岩手県立大学
ソフトウェア情報学部
Faculty of Software and Information Science

eTELEMED2021, The 13th International Conference on eHealth, Telemedicine, and Social Medicine

3D Human Pose Estimation of a Partial Body from a Single Image and Its Application in the Detection of Deterioration in Sitting Postures

Oky Dicky Ardiansyah Prima and Kazuki Hosogoe

eTELEMED2021 | July 2021

prima@iwate-pu.ac.jp

About Me

Assoc. Prof. Dr. Oky Dicky Ardiansyah Prima

Iwate Prefectural University, Department of Software and Information Science, Japan.
prima@iwate-pu.ac.jp

Research Interests

- Human Computer Interaction
- Behavior Analysis
- Eye Tracking
- Virtual Reality
- 3D Spherical Display

Iwate Prefectural University

Agenda

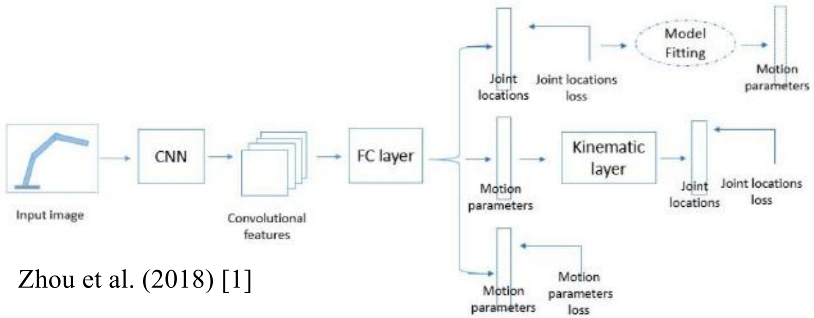
- Background
- Research Aims
- Building 3D Human Pose Models
- Evaluation of 3D Human Pose with/without Occlusion
- Detection of Deterioration in Sitting Postures
- Conclusion

■ Background :: 3D human pose estimation from a single image

Estimating 3D human pose from a single image is an active area of research in the field of computer vision because it has a wide range of potential applications.

Direct 3D Pose Estimation

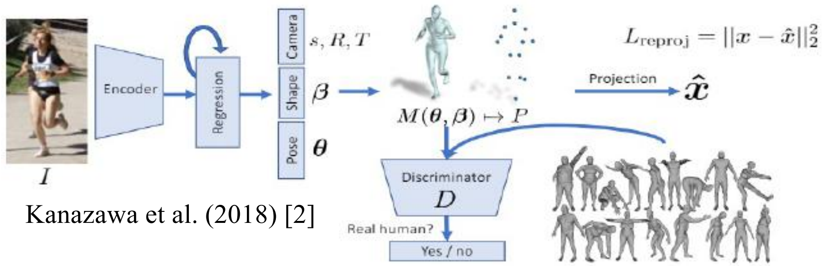
The 3D coordinates of the joints for a pose are predicted using an end-to-end network [1].



Zhou et al. (2018) [1]

SMPL-based Estimation

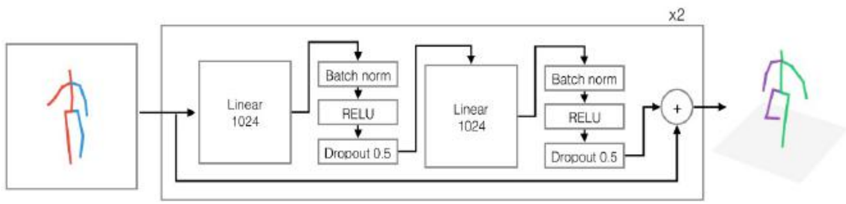
3D model of the human body (SMPL) is fitted to 2D body joints to calculate the 3D coordinates of the joints [2].



Kanazawa et al. (2018) [2]

Lifting 2D Pose to 3D Pose

The 3D coordinates of the joints for a pose are predicted by utilizing 2D pose estimation results [3].

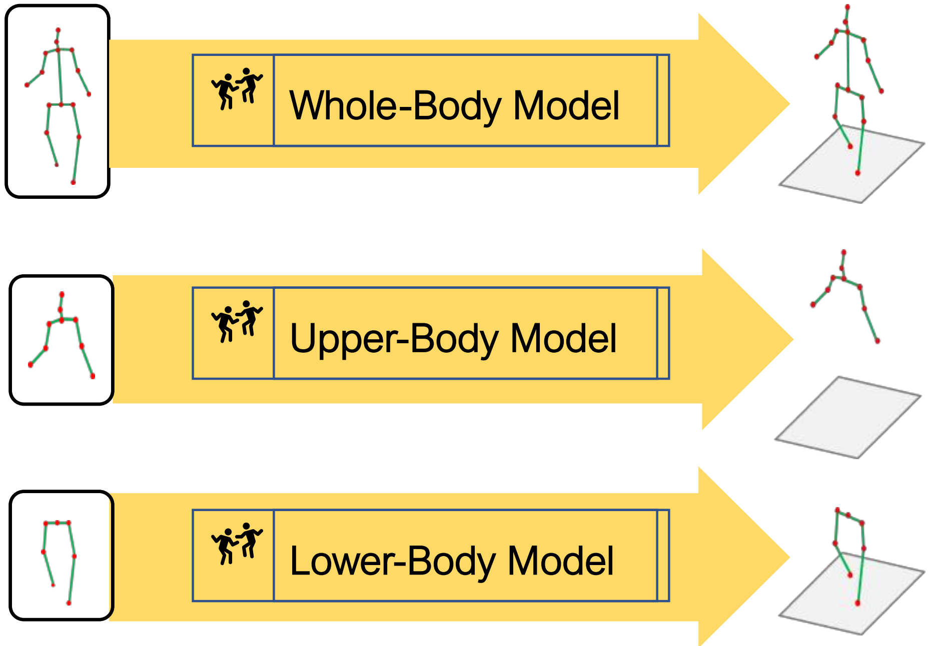


Martinez et al. (2017) [3]

■ Background :: How does occlusion affect the estimation results?

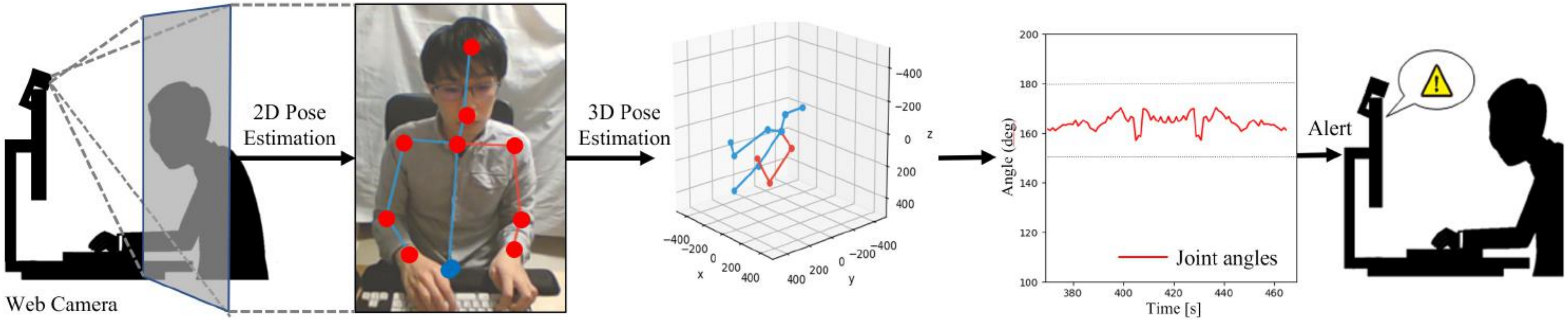
Most of the research on 3D human pose estimation assumes that the whole body can be captured, but there are many cases where parts of the body are occluded.

Could the estimation be improved by using a model that focuses only on the visible parts?



■ Research Aim

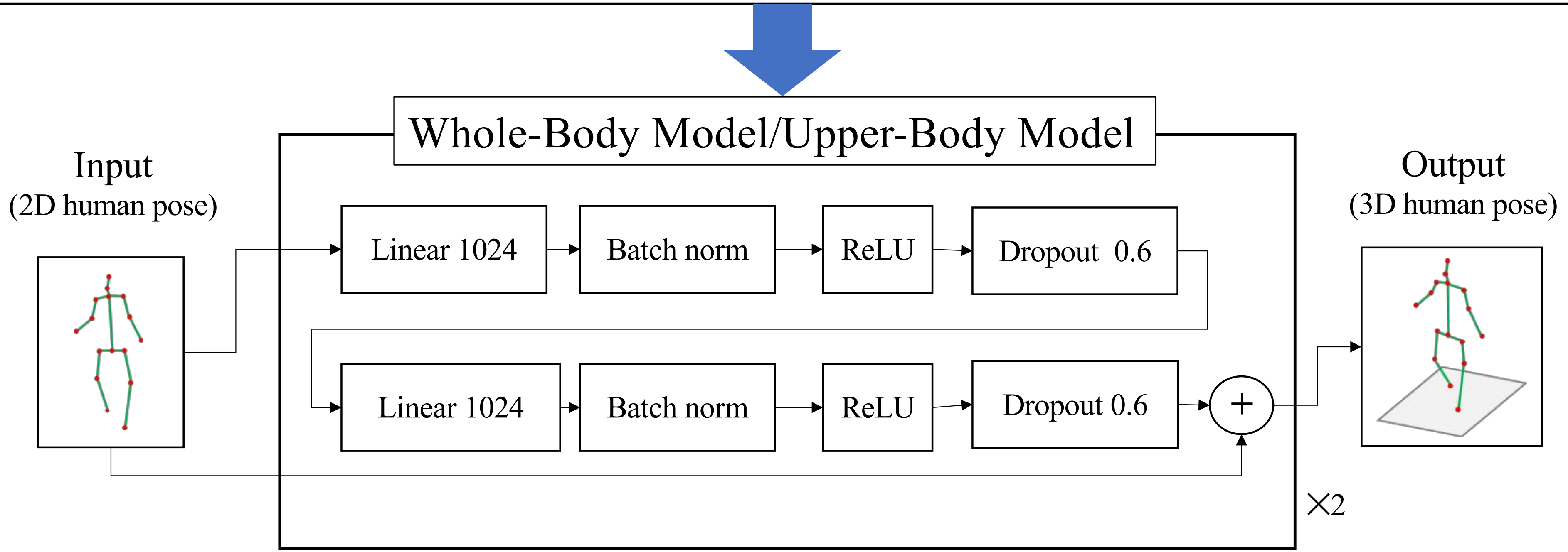
- To evaluate the accuracy of 3D human pose with/without occlusion of body parts.
- To detect deterioration in sitting postures using 3D human pose estimation.



Iwate Prefectural University

■ Building 3D Human Pose Models

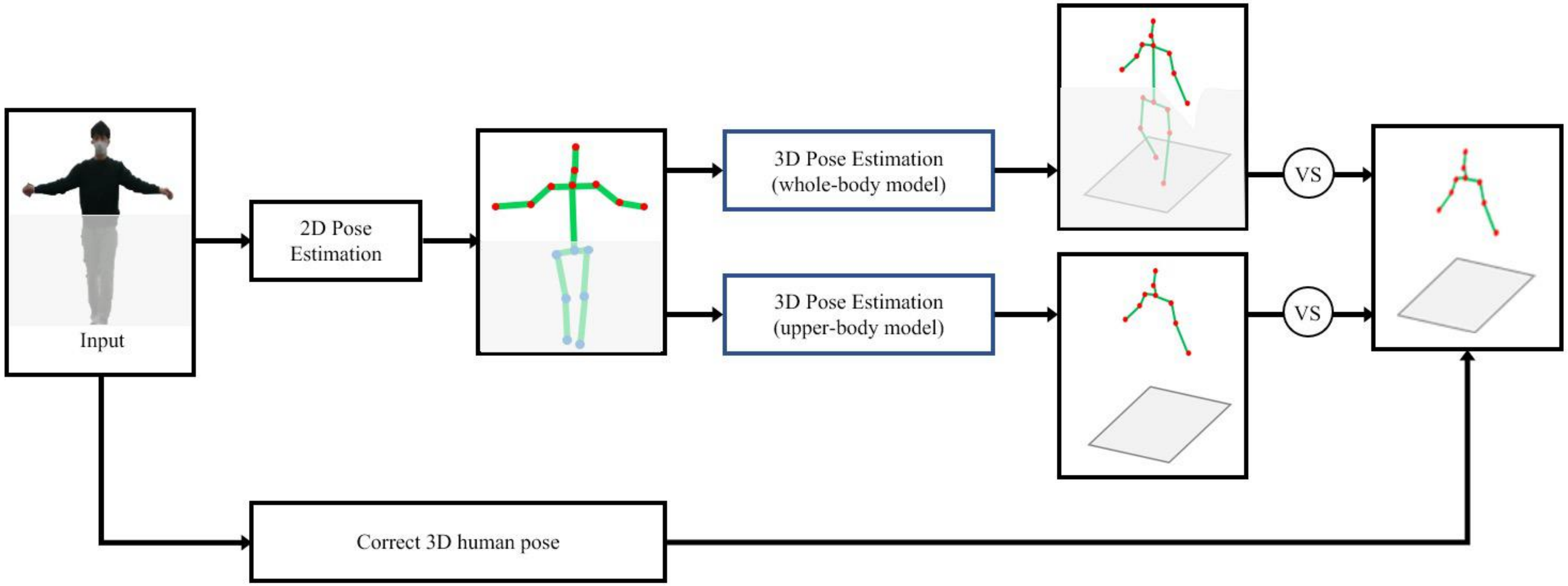
Based on Human3.6M [4], we created 3D human pose models using the 3D baseline method, which is a relatively simple deep feed-forward neural network that can efficiently perform 3D human pose estimation.



Martinez et al. (2017) [3]

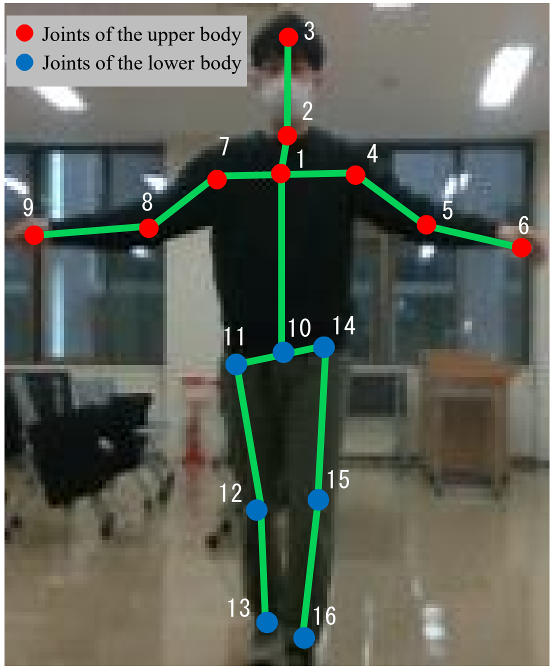
■ Evaluation of 3D Human Pose with/without Occlusion

When only the upper body is captured, how does the accuracy of 3D human pose estimation differ between a full body model and an upper body model?



■ Evaluation of 3D Human Pose with/without Occlusion

A. Randomly selected 548,800 human poses from the Human3.6M, which were not used for training and validation of the models.

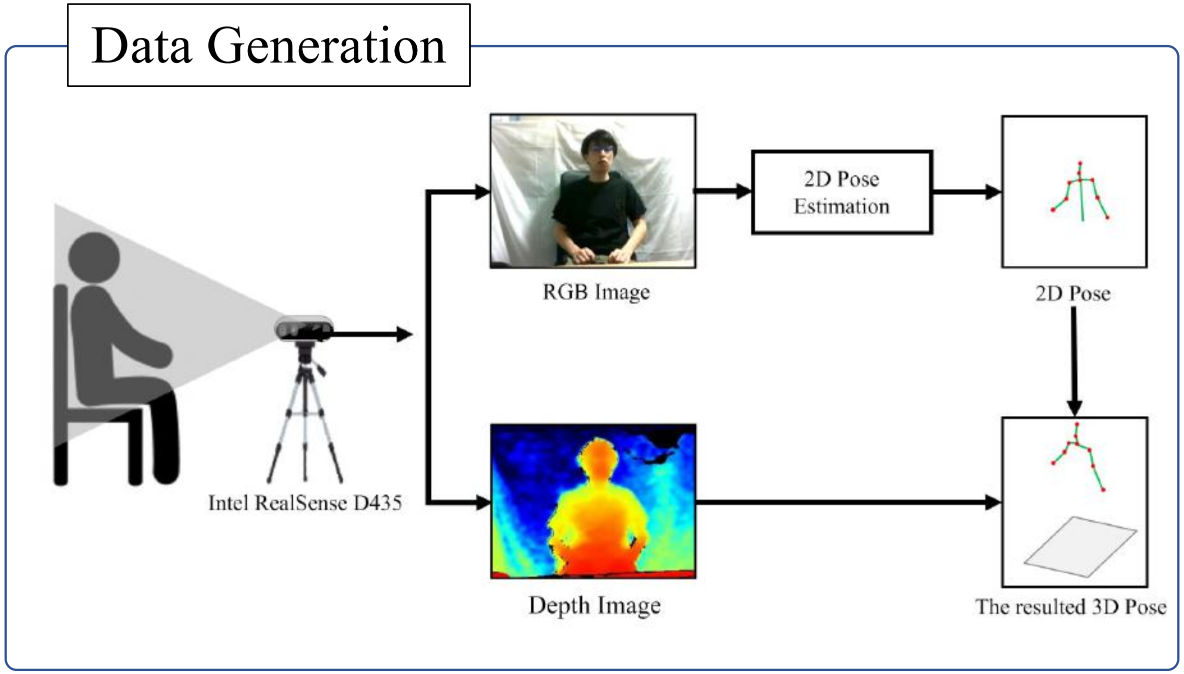


No.	Joints	Model [mm] Whole-Body	Model [mm] Upper-Body
1	Neck	162.79	51.24
2	Nose	38.39	26.71
3	Head	80.75	45.65
4	Left Shoulder	81.05	37.39
5	Left Wrist	91.10	45.13
6	Left Elbow	109.46	69.94
7	Right Shoulder	72.66	35.43
8	Right Wrist	89.63	46.64
9	Right Elbow	105.65	69.50
Mean (M)		92.387	47.514
Standard deviation (SD)		33.5448	14.5406

The accuracy of pose estimation significantly improved by about 45mm when the upper-body model was used to estimate the pose of the upper body ($M = 92.387$, $SD = 33.5448$, $t(8) = 4.99$, $p < .001$).

■ Evaluation of 3D Human Pose with/without Occlusion

B. 7,350 human poses from a subject measured by the Intel RealSense D435.

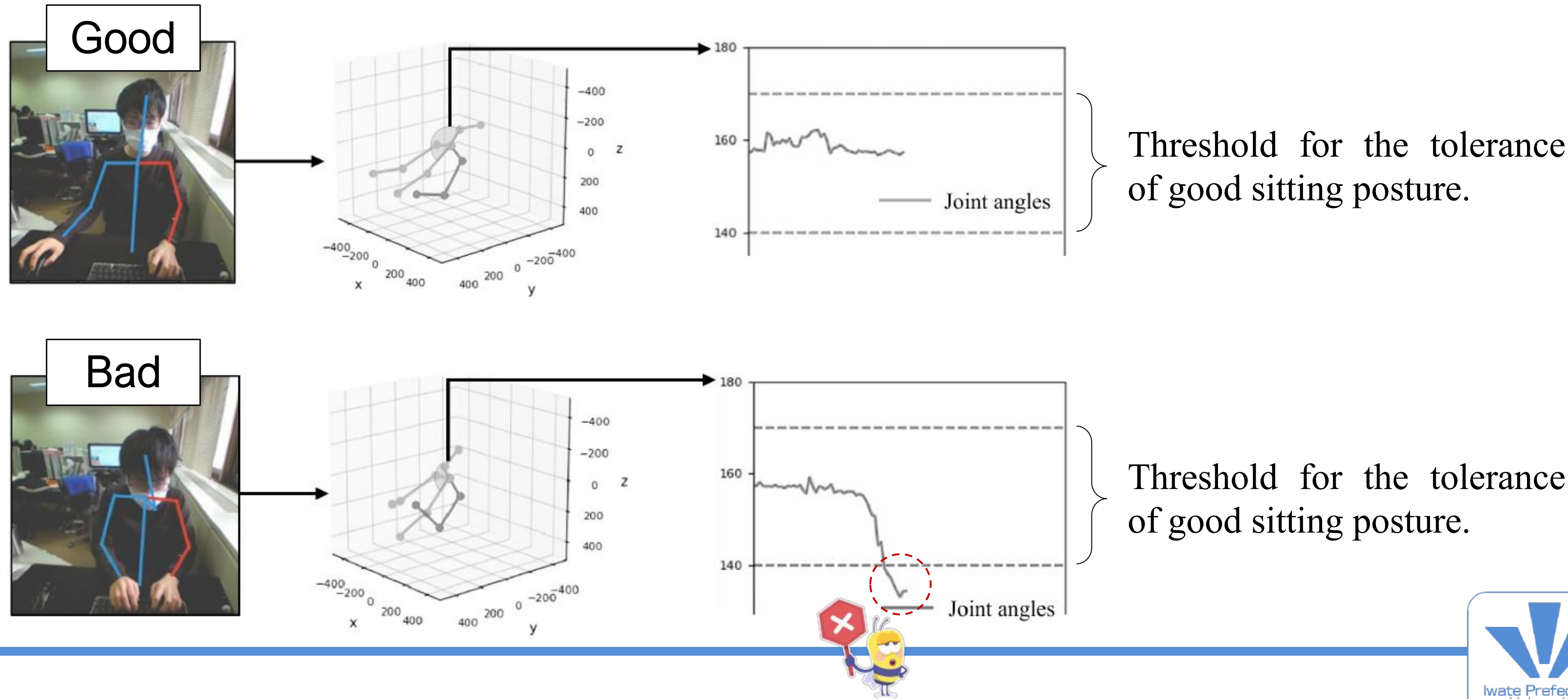


No.	Joints	Model [mm] Whole-Body	Upper-Body
1	Neck	114.70	74.87
2	Nose	181.63	145.16
3	Head	96.70	$\boxed{\mathbf{122.22}}$
4	Left Shoulder	103.78	91.44
5	Left Wrist	102.63	92.35
6	Left Elbow	148.85	107.67
7	Right Shoulder	107.51	$\boxed{\mathbf{113.34}}$
8	Right Wrist	102.86	$\boxed{\mathbf{108.82}}$
9	Right Elbow	158.45	111.21
Mean (M)		124.123	107.453
Standard deviation (SD)		30.7033	20.1145

The accuracy of pose estimation significantly improved by about 17mm when the upper-body model was used to estimate the pose of the upper body (M = 124.123, SD = 30.7033, $t(8) = 1.94$, $p < .05$).

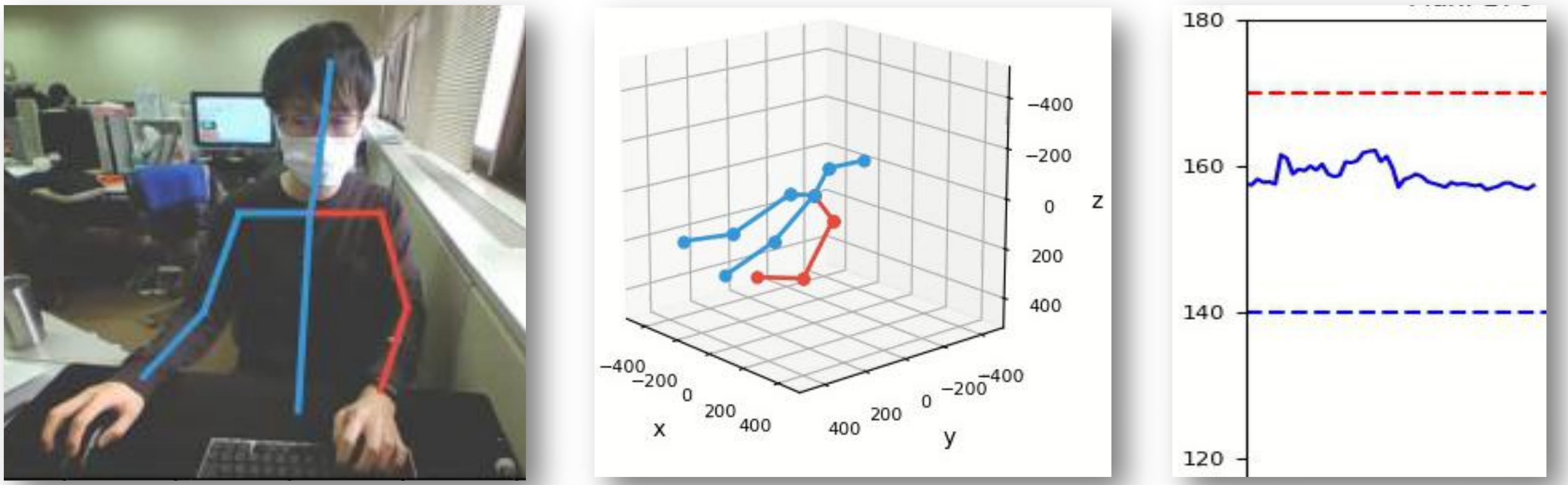
■ Detection of Deterioration in Sitting Postures

Changes in the angles of the nose, neck, and pelvis determine good or bad sitting posture.



■ Detection of Deterioration in Sitting Postures

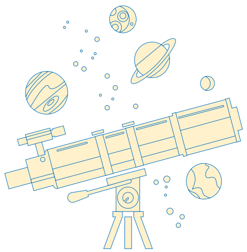
Monitoring of Sitting Postures (Movies)



Conclusion

What we have accomplished:

- We confirmed that the accuracy of 3D human pose estimation by the model specialized for given parts of the body was higher than that by the model for the whole body.
- The deterioration of the sitting posture can be detected by the change in the angle between the nose, neck and pelvis.



In the future, we would like to further improve the reliability of the posture deterioration detection by combining multiple joint angles.

■ References

1. X. Zhou, X. Sun, W. Zhang, S. Liang, and Y. Wei, "Deep kinematic pose regression," Lecture Notes in Computer Science, 9915, LNCS, pp. 186–201, 2016.
2. A. Kanazawa, M. J. Black, D. W. Jacobs, and J. Malik, "End-to-End Recovery of Human Shape and Pose," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7122–7131, 2018.
3. J. Martinez, R. Hossain, J. Romero, and J. J. Little, "A Simple Yet Effective Baseline for 3d Human Pose Estimation," 2017 IEEE International Conference on Computer Vision, pp. 2659–2668, 2017.
4. 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 Transactions on Pattern Analysis and Machine Intelligence, 36(7), pp. 1-15, 2014.