Comparison Semiautomatic NIOSH Lifting Equation: AzKNIOSH versus RGB-based Machine Vision Algorithm

Chiara Forgione*, Antonio Maria Coruzzolo, Francesco Lolli**,***, Elia Balugani**,*****

** Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Via Giuseppe Campi 287, 41125 – Modena, Italy.*
***Department of Sciences and Methods for Engineering, University of Modena and Reggio Emilia, Via Amendola 2, 42122, Reggio Emilia, Italy.*
****Interdepartmental Centre En&Tech, Piazzale Europa, 1, 42124, Reggio Emilia, Italy.*

Abstract: Work related musculoskeletal disorders (WMDs) have a significant impact on industrial productivity and society. With the advent of Industry 5.0, the safety and well-being of human operators are back to being crucial for each modern production system. In this context, many innovative technologies have been developed for ergonomic purposes. Motion Capture (MOCAP) technologies are applied to semi automatically calculate the ergonomic risk in a faster and less expensive way. In the other hand, the usage of MOCAP is not always recommended and data collection with common devices is preferred in industrial environment. For this scope, we compared the effectiveness of a commercial machine vision algorithm (ErgoEdge) based on RGB camera against a developed application based on the depth camera Microsoft Azure Kinect (AzKNIOSH) for NIOSH Lifting Equation computation. Fifty-two tasks in which volunteers performed manual handling of loads were evaluated with both systems, showing a good agreement.

Keywords: Depth Camera, Machine Vision, Ergonomics, Kinect, Picking.

1.Introduction

Order picking is considered the highest-priority area for productivity improvements because it is the most labour-intensive operation in warehouse and it is too difficult to automate (de Koster et al., 2007). Manual material handling (MMH) is one of the most physically intensive activity, often characterized by high load weight, high repetitive and awkward body postures (Weisner & Deuse, 2014). The described conditions are risk factors for low back pain (Hoogendoorn et al., 2000) and, generally, for the musculoskeletal system. In order to manage work related musculoskeletal disorders (WMSDs), the National Institute for Occupational Safety and Health (NIOSH) has proposed the NIOSH Lifting Equation (NLE) (National Institute for Occupational Safety and health, 1981), subsequently updated by the Revised NIOSH Lifting Equation (RNLE) (Waters et al., 1993). The equation is widely used by occupation health practitioners, providing an empirical method to determine a weight limit for manual lifting. The RNLE calculates the recommended weight limit (RWL) that almost all healthy workers may handle without risk. This limit is useful to identify the risk of developing WMSDs. The Lifting Index (LI) is defined as the ratio of the lifted load to the RWL. For the RWL calculation different characteristics of the MMH activity are considered: weight of the load, horizontal and vertical location, vertical travel distance, asymmetry angle, lifting frequency and duration and coupling classification (Waters et al., 1993). The RNLE is classified as an observational method based on direct observation of the worker during the performance of the analysed activity (Diego-Mas & Alcaide-Marzal, 2014). Observational methods are straightforward to use, applicable to different working situations and industrial environments (Diego-Mas & Alcaide-Marzal, 2014). However, they are characterized by the high amount of time to perform the video analysis, the low accuracy and high-and-inter observer variability (Lolli et al., 2022). Due to recent advancements in technology, ergonomic risk assessment (ERA) can be supported by innovative systems that automatically detect body postures and movements (Lunin & Glock, 2021). Specifically, markerless Motion Capture (MOCAP) technologies enable the recording of human's movements and their digital representation. The most widely used MOCAP for ERA are camera-based systems due to their low cost and ease to use in industrial environments (Lunin & Glock, 2021). One example is Microsoft Kinect that is equipped with an RGB camera and a depth sensor. Provided data can be processed by Kinect Software Development Kit (SDK) to obtain digitalized human motions (Microsoft, 2021). Kinect's output is easily use to generate postural data for ERA and it is perfectly suitable for the application of ergonomic methods in a semi-automatic way (Lunin & Glock, 2021). Azure Kinect is the last version of Kinect and it outperforms other versions (Pilati et al., 2023). However, to the best of our knowledge, there is just one study that compares Azure Kinect with an RGB cameras for semi-automatic ERA. Coruzzollo et al. (2022) proposed the comparison of Azure Kinect-based Rapid Upper Limb Assessment (RULA) calculation with a Machine Vision (MV) software based on a RGB camera. Since RGB cameras are increasingly used and Azure Kinect is one of the most used technology for ERA, we propose the comparison of RNLE calculated by an Azure Kinect-based tool and a MV software based on RGB camera.

The paper is organized as follows: Section 2 describes state-of-the-art literature in the field of automatic ERA; Section 3 details the procedure for conducting assessment with AzKNIOSH and ErgoEdge while Section 4 outlines experiment setup and results; Section 5 discusses the conclusions; Section 6 presents practical implications.

2.State-of-the-art

In the field of ergonomics, there are an increasing number of studies that integrate MOCAP systems for ERA (Abobakr et al., 2019; Altieri et al., 2020; Manghisi et al., 2017; Yan, Li et al., 2017). Specifically, optical systems are less intrusive and they are a viable alternative to traditional methods in industrial field (Manghisi et al., 2017). The most commonly used cameras in the field of ergonomic is Microsoft Kinect (Lunin & Glock, 2021): Diego-Mas et al. (2014) calculated Ovako Posture Analysis System (OWAS) using Kinect v1 and compared results with an expert evaluation; Manghisi et al. (2017) developed a tool for the semi-automatic calculation of the RULA score with a Kinect v2 and compared the results with an ergonomist, an optical MOCAP and a commercial software based on Kinect v1; Coruzzolo et al. (2022) calculated RULA score with an Azure Kinect-based tool and compared it with a commercial MV algorithm, named ErgoEdge. To the best of our knowledge, there are few studies on ERA using the new Azure Kinect (Coruzzolo et al., 2022; Lolli et al., 2022). Tolgyessy (Tölgyessy et al., 2021) proved that Azure Kinect outperforms Kinect v1 and Kinect v2 in terms of repeatability in the number of joints tracked and body segmentation. However, Azure Kinect has some limitations: occlusions and the limited working range of maximum 5.46 m (Tölgyessy et al., 2021). On the other hand, a normal camera has a greater vision field depth. RBG cameras are widely used in combination with the MV algorithm to extract the body segmentation and analyse posture (Li et al., 2020). Yan et al. (2017) developed an RGB cameras-based system to calculate OWAS. Ding et al. (2019) implemented a vision-based method for assessing the upper body posture of a work sitting in front of a desk, directly classifying them into pre-defined classes based on the scoring method. Altieri et al. (2020) used a network of RGB cameras that exploit "OpenPose" to calculate the joint angles and automatically calculate Occupational Repetitive Actions Index (OCRA) index. As mentioned, RGB camera for ERA are increasingly used but, to the best of our knowledge, no studies have performed a comparison between the RNLE derived by a RGB camera-based MV algorithm and Azure Kinect. In our work we propose the comparison of RNLE between Azure Kinect-based tool (AzKNIOSH) and an existing MV algorithm (ErgoEdge). We also analyse cases with occlusions and frontal-views here the capability of ErgoEdge could have some problems to calculate RNLE factors.

3.Methods

3.1 AzKNIOSH

AzKNIOSH is an Azure Kinect-based tool for the semi-automatic risk assessment during manual material handling of loads. Using Azure Kinect SDK 32 body joints have been tracked. Using 3D coordinates of body joints, useful angles and distances can be measured through a geometrical informatic model that was built on Python. Analysed activity must be recorded with the depth camera and after post-processing data can be used for semi-automatic RNLE. At the beginning of the assessment, it is necessary to indicate the beginning frame of the picking and the ending one. AzKNIOSH is based on RNLE (Waters et al., 1993) and its output is the LI for origin and destination of analysed activity. For both origin and destination of the picking RWL is calculated with the Equation (1):

(1)

$$RWL = LC * HM * VM * DM * AM * FM * CM$$

Calculation of each component is described in Table 1.

Table 1: Calculation of RNLE components

Component	Metric	Informatic model
LC = load constant	23 kg	Fixed to 23 kg
HM = horizontal multiplier	25/H	H = horizontal distance between the hands midpoint and the ankles midpoint
VM = vertical multiplier	1-(0.003*|V -75|)	V = vertical distance between the hands midpoint and the floor
DM = distance multiplier	0.82+(4.5/D)	D = difference between vertical coordinate of hands midpoint at the origin and at the destination of picking
AM = asymmetry multiplier	1-(0.0032*A)	A = angle between the intersection of the sagittal plane and the floor plane and the vector joining the hands midpoint and ankles midpoint
FM = frequency multiplier	From table by Water et al. (1993)	F = lift per minute to be manually entered; D(h) = duration in hours to be manually entered
CM = coupling multiplier	From table by Waters et al. (1993)	Coupling judgement to be manually entered

From the RWL calculation it is possible to calculate the LI as the ratio of the load lifted to RWL (Waters et al., 1993).

3.2 ErgoEdge

ErgoEdge is a commercial solution that allows to process acquisitions from smartphones through the deep learning models to assess worksite. Specifically, ErgoEdge provides the calculation of RULA, NIOSH, REBA, OWAS. It takes in input an RGB video then the joints are detected and tracked in 2D. From 2D to 3D joints are inferred and used to calculate RNLE with some manual inputs. At the beginning of the assessment, it is necessary to indicate the beginning frame of the picking and the ending one. Other manual inputs that have been manually insert are:

- Approximate height of the human in cm.
- Unit: metric or US Customary.
- Sig control: whether it requires "precision placement" of the load at destination.
- Load weight in kg.
- Duration in hours.
- Coupling judgment.
- Frequency.

There is a difference between the manual inputs for AzKNIOSH and ErgoEdge. Both require manual input for frames, lifted load, frequency, duration and coupling judgment. Conversely, while AzKNIOSH is able to automatically detect floor plane and vertical distances, ErgoEdge calculates vertical distances based on human height. Horizontal distance is automatically calculated only when both feet are visible, checking the probability score. If they are visible, mean point between ankles is calculate. For hands, if both wrists are visible, mean point is calculated, otherwise, only the wrist with a higher probability is considered. Other difference between the two investigated systems is related to the number of detected joint: ErgoEdge tracks fewer points on the head and fails to track fingers due to the distance from the camera, occlusion, and hand gloves used by workers. In Figure 1 a comparison of both joint hierarchy is shown.

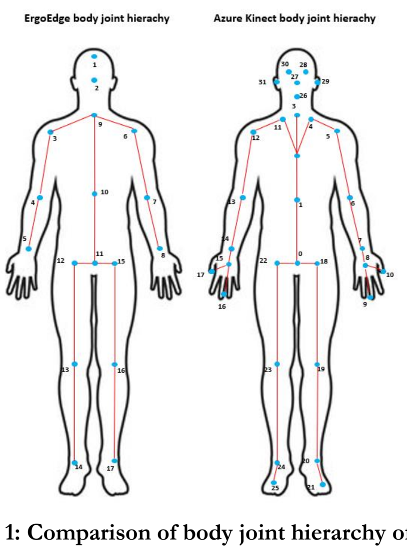

Figure 1: Comparison of body joint hierarchy of ErgoEdge and Azure Kinect.

4.Experiment

4.1 Equipment

For the video recording we used an Azure Kinect with the following settings: Colour mode on 720p, Depth mode On FOV 2x2 binned, no depth delays, 15 frames per second (fps), IMU on, External sync stand alone, Sync delay 0, Auto exposure, Auto gain. The Kinect was placed at a height of 100 cm and at a distance of 220-250 cm from the reordered subject. The PC connected to the Azure Kinect used to run AzKNIOSH has a CPU Intel® CoreTM i9-10900K 3.7 GHz, 32 GB RAM, GPU NVIDIA GeForce RTX 2070 Super, OS Windows 10. To run ErgoEdge we used the RGB images from Azure Kinect. To perform picking activities, laboratory was equipped with the following items:

- 1 shoebox weighing 1 kg;
- Industrial table of 90 cm high;
- 5 shelves at different height: 11 cm, 47 cm, 83 cm, 119 cm, 155 cm.

4.2 Procedure

In the experiment we evaluated the RNLE using AzKNIOSH and ErgoEdge on 52 picking activities performed by 2 different subjects. The first subject is a 22-year-old male, 170 cm tall. The second one is a 28-years-old male, 190 cm tall. Coupling judgement is set as "fair" because the box did not have handles but it was 32x22x13 (h) cm, so easy to handle. The analysed activities consisted of moving the box from one surface to another. The surfaces were at different heights. Frequency was set at 3 lifts/min with a duration of 2 hours. Only one picking action is made in each acquisition and the mean duration of the video is of 6 seconds. Table 2 summarizes the main objective and parameters of the experiments.

Table 2: Objective and parameters of the experiment.

Experiment goal	Comparison of NIOSH evaluation calculated by Azure Kinect-based tool and a MV software based on RGB camera.
Motion Capture technology used	Microsoft Azure Kinect
Compared software	AzKNIOSH, ErgoEdge
Output values analysed for benchmarking	LI, KM, HM, VM, AM, DM
Participants	2
Number of picking activities performed by each subject	52
Total acquisitions	104
Shoebox characteristics	32x22x13 (h) cm, 1 kg weight, no handles
Coupling judgement	Fair
Picking frequency	3 lifts/min
Duration of the task	2 hours

Figure 2 shows some examples of the analysed acquisitions.

Figure 2: Examples of origin and destination frames of analysed picking activities.

4.3 Results

Fifty-two acquisitions were analysed benchmarking LI obtained from AzKNIOSH and ErgoEdge. The Mean Bias Error (MBE) resulting from the difference of AzKNIOSH and ErgoEdge is of 0.0567. From data analysis we observed that, using ErgoEdge, in frontal-view acquisitions the joint detection was not very accurate and in some cases the body was not detected, as shown in Figure 3.

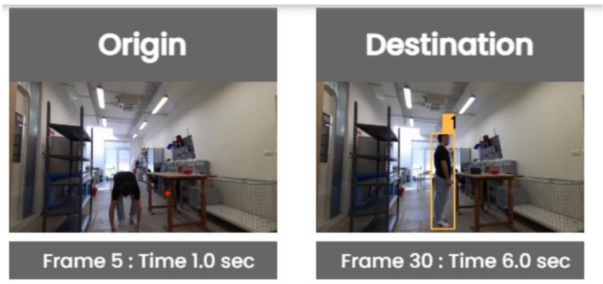

Figure 3: Examples of frontal-view in the origin of the lift. Image from ErgoEdge body detection.

For this reason, frontal-view videos were excluded (twelve videos) from the analysis, including forty evaluations. Without frontal-views, we obtained an MBE of 0.0515, with a boxplot and Bland-Altman Plot in Figure 4 and Figure 5.

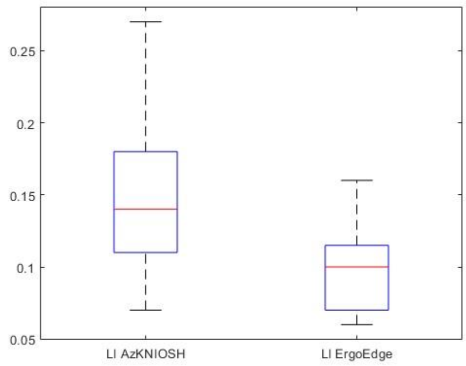

Figure 4: Boxplot representing LI calculation by AzKNIOSH and ErgoEdge.

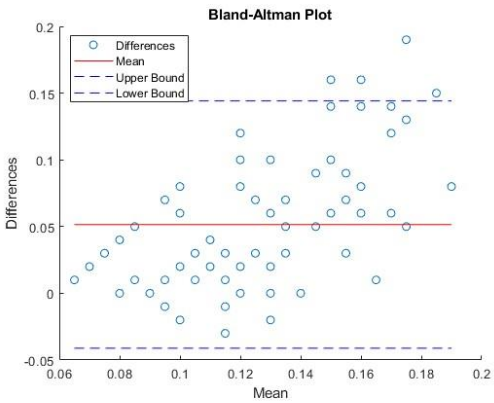

Figure 5: Bland-Altman Plot between LI from AzKNIOSH and ErgoEdge.

To investigate how measurements change between the two evaluation methods, Kinematics Multiplier was calculated (Patrizi et al., 2016). KM summarizes all the postural aspects in a single variable:

(2)

$$KM = HM * VM * DM * AM$$

HM, VM, AM, and DM require the evaluation of distances and angles and they are directly involved in computation of the postural kinematics. These measurements are computed using vectors built from the joint locations using geometric functions in AzKNIOSH and ErgoEdge. The MBE is of -0.168, showing a slight underestimation by AzKNIOSH, as shown in Figure 6.

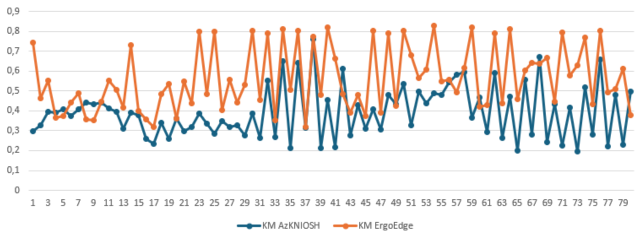

Figure 6: Benchmarking of KM calculated by AzKNIOSH and ErgoEdge.

Due to the difference between KM calculations, a detailed analysis of each multiplier was done. MBE and Mean Absolute Error (MAE) for each analyzed multiplier is reported in Table 3.

Table 3: Benchmarking results.

Component	MBE	MAE
LI	0.0515	0.0543
KM	-0.1679	0.1841
HM	-0.1551	0.1986
VM	-0.0189	0.0606
DM	-0.0245	0.0250
AM	-0.0565	0.0565
KM/AM	-0.1447	0.1703

The highest MBE corresponds to HM calculation, due to different method of horizontal distance measurements, highlighted in 3.2. The analysed difference leads to an overestimation of HM by ErgoEdge with the 19% of total HM equal to 1 (optimal conditions). While, for AzKNIOSH the horizontal factor is a critical issue because for the 50% of the analysed tasks it is less than 0.5, corresponding to 50 cm. Other significant factor, highlighted during a detailed analysis, is the AM detection. The 98.75% of total AM calculated by ErgoEdge are equal to 1 (optimal conditions), shown in Figure 7.

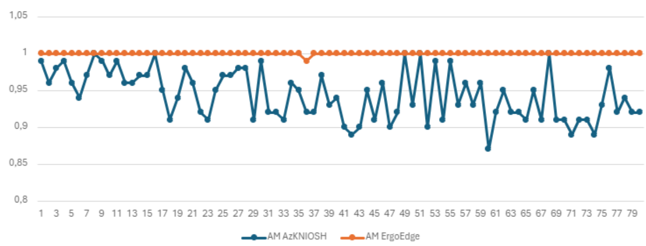

Figure 7: Benchmarking of AM calculated by AzKNIOSH and ErgoEdge.

The twisting angle calculated by ErgoEdge is not accurate, so users can manually input the approximated value for this field. Due to the inaccurate AM calculation, KM/AM was analysed resulting an MBE of -0.145 and a MAE of 0.170.

5.Conclusions and discussions

To the best of our knowledge, this study is the first benchmarking of RNLE automatic calculation between Azure Kinect and a MV algorithm. 52 different acquisitions were analysed, in which two volunteers perform lifting actions. LI is calculated for both origin and destination of the lift, so there are 104 different results to comparing. AzKNIOSH and ErgoEdge are two different methods to automatically calculate RNLE. AzKNIOSH is based on Azure Kinect, while ErgoEdge is a MV software based on RGB camera. For data analysis, 12 frontal-view acquisitions have been excluded due to inaccurate body detection for ErgoEdge. Difference between LI from the two methods is very low, mean of 0.0515. However, analysing KM a substantial difference was found. So, we analysed more in details each multiplier, founding a significant error for HM and AM. AzKNIOSH calculates HM considering the distance between the midpoint of the ankles and the midpoint of the hands. While ErgoEdge evaluates the probability of joint detection and takes in input the midpoint of the wrist or the only wrist with higher probability. This entails a MAE for HM of 0.1986. For AM calculation, AzKNIOSH provides plausible measurements while ErgoEdge assigns a value of 1 to the 98.75% of the total acquisitions. This evaluation is not so accurate, because not in each acquisitions the asymmetry dislocation corresponds to AM equal to 1 (optimal conditions). To overcome this problem, users have to manually insert the approximated value of asymmetry multiplier. AzKNIOSH is more accurate then ErgoEdge also in semi-occluded postures but ErgoEdge can use RGB video in input so is easily usable in different contest.

5.1 Limitations and future research

This study has few limitations. First, using optical sensors there is the common problem related to occlusion for both software. However, ErgoEdge is particularly susceptible: frontal-view have been excluded from the research and the comparison with AzKNIOSH results was not possible. Second, the sample size is reduced because only 2 subjects were evaluated and only one shoebox was used. This might compromise the generalisability of the experiment to a real context. To bridge the highlighted gaps, our next steps involve the usage of multiple optical sensors to avoid occlusion problems and an extended dataset with more subjects and boxes with heterogenous characteristics.

6. Practical implications

ErgoEdge can perform ergonomic risk assessment from RGB video, recorded with a common device. On the other hand, to record video with Azure Kinect it is necessary to design a small area dedicated to the hardware equipment, and it could be difficult in industrial environment. RGB data can be acquired from common device so there is no design phase for the positioning of equipment. ErgoEdge could be used for a preliminary phase, maybe to identify critical tasks in an industrial environment through a fast ergonomic-assessment. While, AzKNIOSH could be use in a more detailed assessment on critical postures previously identified with ErgoEdge. Specifically, the detection of horizontal distance is the most critical aspects for ergonomic risk: often the load is too far from the body and the posture is dangerous. For a detailed analysis and a re-design of the tasks, using ErgoEdge could lead to a wrong detection of this critical measurement.

References

Abobakr, A., Nahavandi, D., Hossny, M., Iskander, J., Attia, M., Nahavandi, S., & Smets, M. (2019). RGB-D ergonomic assessment system of adopted working postures. *Applied Ergonomics*, *80*, 75–88. https://doi.org/10.1016/j.apergo.2019.05.004

Altieri, A., Ceccacci, S., Talipu, A., & Mengoni, M. (2020). A low cost motion analysis system based on RGB cameras to support ergonomic risk assessment in real workplaces. *Proceedings of the ASME Design Engineering Technical Conference*, *9*. https://doi.org/10.1115/DETC2020-22308

Coruzzolo, A. M., Lolli, F., Amicosante, N., Kumar, H., Thupaki, P., & Agarwal, S. (2022). *Comparing semiautomatic Rapid Upper Limb Assessments (RULA): Azure Kinect versus RGB-based machine vision algorithm*. https://doi.org/10.54941/ahfe1002596

de Koster, R., Le-Duc, T., & Roodbergen, K. J. (2007). Design and control of warehouse order picking: A literature review. *European Journal of Operational Research*, *182*(2), 481–501. https://doi.org/10.1016/j.ejor.2006.07.009

Diego-Mas, J. A., & Alcaide-Marzal, J. (2014). Using Kinect™ sensor in observational methods for assessing postures at work. *Applied Ergonomics*, *45*(4), 976–985. https://doi.org/10.1016/j.apergo.2013.12.001

Ding, Z., Li, W., Ogunbona, P., & Qin, L. (2019). A real-time webcam-based method for assessing upper-body postures. *Machine Vision and Applications*. https://doi.org/10.1007/s00138-019-01033-9

Hoogendoorn, W. E., Bongers, P. M., De Vet, H. C. W., Douwes, M., Koes, B. W., Miedema, M. C., Ariëns, G. A. M., & Bouter, L. M. (2000). Flexion and Rotation of the Trunk and Lifting at Work Are Risk Factors for Low Back Pain Results of a Prospective Cohort Study. In *SPINE* (Vol. 25, Issue 23).

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*, *87*. https://doi.org/10.1016/j.apergo.2020.103138

Lolli, F., Coruzzolo, A. M., Forgione, C., Gonçalves, P., & Neto, T. (2022). Ergonomic risk reduction: a height-adjustable mesh truck for picking activities evaluated with a depth camera. *AHFE*, *2023*, 2023. https://doi.org/10.54941/ahfeXXXX

Lunin, A., & Glock, C. H. (2021). Systematic review of Kinect-based solutions for physical risk assessment in manual materials handling in industrial and laboratory environments. *Computers & Industrial Engineering*, *162*, 107660. https://doi.org/10.1016/j.cie.2021.107660

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*, *65*, 481–491. https://doi.org/10.1016/j.apergo.2017.02.015

Microsoft. (2021). *About Azure Kinect DK*. https://learn.microsoft.com/en-gb/azure/kinect-dk/about-azure-kinect-dk

National Institute for Occupational Safety and health. (1981). *Work practices guide for manual lifting*.

Patrizi, A., Pennestrì, E., & Valentini, P. P. (2016). Comparison between low-cost marker-less and high-end marker-based motion capture systems for the computer-aided assessment of working ergonomics. *Ergonomics*, *59*(1), 155–162. https://doi.org/10.1080/00140139.2015.1057238

Pilati, F., Sbaragli, A., Tomelleri, F., Picariello, E., Picariello, F., Tudosa, I., & Nardello, M. (2023). Operator 5.0: Enhancing the Physical Resilience of Workers in Assembly Lines. *2023 IEEE International Workshop on Metrology for Industry 4.0 and IoT, MetroInd4.0 and IoT 2023 - Proceedings*, 177–182. https://doi.org/10.1109/MetroInd4.0IoT57462.2023.10180145

Tölgyessy, M., Dekan, M., Chovanec, Ľ., & Hubinský, P. (2021). Evaluation of the azure kinect and its comparison to kinect v1 and kinect v2. *Sensors (Switzerland)*, *21*(2), 1–25. https://doi.org/10.3390/s21020413

Waters, T. R., Putz-Anderson, V., Garg, A., & Fine, L. J. (1993). Revised NIOSH equation for the design and evaluation of manual lifting tasks. *Ergonomics*, *36*(7), 749–776. https://doi.org/10.1080/00140139308967940

Weisner, K., & Deuse, J. (2014). Assessment Methodology to Design an Ergonomic and Sustainable Order Picking System Using Motion Capturing Systems. *Procedia CIRP*, *17*, 422–427. https://doi.org/10.1016/j.procir.2014.01.046

Yan, X., Li, H., Li, A. R., & Zhang, H. (2017). Wearable IMU-based real-time motion warning system for construction workers' musculoskeletal disorders prevention. *Automation in Construction*, *74*, 2–11. https://doi.org/10.1016/j.autcon.2016.11.007

Yan, X., Li, H., Wang, C., Seo, J., Zhang, H., & Wang, H. (2017). Development of ergonomic posture recognition technique based on 2D ordinary camera for construction hazard prevention through view-invariant features in 2D skeleton motion. *Advanced Engineering Informatics*, *34*, 152–163. https://doi.org/10.1016/j.aei.2017.11.001