HIDDEN MARKOV MODELLING AND RECOGNITION OF EULER-BASED MOTION PATTERNS FOR AUTOMATICALLY DETECTING RISKS FACTORS FROM THE EUROPEAN ASSEMBLY WORKSHEET

Brenda Elizabeth Olivas-Padilla[1] *Dimitrios Menychtas*[1] *Alina Glushkova*[1] *Sotiris Manitsaris*[1]

[1]Centre for Robotics
MINES ParisTech, PSL Université Paris, France
{brenda.olivas, dimitrios.menychtas, alina.glushkova, sotiris.manitsaris}@mines-paristech.fr

ABSTRACT

To prevent work-related musculoskeletal disorders (WMSD) the ergonomists apply manual heuristic methods to determine when the worker is exposed to risk factors. However, these methods require an observer and the results can be subjective. This paper proposes a method to automatically evaluate the ergonomic risk factors when performing a set of postures from the ergonomic assessment worksheet (EAWS). Joint angle motion data have been recorded with a full-body motion capture system. These data modeled the motion patterns of four different risk factors, with the use of hidden Markov models (HMMs). Based on the EAWS, automated scores were assigned by the HMMs and were compared to the scores calculated manually. Because the method proposed here is intrusive and requires expensive equipment, kinematic data from a reduced set of two sensors was also evaluated.

Index Terms**—** Hidden Markov Models, risk factors, wearables, gesture recognition, work-related musculoskeletal disorders

1. INTRODUCTION

The work-related musculoskeletal disorders (WMSD) in the industry are becoming increasingly common. These disorders are caused from the execution of activities that are repetitive, exert forces or require awkward postures [1]. Treatment and recovery of WMSDs is often unsatisfactory, resulting in a temporal or permanent disability, affecting the industrial worker's quality of life and increases company's costs. Experts have developed ergonomic assessment methods to prevent WMSD-related hazards. These methods are based on theoretical knowledge of human physical limitations and abilities [2] and define accepted standards (e.g ISO 11226:2000 and EN 1005-4). Some of the most used methods in industry are the Rapid Upper Limb Assessment (RULA) [3], Ovako Working Posture Analysing System (OWAS) [4], and the European Assembly Worksheet (EAWS) [5], which consists of four sections for the evaluation: working postures, action forces, manual materials handling, and repetitive loads of the upper limbs. To carry out an assessment with these methods, it is necessary that the ergonomist fill manually their respective worksheet. This worksheet evaluates the exposure of the worker to ergonomic risk factors. These factors are mainly related to the working posture, action forces of the whole body, manual material handling and task repetitiveness. However, the results of current ergonomic assessments are subjective since they rely on the ergonomist perception and experience. Moreover, the frequency of evaluation and monitoring of WMSD risk factors is limited since the evaluations are time-consuming and need to be applied by an ergonomist. To overcome some of these limitations, motion capture technology is used for a more objective ergonomic evaluation. Yan et al. [6] used inertial measurement units (IMU) to measure torso inclination for a monitoring system for construction workers' WMSD prevention. Busch et al. [7] used optical markers to track the upper body segments (head, hands, elbows, torso, and waist) and fill in automatically the REBA ergonomic assessment worksheet. Such approaches still present issues that made them impractical to implement in industry. Only a few postural risk factors are screened accurately, and vision approaches are costly and face occlusions issues. In this paper, a pipeline for automatic recognition of four postural risk factors and the computation of ergonomic score is proposed based on the evaluation protocol of EAWS, where the first risk factor is the posture of the legs (F1) with three possible motion patterns: standing, seating and kneeling. The second factor focuses on the torso inclination (F2) with 2 patterns: bending forward or not. The third risk factor is the lateral bending and torso rotation (F3). Finally the fourth one, is the elevation of the arms (F4). Depending on which factor is present during the performance, an automatic EAWS-related score is assigned on a scale from 0.5 to 26.5, with the larger values assigned to the more dangerous postures. Motion capture (MoCap) from an IMU suit was used

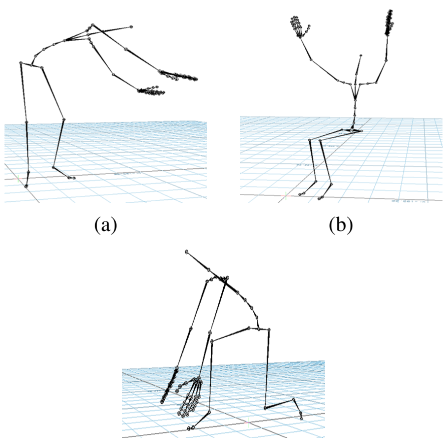
The research leading to these results has received funding by the EU Horizon 2020 Research and Innovation Programme under grant agreement No. 820767, project CoLLaboratE.

to record the postures this paper work examined. For the risk factors modeling, Hidden Markov Models (HMM) were used for automatic recognition. However, to implement the proposed system in an industrial environment, it is necessary to use less intrusive technologies and to minimize the number of sensors placed on the human body. To address this limitation, this paper also evaluates the performance of the pipeline with only two IMU sensors. The results indicate that it is possible to monitor the exposure to ergonomic risk factors using two accelerometers, potentially from a smatphone and a smartwatch, which is a more realistic attempt for wide industrial implementation.

2. PIPELINE

2.1. Generation of the data set

Ten healthy subjects were recorded performing 28 gestures, with three repetitions, six seconds each. During these gestures, the subject could be exposed to any combination of risk factors. The motion capture technology (MoCap) used was an IMU full-body suit (NANSENSE-BioMed Bundle, Baranger Studios, Los Angeles, CA, USA). The output was joint angles of the full body as a BVH file. For data processing, only a low-pass Butterworth filter was applied to remove noise in the MoCap data. In Fig. 1 it is shown three examples of gestures that were recorded, each one exposing to a different combination of risk factors.



(a) (b) (c)

Fig. 1: Example of three different awkward gestures. (a) Standing while bending forward and rotating the torso; (b) Sitting while rising arms above shoulder level; (c) Kneeling while bending forward.

2.2. Recognition of risk factors on awkward gestures by using Hidden Markov Models

For the recognition of the postural risk factors, four sets of Hidden Markov Models (HMM) were used, one for each factor. HMMs were used because they have proved to be a prominent tool for gesture recognition [8, 9]. In Fig. 2 it is shown the scheme for the recognition of the four factors. For the recognition of F1, three HMMs were trained using only the joint angles from the lower body. Each HMM modeled one of the three possible postures of the legs (standing, sitting, and kneeling). The HMM that provided the maximum probability indicated the posture recognized. E.g. if HMM F1.1 has the highest probability then the posture recognized is standing. Two HMMs were trained for the recognition of F2, using only the data from sensors located on the spine. One HMM modeled the gestures where the subjects were upright and the other where they were bending forward. Another two HMM were trained for the recognition of F3, using the data from the spine and arms, hence subjects moved both body regions to execute the gestures involving the risk factor F3. One HMM corresponded to the gestures where the subjects were rotating and lateral bending their torso and the other where they were not. The recognition of F4 was done with another two HMM trained with the data from the arms and shoulders. One HMM modeled the gestures where the subjects raised their arms above shoulder level and the other gestures where they kept their arms low. An ergodic HMM learned the hidden states given the observation sequence (joint angles) of each gesture, by using the Baum Welch algorithm. The ergodic models were selected since for all gestures the subjects returned to the initial posture. The gestures were discretized using K-means clustering. The number of states for each model and the number of clusters for discretization were chosen by applying a stratified 10-fold cross-validation. The centroids that produced the best results were retained to quantize new test gestures. For every new test gesture, the L2-norm was computed with each centroid and the cluster that had the minimum distance was the one where the gesture was assigned.

To evaluate the possibility to implement the proposed pipeline with IMUs from smartphones and smartwatches, a configuration of two sensors was also evaluated. The sensors used for this configuration were the sensor located at the right forearm, representing the IMU of a smartwatch, and the sensor located on the hips, representing the IMU from a smartphone. The right forearm was chosen since most of the subjects are right-handed, and the sensor of the hips because the origin of the movement for bending forward and rotating the torso starts from the hips.

2.3. Automatic computation of an EAWS-related ergonomic score

For the computation of the automatic EAWS-related score, four equations were designed based on the tables provided by the EAWS worksheet in the working posture assessment section [5]. Note that the numbers used for awkward postures that are assumed for a lapse around six seconds. The automatic EAWS-related score defined as $S \in \big[0.5, 26.5\big]$, is the

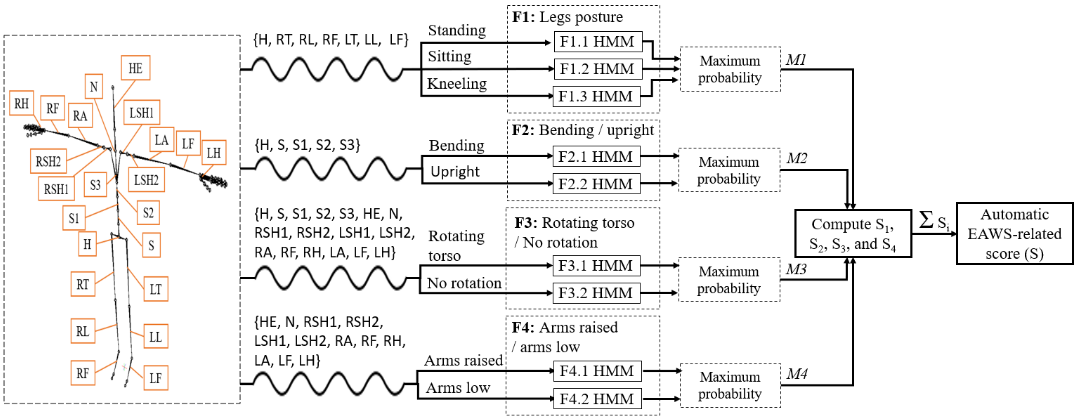


Fig. 2: Pipeline of the modeling of the risk factors using motion data and the computation of the EAWS-related score.

sum of the scores S_1, S_2, S_3, and S_4 as shown in (1).

$$S = \sum_{i=1}^{4} S_i \tag{1}$$

S_1 was computed by using (4):

$$S_1 = L_{M1}, \quad L = \begin{bmatrix} 1.5 \\ 0.5 \\ 7 \end{bmatrix} \tag{2}$$

where $M1$ was used as index of vector L, which consists of the initial scores defined by EAWS for standing, sitting, and kneeling respectively. For example, 0.5 a low risk value is assigned when the subject is sitting and a seven a higher risk value, when the subject is kneeling. The second score S_2 was computed with the following equation:

$$S_2 = (M2 - 1)B_{M1}, \quad B = \begin{bmatrix} 7 \\ 1 \\ 3 \end{bmatrix} \tag{3}$$

where $M2$ is two if the subjects are bending and one if not; B is the scores for bending forward, depending if the subjects are standing, sitting or kneeling, which is indicated by $M1$. In this case, if the subjects are bending forward, $(M2 - 1)$ will be one and a score from the vector B will be obtained. If the subject is upright, the subtraction $(M2 - 1)$ will be equal to zero as S_2. The next score S_3 was computed as:

$$S_3 = 7.5(M3 - 1) \tag{4}$$

where $M3$ is two if the subjects are rotating their torso and one if they are not. If the subjects are rotating their torso then S_3 will be equal to 7.5, if not is equal to zero. Finally, S_4 was computed using (5) and (6).

$$S_4 = (M4 - 1)(2 - M2)A_{M1} + 5(M4 - 1)(M2 - 1) \tag{5}$$

$$A = \begin{bmatrix} 7 \\ 6.5 \\ 9 \end{bmatrix} \tag{6}$$

where $M4$ is two if the subjects are rising their arms and one if they are not, if they are not $(M4 - 1)$ will be zero as S_4. If the subjects are rising their arms, then the score of S_4 would depend if the subjects are bending forward too and if they are standing, sitting or kneeling. For example, if the subjects are not bending, a score for having the arms raised will be obtained from the vector A, this score will depend if the subjects are standing, sitting or kneeling, which is indicated by $M1$. If the subjects is bending forward too (indicated by $M2$), S_4 will be equal to five.

3. RESULTS

For the evaluation, a stratified cross-validation (CV) procedure with ten iterations was followed. The data set was randomly partitioned in ten parts of equal size. Then nine of them were used for training of the models and the remaining was used for testing. The process was repeated for all ten parts. Since the data set had a less number of gestures where the subject was kneeling or rising his arms, a stratified CV was chosen to keep the same proportion of gestures with different factors for each iteration. Therefore, only 180 gestures for each class were used for F1 (standing, sitting, and kneeling), 90 per class for F2 (upright and bending), 90 per class for F3 (no rotating the torso and rotating torso), and 90 per class for F4 (arms low and arm raised). Table 1 illustrate the resulting confusion matrices after the ten iterations.

From these confusion matrices the F1-score was computed by using the following equation:

$$F1_{score} = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{7}$$

The set composed of all joint angles achieved an F1-score of 0.9499 for F1, 0.9443 for F2, 0.9165 for F3, and 0.9272 for F4, thus an overall F1-score of 0.9345. These recognition performances are compared with the ones achieved with the minimum set of sensors in Table 2.

Table 1: Confusion matrices for the recognition of F1, F2, F3, and F4, using all joint angles. Note that U: Upright; B: Bending forward; NTR: No rotating the torso; TR: Rotating the torso; AL: Arms low; AR: Arms raised

		Motion data		
HMM		*Standing*	*Seated*	*Kneeling*
F1	F1.1	60	0	0
	F1.2	0	47	6
	F1.3	0	2	53

		Motion data	
HMM		*U*	*B*
F2	F2.1	73	9
	F2.2	0	86

		Motion data	
HMM		*NTR*	*TR*
F3	F3.1	69	12
	F3.2	2	85

		Motion data	
HMM		*AL*	*AR*
F4	F4.1	73	12
	F4.2	0	83

Table 2: Overall recognition performance with each configuration of sensors for F1, F2, F3, and F4. Note that ALL: Configuration with all the sensors; H and RF: Configuration with two sensors

Risk factor	**Sensors**	**F1-scores**
F1	ALL	0.9505
	H	0.7927
F2	ALL	0.9461
	H	0.8593
F3	ALL	0.9159
	H and RF	0.9272
F4	ALL	0.9283
	H and RF	0.9451

Table 3: Mean absolute errors and the absolute error standard deviation with each configuration of sensors. Note that ALL: Configuration with all the sensors; H and RF: Configuration with two sensors

Sensors	**MEA**	**Std**
ALL	1.5206	0.6337
H and RF	1.9496	0.4005

The proposed minimum set achieved an F1-score of 0.8811. Hence by using all joint angles, there is an improvement of only 0.0534 over the minimum set of two sensors. The factor that was the most challenging for the minimum set was F1. Because there is only one sensor on the hips, there are not enough data to discriminate between the three different posture of the legs. Therefore, the minimum set proposed is recommended for upper body ergonomics monitoring. For the three out of four risk factors, when using only two sensors, satisfying recognition results are achieved (F1-Score $>$ 85%). These results are promising and open perspectives for the use of this pipeline in industrial environments by using less invasive technologies such as smartphones/smartwatches, etc.

The absolute difference between the computed automatic EAWS-related score and the manually assigned EAWS score was calculated per each EAWS-related score prediction in the 10-fold cross-validation. After the cross-validation, the mean of all absolute differences was computed, this corresponds to the mean absolute error (MAE). The MAE of each configuration of sensors and the standard deviation of the absolute differences are shown in Table 3.

4. CONCLUSION

In this paper, a methodology for the recognition of postural risk factors on ergonomically hazardous gestures is proposed. Wearables IMUs were used for the data collection, where ten subjects executed 28 gestures, with different levels of ergonomic risk according to the EAWS. From the data, joint angles were obtained, from which motion patterns were successfully recognized using models based on HMMs. By using only two sensors placed on the right forearm and on the hips and following the pipeline proposed, it was possible to compute the automatic EAWS-related score with an MAE of 1.9496 and small standard distribution of the error. This indicates that it is possible to use a minimum set of sensors for the automatic computation of the EAWS-related score. This can potentially allow the use of smartwatches and smartphones for ergonomic assessment for the industry in a day-to-day basis. Wearables measuring working postures have the potential to reduce the prevalence of WMSD. High frequency and easily-accessible monitoring technology can help give real-time feedback to workers. Therefore, for future research, the recognition algorithm presented in this paper will be tested with data from less specialized technologies to design a module for real-time ergonomic feedback.

5. REFERENCES

[1] D. Ranney, R. Wells, and A. Moore, "Upper limb musculoskeletal disorders in highly repetitive industries: Precise anatomical physical findings," Ergonomics, vol. 38, no. 7, pp. 1408–1423, 1995.

[2] C. Berlin and C. Adams, "Ergonomics Evaluation Methods," in Production Ergonomics: Designing Work Systems to Support Optimal Human Performance, 2017, pp. 139–160.

[3] M. Lynn and N. Corlett, "RULA: A survey method for the investigation of work-related upper limb disorders," Appl. Ergon., vol. 24, no. 2, pp. 91–99, 1993.

[4] O. Karhu, P. Kansi, and I. Kuorinka, "Correcting working postures in industry: A practical method for analysis," Appl. Ergon., vol. 8, no. 4, pp. 199–201, 1977.

[5] K. Schaub, G. Caragnano, B. Britzke, and R. Bruder, "The European Assembly Worksheet," Theor. Issues Ergon. Sci., vol. 14, no. 6, pp. 616–639, 2013.

[6] X. Yan, H. Li, A. R. Li, and H. Zhang, "Wearable IMU-based real-time motion warning system for construction workers' musculoskeletal disorders prevention," Autom. Constr., vol. 74, pp. 2–11, 2017.

[7] B. Busch, G. Maeda, Y. Mollard, M. Demangeat, and M. Lopes, "Postural optimization for an ergonomic human-robot interaction," IEEE Int. Conf. Intell. Robot. Syst., vol. 2017–September, pp. 2778–2785, 2017.

[8] C. Volioti, S. Manitsaris, and A. Manitsaris, "Offline statistical analysis of gestural skills in pottery interaction," ACM Int. Conf. Proceeding Ser., pp. 172–173, 2014.

[9] A. Malaise, P. Maurice, F. Colas, and S. Ivaldi, "Activity Recognition for Ergonomics Assessment of Industrial Tasks with Automatic Feature Selection," IEEE Robot. Autom. Lett., vol. 4, no. 2, pp. 1132–1139, 2019.