International Journal on Advances in Life Sciences, vol 13 no 1 & 2, year 2021, http://www.iariajournals.org/life_sciences/

Evaluations and Applications of Partial Body Joint Model in 3D Human Pose Estimation from Single Image

Yuta Ono, Oky Dicky Ardiansyah Prima, Kazuki Hosogoe
Graduate School of Software and Information Science, Iwate Prefectural University
152-52 Sugo Takizawa, Japan
e-mail: g236s001@s.iwate-pu.ac.jp, prima@iwate-pu.ac.jp, g231r026@s.iwate-pu.ac.jp

***Abstract*—Human pose estimation has been used to perform human motion analysis in widespread applications. Three-dimensional (3D) human pose estimation from single image has attracted much attention because of its ease of measurement. Methods of this approach have become more accurate with the introduction of deep neural networks. Most of these methods are trained to estimate the body joints of the whole human body. However, when a part of the body joints is obscured by the presence of other objects or the camera position and angle, the estimation accuracy of the overall body joints may be degraded. In this study, we attempt to experimentally construct a 3D human pose estimation model for partial body joints to accurately estimate the pose of a partially human body that can be visibly measured. To evaluate the performance of the proposed model, we construct a neural network model that estimates only the 3D position of the visible upper body joints, assuming that only those joints are visible. Our evaluations showed that the partial body joint model was more accurate in estimating the posture from frontal human images. However, there was no significant difference in the accuracy between our model and previous pose estimation model when the posture was estimated from images of people taken from extreme angles. Finally, we attempt to extend our model to a system for detecting deterioration in sitting postures to verify the effectiveness of the model.**

Keywords-3D human pose estimation; partial body joint; RGB-D camera; computer vision; sitting posture.

I. INTRODUCTION

Human pose estimation is a task that uses computer vision technology to estimate the location of body joints of a human body from a given image. Over the past five years, research in human pose estimation has shifted from 2D to 3D. 3D human pose estimation has attracted a significant interest from the scientific community. The technique has made significant progress by introducing methods based on deep learning. However, it still has some problems, such as depth ambiguities and lack of in-the-wild datasets. This study extends our previous work on the influence of occlusion on 3D human pose estimation [1].

3D human pose estimation has been applied in a wide range of fields including human-computer interaction, games, sports performance analyses, and other motion analyses [2][3]. Liu et al. (2013) extracted skeletal information of seven basic human actions using Microsoft Kinect v1 and classified them using k-means clustering and Hidden Markov Models (HMMs) [4]. Ono et al. (2021) used Microsoft Azure Kinect

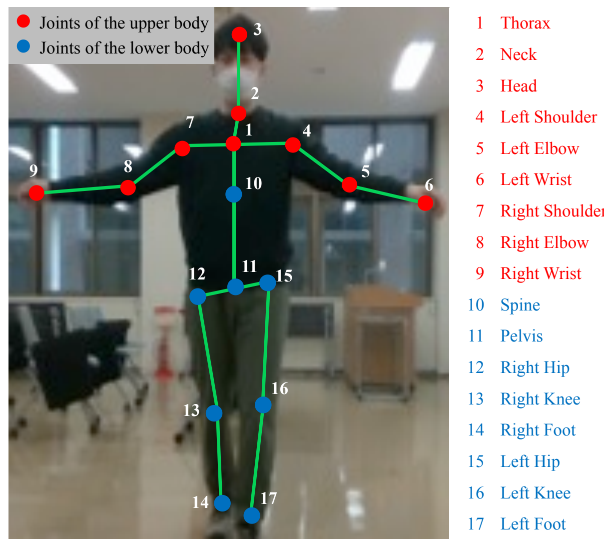


Figure 1. Body joints used in this study. 9 joints were used for the upper-body model, and 17 joints for the whole-body model.

[5] to measure the hand movements of pharmacists to monitor drug picking operations in pharmacies [6]. Their method uses finger landmark detection with MediaPipe [7] for more accurate detection of the picking tasks. 3D human pose estimation has also been used for a motion analysis to evaluate the effectiveness of rehabilitation. Prima et al. (2019) demonstrated the usability of 3D human pose estimation using a vision camera for measuring the range of motion of joints to promote self-rehabilitation by patients [8]. Their experiments show that the resulting 3D human pose estimated from a single image is more advantageous for estimating semi-occluded body joint locations than those estimated by a depth sensor.

3D human pose estimation methods can be broadly classified into two categories: a method using multiple cameras and a method using a single camera. Methods using multiple cameras are advantageous for depth measurement and occlusion avoidance. Ziegler et al. (2006) proposed a method to track an articulated upper body using four stereo cameras [9]. A point cloud of the synthesized body model was fitted to the measured 3D data using an iterative closest point (ICP) registration algorithm. Nakano et al. (2020) estimated the 2D human pose from images by OpenPose library [10] and estimated the 3D human pose based on triangulation of these 2D body joints [11]. However, these methods suffered from the difficulty of camera calibration. In contrast, 3D human pose estimation from a single camera is based on an estimation

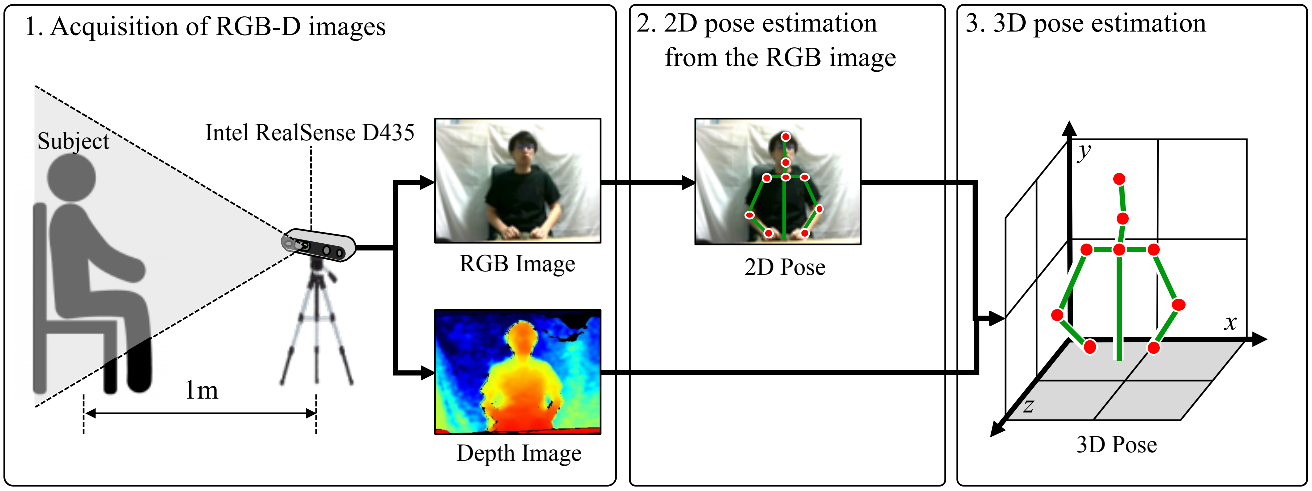


Figure 2. 2D and 3D human pose estimation using an Intel RealSense D435 to build the sitting posture dataset in this study.

method that uses 3D human pose datasets. Chen and Ramanan (2017) proposed a method to generate 3D human poses using a 3D human pose library consisting of pairs of 2D and 3D human poses [12]. Their results suggest that such a simple baseline should be used as a benchmark for future work in 3D human pose estimation. Martinez et al. (2017) constructed a relatively simple deep neural network that converts 2D human pose data to 3D human pose data [13]. Moon et al. (2019) used the correlation between 2D and 3D human poses to estimate the position and posture of the human body in real environment [14]. However, since the 3D human pose estimation from a single camera assumes that the whole-body joints are completely visible in the input image, if any part of the body joints is hidden, the estimation accuracy for the whole-body joints may be degraded.

Methods for 3D human pose estimation considering occlusion have been proposed. Vosoughi et al. (2018) proposed a deep Convolutional Neural Network (CNN) that regresses 3D human pose from an RGB image and a CNN that detects the presence of human body joints from an RGB image [15]. Sárándi et al. (2018) evaluated the robustness for occlusion in 3D human pose estimation using Human3.6M dataset [16] with synthetic occlusions [17]. However, these methods have been evaluated using only existing datasets, and have not been tested for the accuracy of 3D human pose estimation in the case where occlusion is caused in real world. In addition, these methods have not been tested for the case where self-occlusion occurs.

This study attempts to experimentally construct a partial 3D human pose estimation model to accurately estimate only the visible body joints of the human body even if some joints of the human body are hidden by other objects or self-occlusion. To verify the effectiveness of the model, we constructed a model that estimates only the body joints of the upper body and verified the estimation accuracy of the model using a modified Human3.6M dataset and two datasets originally created using RGB-D cameras. In addition, we developed a system for detecting deterioration in sitting posture using the model and verify the usability of the system.

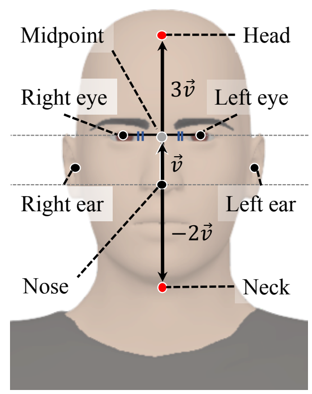


Figure 3. Calculation of Head and Neck points (red dots) using OpenPose face landmarks (black dots) in the sitting posture datasets.

This paper is organized as follows. Section II describes the methodology of constructing the 3D human pose estimation model for partial body joints and three human pose datasets to verify the estimation accuracy of the model. In Section III, we present our evaluation results evaluated using these datasets. Section IV describes how to construct a posture deterioration detection system using the model and verify the usability of the system. Finally, Section V summarizes the results of this study.

II. Methodology

In this study, we attempt to experimentally construct a partial 3D human pose estimation model to accurately estimate only visible body joints of the human body, even if some of them are hidden. Our model is constructed by improving the existing 3D human posture estimation model [13]. Figure 1 shows the nine upper body joints of the human body to be estimated by the partial 3D human pose estimation model in this study.

To evaluate the performance of our model, we compare its estimation accuracy with that of a whole-body joints model

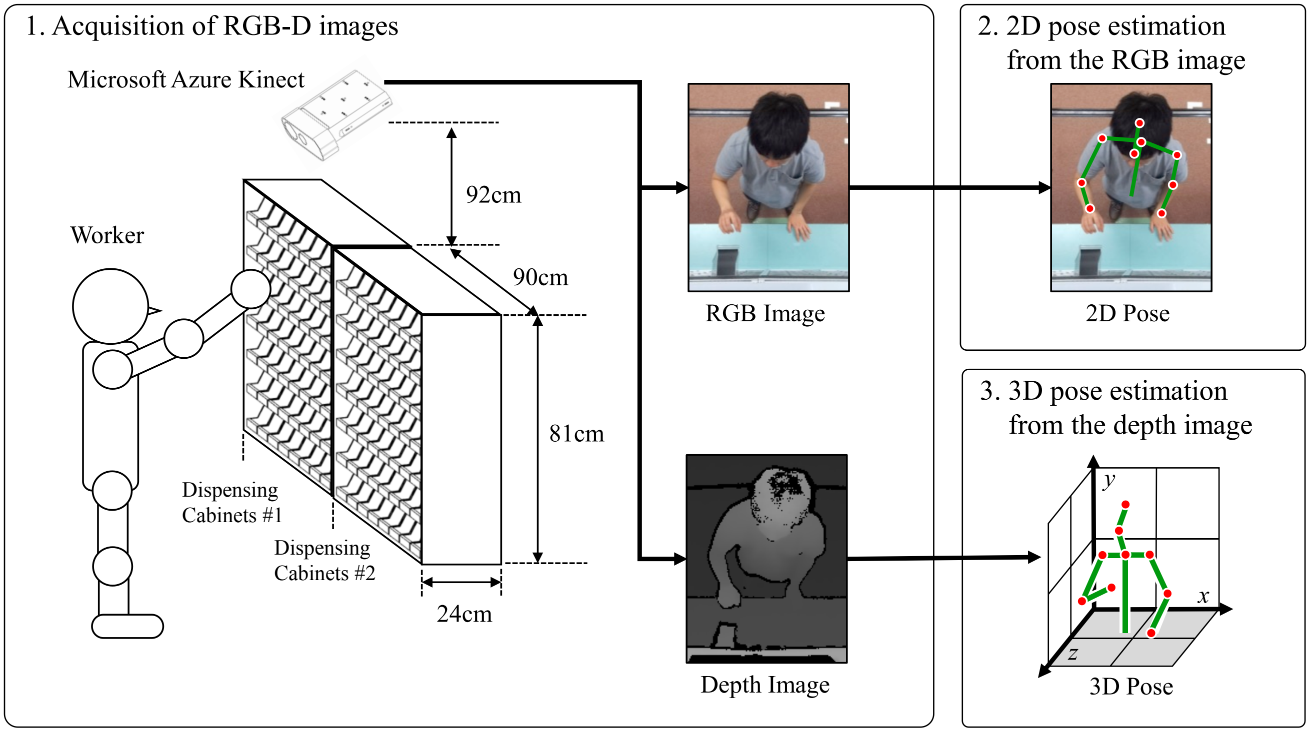


Figure 4. 2D and 3D human pose estimations using a Microsoft Azure Kinect to build the high-angle captured standing posture dataset in this study.

[13] using three different 3D human pose datasets. For convenience, the model of the body joints of the whole body is referred to as the "whole-body model" and the model of only the body joints of the upper body as the "upper-body model". The evaluation involves the following procedures. First, we simulate a scene in which only the upper body is visible using the Human3.6M dataset to evaluate the estimation accuracy of each model. Second, in order to evaluate the estimation accuracy of the 3D human pose of the upper body in the real world, two 3D human pose datasets, such as sitting postures and high-angle captured standing postures are constructed independently using RGB-D cameras.

A. Building a Partial 3D Human Pose Estimation Model

To train the upper-body model, we use the Human3.6M dataset. The dataset consists of 3.6 million 3D human poses and their corresponding images measured by 11 professional actors (5 females and 6 males) in 17 different daily activities. We adopted the method of Martinez et al. [13] to build the upper-body model. This method uses a relatively simple deep feed-forward neural network to efficiently estimate the 3D human pose of the whole body. Following Martinez et al., the model is trained from the measurement data of subjects 1, 5, 6, 7 and 8, and validated from the measurement data of subjects 9 and 11. The neural network receives the 2D human pose as input, and the input data is extended to 1024 dimensions by the fully connected layers. The weights of all the layers are initialized using the method of He et al. (2015) [18]. After that, the network performs batch normalization, Rectified Linear Unit (ReLU), dropout rate of 0.5, and a residual connection. The model is trained for 200 epochs with a learning rate of 0.001 and a batch size of 64.

B. Building Human Pose Datasets Using RGB-D cameras

We create two 3D human pose datasets, as ground truth data, using RGB-D cameras in real-world situations where only the upper body is visible due to the presence of other objects or due to the extreme capture angle and evaluate the performance of models using these datasets. RGB images and depth data of the participants in their postures are recorded.

1. Sitting Posture Dataset

An Intel RealSense D435 [19] was used to measure the sitting posture. The resolution of the RGB-D camera is 640×480px and the number of Frame Per Second (FPS) is 30. Figure 2 shows the procedure for generating human pose data in the sitting posture dataset. Participants were asked to sit 1m from the device and move their hands and bodies.

The procedure for generating 3D human pose data in a sitting posture is shown as follows. First, using the Intel RealSense D435 SDK, we calculated the depth value corresponding to each pixel of the measured RGB image. Then OpenPose library was used to estimate the 2D body joints from RGB images as shown in Figure 2. We also calculated head and neck points using OpenPose facial landmarks. Figure 3 shows calculations of these points. First, we calculated the 2D vector from nose to the midpoint of both eyes. Then the head point was obtained by extending this vector upward by three times its length. Similarly, the neck point was obtained by extending this vector downward by

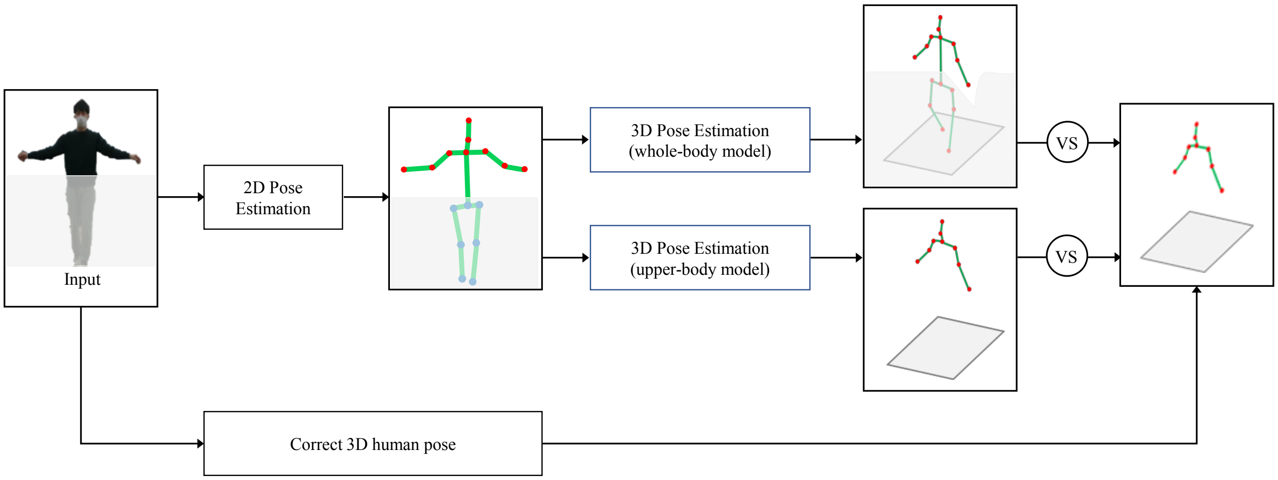


Figure 5. The process of evaluating the estimation accuracy of the upper-body model and the whole-body model in this study.

TABLE I. THE ESTIMATION ACCURACY OF THE UPPER-BODY MODEL AND THE WHOLE-BODY MODEL FOR THREE HUMAN POSE DATASETS.

No.	Joints	Human3.6M [cm]		Sitting Posture [cm]		High-Angle Captured Standing Posture [cm]	
		Whole-Body	Upper-Body	Whole-Body	Upper-Body	Whole-Body	Upper-Body
1	Thorax	16.3	5.0	11.4	7.5	8.8	9.6
2	Neck	3.8	2.7	18.2	14.5	16.3	11.2
3	Head	8.1	4.6	9.7	12.2	19.6	12.4
4	Left Shoulder	8.1	3.7	10.4	9.1	6.9	10.5
5	Left Elbow	9.1	4.5	10.3	9.2	14.2	11.8
6	Left Wrist	10.9	7.0	14.9	10.8	14.2	14.1
7	Right Shoulder	7.3	3.5	10.8	11.3	4.7	8.1
8	Right Elbow	9.0	4.7	10.3	10.9	12.3	11.0
9	Right Wrist	10.6	7.0	15.8	11.1	17.7	19.0
Mean (M)		9.24	4.74	12.42	10.73	12.74	11.97
Standard Deviation (SD)		3.363	1.464	3.064	2.009	5.038	3.132

twice its length. The corresponding 3D body joints might not be measured accurately due to missing depth values. In such a case, the Hampel filter [20] was applied to the time-series values to interpolate the concerned joints by computing the Median Absolute Deviation (MAD) over a specified range of analysis targets. For this study, the window size was set to 30. These 3D body joints are treated as ground truth data.

2. High-Angle Captured Standing Posture Dataset

To evaluate the estimation accuracy of the upper body model when the human body is measured from extreme angles, the motion of a worker picking drugs was captured from the ceiling. A Microsoft Azure Kinect was used to capture the motion of a worker. 2D body joints are estimated by OpenPose library from RGB images. In the standing posture dataset, the head point was obtained as the midpoint of both ears. The neck point was the midpoint of nose and neck estimated by OpenPose library. 3D body joints were measured using the Software Development Kit (SDK) of the Azure Kinect Body Tracking. The head point was derived from the midpoint of both ears and the neck point was derived from the midpoint of the nose and neck measured by the Azure Kinect SDK.

Figure 4 shows the procedure for generating human pose data in the standing posture dataset. Two dispensing cabinets were used to simulate the drug-picking environment. These cabinets can hold 63 shelves (7 rows by 9 columns). These cabinets were placed side-by-side at a height of 85.5 cm from the floor, and the Microsoft Azure Kinect was placed 92 cm above the dispensing cabinet. The resolution of the color camera is 1920×1080px and the angle of view is 90°×59°. The depth sensor has a resolution of 512×512px and the angle of view is 120°×120°.

The procedure for generating the dataset of the standing human posture is as follows. First, the subject stands in front of two dispensing cabinets and manipulates the shelves

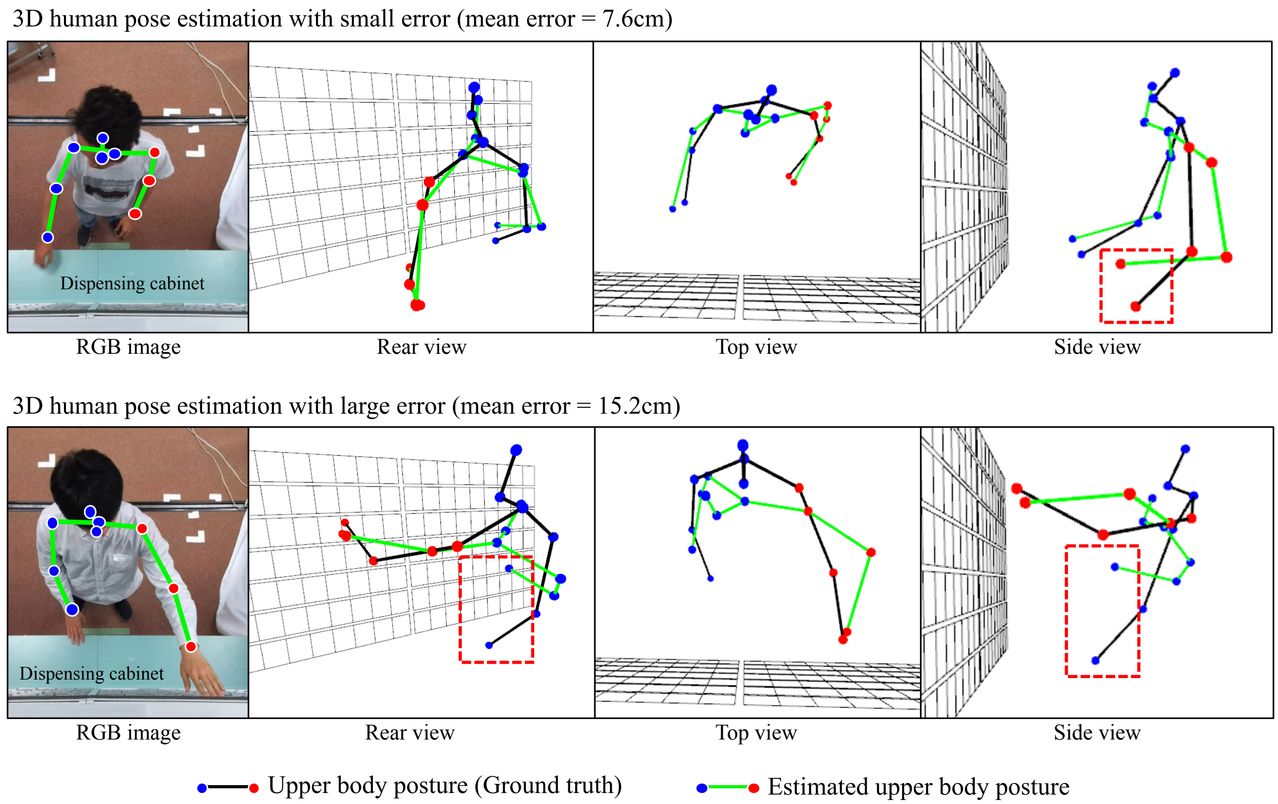


Figure 6. Examples of a 3D human pose estimated by the upper-body model on the standing posture dataset.

indicated by the experimenter. This operation is a series of movements from pulling out the target shelf to putting it back. The subject should operate the shelves on the left side with the left hand and the shelves on the right side with the right hand. After this series of operations, the subject returns to the initial standing position. The subject repeats the above procedure until all the shelves have been manipulated.

C. Evaluation

Using the Human3.6M, the sitting posture dataset, and the standing posture dataset, we evaluate the estimation accuracy of the upper-body model against the whole-body model. Figure 5 shows the process of evaluating the estimation accuracy of these models in this study. 3D human pose of the upper body is estimated using the 2D body joints from each dataset. For the upper-body model, all these body joints of the upper body were used, but for the whole-body model, the coordinates of the body joints corresponding to the lower body were set to (0,0). Finally, the error between the 3D human pose estimated by each model and the corresponding ground truth was calculated.

In this study, Procrustes Analysis, a shape-preserving Euclidean transform, is used to evaluate the differences in the pose data. This analysis eliminates the variation in movement, rotation, and scaling between the pose data while preserving the shape.

III. Result

We evaluated the difference in estimation accuracy between the whole-body model and the upper-body model using the 3D human pose datasets. For this evaluation, we randomly selected 548,800 human poses from the Human3.6M datasets, which were not used for training and validation of these models. For the sitting posture datasets, we generated 7,350 3D human pose from a subject using the Intel RealSense D435. For the standing posture dataset, we randomly selected 10,000 3D human pose from five subject's picking movement data.

Table 1 shows the estimation accuracy of the upper-body model and the whole-body model for the Human3.6M dataset, the sitting posture dataset, and the standing posture dataset. In the Human3.6M dataset, the error from the upper-body model was significantly smaller than that from the whole-body model ($M = 9.24$, $SD = 3.363$, $t(8) = 4.91$, $p < 0.001$), resulting in an improvement of about 4.5cm in estimation accuracy. Likewise, for the sitting posture dataset, the upper-body model improved the estimation accuracy by about 2 cm compared to the whole-body model ($M = 12.4$, $SD = 3.064$, $t(8) = 1.98$, $p < 0.05$). However, the error of the upper-body model was larger than that of the whole-body model at some body joints, such as head and right shoulder. The reason for this may be due to the accuracy of the constructed ground truth data. Finally, for

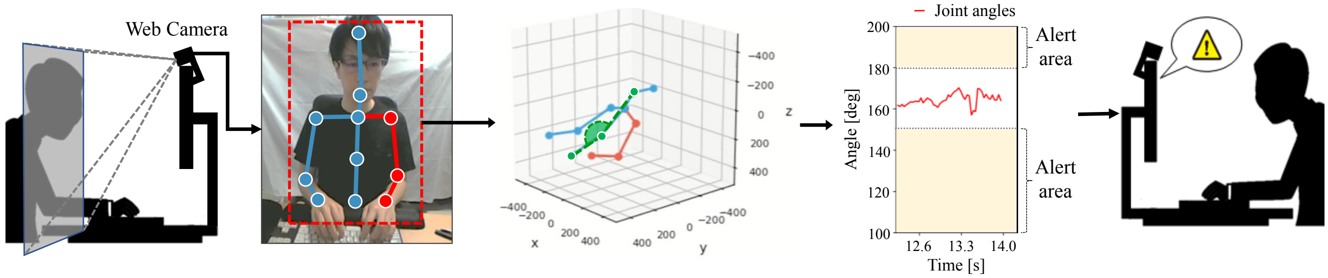


Figure 7. The illustration of the posture deterioration detection system using the upper-body model.

the standing posture dataset, the upper-body model improved the estimation accuracy by 0.8cm compared to the whole-body model, but the difference was not significant ($M = 12.7$, $SD = 5.038$, $t(8) = 0.63$, $p > 0.05$).

Figure 6 shows examples of a 3D human pose estimated by the upper-body model on the standing posture dataset. The top row shows an example where the estimation error is small for all joints of the upper body. The average error for the body joints of the upper body is 7.6 cm. However, the error is particularly large for the wrist joint. This reason for this error may be that the upper-body model is not sufficiently trained to estimate the body joints of the human body measured by the high-angle camera. The bottom row shows an example where the estimation error is large for all joints of the upper body. In this example, because the body joints estimated by the upper-body model differed significantly from the ground truth, the errors are particularly large for the left elbow and the right wrist joints. The dotted box in the figure indicates the body joint with large estimation error in each example.

IV. Detection of Deterioration in Sitting Postures

Workers tend to unconsciously hunch over while seated at a desk or other workstation. Such a posture may cause physical health problems because it puts a burden on the shoulders and hips. Techniques using pressure sensors [21] [22] and IMU sensors [23] have been proposed to detect such an improper posture. However, these techniques require these sensors to be attached to a human body or a chair.

In this study, we attempt to extend an upper-body model to a posture deterioration detection system to solve these problems. This system estimates the 3D human pose of a person using an upper-body model from RGB images measured by a web camera. The system then detects the deterioration of posture based on the body angle calculated from the estimated 3D human pose. In addition to the body joints of the upper body shown in Figure 1, the upper-body model estimates the Spine and the Pelvis. Our system detects deterioration in sitting posture without contact and encourages improvement in posture.

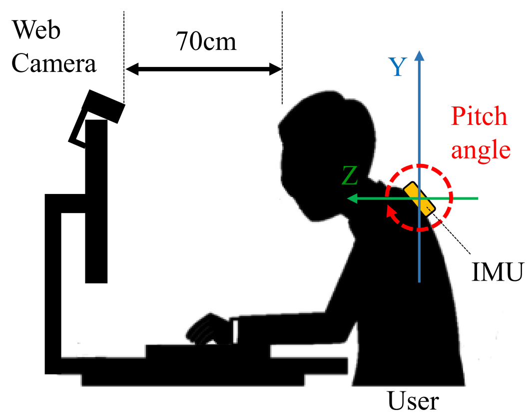
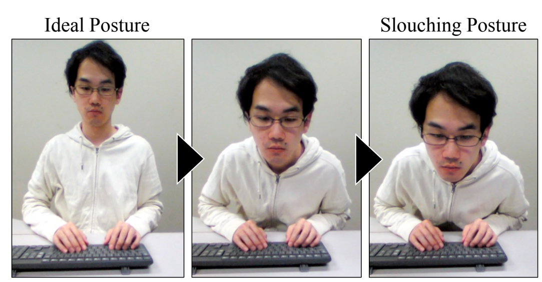
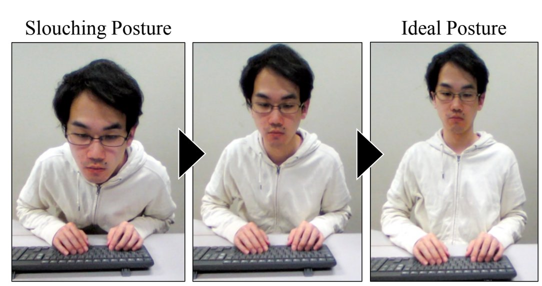


Figure 8. The placement of the web camera and the IMU.

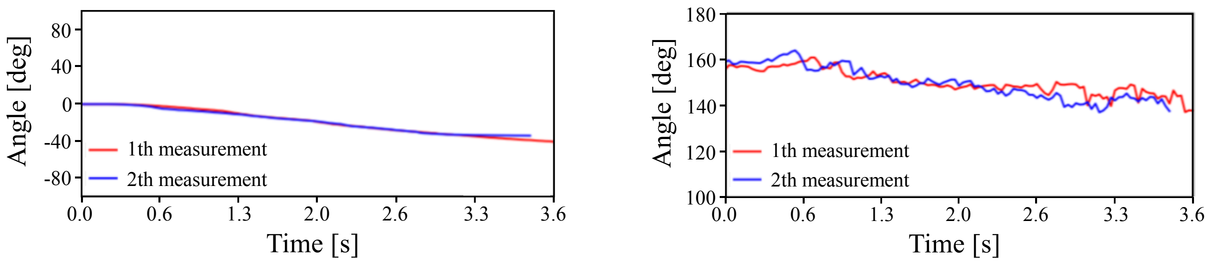


(a) Forward leaning movement



(b) Raising movement

Figure 9. Two motions measured for evaluating the posture deterioration detection system in this study.



(a-1) The IMU measurement

(a-2) The proposed system

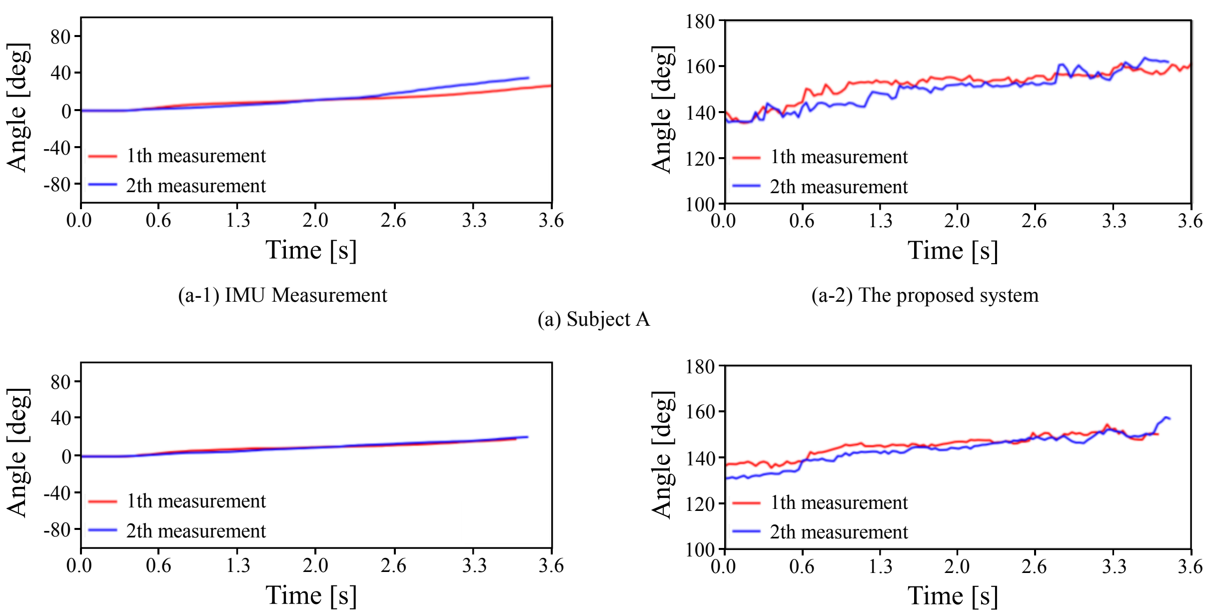
(a) Subject A

(b-1) The IMU measurement

(b-2) The proposed system

(b) Subject B

Figure 10. The changes in the body angle of two subjects measured by the proposed system and the IMU during forward leaning movement.



(a-1) IMU Measurement

(a-2) The proposed system

(a) Subject A

(b-1) IMU Measurement

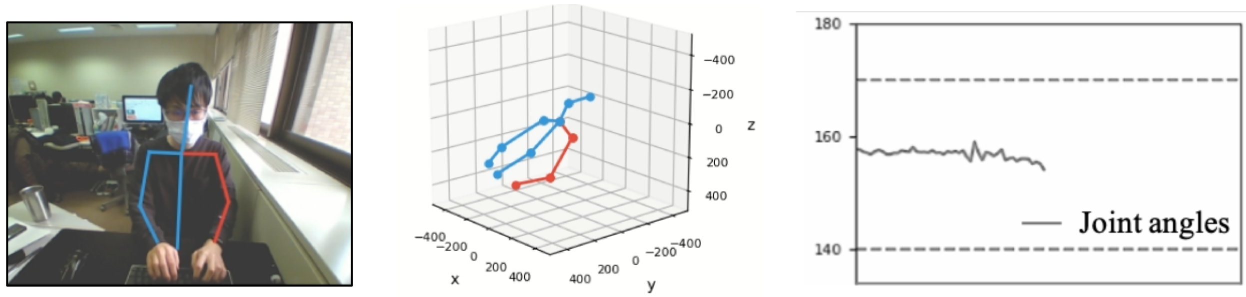
(b-2) The proposed system

(b) Subject B

Figure 11. The changes in the body angle of two subjects measured by the proposed system and the IMU during raising movement.

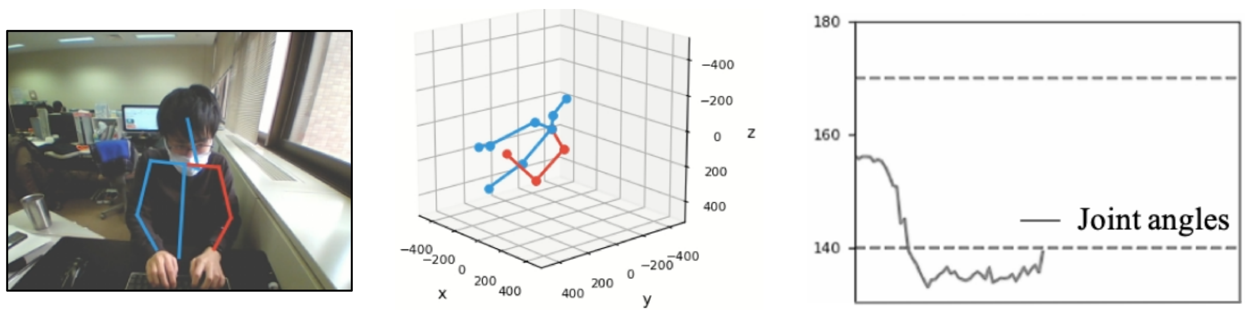
A. Posture Deterioration Detection System

Figure 7 shows a schematic diagram of the posture deterioration detection system using the upper-body model. First, we use a web camera to capture the frontal view of a worker at a desk. Next, the 2D body joints of the worker is estimated from the RGB images using the OpenPose library. Then, we estimate the 3D human pose of the upper-body from these body joints using the upper-body model. Finally, we detect the deterioration of the posture based on the calculation of the body angle calculated using the estimated 3D body joints.



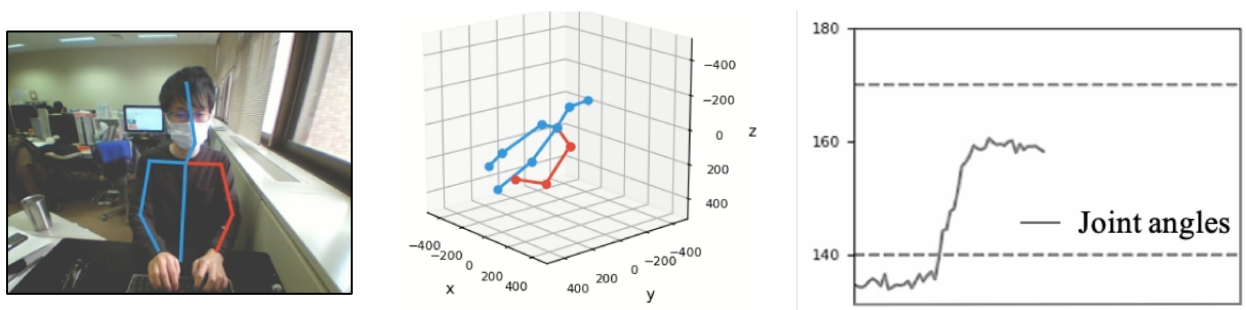
(a-1) The 2D pose (a-2) The predicted 3D pose (a-3) The body angle

(a) Ideal Posture



(b-1) The 2D pose (b-2) The predicted 3D pose (b-3) The body angle

(b) Forward Leaning Movement



(c-1) The 2D pose (c-2) The predicted 3D pose (c-3) The body angle

(c) Raising Movement

Figure 12. The working example of our system.

Our system focuses on the changes in the body angles formed by neck, spine, and pelvis to detect the deterioration of the posture. In this system, the user decides the upper and lower limits of the body angle to be used for judging the deterioration of posture. When the body angle measured by the system exceeds the threshold value, the system detects the posture as a deteriorated posture.

B. *Evaluation of Posture Deterioration Detection System*

To evaluate the usability of the posture deterioration detection system, we evaluate the system from two perspectives: whether the system can measure the change in body angle appropriately and whether the system can properly detect a deteriorated posture based on the threshold value set from the measured change in body angle.

For this evaluation, we assume that the system is used during desk work and detects posture deterioration in the sitting posture. We use a 57cm×43.5cm display and place it on a desk. The web camera is Buffalo BSW200MBK. This camera is placed at the center of the upper part of the display to measure the subject from the front. The angle of view of the camera is 120°×67°, the resolution is 640×480px, and the frame rate is set to 30FPS.

To evaluate the angle of the subject's body, an Inertial Measurement Unit (IMU) is used as reference data. We use the Adafruit BNO05 (100Hz) IMU, which is attached to the back of the subject's neck. Figure 8 shows the placement of the web camera and the IMU. Five subjects participated in this evaluation.

The procedure for evaluating the posture deterioration detection system is as follows. Figure 9 shows the two motions measured in this evaluation. First, the subject is seated 70cm from the web camera. At this point, the subject is instructed to put his hands on the desk, raise his head, and straighten his back. We refer to this posture as the ideal posture. Next, the subject takes 5 seconds from the ideal posture to tilt the upper body forward at a certain speed. After that, the subject returns to the ideal posture at a certain speed for 5 seconds. This series of motions is performed twice. In this evaluation, we refer to the movement from the ideal posture to the forward leaning posture as the forward leaning movement. In addition, we refer the movement from the forward leaning posture to the ideal posture as the rising movement. The body angle is measured by the web camera and the IMU. Finally, we calculate the correlation between the changes in the body angles measured by the web camera and the IMU.

The correlation coefficients of the body angle changes measured by the web camera and the IMU were 0.933 on average. Figures 10 and 11 show the changes in the body angle measured by the web camera and the IMU during forward leaning and rising movements. Figure 12 shows the working example of our system. As shown in Figure 12(b), when the worker's head moves forward and the posture becomes hunched, the body angle changes significantly. The dotted line in the figure indicates an empirical threshold that indicates the acceptable range of an appropriate sitting posture. When the seating posture deteriorates, the measured angle does not fall within this range.

V. Conclusion and Future Work

The purpose of this study is to accurately estimate the body joints of the human body that can be measured if a part of body joints is hidden due to the presence of other objects or the position and angle of the vision camera. For this purpose, we developed a partial 3D human pose estimation model that estimates the upper body joints, assuming that only the upper body is visible. We evaluated the performance of the model using three different 3D human pose datasets to examine the estimation accuracy of the model. And we also attempted to extend the model to a system for detecting deterioration in sitting posture.

To verify the estimation accuracy of the model, we evaluated the estimation accuracy of the model against the whole-body model using three different 3D human pose datasets. At first, we conducted an evaluation using the Human3.6M dataset. Next, we constructed two datasets independently using RGB-D cameras and evaluated the estimation accuracy of both models to confirm the estimation results of 3D human pose with occlusion in the real world. The first dataset is a sitting posture dataset, which was constructed to evaluate the pose of the human body when occlusion by other objects occurs. The second dataset is a high-angle captured standing posture dataset. This dataset was constructed to evaluate the estimation accuracy of the pose of a human body measured by a camera from extreme angles. The results show that the proposed model can estimate the 3D human pose more accurately from frontal human images than the whole-body model. However, when the human body was measured from extreme angles, there was no significant difference in the estimation accuracy between the two models. The reason for this is that the model did not learn sufficiently about the pose of the 3D human body from extreme angles.

We attempted to extend the partial 3D human pose estimation model to a posture deterioration detection system for sitting posture in order to promote the improvement of the sitting posture. In this system, the posture of the person is measured using a web camera, and the pose of the upper body is estimated by the model. The system then calculates the body angle to detect the deterioration of the posture. We confirmed that the system can measure changes in the body angle and is effective in detecting posture deterioration in a sitting posture. In addition, the system has the advantage of being able to detect posture deterioration without contact using only a widely used web camera.

In the future, we will construct other partial body joint models, such as the left half of the body and the lower half of the body and verify the estimation accuracy for each model. In addition, we will study the improvement of the model to increase the accuracy of human pose estimation from extreme measurement angles. Furthermore, we will improve the reliability of posture deterioration detection to build a more practical posture deterioration detection system.

References

[1] O.D.A. Prima and K. Hosogoe, “3D Human Pose Estimation of a Partial Body from a Single Image and Its Application in the Detection of Deterioration in Sitting Postures,” The Thirteenth International Conference on eHealth, Telemedicine, and Social Medicine, eTELEMED2021, pp. 1-5, 2021.

[2] N. Sarafianos, B. Boteanu, B. Ionescu, and I. A. Kakadiaris, “3D Human Pose Estimation: A Review of the Literature and Analysis of Covariates,” Computer Vision and Image Understanding, vol. 152, pp. 1-20, Nov. 2016, doi: 10.1016/j.cviu.2016.09.002.

[3] J. Wang et al., “Deep 3D human pose estimation: A review,” Computer Vision and Image Understanding, vol. 210, pp. 1-21, Sept. 2021, doi: 10.1016/j.cviu.2021.103225.

[4] T. Liu, Y. Song, Y. Gu, and A. Li, “Human Action Recognition Based on Depth Images from Microsoft Kinect,” 2013 Fourth Global Congress on Intelligent Systems, vol. 1, pp. 200-204, 2013, doi: 10.1109/GCIS.2013.38.

[5] Microsoft Azure, “Azure Kinect DK,” https://azure.microsoft.com/en-us/services/kinect-dk/ [retrieved: Nov, 2021]

[6] Y. Ono and O.D.A. Prima, “Assessment of Drug Picking Activity using RGB-D Camera,” The Fourteenth International Conference on Advances in Computer-Human Interactions, ACHI 2021, pp. 6-11, 2021.

[7] F. Zhang et al., “MediaPipe hands: on-device real-time hand tracking,” CVPR Workshop on Computer Vision for Augmented and Virtual Reality, pp. 1-5, Jun. 2020, arXiv:2006.10214.

[8] O.D.A. Prima et al., “Evaluation of Joint Range of Motion Measured by Vision Cameras,” International Journal on Advances in Life Sciences, 11, 3 & 4, pp. 128-137, 2019.

[9] J. Ziegler, K. Nickel, and R. Stiefelhagen, “Tracking of the Articulated Upper Body on Multi-View Stereo Image Sequences,” Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’06), pp. 1-8, 2006, doi: 10.1109/CVPR.2006.313.

[10] Z. Cao, G. Hidalgo, T. Simon, S. Wei, and Y. Shikh, “OpenPose: Realtime Multi-Person 2D Pose Estimation using

Part Affinity Fields," arXiv preprint, pp. 1-14, 2018, arXiv:1812.08008v2.

[11] N. Nakano et al., "Evaluation of 3D Markerless Motion Capture Accuracy Using OpenPose With Multiple Video Cameras," Front. Sports Act. Living, May. 2020, doi: 10.3389/fspor.2020.00050.

[12] C. Chen and D. Ramanan, "3D Human Pose Estimation = 2D Pose Estimation + Matching," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7035-7043, 2017.

[13] J. Martinez, R. Hossain, J. Romero, and J. J. Little, "A Simple Yet Effective Baseline for 3d Human Pose Estimation," arXiv preprin, pp. 1-10, 2017, arXiv:1705.03098.

[14] G. Moon, J. Y. Chang, and K. M. Lee, "Camera distance-aware top-down approach for 3d multi-person pose estimation from a single RGB-image," pp. 1-15, Aug. 2019.

[15] S. Vosoughi and M. A. Amer, "Deep 3D Human Pose Estimation Under Partial Body Presence," 2018 25th IEEE International Conference on Image Processing (ICIP), Oct. 2018, doi: 10.1109/ICIP.2018.8451031.

[16] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu, "Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing in Natural Environments," IEEE Transaction on Pattern Analysis and Machine Intelligence, 36, 7, pp. 1325-1339, Dec. 2014, doi: 10.1109/TPAMI.2013.248.

[17] I. Sárándi, T. Linder, K. O. Arras, and B. Leibe, "How Robust is 3D Human Pose Estimation to Occlusion?," IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1-5, Aug. 2018, arXiv:1808.09316.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on Imagenet Classification," Proceedings of the IEEE International Conference on Computer Vision, pp. 1026-1034, 2015.

[19] Intel Corporation, "Intel RealSense Depth Camera D435," https://www.intelrealsense.com/depth-camera-d435/ [retrieved: Nov, 2021]

[20] F. Hampel, "The Influence Curve and Its Role in Robust Estimation," Journal of the American Statistical Association, Vol. 69, pp. 383-393, Jun. 1974, doi: 10.2307/2285666.

[21] B. Mutlu, A. Krause, J. Forlizzi, C. Guestrin, and J. Hodgins, "Robust, Low-cost, Non-intrusive Sensing and Recognition of Seated Postures," Proceedings of the 20th annual ACM symposium on User interface software and technology, pp. 149-158, Oct. 2007, doi: 10.1145/1294211.1294237.

[22] L. Martins et al., "Intelligent Chair Sensor – Classification and Correction of Sitting Posture," XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013, pp. 1489-1492, 2014, doi: 10.1007/978-3-319-00846-2_368.

[23] A. Petropoulos, D. Sikeridis, and T. Antonakopoulos, "SPoMo: IMU-based Real-time Sitting Posture Monitoring," 2017 IEEE 7th International Conference on Consumer Electronics, pp. 5-9, Sept. 2017, doi: 10.1109/ICCE-Berlin.2017.8210574.