3D Posture Estimation from 2D Posture Data for Construction Workers

Y. Yu[a] , H. Li[a], and X. Yang[a]

[a]Department of Building and Real Estate, The Hong Kong Polytechnic University, HKSAR, China
E-mail: yt.yu@connect.polyu.hk, xincong.yang@outlook.com, heng.li@polyu.edu.hk

Abstract –
Construction workers' behaviour is important for safety, health and productivity management. Workers' 3D postures are the data foundation of their behaviours. This paper established a preliminary 3D posture dataset of construction tasks and provided a 3D posture estimation method based on 2D joint locations. The results showed that the method could estimate 3D postures accurately and timely. The mean joint error and estimation time of each frame were 1.10 cm and 0.12 ms respectively. This method makes it possible to estimate construction workers' 3D postures from construction site images and contributes to a data-based construction workers' behaviour management.

Keywords –
Posture estimation; Construction worker; Behavior management

1 Introduction

Construction workers' behavior is an important factor in construction management. Construction workers' behavior is closely related to safety, health and productivities. More than 80% of construction accidents are related to workers' unsafety behaviors [1]. In addition, working behaviors, especially working postures, durations and work-rest schedules, are closely related to musculoskeletal disorders, which are very common in construction workers and have caused extremely negative effects on construction workers' health [2,3]. Finally, construction workers' motions, such as the number of production cycles, can also effect the labor productivity [4]. Therefore, it is important to understand construction workers' behaviors for better performance.

Construction workers' posture data provides a foundation for working behavior analysis. For safety management, working postures could help to identify unsafe behaviors and prevent safety accidents [5,6]. For health and sustainability of construction workers, posture data could help to assess the workloads of different working tasks and mitigate the risk of fatigue and injuries [7,8]. For labor productivity, posture data has been used for working/rest status identification for evaluating work efficiency [9]. The studies have demonstrated the importance of workers' posture data in construction management. However, considering the complexity of construction site environments and the dynamics of construction motions, the posture data collection methods used in previous methodologies cannot support effective behavior-based management due to the inaccurate data, uncomfortableness or limitations on indoor working environments.

This study aims to provide a 3D posture data collection method for construction workers, which (1) could provide 3D joint locations, (2) has no limitation on working environments and (3) doesn't require any wearable sensors and thus will not lead to uncomfortableness of workers. The method could provide the data foundation of posture-based behavior analysis and management for individualized unsafe behavior identification, ergonomic assessments and productivity evaluation.

2 Related work

2.1 Previous posture data collection method in construction industry

There are mainly four categories of objects on construction sites, namely human, materials, machines and environment. For materials and machines, tracking technologies such as radio frequency identification and global positioning system have been widely used [10,11]. For environments, laser scanning techniques were used to collect 3D point data for building construction site model [11]. This section mainly focuses on the approaches for collecting the motion data of construction workers.

Manual observation was commonly-used to collect construction workers' joint angle data for ergonomic assessments [12,13]. Manual observation method usually classifies joint angles in to different categories, such as 0-30 degree, 30-60 degree and 60-90 degree. The results of manual observation depend heavily on the observers' experiences and subjective judgement, thus is not accurate enough. In addition, the manual observation cannot collect data continuously from the whole construction site, thus cannot support timely management [14].

Inertial measurement units (IMUs) have been widely used to collect construction workers' posture data. If attached to the key joints, IMU sensors could provide the three-axis rotation angle of each body segments, which could be used to calculate the 3D joint location if given the length of each body segment [15]. Based on the 3D posture data collected with IMU, previous studies have tried to detect construction workers' unsafe behaviour [16], identify awkward postures [15] and estimate productivity [17]. The main disadvantage of IMU, however, is the requirement of attaching IMU sensors to the human body, which may interfere workers' performance. In addition, such sensors may not be suitable for prolonged usage because they may lead to discomfort [14,18].

Motion capture systems (e.g., the VICON system, OptiTrack and Optotrak) are commonly used in laboratories for 3D motion capture and analysis. To capture motion data, an examiner needs to set up multiple cameras in a laboratory and then put reflective markers on the designated locations of an individual's body. The system estimates the 3D position and movement trajectory of each marker based on the signals of the reflective markers captured by the cameras. The reported accuracy of the VICON system is as high as 2 mm [19]. In construction industry, motion capture systems have been applied to collect 3D joint locations for the detailed biomechanical analysis of working postures. However, since these motion capture systems require the installation of at least 4 cameras within 10 m from the attached reflective markers on the body of target workers in order to capture the whole-body posture, it is impractical to use on construction sites [20].

Depth cameras provides a non-invasive method for 3D posture data collection. Depth camera can provide more information than ordinary 2D cameras [21]. There are mainly two kinds of depth cameras, stereo camera and infrared camera. A stereo camera infers the 3D structure of a scene from tow images from different viewpoints. If applied on construction workers, the method could be used to construction the 3D skeletons with 2D images [22]. Infrared cameras could infer the depth of each RGB pixels and provide 3D skeleton based on machine learning networks [5]. However, as for the application on real construction sites, a search of the relevant products (e.g. ZED, Realsense D435 and Kinect) yielded that depth cameras cannot provide accurate 3D joint locations over a long distance. In addition, the infrared cameras cannot provide accurate depth information in outdoor environments due to the interference of sunlight on infrared signals.

RGB cameras are the most common-seen cameras in daily life. Considering the widely use and low cost of RGB cameras, pervious research has tried to identify construction workers posture motions or cameras based on RGB camera [6,23,24]. The methods successfully recognized construction workers from site pictures and classified postures into squatting, standing or walking. However, these methods could only get 2D joint information from the images, which cannot support 3D posture analysis for accurate behaviour recognition or ergonomic analysis.

In summary, above posture data collection methods have the following limitations if applied on construction sites: (1) intrusiveness: sensor-based methods may make the workers feel uncomfortable and even interfere working performance; (2) possible poor performance on construction sites: depth cameras may not provide accurate 3D pose estimation results over long distances in outdoor environments; (3) the lack of 3D results: 3D poses could provide better support for behaviour-based management. Recent progresses in computer vision provide possible solutions for the above limitations. The following is a review on related computer vision algorithms.

2.2 Pose estimation in computer vision

Pose estimation is a classical problem in computer vision. With the development of deep learning, the performance of pose estimation algorithms has been enhanced a lot [25]. The pose estimation algorithms focus on mainly two tasks: 1) 2D pose estimation, which aims to evaluate 2D joint locations from RGB images, 2) 3D pose estimation, which aims at inferring the depth of each joint based on 2D joint locations. The 2D pose estimation algorithms, open pose [26], has been successfully applied on construction sites to estimate 2D construction postures, which worked well even over long distances or when some parts of the body were obstructed [27]. 3D pose estimation algorithms, however, performed not very well when applied in estimating the postures of construction workers. The gaps are 1) previous 3D pose dataset for training the 3D pose estimation algorithms are mainly daily life postures, such as sitting, taking and calling, which differ a lot with the postures of the construction workers, and 2) the structures and parameters of the algorithms are not suitable for estimating the 3D postures of construction workers.

3 Research aim and contribution

Considering above research gaps and limitations, this study tries to develop 3D pose estimation algorithm that is suitable for the postures of construction workers. The method could estimate the 3D joint locations based on RGB images in near real time. This method makes it possible to continuously collect 3D pose data from construction site videos and contributes to 3d-pose-data-based behavior management, such as identifying unsafe behavior postures, estimating joint workloads and assessing labor productivity.

4 Methodology

This study aims to train a 3D pose estimation according to 2D pose with transfer learning [25]. To reach the aim, a 3D database of the postures in construction tasks was firstly built, then a deep learning network was trained based on the dataset.

4.1 Establish the training database

The training dataset includes the 3D joint location data of construction tasks and the corresponding 2D joint locations. A laboratory experiment was performed to establish the dataset with an IMU system (3-Space™ Wireless 2.4GHz DSSS, OH, USA).

4.1.1 Collecting 3D posture data

Participants: A healthy male graduate student, aged 27 years, was recruited to perform a simulated plastering task in a laboratory.

Equipment: The participant was required to wear the IMU system to collect 3D posture data. The IMU sensor has an accuracy of 1° and an frequency of 50 Hz [28]. The IMU system includes 13 IMU sensors. They were tightly tied to the head, chest, back, waist, upper arms, forearms, thighs and shanks.

Simulated plastering task: After putting on the IMU sensors, the participant was instructed to perform a simulated plastering task. The participant mimiced the motion of plastering an area of 5 meters width and 2 meters height. To calibrate the IMU system before the task, the paricipant was requried to stand with both feet closed together and both arms stretched out to the sides and held parallel to the ground to form a T shape.

4.1.2 Data processing

The results of the IMU system were 1397 frames of postures. The data was stored in a BVH file, which includesdthe three-axis rotations of each body segments in each frame. The 3D joint locations were calculated based on the three-axis rotations angles and the length of each body segment with Denavit-Hartenberg matrix [29].

Then the 2D joint locations were calculated based on the 3D joint locations with projection matrix. The generated 2D joint locations are related to the location of the camera. In this study, given the hip joint as the origin, the spherical coordinate of the camera is (0°, 75°, 20 m).

4.1.3 Dataset structure

The dataset includes input dataset and target dataset. The input dataset is a matrix with 1397 rows and 48 columns. Each row stores the 3D Cartesian coordinates of 16 joints (head, neck, chest, waist, trunk, central hip, bilateral shoulders/elbows/wrists/hips/knees/ankles). The target dataset is matrix with 1397 rows and 32 columns. Each row stores the 2D Cartesian coordinates of the 16 joints.

For training and testing the algorithms, the dataset was divided into two parts. 1000 rows were randomly selected from the input dataset and target dataset respectively to form the training dataset. The rest rows of the input dataset and target dataset were used for testing the performance of the algorithm.

4.2 Network architecture

The network is composed of several basic network unit. Each unit includes a linear layer and an RELU layers. The linear layer aims to increase the dimensions of the input data to ensure the depth of the networks. The RELU layer next to the linear layer could add non-linearities to the deep neural networks [30]. The residual connections could improve generalization performance and reduce training time [25]. Figure 1 shows the network architecture.

The complexity of the network is decided on the width of linear layers and the numbers of the basic network units. Complex network could increase the accuracy but is prone to overfitting and computationally expensive. In the experiments, various combinations of the width of linear layers and the numbers of the basic network units were tested to decide the proper complexity.

In addition, batch normalization, dropout and max-norm constraints were applied to prevent overfitting and speed up training. Batch normalization allows us to use larger learning rates to accelerate the learning process [31]. Dropout randomly drops components from a layer of neural network, and thus could prevent overfitting and improve the generalization performance [32]. Max-norm constraints enforce an absolute upper bound on the norm of the weight of every neuron, which helps to prevent overfitting.

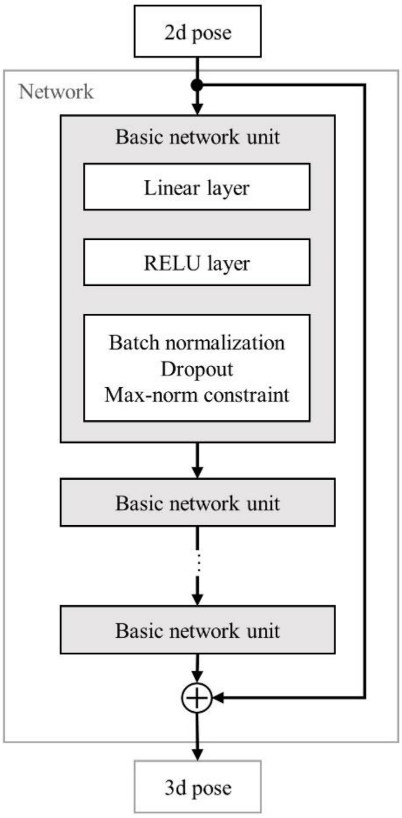

Figure 1. The convergence of training loss (learning rate = 10^{-3}, batch size = 32)

5 Experiment and results

The aim of the experiment is to decide the proper network complexity and the weights of every nodes in the network. In the following training process, the loss function is defined as mean-squared loss and optimized with Adam algorithm [33]. The dropout rate is 0.5. The max-norm constraint is 1.

5.1 Network complexity

As aforementioned, the complexity of the network has a great influence on the network performance. This experiment aims to find the proper network complexity. We change the depth and width of the network through edit the number of basic units (2, 3, 4) and the number of nodes in each linear layer (512, 1024, 2048). As a result, nine different networks were generated. They were trained based on the training dataset, and the accuracy was tested on the testing dataset. In this experiment, the initial learning rate was set as 1e-4, and the batch size was set as 32. Each network was trained for 400 epochs. Table 1 provides the comparison of the performance of the nine networks.

Table 1. The comparison of the performance of different network structures

No. of basic units	No. of layer nodes	Training loss	Test error [cm]	Testing time [ms/frame]
2	512	376.99	24.77	0.09
3	512	78.29	7.76	0.09
4	512	37.67	7.64	0.10
2	1024	29.18	7.16	0.09
3	1024	21.11	5.42	0.09
4	1024	16.09	5.39	0.10
2	2048	13.89	3.46	0.10
3	**2048**	**9.52**	**3.23**	**0.11**
4	2048	8.65	3.42	0.12

Table 1 compares the nine networks according to training loss, testing error and testing time. Training loss represents the final value of loss function. A smaller training loss is preferable. The trained network was then used to estimate the 3D joint locations according to the inputs data in the test database. The estimation results were then compared with the target data. The testing error is defined as the mean of the distances between the estimated 3D location and the target 3D locations of the 16 joints. The last column of Table 1 is the time spent on estimating the 3D joint locations in one frame.

Based on the comparison of the nine networks in Table, it could be found that the increasing the width and depth of the network could significantly decreases the training loss and the testing error. The network with three basic units and 2048 nodes of each layer was selected for the lowest testing error.

5.2 Training the network

Learning rate and batch size decide the step and direction of the training loss decrease, thus are important for the convergence of the loss function. This experiment tried different combinations of learning rate and batch size for training loss convergence and low testing error. Based on the comparison of different network structures in section 5.1, the network with three basic units and 2048-node-layers were used in this experiment. The combination of different learning rates (1e-3, 1e-4, 1e-5) and different batch sizes (2, 4, 8, 16, 32, 64, 128, 256, 512) were tested. Table 2, Table 3 and Table 4 show the comparison of training loss, testing error and testing time for each

Table 2. The comparison of the training loss with different learning rates and batch sizes

BS** \ LR*	10^{-3}	10^{-4}	10^{-5}
2	155.02	2348.89	9583.45
4	50.81	85.68	7113.02
8	19.60	42.38	6772.76
16	7.92	23.43	6494.98
32	**5.44**	13.76	6175.47
64	10.11	9.52	5783.18
128	12.71	8.87	6016.59
256	21.66	9.44	6953.94
512	36.92	116.06	8243.36

*LR represents learning rate.
** BS represents batch size.

Table 3. The comparison of the test error with different learning rates and batch sizes

BS \ LR	10^{-3}	10^{-4}	10^{-5}
2	62.53*	118.41	169.39
4	4.88	11.25	153.75
8	2.51	5.55	148.94
16	1.44	3.99	142.30
32	**1.10**	3.61	134.36
64	2.76	3.23	125.55
128	3.60	2.87	126.82
256	7.60	2.80	136.83
512	24.46	16.61	150.19

*The unit is cm.

Table 4. The comparison of the testing time with different learning rates and batch sizes

BS \ LR	10^{-3}	10^{-4}	10^{-5}
2	0.53*	0.53	0.53
4	0.31	0.31	0.31
8	0.20	0.20	0.20
16	0.14	0.14	0.14
32	0.12	0.11	0.11
64	0.11	0.11	0.11
128	0.10	0.10	0.10
256	0.09	0.09	0.09
512	0.09	0.09	0.09

*The time spent in estimating 3D posture of one frame of 2D posture. The unit is ms.

According to Table 2 and Table 3, both the training loss and testing error reached the minimum when the learning rate was 10^{-3} and the batch size was 32. Table 4 shows that the above combination is also time-saving. In addition, Figure 2 shows the process of training loss convergence under above learn rate and batch size.

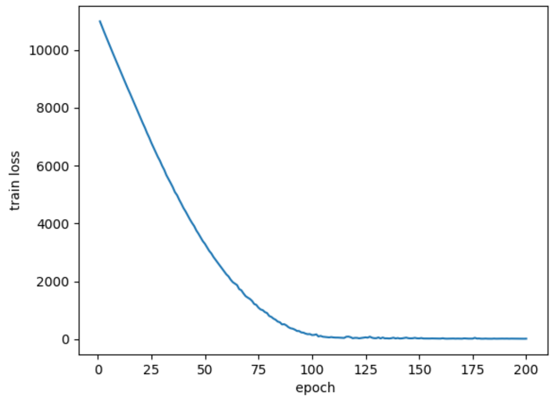

Figure 2. The convergence of training loss (learning rate = 10^{-3}, batch size = 32)

5.3 Testing the results

To this end, the network with 3 basic units, 2048-node-layers, and trained learning rate = 10^{-3}, batch size = 32 was selected. Table 5 shows the mean error of each joints. The mean error of all the joints is 1.10 cm, and the standard deviation is 0.45 cm.

Table 5. The comparison of the performance of different network structures

Joint	Mean error [cm]	Standard deviation [cm]
Waist	0.27	0.28
Right hip	0.46	0.30
Right knee	1.25	1.07
Right ankle	1.51	1.27
Left hip	0.46	0.25
Left knee	1.05	0.99
Left ankle	1.55	1.61
Chest	0.52	0.32
Neck	0.91	0.52
Head	0.94	0.50
Left shoulder	1.08	0.66
Left elbow	1.38	0.78
Left wrist	1.54	0.81
Right shoulder	0.99	0.63
Right elbow	1.60	0.78

Right wrist	2.16	1.08
Mean	1.10	0.45

Figure 3 is the histogram of the mean error of the joints in each frame. The maximum error is about 3.0 cm. Most of the errors are between 0.5 cm and 1.5 cm. The mean error is 1.10 cm, and the standard deviation is 0.45 cm.

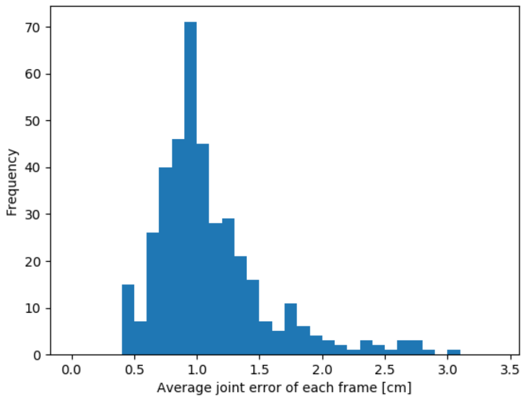

Figure 3. The histogram of the mean error of the joints in each frame

Figure 4 is an intuitive presentation of the estimation results. It could be found that the estimation postures are nearly the same with the ground truth.

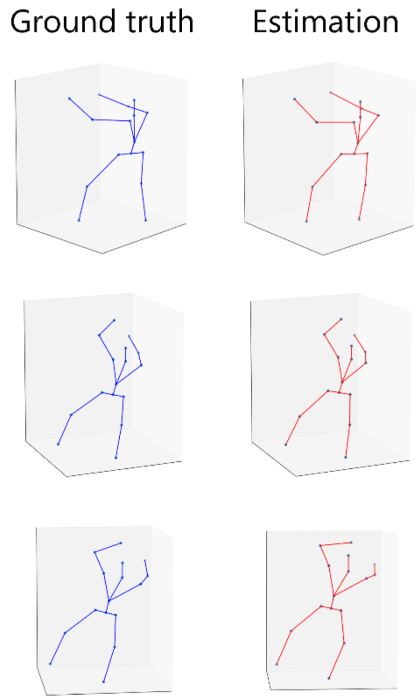

Figure 4. The estimated 3D postures and the ground truth data

6 Discussion

3D postures of construction workers are very important to safety, health and productivity management. This method provides a 3D posture estimator based on workers' 2D postures. The results show that the method could provide accurate 3D posture estimations in nearly real time. The latency time for testing one frame on a GTC 1080Ti GPU was 0.12 ms. The mean error of each joint was 1.10 cm. The accuracy was significantly improved compared with previous computer vision 3D pose estimation methods in construction industry, the mean joint error of which was 3.5 cm [34].

Compared with previous motion capture methods in construction industry, the proposed method was non-intrusive and could work well in outdoor environments. In addition, if combined with 2D posture estimation programs, such as Open Pose [35], the method could make it possible to collect construction workers' 3D postures continuously and timely, providing the data foundation for behavior-based safety, health and production management.

The method has the following limitations and could be improved in the future. First of all, the training dataset is not large enough. The dataset used in this study only includes the postures of one participant during plastering, which may limit the generalization performance of the method. In future, a more diversified training database should be established, which includes the posture data of different construction tasks collected from participants of different heights, weights and BMIs.

Secondly, the 2D postures were generated from 3D posture based on projection matrix. In the current study, only one projection view was used. In the following studies, the 2D postures generated in different views could be used to train the 3D posture estimation method, so that it could be more applicable on construction sites.

Finally, the current study aims to estimate 3D postures from 2D postures. Future studies could try to combine it with 2D posture estimation method from RGB images, so that the 3D postures could be directly inferred from construction site videos or images.

7 Conclusion

Construction workers' posture data provides the foundation for working behavior analysis, such as unsafe behavior identification, ergonomic assessment and labor production evaluation. This paper established a preliminary 3D posture dataset of construction tasks and provided a 3D posture estimation method based on 2D joint locations. The results showed that the method could estimate 3D postures accurately and timely. The mean joint error and estimation time of each frame were 1.10 cm and 0.12 ms respectively. This method makes it possible to estimate construction workers' 3D postures from the images of construction sites and contributes to a data-based construction workers' behavior management.

References

[1] H.W. Heinrich, D. Petersen, N. Roos, Industrial accident prevention: a safety management approach, McGraw-Hill, Tokyo, 19800070280614.

[2] J. Seo, M. Moon, S. Lee, Construction Operation Simulation Reflecting Workers' Muscle Fatigue, in: Computing in Civil Engineering, American Society of Civil Engineers, Reston, VA, 2015: pp. 515–522. doi:10.1061/9780784479247.064.

[3] D. Wang, F. Dai, X. Ning, Risk Assessment of Work-Related Musculoskeletal Disorders in Construction: State-of-the-Art Review, Journal of Construction Engineering and Management. 141 (2015) 04015008. doi:10.1061/(ASCE)CO.1943-7862.0000979.

[4] S.P. Dozzi, S.M. Abourizk, Productivity in Construction, Institute for Research in Construction, National Research Council, Ottawa, ON, Canada, 19930662211340.

[5] Y. Yu, H. Guo, Q. Ding, H. Li, M. Skitmore, An experimental study of real-time identification of construction workers' unsafe behaviors, Automation in Construction. 82 (2017) 193–206. doi:10.1016/j.autcon.2017.05.002.

[6] L. Ding, W. Fang, H. Luo, P.E.D. Love, B. Zhong, X. Ouyang, A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory, Automation in Construction. 86 (2018) 118–124. doi:10.1016/j.autcon.2017.11.002.

[7] Y. Yu, X. Yang, H. Li, X. Luo, H. Guo, Q. Fang, Joint-Level Vision-Based Ergonomic Assessment Tool for Construction Workers, Journal of Construction Engineering and Management. 145 (2019) 04019025. doi:10.1061/(ASCE)CO.1943-7862.0001647.

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

[9] X. Luo, H. Li, X. Yang, Y. Yu, D. Cao, Capturing and Understanding Workers' Activities in Far-Field Surveillance Videos with Deep Action Recognition and Bayesian Nonparametric Learning, Computer-Aided Civil and Infrastructure Engineering. 34 (2019) 333–351. doi:10.1111/mice.12419.

[10] H. Guo, Y. Yu, M. Skitmore, Visualization technology-based construction safety management: A review, Automation in Construction. 73 (2017) 135–144. doi:10.1016/j.autcon.2016.10.004.

[11] S.M.E. Sepasgozar, P. Forsythe, S. Shirowzhan, Evaluation of Terrestrial and Mobile Scanner Technologies for Part-Built Information Modeling, Journal of Construction Engineering and Management. 144 (2018) 04018110. doi:10.1061/(ASCE)CO.1943-7862.0001574.

[12] W.P. Neumann, R.P. Wells, R.W. Norman, J. Frank, H. Shannon, M.S. Kerr, A posture and load sampling approach to determining low-back pain risk in occupational settings, International Journal of Industrial Ergonomics. 27 (2001) 65–77. doi:10.1016/S0169-8141(00)00038-X.

[13] M. Otsuka, T. Nishimura, A. Seo, K. Doi, Relationship between Work Position and Physical Work Load during Insertion of Pin Connectors(Theory and Methodology), Journal of Japan Industrial Management Association. 61 (2010) 275–283.

https://ci.nii.ac.jp/naid/10027761503/en/.

[14] E. Valero, A. Sivanathan, F. Bosché, M. Abdel-Wahab, Musculoskeletal disorders in construction: A review and a novel system for activity tracking with body area network, Applied Ergonomics. 54 (2016) 120–130. doi:https://doi.org/10.1016/j.apergo.2015.11.020.

[15] J. Chen, J. Qiu, C. Ahn, Construction worker's awkward posture recognition through supervised motion tensor decomposition, Automation in Construction. 77 (2017) 67–81. doi:10.1016/j.autcon.2017.01.020.

[16] H. Jebelli, C.R. Ahn, T.L. Stentz, Fall risk analysis of construction workers using inertial measurement units: Validating the usefulness of the postural stability metrics in construction, Safety Science. 84 (2016) 161–170. doi:10.1016/j.ssci.2015.12.012.

[17] A. Alwasel, M. Nahangi, C. Haas, E. Abdel-Rahman, Level-of-Expertise Classification for Identifying Safe and Productive Masons, in: Computing in Civil Engineering 2017, 2017. doi:10.1061/9780784480823.043.

[18] A. Golabchi, S. Han, J. Seo, S. Han, S. Lee, M. Al-Hussein, An Automated Biomechanical Simulation Approach to Ergonomic Job Analysis for Workplace Design, Journal of Construction Engineering and Management. 141 (2015) 04015020. doi:10.1061/(ASCE)CO.1943-7862.0000998.

[19] VICON Motion Capture System, (2018). https://www.vicon.com (accessed April 18, 2018).

[20] J. Hicks, Preparing your data for OpeinSim 3.3, National Center for Simulation in Rehabilitation Research (NCSRR), US. (2018). https://simtk-confluence.stanford.edu:8443/display/OpenSim33/Preparing+Your+Data (accessed December 24, 2018).

[21] C. Liu, S. Shirowzhan, S.M.E. Sepasgozar, A. Kaboli, Evaluation of Classical Operators and Fuzzy Logic Algorithms for Edge Detection of Panels at Exterior Cladding of Buildings, Buildings. 9 (2019) 40. doi:10.3390/buildings9020040.

[22] M. Liu, S. Han, S. Lee, Tracking-based 3D human skeleton extraction from stereo video camera toward an on-site safety and ergonomic analysis, Construction Innovation. 16 (2016) 348–367. doi:10.1108/CI-10-2015-0054.

[23] X. Luo, H. Li, D. Cao, Y. Yu, X. Yang, T. Huang, Towards efficient and objective work sampling: Recognizing workers' activities in site surveillance videos with two-stream convolutional networks, Automation in Construction. 94 (2018) 360–370. doi:10.1016/j.autcon.2018.07.011.

[24] H. Zhang, X. Yan, H. Li, Ergonomic posture recognition using 3D view-invariant features from single ordinary camera, Automation in Construction. 94 (2018) 1–10. doi:10.1016/j.autcon.2018.05.033.

[25] J. Martinez, R. Hossain, J. Romero, J.J. Little, A simple yet effective baseline for 3d human pose estimation, ArXiv Preprint ArXiv:1705.03098. (2017).

[26] Z. Cao, T. Simon, S.-E. Wei, Y. Sheikh, Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields, CVPR2017. (2016). http://arxiv.org/abs/1611.08050.

[27] X. Yan, H. Li, C. Wang, J. Seo, H. Zhang, 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. 34 (2017) 152–163. doi:10.1016/j.aei.2017.11.001.

[28] Yost Labs, 3-Space™ Sensors, Yost Labs. (2017). https://yostlabs.com/3-space-sensors/ (accessed April 19, 2018).

[29] G. Legnani, F. Casolo, P. Righettini, B. Zappa, A homogeneous matrix approach to 3D kinematics and dynamics — I. Theory, Mechanism and Machine Theory. 31 (1996) 573–587. doi:10.1016/0094-114X(95)00100-D.

[30] V. Nair, G.E. Hinton, Rectified Linear Units Improve Restricted Boltzmann Machines, in: Proceedings of the 27th International Conference on International Conference on Machine Learning, Omnipress, USA, 2010: pp. 807–814. http://dl.acm.org/citation.cfm?id=3104322.3104425.

[31] S. Ioffe, C. Szegedy, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, (2015). http://arxiv.org/abs/1502.03167.

[32] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: A Simple Way to Prevent Neural Networks from Overfitting, J. Mach. Learn. Res. 15 (2014) 1929–1958. http://dl.acm.org/citation.cfm?id=2627435.2670313.

[33] D.P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization, (2014). http://arxiv.org/abs/1412.6980.

[34] Y. Yu, H. Li, X. Yang, W. Umer, Estimating Construction Workers' Physical Workload by Fusing Computer Vision and Smart Insole Technologies, in: 2018 Proceedings of the 35th ISARC, International Association for Automation and Robotics in Construction, Berlin, Germany, 2018: pp. 1212–1219. doi:10.22260/ISARC2018/0168.

[35] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, Y. Sheikh, OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields, (2018). http://arxiv.org/abs/1812.08008.