Monocular Vision-Based 3D Human Pose Estimation and Cumulative Damage Assessment at Industrial Workplaces

W.S. Lor[1], J.W. Kim[2]
[1]School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore
[2]School of Civil and Environmental Engineering, Nanyang Technological University, Singapore (Corresponding author)
(WLOR001@e.ntu.edu.sg[1], jinwoo.kim@ntu.edu.sg[2])

Abstract - **Although work-related musculoskeletal disorders (WMSDs) have been a major concern in physically demanding industries, ergonomic risk assessment often lacks comprehensiveness in considering activity duration and an effective way to monitor industrial workers' postures. Furthermore, they are not integrated due to the cumbersome operational process when using biomechanical analysis software such as OpenSim and the complexity of estimating total lumbar compressive force. To address this issue, we present a method to estimate the total lumbar compressive force only with a monocular camera by applying a state-of-the-art 3D Human Pose Estimation algorithm; and simplify the operational process with a 'parameter *a* method' for the estimation of total lumbar compressive force, which can be easily adjusted by a professional ergonomist. Results show that the estimated force and ergonomic injury risk fall within a reasonable range when compared to the results obtained from the previous studies, where existing, complex biomechanical analysis was performed. This finding implies an enormous potential for enhancing the prevention of WMSDs by adopting the proposed method, which integrates technologies, simplifies the operational process, and enables comprehensive ergonomic risk assessment.**

Keywords - **Ergonomics, Computer Vision, Cumulative Damage Assessment, Monocular Vision-Based 3D Human Pose Estimation, Workers' Safety in Industrial Workplaces**

I. INTRODUCTION

Work-related musculoskeletal disorders (WMSDs) are major occupational hazards which account for one-third of all injuries and illnesses in industrial workplaces in the United States [1]. In the US, the median days of the workers away from work have increased from 8 days in 1992 to 13 days in 2014 due to WMSDs where this issue causes the inefficient project's progress and the economic burden (i.e., total wage loss of $46 million in 2014) to the injured workers as well [2]. The effects of WMSDs on the muscles may cause permanent, irreversible consequences [3].

There are several existing standard ergonomic risk assessments to deal with this problem from an ergonomics perspective, National Institute for Occupational Safety and Health (NIOSH) Lifting Equation is one of them [4]. NIOSH is able to classify workers' postures into several categories according to the risk level as low or high. NIOSH can be automated using wearable sensors to resolve the accuracy of manual observation to resolve the accuracy of manual observation [5], [6] due to the subjective decisions by the designers while doing the assessment.

However, with the development of vision-based Human Pose Estimation (HPE), wearables are considered too bulky for workers while working [7]–[9]. Hence, multiple-camera vision-based HPE becomes the alternative until there is a breakthrough in monocular vision-based HPE studies, resolving the complex-setup issues for multiple-camera vision-based HPE.

However, there are at least two main challenges for the existing solutions. First, existing automated methods equipped with NIOSH cannot quantify industrial workers' physical fatigue that accumulates over time while conducting a series of manual tasks. Second, the cumbersome operational process and the complexity of biomechanical analysis limit the practical usage of monocular vision-based 3D HPE in industrial workplaces. To address these issues, our research objective aims to integrate a more comprehensive ergonomic risk assessment at the lumbar and a state-of-the-art monocular vision-based 3D HPE architecture (i.e., Videopose3D) [10] to quantify the ergonomic risk of a subject, answering the following research question: Can we effectively introduce a more comprehensive ergonomic risk assessment in industrial workplaces by simplifying the complex HPE setup? Answering this question helps us discover a practical way to apply a better quantification of ergonomic risk, which eventually may help reduce WMSDs with proper interventions.

II. METHODOLOGY

Fig. 1 shows the visualization of the overall pipeline of this research's ergonomic injury risk assessment. The entire pipeline can be decomposed into two models, namely, the biomechanical analysis model and the ergonomic risk assessment model based on cumulative fatigue theory. The details for each model are stated in the subsequent sections.

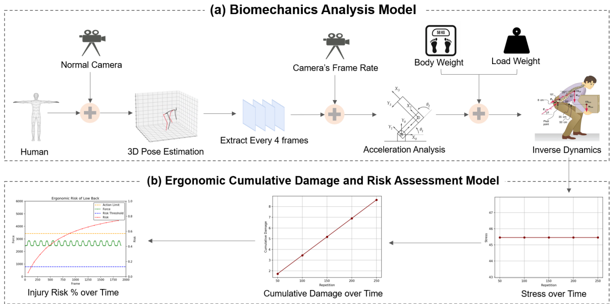


Fig. 1. Overall pipeline of ergonomic injury risk assessment

A. Biomechanical Analysis Model:

1. Data pre-processing

Since this research aims to perform a vision-based HPE technique, no motion-capture camera (e.g., Vicon camera) is required. All we need is just a regular camera with any frame rate. First, we applied the state-of-the-art computer vision architecture (i.e., Videopose3D) to perform the monocular vision-based 3D pose. After extracting the 3D coordinates, we filtered out some frames to reduce the noise since results are affected if the frame extraction frequency was too high. However, it might also lose accuracy if the frame extraction frequency was too low. Hence, we experimented to find a reasonable balance point.

2. Estimating joints acceleration using 3D coordinates

Three frames were required to estimate the linear and angular acceleration based on the central difference scheme [11]. However, one more additional frame was required as an initial condition when starting at the beginning. The reason is explained more intuitively in the following mathematical expression formulated based on the coordinate system in Fig. 2(a) and the free body diagram in Fig. 2(b) by taking point B as an example, where Table 1. shows the interpretation of the important symbols in Fig. 2.

Linear Acceleration:

$$(\overrightarrow{v_B})_2 = \frac{(\overrightarrow{r_B})_3 - (\overrightarrow{r_B})_1}{2dt} \tag{1}$$

$$(\overrightarrow{v_B})_1 = \frac{(\overrightarrow{r_B})_2 - (\overrightarrow{r_B})_0}{2dt} \tag{2}$$

$$(\overrightarrow{a_B})_2 = \frac{(\overrightarrow{v_B})_2 - (\overrightarrow{v_B})_1}{dt} \tag{3}$$

Angular acceleration:

$$|(\overrightarrow{v_{BA}})_2| = |(\overrightarrow{\omega_2})_2| * |(\overrightarrow{r_{BA}})_2| \tag{4}$$

$$|(\overrightarrow{\omega_2})_2| = \frac{|(\overrightarrow{v_{BA}})_2|}{|(\overrightarrow{r_{BA}})_2|} \tag{5}$$

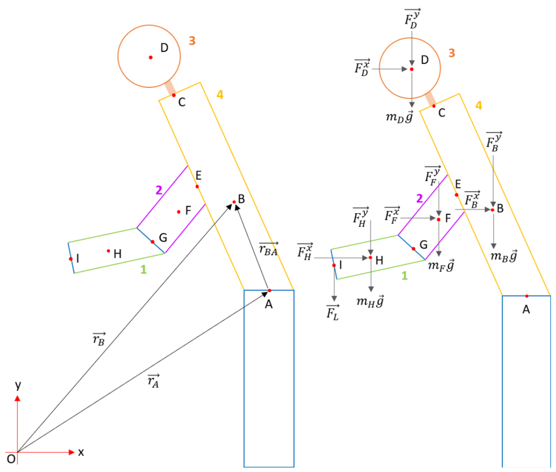
$$(\overrightarrow{R_2})_2 = (\overrightarrow{r_{BA}})_1 \times (\overrightarrow{r_{BA}})_2 \tag{6}$$

$$(\widehat{\omega_2})_2 = \frac{(\overrightarrow{R_2})_2}{|(\overrightarrow{R_2})_2|} \tag{7}$$

$$(\overrightarrow{\omega_2})_2 = |(\overrightarrow{\omega_2})_2|(\widehat{\omega_2})_2 \tag{8}$$

Same method for $(\overrightarrow{\omega_2})_1$,

$$(\overrightarrow{\alpha_2})_2 = \frac{(\overrightarrow{\omega_2})_2 - (\overrightarrow{\omega_2})_1}{dt} \tag{9}$$



(a) (b)

Fig. 2. (a) Coordinate system and (b) free body diagram

A general idea for the mathematical expression of the derivation of linear acceleration was as follows: we only needed frame n-2 to frame n to estimate the acceleration of frame n-1 since the information of frame n-3 had been obtained from the previous case. The same concept applied to angular acceleration as well. In the derivation of angular acceleration, point A was also one of the body joints considered the reference point (i.e., lumbar). The capital letter 'R' in (6) was just an auxiliary vector to find the direction of the angular velocity since they shared the same direction, where its magnitude was not essential.

3. Estimating joints forces using inverse dynamics and parameter 'a' method

The last step of the biomechanics analysis was the inverse dynamics, namely, to find the lumbar axial compressive force and the bending moment about the lumbar in this case. Lumbar axial compressive force was affected by the load's weight, the gravitational force exerted on each body joint, and the inertial force exerted on each body joint. The impact of the load's weight was complicated in biomechanics since it was not linearly proportional to its actual weight; sometimes, it was causing an impact of up to ten times its actual weight. Hence, we decided to first deal with the forces caused by the gravitational force exerted on each body joint and the inertial force exerted on each body joint before considering the impact of the load's weight [8]

According to the research in [12], the lumbar will experience a 1000N of force for every additional 5kg of load's weight. While for the bending moment, it is experiencing a 20Nm bending moment for every additional 5kg of load's weight. When dealing with the bending moment about the lumbar, we picked an arbitrary point as the reference point to define the position vectors from the reference point to the trunk, head, upper arm, and forearm, similarly for the lumbar axial compressive force. However, to obtain the inertial moment of each body joint, the moment of inertia of each body segment of interest was required [13], as well as the body parameters for the estimation of the moment of inertia of each body segment (i.e., trunk, head, upper arm, forearm) such as body segment mass, body segment length, and radius of gyration about the sagittal axis, longitudinal axis, and transverse axis respectively [14].

TABLE 1

Definition of each important symbol in Fig. 2

Symbol of Body Part	**Name of Body Part**	**Location of Center of Mass of Body Part**
1	Forearm	H
2	Upper Arm	F
3	Head	D
4	Trunk	B

After understanding the relationship between the axial compressive force $|\vec{F'}|$, the bending moment, $|\vec{M'}|$, and the load's weight, we were interested in the sum of the axial compressive force and the equivalent force of the bending moment, $|\vec{F_b}|$ to obtain the total lumbar compressive force, $|\vec{F_c}|$. To do so, the parameter a method was introduced, where a was equal to the product of the cross-sectional area of the muscles at the lumbar, A and the distance from the point of interest to the neutral axis, y over the second moment of the area at the lumbar, I. The parameter a allowed professional ergonomists to fine-tune the equivalent force of the bending moment. We calibrated the model by fine-tuning parameter a's value using the cross-validation technique. The participant was requested to stand upright and fine-tune the value of parameter a until the value falls in a reasonable range compared to the results in the following research [15], [16]. The participant was then asked to carry 160N of load's weight, with his back was bent and straight, respectively. The total lumbar compressive forces at 160N were compared to the results in [16]. The calibration process is done if the total lumbar compressive forces tally nicely with the results in [15]. Otherwise, refer to [16] and fine tune the parameter a again. We repeated this process, and eventually, parameter a equaled four was found to be a reasonable value. According to (15) it explained why 920N was added to the axial lumbar compressive force in (10). Note that in this research, we only extracted the peak lumbar axial compressive force and peak bending moment for each repetition of the activity to estimate the ergonomic risk with a higher safety factor conservatively.

$$|\vec{F'}| = |\vec{F}| + 920 \tag{10}$$
$$|\vec{M'}| = |\vec{M}| + 20 \tag{11}$$
$$\sigma = \frac{|\vec{M'}|y}{I} \tag{12}$$
$$\sigma A = \frac{|\vec{M'}|y}{I}A \tag{13}$$
$$|\vec{F_b}| = a|\vec{M'}| \tag{14}$$
$$|\vec{F_c}| = |\vec{F'}| + |\vec{F_b}| \tag{15}$$
$$|\vec{F_c}| = |\vec{F}| + 4|\vec{M}| + 1000 \tag{16}$$

B. Ergonomic Risk Assessment Model:

1. Estimating ultimate strength percentile at lumbar

The ergonomic risk assessment model based on cumulative fatigue theory is illustrated in Fig. 1(b). This model aims to quantify the ergonomic risk by taking the cumulative damage theory into account by performing the subsequent steps.

In (17), Ultimate Strength Percentile (USP), S was obtained from the ratio of the stress of total lumbar compressive force, σ_c to the ultimate strength of lumbar, σ_u. Referring to the data of blended males and females for specimens aged from 20 to 60 in [17], the average ultimate force of lumbar, $|\vec{F_u}|$ was approximately 6000N. However, effective cross-sectional area, A of the lumbar was required before calculating the stress, where it was complicated to estimate a reliable, effective cross-sectional area [18]. Therefore, to simplify the estimation of USP, we used the ratio of forces instead of the ratio of stresses by cancelling off the A referring to (18).

$$S = \frac{\sigma_c}{\sigma_u} \times 100\% \tag{17}$$

$$S = \frac{|\vec{F_c}|}{|\vec{F_u}|} \times 100\% \tag{18}$$

2. Estimating cumulative damage at lumbar

Substituting the value S of USP into (19), the number of cycles to failure was estimated. Then, we obtained the damage per repetition or cycle by taking the reciprocals of the number of cycles to failure. Taking the product of the damage per repetition and the number of repetitions of activities provided us the cumulative damage, CD at lumbar as shown in (20) [17].

$$N = 902416e^{-0.162S} \tag{19}$$

$$CD = \frac{1}{N} \times no.of\ repetitions \tag{20}$$

3. Estimating injury risk probability at lumbar

Lastly, the injury risk probability at the lumbar was estimated by substituting the value of CD obtained from (20) into (21) [17]. The cumulative damage and ergonomic risk assessment has been completed until this step [17].

$$p = \frac{e^{Y'}}{1+e^{Y'}}, where\ Y' = 1.72 + 1.03log(CD) \tag{21}$$

III. RESULTS

In this research, we assume there is a constant stress, which led to a linear cumulative damage versus time (see Fig. 1(b)). Based on this result, the exponential relationship between the injury risk and frame number (i.e., red curve), as well as the relationship of the total lumbar compressive force versus frame number (i.e., green curve) under no-load condition was illustrated in Fig. 3, where each red dot in the red curve stood for one repetition. The amber-dotted line referred to the force action limit according to NIOSH. In contrast, the blue-dotted line referred to the threshold of the injury risk probability, converted from the threshold of cumulative damage (i.e., 0.03) [17]. From Fig. 3, the risk was accumulated all the time based on the total lumbar compressive force using the cumulative fatigue theory.

Furthermore, the total lumbar compressive force varied according to the spine angle, compared to Fig. 3 and 4. By comparing the valley values in Fig. 3 to the results of zero degree of spine angle in the no-load condition in [15], approximately 200N of deviation was noticed. This result was satisfying to us, and hence, we moved on to comparing the valley and peak values in Fig. 5 to the force-estimation results of building the level-two and level-four walls using a 160N of Concrete Masonry Unit (CMU), respectively, in [16]. Note that for the 160N loading condition, we were referring to the results of different papers from the no-load condition as cross-validation to avoid bias. Referring to [16], at course two, where the workers needed to work in the most critical position (i.e., largest flexion angle), the range of total lumber compressive force was 2800N to 3500N (we obtained 4050N). Next, we observed that the total lumber compressive force range was 2200N to 3300N (we obtained 3700N) at course four, where the workers were standing straight. The difference between our force-estimation results and the results in [16] ranges from 700N to 900N.

IV. Discussion

The results reveal that the biomechanical analysis outcome falls within the same range stated in previous studies for no-load. In contrast, for 160N loading condition, the biomechanical analysis result is larger than the maximum value by hundreds of Newtons of force.

Our biomechanical analysis has two different loading conditions (i.e., no-load, 160N). Under a no-load condition in Fig. 3, we have an approximately 200N deviation compared to the results in [15]. A deviation is very typical in statistics. It may be caused by several factors such as different physical conditions of different people or 3D HPE accuracy issues. However, this previous research does not state the variance. Therefore, referring to [16], we notice that around 500N deviation from the average value is still acceptable due to the authors in [16] only considering quasi-static postures instead of dynamic postures in their biomechanical analysis. On the contrary, in our research, we take account of the inertial force referring to the free body diagram in Fig. 2 as the effect of the inertial force plays a significant role [19]. Even a 2.5kg of load will make the valley values from lower than the action limit to higher than the action limit (i.e., amber dotted line) by approximately 500N of force which implies from a low-risk category to a high-risk category (see Fig. 6).

From the analysis above, our proposed framework provides reasonable and expected results. In other words, the ergonomic risk assessment based on cumulative damage is vital because it classifies the risk into two categories (i.e., low risk and high risk) and quantifies it into a percentage. Moreover, it helps us to understand that the current damage is the outcome of accumulating the previous damage so that we will notice the hazards consciously when doing any activity at the workplace. In addition, we should also recognize the contributions of the monocular vision based HPE and our biomechanical analysis algorithm using parameter a method for providing us the satisfying ergonomic risk assessment results. Hence, this research has the same perspective as [13], [14] for the parameters of human body; [20] for the ergonomic risk assessment based on cumulative fatigue theory; [16], [21] for the conclusion of awkward postures (e.g., larger spine flexion angle) lead to higher risk and lower productivity; and [8] for the effectiveness of monocular vision-based HPE compared to the existing HPE methods. Overall, the success of the integration of ergonomic risk assessment based on cumulative fatigue theory and monocular vision-based HPE will decrease the case of WMSDs by widely adopted at workplaces.

However, several improvements still remain. For future work, we can train a Deep Neural Network (DNN) model so that the machine-learning agent can decide the value of parameter a in different situations and resolve the dataset problem by performing the cross-validation using the same and large enough dataset to consolidate the conclusion of this research.

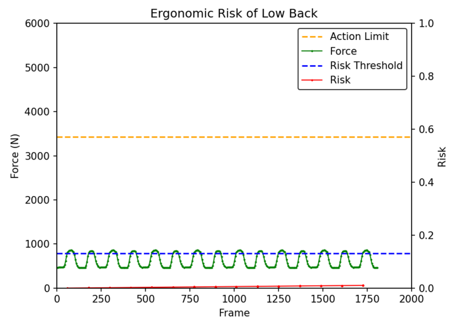


Fig. 3. Injury risk % and total lumbar compressive force versus frame number, respectively under no-load condition

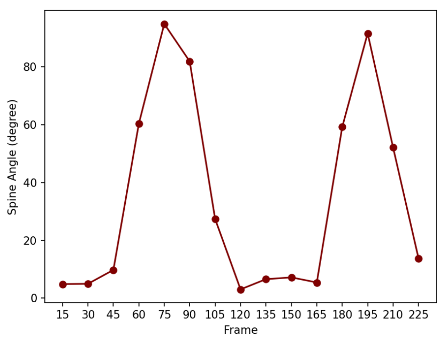


Fig. 4. Spine angle versus time

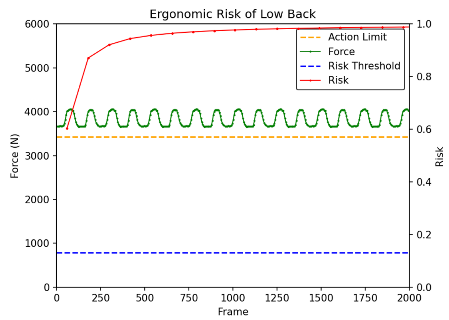


Fig. 5. Injury risk % and total lumbar compressive force versus frame number, respectively with 160N of load

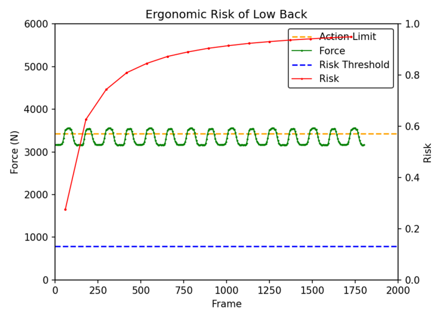


Fig. 6. Injury risk % and total lumbar compressive force versus frame number, respectively with 135N of load

V. Conclusion

As an effort to reduce WMSDs in industrial workplaces for workers' safety, we presented an approach to integrate a more comprehensive ergonomic risk assessment based on cumulative fatigue theory and monocular vision-based HPE. Specifically, we eliminated the necessity of the manual input of inertial force. Instead, we estimated the inertial acceleration based on the rate of change of the position vectors with the assistance of the camera frame rate for the time interval between every two frames. This approach provided us with the inertial force for estimating the total lumbar compressive force, which could be used to estimate the ergonomic risk based on the cumulative fatigue theory. To conclude, this research contributes to introducing an integrated approach to adopt a more comprehensive ergonomic risk assessment in industrial workplaces, leading to a safer working environment for the workers. Based on the findings, it is believed that a thorough comprehension of various human factors plays a crucial role in mitigating the complexity of the environmental conditions found in industrial workplaces. By adopting such an approach, the potential exists to enhance workers' safety, improve productivity within the industry, and optimize individual performance, provided that their physical health is consistently upheld.

Acknowledgment

We would like to acknowledge the funding support from Nanyang Technological University – URECA Undergraduate Research Programme for this research project. This research is supported by the Ministry of Education, Singapore, under its Academic Research Fund Tier 1 (RS16/22). Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not reflect the views of the Ministry of Education, Singapore.

References

[1] "Ergonomics - Overview | Occupational Safety and Health Administration." https://www.osha.gov/ergonomics (accessed May 29, 2023).

[2] X. Wang, S. Dong, S. D. Choi, and J. Dement, "Work-related musculoskeletal disorders among construction workers in the United States," *Environmental Medicine*, vol. 74, no. 5, pp. 374–380, 2017, doi: 10.1136/oemed.

[3] M. Y. Athirah Diyana *et al.*, "Risk factors analysis: Work-related musculoskeletal disorders among male traffic policemen using high-powered motorcycles," *Int J Ind Ergon*, vol. 74, Nov. 2019, doi: 10.1016/j.ergon.2019.102863.

[4] T. R. Waters, V. Putz-Anderson, A. Garg, and L. J. Fine, "Revised NIOSH equation for the design and evaluation of manual lifting tasks," *Ergonomics*, vol. 36, no. 7, pp. 749–776, 1993, doi: 10.1080/00140139308967940.

[5] H. Zhang, X. Yan, and H. Li, "Ergonomic posture recognition using 3D view-invariant features from single ordinary camera," *Autom Constr*, vol. 94, pp. 1–10, Oct. 2018, doi: 10.1016/j.autcon.2018.05.033.

[6] J. Wang, X. Li, S. H. Han, and M. Al-Hussein, "3D standard motion time-based ergonomic risk analysis for workplace design in modular construction," *Autom Constr*, vol. 147, Mar. 2023, doi: 10.1016/j.autcon.2022.104738.

[7] J. Zurada, W. Karwowski, and W. S. Marras, "A neural network-based system for classification of industrial jobs with respect to risk of low back disorders due to workplace design," *Appl Ergon*, vol. 28, no. 1, pp. 49–58, Feb. 1997, doi: 10.1016/S0003-6870(96)00034-8.

[8] H. Wang, Z. Xie, L. Lu, L. Li, X. Xu, and E. P. Fitts, "A Single-Camera Computer Vision-Based Method for 3D L5/S1 MomentEstimation During Lifting Tasks," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 65, no. 1, pp. 472–476, Sep. 2021, doi: 10.1177/1071181321651065.

[9] X. Yan, H. Li, C. Wang, J. O. Seo, H. Zhang, and H. Wang, "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*, vol. 34, pp. 152–163, Oct. 2017, doi: 10.1016/j.aei.2017.11.001.

[10] D. Pavllo, C. Feichtenhofer, D. Grangier, and M. Auli, "3D human pose estimation in video with temporal convolutions and semi-supervised training," Nov. 2018, [Online]. Available: http://arxiv.org/abs/1811.11742

[11] G. Großeholz, D. Soares, and O. von Estorff, "A stabilized central difference scheme for dynamic analysis," *Int J Numer Methods Eng*, vol. 102, no. 11, pp. 1750–1760, Jun. 2015, doi: 10.1002/nme.4869.

[12] M. Jäger, A. Luttmann, and W. Laurig, "Lumbar load during one-handed bricklaying," *Int J Ind Ergon*, vol. 8, no. 3, pp. 261–277, Nov. 1991, doi: 10.1016/0169-8141(91)90037-M.

[13] H. K. Ramachandran, S. Pugazhenthi, R. H. Krishnan, V. Devanandh, and A. K. Brahma, "Estimation of mass moment of inertia of human body, when bending forward, for the design of a self-transfer robotic facility Design and Development an 8 DoF Humanoid and Achieving Human Robot Interaction View project MAV, Low Reynolds number flow View project ESTIMATION OF MASS MOMENT OF INERTIA OF HUMAN BODY, WHEN BENDING FORWARD, FOR THE DESIGN OF A SELF-TRANSFER ROBOTIC FACILITY," 2016. [Online]. Available: https://www.researchgate.net/publication/271292606

[14] P. De Leva, "Adjustments to Zatsiorsky-Seluyanov's segment inertia parameters," *J Biomech*, vol. 29, no. 9, pp. 1223–1230, Sep. 1996, doi: 10.1016/0021-9290(95)00178-6.

[15] X. Xiang, Y. Yamada, Y. Akiyama, Z. Tao, and N. Kudo, "Validation of lumbar compressive force simulation in forward flexion condition," *Applied Sciences (Switzerland)*, vol. 11, no. 2, pp. 1–17, Jan. 2021, doi: 10.3390/app11020726.

[16] J. Ryu, B. Banting, E. Abdel-Rahman, and C. T. Haas, "Ergonomic Characteristics of Expert Masons," *J Constr Eng Manag*, vol. 149, no. 1, Jan. 2023, doi: 10.1061/(asce)co.1943-7862.0002434.

[17] S. Gallagher, R. F. Sesek, M. C. Schall, and R. Huangfu, "Development and validation of an easy-to-use risk assessment tool for cumulative low back loading: The Lifting Fatigue Failure Tool (LiFFT)," *Appl Ergon*, vol. 63, pp. 142–150, Sep. 2017, doi: 10.1016/j.apergo.2017.04.016.

[18] A. Keller, R. Gunderson, O. Reikeras, and J. Brox, "Reliability of Computed Tomography Measurements of Paraspinal Muscle Cross-Sectional Area and Density in Patients With Chronic Low Back Pain," *Spine (Phila Pa 1976)*, vol. 28, no. 13, pp. 1455–1460, Jul. 2003, doi: 10.1097/00007632-200307010-00017.

[19] H. Schechtman and D. L. Bader, "In vitro fatigue of human tendons," *J Biomech*, vol. 30, no. 8, pp. 829–835, Aug. 1997, doi: 10.1016/S0021-9290(97)00033-X.

[20] S. Gallagher and M. C. Schall, "Musculoskeletal disorders as a fatigue failure process: evidence, implications and research needs," *Ergonomics*, vol. 60, no. 2, pp. 255–269, Feb. 2017, doi: 10.1080/00140139.2016.1208848.

[21] J. Ryu, A. Alwasel, C. T. Haas, and E. Abdel-Rahman, "Analysis of Relationships between Body Load and Training, Work Methods, and Work Rate: Overcoming the Novice Mason's Risk Hump," *J Constr Eng Manag*, vol. 146, no. 8, Aug. 2020, doi: 10.1061/(asce)co.1943-7862.0001889.