Physical Ergonomics and Human Factors, Vol. 63, 2022, 54–61
https://doi.org/10.54941/ahfe1002596

Comparing Semiautomatic Rapid Upper Limb Assessments (RULA): Azure Kinect Versus RGB-Based Machine Vision Algorithm

Antonio Maria Coruzzolo[1], Francesco Loll[1,2], Nazareno Amicosante[1], Hrishikesh Kuma[3], Pramod Thupaki[3], and Saurav Agarwal[3]

[1]Department of Sciences and Methods for Engineering, University of Modena and Reggio Emilia, Reggio Emilia, RE 42122, Italy
[2]En&Tech Interdipartimental Center of the University of Modena and Reggio Emilia, Reggio Emilia, RE 42122, Italy
[3]CerebrumEdge Technologies, Bangalore, India

ABSTRACT

Correctly using a rapid upper limb assessment is crucial to avoid musculoskeletal disorders. Although motion capture technologies are widely used, they cannot be used in large-scale industrial environments due to their high cost and performance impacted by the surrounding environment. We thus compared the effectiveness of a commercial machine vision (MV) algorithm (ErgoEdge) based on an RGB camera against a developed application based on the depth camera Microsoft Azure Kinect (AzKRULA) for RULA evaluation. Fifteen static postures were evaluated with both systems and compared with those of an ergonomics expert showing a substantial agreement between the three evaluations. At the same time, it showed that RGB camera must be placed on the side of the worker due to the difficulties of the MV algorithm in reconstructing from a frontal view, important joint angles (e.g., to evaluate the neck), which can invalidate the RULA evaluation provided by ErgoEdge.

Keywords: Exemplary paper, Human systems integration, Systems engineering

INTRODUCTION

The fourth industrial revolution has changed organizations and processes as well as human work (Neumann and Village, 2012). Although there have been continuous improvements in working conditions with the automation of manual activities, several tasks are still difficult to automate (Bortolini *et al.*, 2017). In addition, more than a half of the tasks performed are short and repetitive (Bortolini *et al.*, 2018), conditions leading to work related musculoskeletal disorders (WMSDs). Repetitive hand and arm movements are the most prevalent in Europe being reported by 61% of workers and play a primary role in causing WMSDs (Eurofound, 2017). Consequently, it is crucial to take into account the health of the workers by applying policies that minimize the risk of WMSDs (Silverstein and Clark, 2004).

Firstly, exposure to risk factors needs to be assessed and any subsequent ergonomic interventions need to be planned (Manghisi *et al.*, 2017) (e.g., workplace redesign). There are three methods for the evaluation: self-reports, direct measurements and observational methods (Li and Buckle, 1999). Self-reports are negatively affected by subjectivity (David, 2005), while direct methods require sensors attached to the worker's body that are usually expensive, highly intrusive and not precise (Kowalski *et al.*, 2012). Among the observational methods, the most commonly used are: RULA (Lynn and Corlett, 1993), Rapid Entire Body Assessment (REBA) (Hignett and McAtamney, 2000), NIOSH lifting equation (Snook and Ciriello, 1991), Strain Index (Moore and Garg, 1995), OVAKO Working posture Analysing System (OWAS) (Karhu, Kansi and Kuorinka, 1977) and the concise exposure index (OCRA) (Occhipinti, 1998). All these methods require an ergonomic practitioner who has to perform a posture evaluation with a time-consuming video analysis that leads to low accuracy and high intra-and-inter-observer variability. To lower these weaknesses many studies now integrate different motion capture (MOCAP) technologies in the ergonomic assessment (Manghisi *et al.*, 2017) (Abobakr *et al.*, 2019) (Yan, Li, Li, *et al.*, 2017) (Altieri *et al.*, 2020). MOCAP systems can be classified into sensor-based and optical. Although some studies (Battini, Persona and Sgarbossa, 2014) (Hsu and Lin, 2019) have used Inertial Measuraments Units (IMUs) as direct methods in real working environments, they are still challenging for the discomfort caused to the worker that is in contrast with a human-centred design. While optical systems are less intrusive and provide a valid solution for an ergonomic risk assessment (ERA) (Manghisi *et al.*, 2017). The most commonly cameras for these applications are Kinects: a ready-to-use technology that provides a real-time segmentation using depth-RGB data (Microsoft, 2021). The third and current Kinect generation is called Azure Kinect, which was preceded by Kinect v1 and v2 (Tölgyessy *et al.*, 2021). Many studies exploited Kinects for ERA. Diego-Mas et al. (Diego-Mas and Alcaide-Marzal, 2014) calculated OWAS with Kinect v1, Manghisi et al. (Manghisi *et al.*, 2017) semi-automatically calculates RULA with a Kinect v2, Plantard et al. (Plantard *et al.*, 2017) calculated RULA from Kinect v2, Bortolini et al. (Bortolini *et al.*, 2018) utilised four Kinect v2, creating a motion capture system (MAS) overcame one of the main drawbacks of depth cameras, i.e., the inaccuracy due to occluded postures (Plantard *et al.*, 2015). The most two commonly used MOCAP systems for ERA are depth cameras and IMUs (Wilhelm *et al.*, 2020). However, to the best of our knowledge and as confirmed in (Lunin and Glock, 2021), there have been no studies on ERA using the new Azure Kinect even if it has been proved to outperforms Kinect v1 and v2 in terms of precision, in the number of joints tracked and in terms of the body segmentation accuracy (Tölgyessy *et al.*, 2021). This is the first gap we want to fulfil with the development of an application based on Azure Kinect for semi-automatically calculate RULA (AzKRULA). Azure Kinect nevertheless still has some of the limitations of depth cameras: occlusion and the limited working range which is 3.86 m for Azure Kinect in NFOV unbinned, i.e., the set-up here used. A normal camera has a greater vision field depth and can be used to overcame these limitations combining it with a MV algorithm to extract the body segmentation (Li, Martin and Xu, 2020). Reconstructing postures from a 2D image is challenging (Toshev and Szegedy, 2014) but new powerful GPUs enable the use of convolutional neural networks (CNNs) to speed-up pose reconstruction (Newell, Yang and Deng, 2016) . Yan et al. (Yan, Li, Wang, *et al.*, 2017) evaluated OWAS based on RGB cameras testing several types of machine learning models. Ding et al. (Ding *et al.*, 2019) presented a vision-based method for assessing the upper body postures of a sitting worker. Altieri et al. (Altieri *et al.*, 2020) developed a MAS using a network of RGB-cameras that exploits the key-point detection library "OpenPose" to calculates OCRA. Li, Martin and Xu (Li, Martin and Xu, 2020) exploited a RGB camera for assessing RULA by training a CNN on a sample of labelled postures from (Xu *et al.*, 2011) and (Ionescu *et al.*, 2014). Postural risk assessments using normal RGB cameras are increasingly used but have not already been compared with depth cameras. Thus, as second contribution we compared our new Azure Kinect-based RULA calculation (AzKRULA) with a MV software based on a normal RGB camera (ErgoEdge) on 15 static postures well known in literature (Manghisi *et al.*, 2017). Section 2 describes the two methods for semi-automatically calculating RULA; Section 3 reports the experimental results; and, finally, Section 4 contains the conclusions and some suggestions for the further research.

METHODS

The RULA score comes from an assessment grid which is filled in using joint angles. For each joint in each section, Section A (upper arm, lower arm, and wrist) and Section B (neck, trunk, and legs), a series of principal factors, based on the joint angles, along with secondary factors here defined as factors in the RULA calculation for which there are not prescribed threshold (e.g., abduction or trunk twist) are evaluated and integrated in a score (Lynn and Corlett, 1993). The final RULA score ranges from 1 to 7 and is classified into four levels of actions. In both the methods, RULA is semi-automatically calculated since the muscle use and force are assumed to be fixed and are assigned manually as well as some minor factors that will be detailed.

AzKRULA

We implemented AzRULA on Python 3.8 with an application that calculates the RULA score for each frame passed to it exploiting the Azure Kinect Body Tracking SDK (Microsoft, 2021a), which is an algorithm based on CNNs that retrieved 32 joints in the 3D space as reported in Figure 1.

Not all the retrieved joints were needed to calculate the angles for the RULA tables. However, additional processing was required on retrieved joints. We followed the procedure developed in (Manghisi *et al.*, 2017) modifying it accordingly with the new joints provided by Azure Kinect (with respect to Kinect v2). As they done for the *leg score*, we carried out a manual evaluation like we had done for *muscle* and *force factors* and as we also did for *wrist rotation*. However, with respect to (Manghisi *et al.*, 2017), where a manual evaluation of the neck twist was made, we calculate it automatically thanks to the new joints tracked by the Azure Kinect.

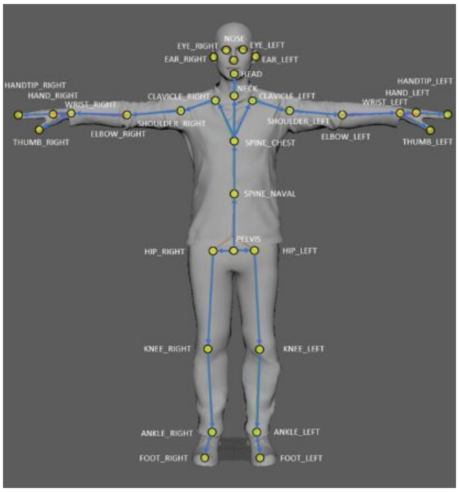

Figure 1: Azure kinect joints detected, and nomenclature.

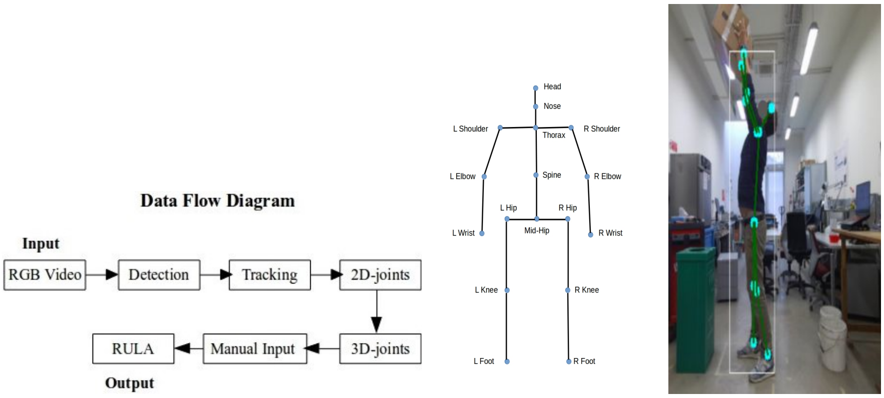

Figure 2: ErgoEdge joint detected; nomenclature and example of a segmentation.

ErgoEdge

ErgoEdge is a commercial solution designed for ease-of-use and comprehensive assessment at worksites. A normal smartphone video is processed through the deep learning models to estimate RULA. A demo of the software can be found at (CerebrumEdge, 2021). Figure 2 shows the data flow diagram, the joints detected by ErgoEdge and their nomenclature as well as an example of a segmentation performed on a well-known static posture. From the RGB video row, the joint is detected and tracked than from the 2D joints the 3D ones are inferred and used to calculate the final RULA score along with some manual input. ErgoEdge automatically detects the leg score, which is a manual input for AzKRULA. Conversely, while ErgoEdge requires the wrist flexion and twist to be inputted manually, AzKRULA automatically calculates them thanks to the additional points on the hands provided by Azure Kinect.

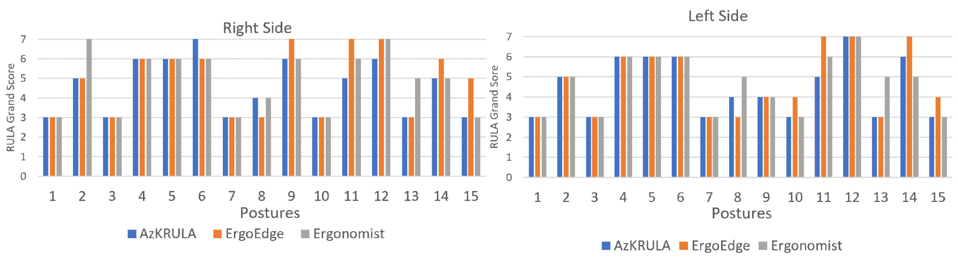

Figure 3: RULA score derived with AzKRULA, ErgoEdge and the expert.

EXPERIMENT

For the video recording we used an Azure Kinect with the following settings: Colour mode➔ On 720p, Depth mode➔ On NFOV 2x2 binned, No depth delays, Frames per second (fps)➔15, IMU➔ON, External sync➔Standalone, Sync delay➔0, Exposure➔ Auto, Gain➔Auto. The Kinect was always placed at a height of 110 cm and at 220-250 cm from the subject. We evaluated RULA using AzKRULA and ErgoEdge on 15 different static postures, the same as those investigated in (Manghisi *et al.*, 2017) numbered in the same way, that were also evaluated by an ergonomist expert with more than 10 years of experience. The results of the experiment are shown in Figure 3. The file with the videos (in mkv format) from which the frames were extracted and the exact frame used for each posture can be found here: https://github.com/amcorGit/AzKRULA-experimental-files.

The hypothesis we tested in this experiment was that with regard to the 15 static postures, AzKRULA and ErgoEdge would agree in the RULA calculation and that both would also agree with the expert. The agreement between AzKRULA and the ergonomic expert calculated in terms of linear weighted Cohen's Kappa (Fleiss, Cohen and Everitt, 1969) using the Landis and Koch Scale (Landis and Koch, 1977) result in 0.78 for the right side and 0.71 for the left one indicating a substantial agreement. This result is confirmed by the statistically significant p-values (<0.001) and by the high proportional agreement (0.95 and 0.93). AzKRULA has a mean error of -0.26 respect to the ergonomist and this slightly underestimation is mainly caused by secondary factors for which minor changes can lead to high differences in RULA. Secondary factors for which are not available standardise thresholds (e.g. how much the trunk have to be twisted for assessing a trunk twist?). Similarly, ErgoEdge and the expert have a substantial agreement in both the side of the body with Cohen's Kappa that result in 0.63 for the left side and 0.61 for the right with a p-value less than 1% less than 5% respectively. ErgoEdge showed a mean error near to zero while the major differences were given by secondary factors (e.g. a neck flection not individuated by ErgoEdge in posture 8). Likewise, AzKRULA and ErgoEdge have a substantial agreement in both side of the body. It is worth focusing on the results for the frontal postures. In fact, ErgoEdge did not agree with the expert in all six evaluations of frontal postures. These differences are due to the challenge in calculating angles in a 3D space from a 2D image. The use of ErgoEdge at this development phase is thus recommended with a side-view.

CONCLUSION

We have presented a new semi-automatic evaluation of RULA, AzKRULA, based on a the new delpth camera Azure Kinect. We compared our new application with a commercial one, ErgoEdge, which exploits simple RGB images with a MV algorithm. The following conclusions can be drawn together with some ideas for future research:

- The slight underestimation provided by AzKRULA is due to secondary factors. Thus, thresholds on secondary factors need to be standardized (e.g., how many twisting degrees indicate a trunk twist?) to create a standard reference for practitioners in similar situation.
- The MV algorithm based on simple RGB data (i.e., ErgoEdge) can be exploited for RULA evaluation as both the expert and AzKRULA were in substantial agreement with it. However, ErgoEdge has limitations in calculating the joint angles from frontal postures suggesting that when using ErgoEdge side views should be used of the principal parts of the body.
- Extended tests are needed in real working environments for both AzKRULA and ErgoEdge as well as for other methods available in literature (e.g. (Manghisi *et al.*, 2017)), for this reason applications by practitioners of the two applications would be beneficial.

Finally, we encourage academics and practitioners to exploit the supplementary materials within this paper, to test different software and compare results. This to construct a large database of postures and relative ergonomic assessments aimed at training, validating, and testing further solutions.

REFERENCES

Abobakr, A. *et al.* (2019) 'RGB-D ergonomic assessment system of adopted working postures', *Applied Ergonomics*, 80(April 2018), pp. 75–88. doi: 10.1016/j.apergo.2019.05.004.

Altieri, A. *et al.* (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(August). doi: 10.1115/DETC2020-22308.

Battini, D., Persona, A. and Sgarbossa, F. (2014) 'Innovative real-time system to integrate ergonomic evaluations into warehouse design and management', *Computers and Industrial Engineering*, 77, pp. 1–10. doi: 10.1016/j.cie.2014.08.018.

Bortolini, M. *et al.* (2017) 'system the Industry a framework Assembly system the Industry a framework Assembly system the Industry a', *IFAC-PapersOnLine*, 50(1), pp. 5700–5705. Available at: https://doi.org/10.1016/j.ifacol.2017.08.1121.

Bortolini, M. *et al.* (2018) 'Automatic assessment of the ergonomic risk for manual manufacturing and assembly activities through optical motion capture technology', *Procedia CIRP*, 72, pp. 81–86. doi: 10.1016/j.procir.2018.03.198.

CerebrumEdge (2021) *ErgoEdge - Advanced and Comprehensive Ergonomics Assessment Models - CerebrumEdge - AI and Computer Vision Solutions for Occupational Health, Safety, and Digital Ergonomics | CerebrumEdge*. Available at: https://cerebrumedge.com/pages/ergoedge (Accessed: 28 September 2021).

David, G. C. (2005) 'Ergonomic methods for assessing exposure to risk factors for work-related musculoskeletal disorders', *Occupational Medicine*, 55(3), pp. 190–199. doi: 10.1093/occmed/kqi082.

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

Ding, Z. *et al.* (2019) 'A real-time webcam-based method for assessing upper-body postures', *Machine Vision and Applications*. doi: 10.1007/S00138-019-01033-9.

Eurofound (2017) *6th European Working Conditions Survey: 2017 update*, *European Union*.

Fleiss, J. L., Cohen, J. and Everitt, B. S. (1969) 'Large sample standard errors of kappa and weighted kappa', *Psychological Bulletin*, 72(5), pp. 323–327. doi: 10.1037/h0028106.

Hignett, S. and McAtamney, L. (2000) 'Rapid Entire Body Assessment (REBA)', *Applied Ergonomics*, 31(2), pp. 201–205. doi: 10.1016/S0003-6870(99)00039-3.

Hsu, C. F. and Lin, T. Te (2019) 'Development of an ergonomic evaluation system based on inertial measurement unit and its application for exoskeleton load reduction', in *2019 ASABE Annual International Meeting*. American Society of Agricultural and Biological Engineers. doi: 10.13031/aim.201901464.

Ionescu, C. *et al.* (2014) 'Human3.6M: Large scale datasets and predictive methods for 3D human sensing in natural environments', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(7), pp. 1325–1339. doi: 10.1109/TPAMI.2013.248.

Karhu, O., Kansi, P. and Kuorinka, I. (1977) 'Correcting working postures in industry: A practical method for analysis', *Applied Ergonomics*, 8(4), pp. 199–201. doi: 10.1016/0003-6870(77)90164-8.

Kowalski, K. *et al.* (2012) 'Direct and indirect measurement of physical activity in older adults: A systematic review of the literature', *International Journal of Behavioral Nutrition and Physical Activity*, 9(August 2014). doi: 10.1186/1479-5868-9-148.

Landis, J. R. and Koch, G. G. (1977) 'An Application of Hierarchical Kappa-type Statistics in the Assessment of Majority Agreement among Multiple Observers', *Biometrics*, 33(2), p. 363. doi: 10.2307/2529786.

Li, G. and Buckle, P. (1999) 'Current techniques for assessing physical exposure to work-related musculoskeletal risks, with emphasis on posture-based methods', *Ergonomics*. Taylor & Francis Ltd, pp. 674–695. doi: 10.1080/001401399185388.

Li, L., Martin, T. and Xu, X. (2020) 'A novel vision-based real-time method for evaluating postural risk factors associated with musculoskeletal disorders', *Applied Ergonomics*, 87(January). doi: 10.1016/j.apergo.2020.103138.

Lunin, A. and 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*, p. 107660. doi: 10.1016/j.cie.2021.107660.

Lynn, M. and Corlett, N. (1993) 'RULA: A survey method for the investigation of work-related upper limb disorders', *Applied Ergonomics*, 24(2), pp. 91–99.

Manghisi, V. M. *et al.* (2017) 'Real time RULA assessment using Kinect v2 sensor', *Applied Ergonomics*, 65, pp. 481–491. doi: 10.1016/j.apergo.2017.02.015.

Microsoft (2021) *Azure Kinect DK documentation | Microsoft Docs*. Available at: https://docs.microsoft.com/en-us/azure/kinect-dk/ (Accessed: 12 October 2021).

Moore, J. S. and Garg, A. (1995) 'The strain index: A proposed method to analyze jobs for risk of distal upper extremity disorders', *American Industrial Hygiene Association Journal*, 56(5), pp. 443–458. doi: 10.1080/15428119591016863.

Neumann, W. P. and Village, J. (2012) 'Ergonomics action research II: A framework for integrating HF into work system design', *Ergonomics*, 55(10), pp. 1140–1156. doi: 10.1080/00140139.2012.706714.

Newell, A., Yang, K. and Deng, J. (2016) 'Stacked Hourglass Networks for Human Pose Estimation', *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9912 LNCS, pp. 483–499. doi: 10.1007/978-3-319-46484-8_29.

Occhipinti, E. (1998) 'OCRA: A concise index for the assessment of exposure to repetitive movements of the upper limbs', *Ergonomics*, 41(9), pp. 1290–1311. doi: 10.1080/001401398186315.

Plantard, P. *et al.* (2015) 'Pose estimation with a kinect for ergonomic studies: Evaluation of the accuracy using a virtual mannequin', *Sensors (Switzerland)*, 15(1), pp. 1785–1803. doi: 10.3390/s150101785.

Plantard, P. *et al.* (2017) 'Validation of an ergonomic assessment method using Kinect data in real workplace conditions', *Applied Ergonomics*, 65, pp. 562–569. doi: 10.1016/j.apergo.2016.10.015.

Silverstein, B. and Clark, R. (2004) 'Interventions to reduce work-related musculoskeletal disorders', *Journal of Electromyography and Kinesiology*, 14(1), pp. 135–152. doi: 10.1016/j.jelekin.2003.09.023.

Snook, S. H. and Ciriello, V. M. (1991) 'The design of manual handling tasks: Revised tables of maximum acceptable weights and forces', *Ergonomics*, 34(9), pp. 1197–1213. doi: 10.1080/00140139108964855.

Tölgyessy, M. *et al.* (2021) 'Evaluation of the azure kinect and its comparison to kinect v1 and kinect v2', *Sensors (Switzerland)*, 21(2), pp. 1–25. doi: 10.3390/s21020413.

Toshev, A. and Szegedy, C. (2014) 'DeepPose: Human pose estimation via deep neural networks', in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society, pp. 1653–1660. doi: 10.1109/CVPR.2014.214.

Wilhelm, M. *et al.* (2020) 'ErgoTakt: A novel approach of human-centered balancing of manual assembly lines', *Procedia CIRP*, 97, pp. 354–360. doi: 10.1016/j.procir.2020.05.250.

Xu, X. *et al.* (2011) 'The Validity and Interrater Reliability of Video-Based Posture Observation During Asymmetric Lifting Tasks.', *http://dx.doi.org/10.1177/0018720811410976*, 53(4), pp. 371–382. doi: 10.1177/0018720811410976.

Yan, X., Li, H., Wang, C., *et al.* (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(November), pp. 152–163. doi: 10.1016/j.aei.2017.11.001.

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