevance to industry: The ISPM method requires a minimal level of training. It uses conventional video recording and imposes almost no interruptions to the workers' performance. The ISPM method showed moderate inter-rater reliability for the trunk, shoulder flexion/extension, shoulder abduction/adduction, and elbow flexion during the joint angle analysis.

1. Introduction

Work-related musculoskeletal disorders (WMSDs) are prevalent in various occupations in the United States (CPWR, 2018; Bureau of Labor Statistics, 2016; Marcum and Adams, 2017; Smith and Anderson, 2017). Between 2006 and 2015, the construction, transportation, utility and warehousing, and health care industries reported high rates of WMSDs claims, which are determined as compensable claims per 10,000 full-time equivalent employees (Howard and Adam, 2018). In the Washington State, between 1999 and 2013, 72.6% of WMSD compensable claims in all industry sectors were related to the back and upper extremities (i.e., shoulder, elbow, hand, and wrist) (Marcum and Adams, 2017). Repetitive motions, heavy lifting, awkward postures, and high-force exertions are, independently or combined, factors contributing to WMSDs of the back and upper extremities (Bernard et al., 1997; Keyserling et al., 1993). Therefore, posture quantification is important for the analysis of trunk and upper extremity loading in order to estimate the risk of musculoskeletal disorders. Such analysis is commonly performed by observers, such as occupational safety and health specialists, or ergonomists.

A checklist-type of observation technique is commonly used to assess the postural demands of jobs associated with a high prevalence of WMSDs (Dempsey et al., 2005; Lowe et al., 2019). The Rapid Upper Limb Assessment (RULA) (McAtamney and Corlett, 1993), Strain Index (Moore and Garg, 1995; Garg et al., 2017), and Quick Exposure Checklist (QEC) (David et al., 2008) are commonly used to assess work-related tasks such as postural risk factors for the trunk and upper extremities. These tools can be implemented by on-site worker observations or office/laboratory assessments based on photos or video recordings obtained at the job site. The Washington State Safety and Health Assessment and Research for Prevention (SHARP) program developed a conventional posture quantification method, which is a posture analysis tool (Bao et al., 2006), to assess the postural risk through observation. Using this tool, analysts can estimate joint angles by observing work postures on a computer screen.

The reliability of observational assessment tools is influenced by the

* Corresponding author.
E-mail address: lwon235@lni.wa.gov (W. Lee).

https://doi.org/10.1016/j.ergon.2020.103025
Received 4 May 2020; Received in revised form 13 July 2020; Accepted 31 August 2020
Available online 10 November 2020

rater's ability to process multiple sources of information simultaneously to estimate posture during on-site assessments (Bao et al., 2009). The rater's experience level, camera angle, body part size, and video quality affect the reliability of deferred (i.e., post-hoc off-site assessment) posture analysis methods using video data collected on-site (Bao et al., 2009; Lowe et al., 2014; Weir et al., 2011). Parallax errors, which are a common cause of individual differences in posture analysis outcomes, lead to wrist posture analysis variations among observers that are dependent on the photography angle and joint direction to be analyzed (Lau and Armstrong, 2011). Parallax errors can occur if the camera (or rater) perspective is not perpendicular to the plane of motion of the joint of interest (Lowe et al., 2014).

Because of the individual differences in implementing observation methods, direct measurement methods have been considered as alternatives to observational methods to objectively obtain body movement details. Instruments often used by professional ergonomists include an electrogoniometer, lumbar motion monitor, inclinometer, and motion capture system (Lowe et al., 2019). The optical motion capture system is considered as the gold standard motion tracking method (Kim and Nussbaum, 2013). Lowe et al. (2019) reported that 35% of ergonomists who responded to a survey in 2017 in the United States used this system. This was greater than the number of users of non-optical motion capture systems, such as the inertial measurement unit (IMU) sensor-based motion capture system. However, the optical motion capture system cannot be applied in the field or office/laboratory environments because it requires the installation of several cameras that surround the worker. These installed cameras only capture the sensors in a limited area surrounded by the cameras. Moreover, this method cannot be used to assess movements behind any obstacles, such as machinery and other objects.

IMU sensor-based (non-optical) motion capture systems were designed to quantify posture and overcome the restrictions of the optical motion capture system. IMU sensors worn by workers on designated areas of the upper limbs allow ergonomists to quickly assess multiple joint angles simultaneously. Quantification of the postural risk using inertial sensors has been evaluated as acceptable in terms of validity and reliability for field application in some industries, such as recycling and dairy farming (Asante et al., 2018; Schall et al., 2015, 2016). However, the sensors interfere with personal protective equipment such as the fall arrest system worn by workers with occupations that involve a high risk of falls (e.g., construction workers), thereby making the use of such systems cumbersome or limited for field-based postural risk assessments. Magnetometer disturbances due to ferrometallic materials on job sites and drifting issues involved with the gyroscope in IMU sensors are known to cause joint angle measurement errors that increase over time (El-Gohary and McNames, 2012). Magnetometer disturbance issues (Roetenberg et al., 2007; Robert-Lachaine et al., 2017, 2020) could be problematic when considering implementing IMU-based motion tracking systems for real-time postural load assessments at job sites because many industries have ferromagnetic materials on-site.

The current study is the first to introduce a posture-matching method as an intermediary solution that combines the benefits of observational and direct measurements (i.e., IMU motion capture system); this method is called the IMU sensor-based posture-matching (ISPM) method. Using this method, observers inspect the site and collect photos or videos containing the most common and worst postures. The ISPM method is designed to allow a rater in an office/laboratory who is wearing the IMU sensors to mimic the observed posture based on the photos or videos obtained at the job site. Therefore, workers are not required to wear sensors on-site. This technique can be incorporated as a routine tool for assessing the risk of WMSDs and implementing interventions to prevent injuries (Fig. 1). Dynamic posture analyses are expected to yield motion artifacts related to the wearable motion capture system (Bergamini et al., 2013). Furthermore, angular velocities also affect joint angle estimation errors (Kim and Nussbaum, 2013). According to a systematic review of wearable IMU system research articles published between 2005 and 2018 conducted by Poitras et al. (2019), the validity and reliability of the system remain uncertain because of conflicting results. They found a low level of accuracy in joint angle calculations in the upper extremities, including the shoulder and elbow, with complex dynamic motions. This is why the ISPM is designed to enable analyses of static postures and minimize motion artifacts when data are collected from the motion capture system. Similar to existing ergonomics assessment tools such as RULA, QEC, and Strain Index, task assessments are usually done in a static fashion. Therefore, the ISPM method is an optimal tool for repetitive tasks involving a cycle that can be sampled during a short period of time and assessed to determine the postural risk. In the case of non-repetitive tasks composed of various sub-tasks, limitations arise because posture analyses can be conducted for only part of the overall task.

The objectives of this proof-of-concept study were to test the inter-rater reliability of the IMU-based posture-matching method, to provide practical applications for the IMU sensor-based motion tracking system, and to identify the risk of WMSDs.

2. Materials and methods

2.1. Participants

Four observers were recruited to test the inter-rater reliability of the ISPM method. Three observers had basic experience with ergonomic risk assessment using observational methods through academic training and one had general experience with health and safety research. However, all were novice-level observers without ergonomic posture analysis field experience. Participation was voluntary. It was not necessary to obtain consent forms from the participants because of the minimal risk of the current study.

2.2. Instruments and experimental design

Ninety-two sample images including 43 different postures were prepared for this proof-of-concept study of the ISPM method. Postures in these photos represent those of typical manual material handling; these include a combination of neutral positions or deviations from the neutral positions of the upper extremities and trunk body segments. A professional ergonomist posed with the prescribed joint angles of the upper extremities in the laboratory and photos were taken from 3 different camera angles (0°, 45°, and 90°, Fig. 2). The prescribed joint angles were verified with a manual goniometer by another professional ergonomist.

The I2M IMU Motion tracking system (NexGen Ergonomics Inc, Pointe Claire, Quebec) was used to automatically track and record joint angles. IMU sensors were placed on the rater's hand, forearm, upper arm, and sternum to track and record the joint angles of the dominant upper extremity (Fig. 3). Sensors and body segments were aligned using an I-pose (also known as N-pose), and subjects maintained a neutral

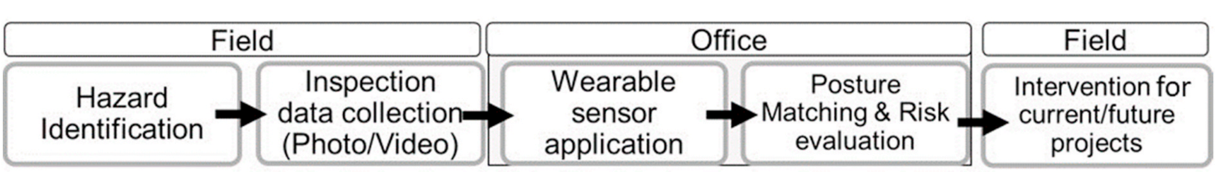

Fig. 1. Concept of the inertial sensor-based posture-matching method.

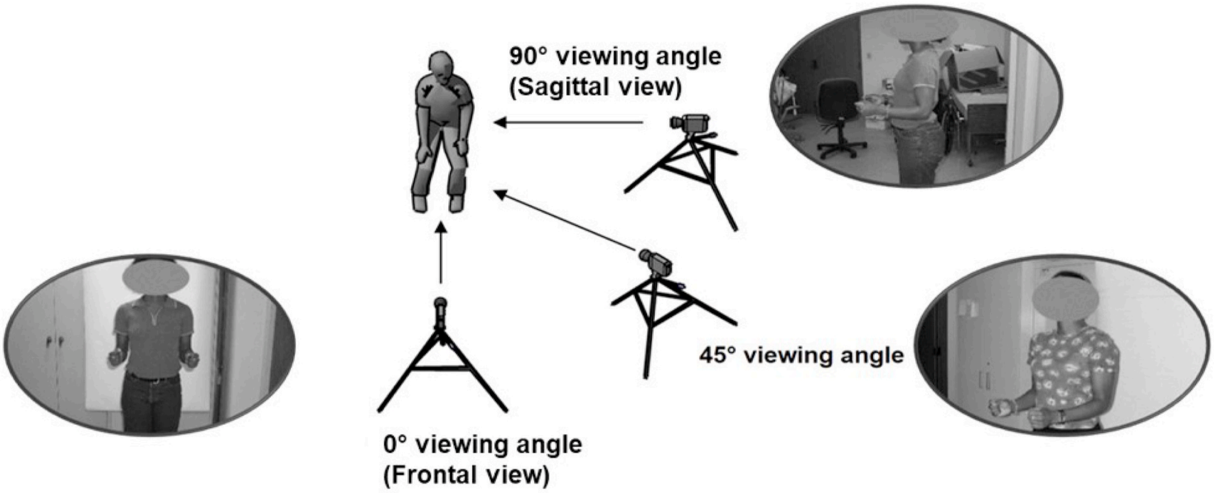

Fig. 2. Camera setup in three different angles.

Fig. 3. Experimental setup.

posture with their arm by the side and palms faced inward. The observers then mimicked the postures displayed in 92 randomly ordered sample images shown on a computer screen. Prior to collecting the data used for analyses, each rater conducted one training session using the same 92 photos to become familiar with the ISPM method and data collection process. The observers used an event marker tool of the I2M IMU Motion tracking system that can input binary on (1) and off (0) data as well as continuous time and angle output data. When the observers decided that a mimicked posture was most similar to the posture in the picture, they pressed a button of the marker and the on (1) signal was stored.

The current study compares the reliability of the ISPM method with that of conventional observational posture analysis methods. In the conventional posture quantification method (SHARP), the rater clicks on an appropriate location of the upper extremity and trunk using a mouse on the computer screen to identify the positions that more closely represent the postures estimated by the rater (Fig. 4).

Three trained observers (new participants who were different from the four observers mentioned previously) analyzed joint angles using the conventional posture analysis method and the same 92 posture frames analyzed using the ISPM method reliability test. We compared the reliability of the ISPM method among novice observers with the reliability of the conventional posture method used by trained observers.

2.3. Data analysis

TK Motion Manager software (version 1.0.0; NexGen Ergonomics Inc, Pointe Claire, Quebec) was used to calibrate the sensors. Human Motion Analyzer software (version 2.5.0; NexGen Ergonomics Inc, Pointe Claire, Quebec) was used to apply a Kalman filter (NexGen Ergonomics, 2016), and convert the raw IMU sensor data (sampling rate: 128 Hz) into joint angle measurements in the local Cartesian coordinate system according to International Society of Biomechanics recommendations (Wu et al., 2002, 2005). Raw joint angle values from the selected data set from each rater were then classified into posture categories pre-defined by Bao et al. (2009) as summarized in Table 1 and transferred to categorical variables for the ISPM method reliability tests.

The inter-rater reliabilities of the ISPM and the conventional posture analysis methods were evaluated using Fleiss kappa (κ) and Conger-weighted kappa (κ_w) coefficients. Prior research (Conger, 1980; McHugh, 2012; Klein, 2017) indicated that κ and κ_w were used for statistics among three or more raters (i.e., observers in the current study), whereas the reliability of assessment tools between two raters was examined using the Cohen kappa and weighted kappa. The weighted kappa test is designed to determine the inter-rater reliability using an ordinal variable as opposed to the evaluation of nominal variables of the simple kappa test. The degree of risk associated with musculoskeletal disorders also varies according to the posture category; therefore, the weighted kappa test also considers these variations. For example, for trunk flexion (Table 1), the incorrect assessment from Bin 3 to Bin 4 is more likely to lead observers (e.g., occupational safety and health specialists or ergonomists) and mangers to make an error when planning interventions than an incorrect assessment from Bin 2 to Bin 3 (i.e., for ordinal variables, a larger bin number represents a higher risk of

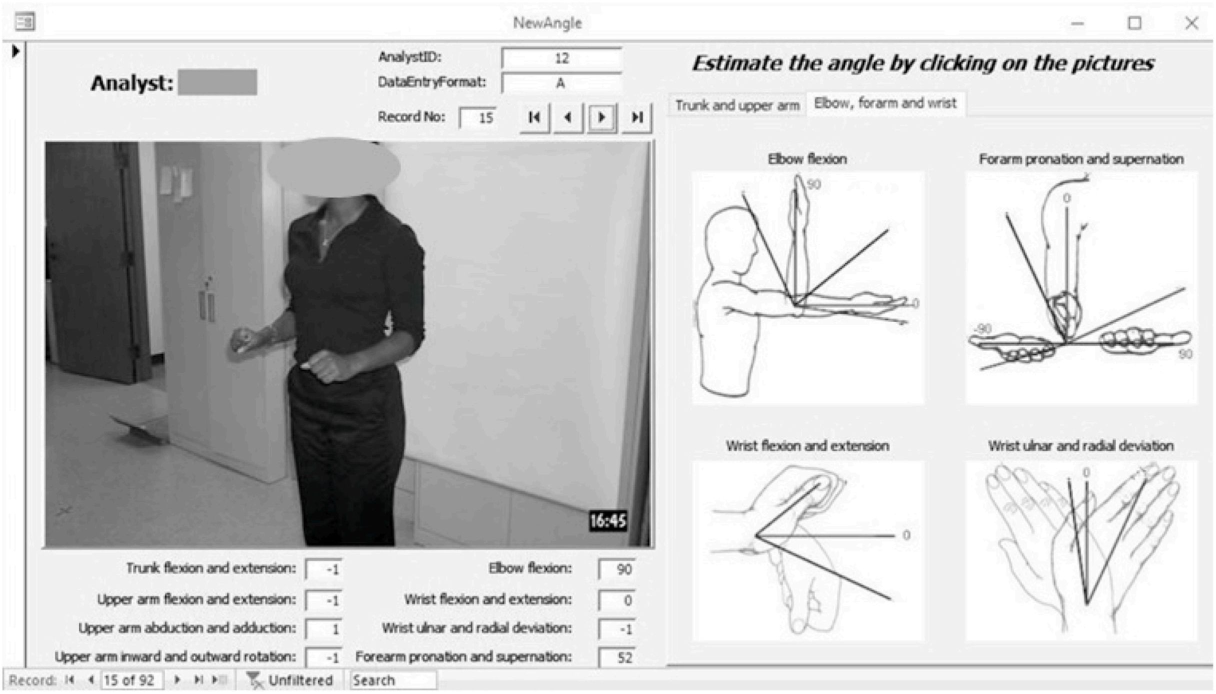

Fig. 4. Washington State SHARP Program posture analysis method.

Table 1
Predefined posture categorization (adapted from Bao et al., 2009).

Posture parameter	Posture category				
	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
Trunk flexion/extension	<−5° Extension	−5°–20° Neutral	20°–60° Flexion	≥60° Flexion	
Upper arm flexion/extension	<−5° Extension	−5°–20° Neutral	20°–45° Flexion	45°–90° Flexion	≥90° Flexion
Upper arm abduction/adduction	<−5° Adduction	−5°–30° Neutral	30°–60° Abduction	60°–90° Abduction	≥90° Abduction
Upper arm rotation	<−5° Outward	−5°–15° Neutral	15°–45° Inward	≥45° Inward	
Elbow flexion	<−5° Hyperextension	−5°–20° Extension	20°–60° Neutral	60°–100° Flexion	≥100° Flexion
Forearm pronation/supination	<−45° Supination	−45°–45° Neutral	≥45° Pronation		
Wrist flexion/extension	<−45° Extension	−45° to −15° Extension	−15°–15° Neutral	15°–45° Flexion	≥45° Flexion
Wrist ulnar/radial deviation	<−15° Radial	−15°–20° Neutral	≥20° Ulnar		

WMSDs). The kappa statistics were interpreted as follows: values larger than 0.8, perfect reliability; 0.61–0.80, excellent reliability; 0.41–0.60, moderate reliability; 0.21–0.41, fair reliability; 0.0–0.20, slight reliability; and less than 0, no reliability (Landis and Koch, 1977). STATA version 15 (StataCorp LP, College Station, TX) was used for statistical analyses.

3. Results

The results of the inter-rater reliability based on κ and $\kappa_w$statistics testing for the ISPM and conventional posture observation analysis methods are summarized in Table 2. The κ_w inter-rater reliability for the ISPM method was 0.52, 0.26, and 0.89 for the shoulder and elbow, wrist, and trunk joint angles, respectively. The corresponding $\kappa_w$scores were 0.68, 0.78, and 0.86 for the conventional posture observation method. The current study also calculated the lump sum reliability score of all upper extremity angles to compare the reliability score of conventional tools that were published in the same lump sum manner (e.g., RULA). The results showed that the inter-rater reliability of the ISPM method was moderate or higher for most posture variables among the 4 different novice observers, especially for trunk flexion/extension, shoulder flexion/extension, shoulder abduction/adduction, and elbow flexion. A low inter-rater reliability was noted for shoulder rotation, wrist flexion/extension, and wrist ulnar/radial deviation. Shoulder flexion/extension and abduction/adduction showed good to excellent inter-rater reliability; the reliability scores of the ISPM method for these body segments were higher than those from the conventional posture analysis method as evaluated by the weighted kappa (Table 2). Kappa statistics for elbow flexion analysis by the ISPM method indicated a moderate inter-rater reliability. The inter-rater reliability of both shoulder flexion and extension was higher for the ISPM method than for the conventional posture observation method. The assessment of forearm supination/protonation and wrist flexion/extension using the ISPM method was fairly reliable according to the weighted kappa statistics. The ISPM method showed slight to no inter-rater reliability for shoulder rotation and wrist ulnar/radial deviation among the 4 observers. The reliability

Table 2
Inter-rater reliability of inertial sensor-based posture matching and event-based posture analysis.

	Inertial sensor-based posture matching		Conventional posture observation method	
Body part	κ	$κ_w$	κ	$κ_w$
Trunk				
Trunk flexion/extension	**0.89**	**0.89**	**0.86**	**0.86**
Shoulder and elbow				
Shoulder flexion/extension	0.72	0.79	0.63	0.76
Shoulder abduction/adduction	0.55	0.70	0.55	0.68
Shoulder rotation	0.00[a]	0.11	0.55	0.58
Elbow flexion	0.49	0.66	0.89	0.92
Forearm supination/pronation	0.31	0.34	0.44	0.47
Mean (shoulder and elbow)	**0.41**	**0.52**	**0.61**	**0.68**
Wrist				
Wrist flexion/extension	0.19	0.28	0.50	0.55
Wrist ulnar/radial deviation	0.21	0.24	1.00[b]	1.00[b]
Mean (wrist)	**0.20**	**0.26**	**0.75**	**0.78**
Mean (upper extremity)[c]	**0.35**	**0.45**	**0.65**	**0.71**

Note: Bold texts represent the lump sum reliability score of a group of body regions compared to the reliability score of conventional tools. [a]Methodological problem estimating κ due to potential sensor errors. [b]Too few rating categories. [c]Lump sum reliability score of the shoulder, elbow, and wrist joint angles.

scores of shoulder rotation and wrist ulnar/radial deviation assessment were higher in the event-based posture analysis method than in the ISPM method (see Table 2).

Elbow flexion showed low reliability results with the 45° camera angle (Fig. 5a); forearm supination/protonation assessments were more reliable at 45° and 90° camera view angles compared to the results at 0° (Fig. 5b). The inter-rater reliability of trunk flexion/extension was consistently high across all camera angles (Fig. 5c). For the shoulder abduction/adduction assessment, inter-rater reliability was excellent based on the weighted kappa statistics estimated in both the frontal (0°) and sagittal views (90°) (Fig. 5d). The highest level of inter-rater reliability for the shoulder was found in the 0° angle view for the flexion/extension assessment (Fig. 5e). The ISPM showed very low reliability for the shoulder rotation assessment for all camera angles (Fig. 5f). The inter-rater reliability showed the greatest variability across camera angles for the wrist flexion/extension angle (Fig. 5g). Wrist flexion/extension also showed the highest inter-rater reliability at 45°, with very low inter-rater reliability when the assessment was performed in the frontal plane view (Fig. 5g). Wrist ulnar/radial deviation showed the highest reliability among the 4 observers for the evaluation performed in the sagittal plane (Fig. 5h).

4. Discussion

The proposed ISPM method is an ergonomics posture quantification technique based on mimicking postures viewed from videos or photos of workers taken at the job site. This method does not require the workers to be interrupted in order to perform the postural assessment and allows for the simultaneous analysis of multiple joint angles in the upper extremities.

The large difference between the Fleiss and weighted kappa (i.e., higher weighted kappa than Fleiss kappa) means that the wrong category is selected more frequently in the lower weighted category, which corresponds to smaller trunk or upper extremity joint angles (Table 1). Observers need to identify awkward postures with higher joint angle deviations. The current study found a higher weighted kappa compared with the unweighted kappa statistic value, which suggests that ISPM is more reliable when the joint angles are in the high-risk zone for WMSD

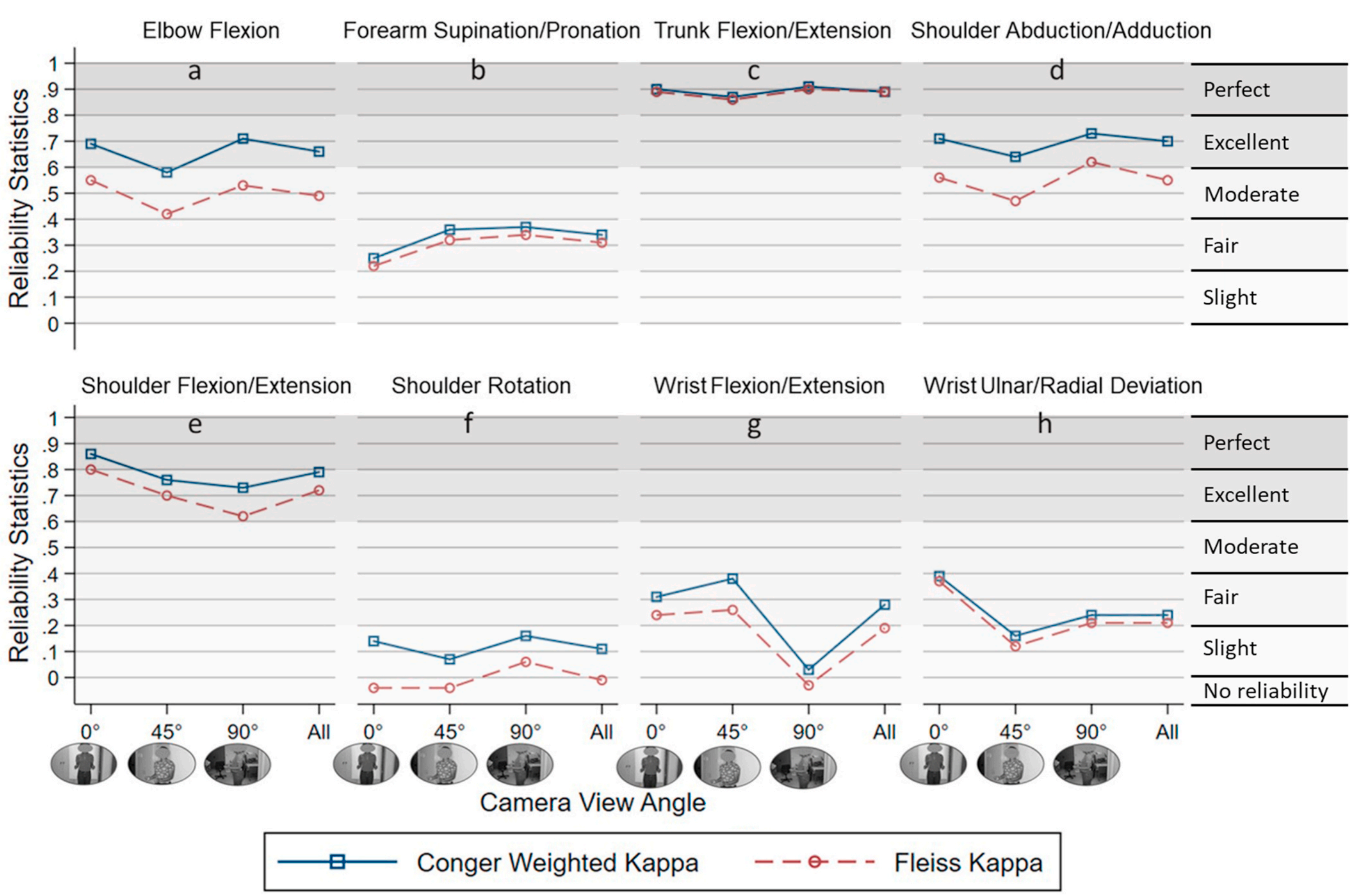

Fig. 5. Comparison of inter-rater reliability statistics for upper extremities and trunk joint angles based on the camera angle.

development (e.g., higher inter-rater reliability in 45°–90° of shoulder flexion than in −5°–20°). The set of 92 test images only contained limited range of joint angles. For instance, wrist ulnar/radial deviation only ranged from −10° to 30° in the photo frames; therefore, only Bin 2 in Table 1 could be consistently observed by observers or estimated by the inertial sensors.

The inter-rater reliability of the conventional observation-based posture assessment tools varied according to the activity being analyzed and the training level of the observers participating in the studies. The κ_w of the ISPM method for wrist angles (0.45) was greater than the value for hand/wrist posture evaluation in the Strain Index (0.34) reported by Spielholz et al. (2008) (Table 3). The inter-rater reliability of the trunk posture assessment using the ISPM method was greater than that of the RULA, while for the assessment of upper extremity angles, the value (κ_w 0.45) was lower than that of OCRA in a study analyzing cheese processing activities [intraclass correlation coefficient (ICC), 0.54]. Comparing the κ_w scores of the ISPM and QEC methods indicated that the ISPM is a more reliable tool for trunk posture analysis, while the upper extremity score was higher in the QEC assessment. Overall, the ISPM upper extremity inter-rater reliability scores were lower than those of other conventional observational posture analysis methods. However, the ISPM included posture assessment for more movements such as shoulder rotation and forearm supination/protonation, which other tools could not assess.

The ISPM method showed higher reliability scores for the trunk, shoulder, elbow, and forearm than for the wrist. Body segment size could affect the observers' visual attention levels when evaluating the photo frames. Larger body segments might be easier to observe and mimic, while smaller segments (i.e., wrist) require closer inspection and careful visual interpretation. For example, a rater would not be required to pay as much attention to effectively mimic shoulder movement in comparison with forearm protonation and wrist extension. Thus, the inter-rater reliability of the ISPM method was higher for larger body segments.

The effect of the camera angle was especially visible in the angles between smaller body segments, including wrist flexion/extension. The camera angles with a higher inter-rater reliability also varied between body segments; when shoulder flexion was analyzed from a frontal view, the inter-rater reliability decreased. However, the inter-rater reliability of shoulder rotation and elbow flexion was higher in the same plane. To achieve optimal reliability and results with the ISPM method, the setup of the camera angle should be determined based on an investigation determining which part of the body is exposed to the highest risk.

This study found that observers must take several pictures of workers' problematic postures on the job site, from multiple camera viewpoints, for the practical application of the ISPM method, as suggested by conventional observational method guidelines (Lowe et al., 2014). Therefore, the ISPM method can be optimally applicable to a repetitive task during which a cycle can be sampled for postural risk assessment without interrupting workers. For the application of the ISPM method, observers should pay more attention to the mimicking of posture angles when they assess jobs require combined/complex body segment movements.

One source of error contributing to the lower reliability of the results was the camera angle (Fig. 5), especially for wrist flexion/extension. The reliability of the rater's posture analysis decreased if the postures were photographed in a parallel direction to the performed movement. The observers did not have a clear perception of wrist and shoulder internal and external rotation from the images. Trunk flexion, shoulder abduction and rotation, elbow flexion, and wrist ulnar/radial deviation showed worse inter-rater reliability with the ISPM because joint rotations are not clearly visible at a 45° angle. Sensor errors, such as gyroscope drift, could contribute to the low reliability. Such an error can be minimized using a Kalman filter (Chen et al., 2018). The potential effect of drifting on the joint angle estimation was previously observed by comparing the joint angles at the beginning and end of a 12–15 min period of data collection in the same static neutral posture; drifting was found to influence the reliability of shoulder rotation measurements (Lee et al., 2019). Chen et al. (2018) tested the IMU system joint angle estimation performance using various filters including Kalman, complementary, and particle filters. The current study only applied the Kalman filter during the joint angle calculation process. Other filter types should be tested to determine the extent to which sensor measurement errors can be minimized.

These results are preliminary due to the limited number of subjects tested during this proof-of-concept study. The current study was

Table 3
Comparison of the inter-rater reliability of posture analysis tools for trunk flexion and upper extremities.

Assessment tool	Activity analyzed	Rater training level	Body part	Inter-rater reliability statistics	Reference
ISPM	Manual material handling	Novice	Trunk	**κ (0.89)**	Present study
				κ_w (0.89)	
			Shoulder, elbow, forearm	**κ (0.41)**	
				κ_w (0.52)	
			Wrist	κ (0.20)	
				κ_w (0.26)	
Conventional posture observation	Manual material handling	Expert	Trunk	**κ (0.86)**	Present study
				κ_w (0.86)	
			Shoulder, elbow, forearm	**κ (0.61)**	
				κ_w (0.68)	
			Wrist	**κ (0.75)**	
				κ_w (0.78)	
Strain Index	Manufacturing, health care	Novice/Expert	Wrist[a]	κ_w (0.34)	Spielholz et al. (2008)
				ICC (0.66)	Stevens et al. (2004)
RULA	Computer use	Novice/Expert	Trunk (incl. neck/legs)	**ICC (0.76)**[c]	Dockrell et al. (2012)
			Shoulder, elbow, and wrist[b]	**ICC (0.61)**[c]	
PATH	Hospital workers	Expert	Trunk	**κ (0.53)**	Park et al. (2009)
			Upper extremities	**κ (0.52)**	
OCRA	Cheese manufacturing	Expert	Upper extremities[d]	**ICC (0.54)**	Paulsen et al. (2015)
QEC	Fifty-one work tasks	Expert	Trunk	**κ_w (0.70)**	Oliv et al. (2019)
			Shoulder, elbow, forearm	**κ_w (0.74)**	

Note: Bold texts represent a greater than moderate level of reliability. [a]Hand/wrist posture. [b]Wrist/arm posture, muscle use, and force. [c]Average value of two trial tests. [d]Awkward posture/movement results.

conducted with novice observers only; further research using expert observers is needed to verify the reliability difference according to the experience level. The 92 posture frames utilized in the current study only included a limited range of movement (e.g., trunk flexion between 0° and 30°). Thus, future studies should include more varied postures with a wide range of movement in order to reflect the task demands of workers exposed to a high risk of WMSDs.

5. Conclusions

Conventional observational ergonomic assessment tools usually require a significant amount of rater training and experience, which poses a huge barrier for their broad implementation and industry adoption. In this study, we introduced an IMU-based posture matching method, an alternative approach to improve the reliability of conventional observation-based posture analysis methods. The ISPM method is a tool that combines the benefits of observational and direct posture analysis tools. It is designed to be used after a minimal level of training even by novice observers. The proposed posture-matching method utilizing the IMU system is a post-hoc ergonomics posture quantification method that industry practitioners can perform in controlled environments without ferrometallic materials based on videos or photos of worker activity recorded on the job site. The outcomes of the ISPM method showed good inter-rater reliability in regards to trunk, shoulder flexion/extension, shoulder abduction/adduction, and elbow flexion. The outcomes for the other joints were not reliable among the different novice observers. The ISPM method was more reliable than conventional observation-based posture assessment tools used for trunk posture analyses. However, the assessment of upper extremity angles indicated that the reliability of ISPM was lower than that of conventional observation-based posture assessment tools.

Author statement

Wonil Lee: Conceptualization, Methodology, Formal analysis, Investigation, Resources, Data curation, Writing - Original Draft, Writing - Review & Editing, Visualization.

Jia-Hua Lin: Conceptualization, Methodology, Investigation, Resources, Data curation, Writing - Review & Editing, Supervision, Project administration.

Stephen Bao: Conceptualization, Methodology, Investigation, Resources, Writing - Review & Editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

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

Acknowledgements

This research was funded in part by the U.S. National Institute for Occupational Safety and Health (OH1007316) and the Washington State Department of Labor and Industries SHARP internship program. The findings and conclusions of this research are those of the authors and do not necessarily represent the views of the U.S. National Institute for Occupational Safety and Health. We also thank Amanda Robinson for her help with data collection.

References

Asante, B.O., Bath, B., Trask, C., 2018. Trunk posture assessment during work tasks at a Canadian recycling center. Int. J. Ind. Ergon. 68, 297–303.

Bao, S., Silverstein, B., Howard, N., Spielholz, P., 2006. The Washington state SHARP approach to exposure assessment. In: Marras, W.S., Karwowski, W. (Eds.), Fundamentals and Assessment Tools for Occupational Ergonomics. Taylor & Francis Group, Boca Raton, FL, USA, pp. 44–1–44–22.

Bao, S., Howard, N., Spielholz, P., Silverstein, B., Polissar, N., 2009. Interrater reliability of posture observations. Hum. Factors 51 (3), 292–309.

Bergamini, E., Guillon, P., Camomilla, V., Pillet, H., Skalli, W., Cappozzo, A., 2013. Trunk inclination estimate during the sprint start using an inertial measurement unit: a validation study. J. Appl. Biomech. 29 (5), 622–627.

Bernard, B.P., Putz-Anderson, V., Burt, S.E., Cole, L.L., Fairfield-Estill, C., Fine, L.J., et al., 1997. Musculoskeletal Disorders and Workplace Factors: a Critical Review of Epidemiologic Evidence for Work-Related Musculoskeletal Disorders of the Neck, Upper Extremity, and Low Back. US Department of Health and Human Services (DHHS) publication no. 97-141 (Cincinnati: National Institute for Occupational Safety and Health).

Bureau of Labor Statistics, 2016. Nonfatal Occupational Injuries and Illnesses Requiring Days Away from Work, 2015. BLS. USDL-16-2130. Available from. https://www.bls.gov/news.release/pdf/osh2.pdf.

Chen, H., Schall Jr., M.C., Fethke, N., 2018. Accuracy of angular displacements and velocities from inertial-based inclinometers. Appl. Ergon. 67, 151–161.

Conger, A.J., 1980. Integration and generalization of kappas for multiple raters. Psychol. Bull. 88 (2), 322–328.

CPWR, 2018. Musculoskeletal disorders in construction and other industries. The Construction Chart Book: the U.S. Construction Industry and its Workers, sixth ed. CPWR—The Center for Construction Research and Training: Silver Spring, MD, USA, p. 24. Available from. https://www.cpwr.com/chart-book-6th-edition-fatal-and-nonfatal-injuries-musculoskeletal-disorders-construction-and-other#Musculoskeleta.

David, G., Woods, V., Li, G., Buckle, P., 2008. The development of the Quick Exposure Check (QEC) for assessing exposure to risk factors for work-related musculoskeletal disorders. Appl. Ergon. 39 (1), 57–69.

Dempsey, P.G., McGorry, R.W., Maynard, W.S., 2005. A survey of tools and methods used by certified professional ergonomists. Appl. Ergon. 36 (4), 489–503.

Dockrell, S., O'Grady, E., Bennett, K., Mullarkey, C., Mc Connell, R., Ruddy, R., et al., 2012. An investigation of the reliability of Rapid Upper Limb Assessment (RULA) as a method of assessment of children's computing posture. Appl. Ergon. 43 (3), 632–636.

El-Gohary, M., McNames, J., 2012. Shoulder and elbow joint angle tracking with inertial sensors. IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 59 (9), 2635–2641.

Garg, A., Moore, J.S., Kapellusch, J.M., 2017. The Revised Strain Index: an improved upper extremity exposure assessment model. Ergon. 60 (7), 912–922.

Howard, N., Adam, D., 2018. Work-Related Musculoskeletal Disorders of the Back, Upper Extremity, and Knee in Washington State, 2006-2015. Washington State Department of Labor and Industries. All Washington Industries. Technical Report Number 40-19-2018 (March 2018). SHARP Program.

Keyserling, W.M., Stetson, D.S., Silverstein, B.A., Brouwer, M.L., 1993. A checklist for evaluating ergonomic risk factors associated with upper extremity cumulative trauma disorders. Ergon. 36 (7), 807–831.

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 (2), 314–326.

Klein, D., 2017. Assessing Inter-rater Agreement in Stata, 15th German Stata Users Group meeting. Berlin, (June 23, 2017). Available from. https://www.stata.com/meeting/germany17/slides/Germany17_Klein.pdf.

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33 (1), 159–174.

Lau, M.H., Armstrong, T.J., 2011. The effect of viewing angle on wrist posture estimation from photographic images using novice raters. Appl. Ergon. 42 (5), 634–643.

Lee, W., Lin, J.H., Bao, S., Lin, K.Y., 2019. Reliability and validity of a posture matching method using inertial measurement unit-based motion tracking system for construction jobs. Computing in Civil Engineering 2019: Data, Sensing, and Analytics. American Society of Civil Engineers, Reston, VA, pp. 589–597.

Lowe, B.D., Weir, P.L., Andrews, D.M., US Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, 2014. Observation Based Posture Assessment: Review of Current Practice and Recommendations for Improvement. U.S. DHHS (NIOSH) Publication 2014-131, Cincinnati, OH.

Lowe, B.D., Dempsey, P.G., Jones, E.M., 2019. Ergonomics assessment methods used by ergonomics professionals. Appl. Ergon. 81, 102882.

Marcum, J., Adams, D., 2017. Work-related musculoskeletal disorder surveillance using the Washington state workers' compensation system: recent declines and patterns by industry, 1999-2013. Am. J. Ind. Med. 60 (5), 457–471.

McAtamney, L., Corlett, E.N., 1993. RULA: a survey method for the investigation of work-related upper limb disorders. Appl. Ergon. 24 (2), 91–99.

McHugh, M.L., 2012. Interrater reliability: the kappa statistic. Biochem. Med. 22 (3), 276–282.

Moore, S.J., Garg, A., 1995. The strain index: a proposed method to analyze jobs for risk of distal upper extremity disorders. Am. Ind. Hyg. Assoc. J. 56 (5), 443–458.

NexGen Ergonomics, 2016. HM-Analyzer™ v2.5 User Manual, HMZ-3-07-2016. NexGen Ergonomics.

Oliv, S., Gustafsson, E., Baloch, A.N., Hagberg, M., Sandén, H., 2019. The Quick Exposure Check (QEC)—inter-rater reliability in total score and individual items. Appl. Ergon. 76, 32–37.

Park, J.K., Boyer, J., Tessler, J., Casey, J., Schemm, L., Gore, R., et al., 2009. Inter-rater reliability of PATH observations for assessment of ergonomic risk factors in hospital work. Ergon. 52 (7), 820–829.

Paulsen, R., Gallu, T., Gilkey, D., Reiser II, R., Murgia, L., Rosecrance, J., 2015. The inter-rater reliability of Strain Index and OCRA Checklist task assessments in cheese processing. Appl. Ergon. 51, 199–204.
Poitras, I., Dupuis, F., Bielmann, M., Campeau-Lecours, A., Mercier, C., Bouyer, L.J., Roy, J.S., 2019. Validity and reliability of wearable sensors for joint angle estimation: a systematic review. Sensors 19 (7), 1555.
Robert-Lachaine, X., Mecheri, H., Larue, C., Plamondon, A., 2017. Effect of local magnetic field disturbances on inertial measurement units accuracy. Appl. Ergon. 63, 123–132.
Robert-Lachaine, X., Larue, C., Denis, D., Delisle, A., Mecheri, H., Corbeil, P., Plamondon, A., 2020. Feasibility of quantifying the physical exposure of materials handlers in the workplace with magnetic and inertial measurement units. Ergon. 63 (3), 283–292.
Roetenberg, D., Slycke, P.J., Veltink, P.H., 2007. Ambulatory position and orientation tracking fusing magnetic and inertial sensing. IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 54 (5), 883–890.
Schall, M.C., Fethke, N.B., Chen, H., Gerr, F., 2015. A comparison of instrumentation methods to estimate thoracolumbar motion in field-based occupational studies. Appl. Ergon. 48, 224–231.
Schall Jr., M.C., Fethke, N.B., Chen, H., Oyama, S., Douphrate, D.I., 2016. Accuracy and repeatability of an inertial measurement unit system for field-based occupational studies. Ergon. 59 (4), 591–602.
Smith, C.K., Anderson, N.J., 2017. Work-related injuries among commercial janitors in Washington State, comparisons by gender. J. Saf. Res. 62, 199–207.
Spielholz, P., Bao, S., Howard, N., Silverstein, B., Fan, J., Smith, C., Salazar, C., 2008. Reliability and validity assessment of the hand activity level threshold limit value and strain index using expert ratings of mono-task jobs. J. Occup. Environ. Hyg. 5 (4), 250–257.
Stevens, E.M., Vos, G.A., Stephens, J.P., Moore, J.S., 2004. Inter-rater reliability of the strain index. J. Occup. Environ. Hyg. 1 (11), 745–751.
Weir, P.L., Andrews, D.M., van Wyk, P.M., Callaghan, J.P., 2011. The influence of training on decision times and errors associated with classifying trunk postures using video-based posture assessment methods. Ergon. 54 (2), 197–205.
Wu, G., Siegler, S., Allard, P., Kirtley, C., Leardini, A., Rosenbaum, D., 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. J. Biomech. 35 (4), 543–548.
Wu, G., Van der Helm, F.C., Veeger, H.D., Makhsous, M., Van Roy, P., Anglin, C., 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. J. Biomech. 38 (5), 981–992.