Thesis

Filip Kronström

November 29, 2020

1 Method

1.1 Pose estimation - 2D

The pose estimation is built around the open source framework MMPose [1] and it uses the *Distribution-Aware coordinate Representation of Keypoint*, **DARK**, method [4] to extract the body joints from the images. The model used is trained on the **COCO** dataset [2]. In this data set the human body is represented by 17 joints which can be seen in Figure 1. In Figure 2 the estimated joints are shown together with joints measured by ... The difference in shoulder location in Figure 2 can be explained by the positions of the joints tracked, which is also illustrated in Figure 1. This is something that is also reflected in the difference between the measured and estimated knee positions, shown in Figure 3.

1.2 Pose estimation - 3D

- [3]. Results not good atm, model needs finetuning with our depth data if used, see Figures 4, 5.

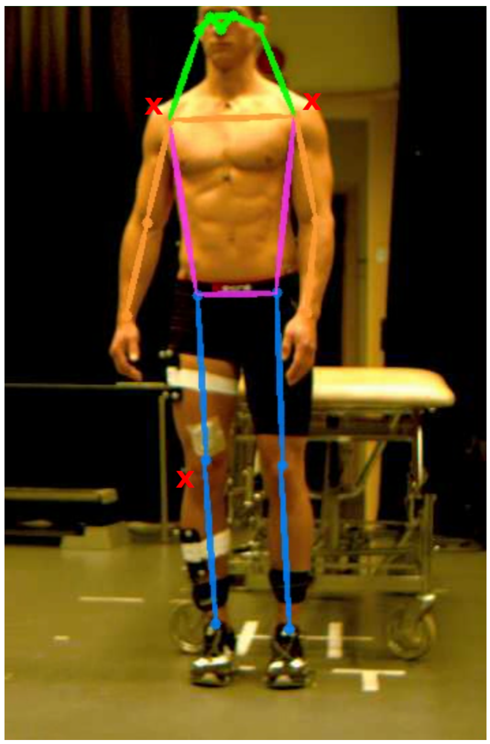

Figure 1: Body joints detected by model trained on the COCO data set. Red crosses indicate where measured joints that deviate from the estimated ones are situated.

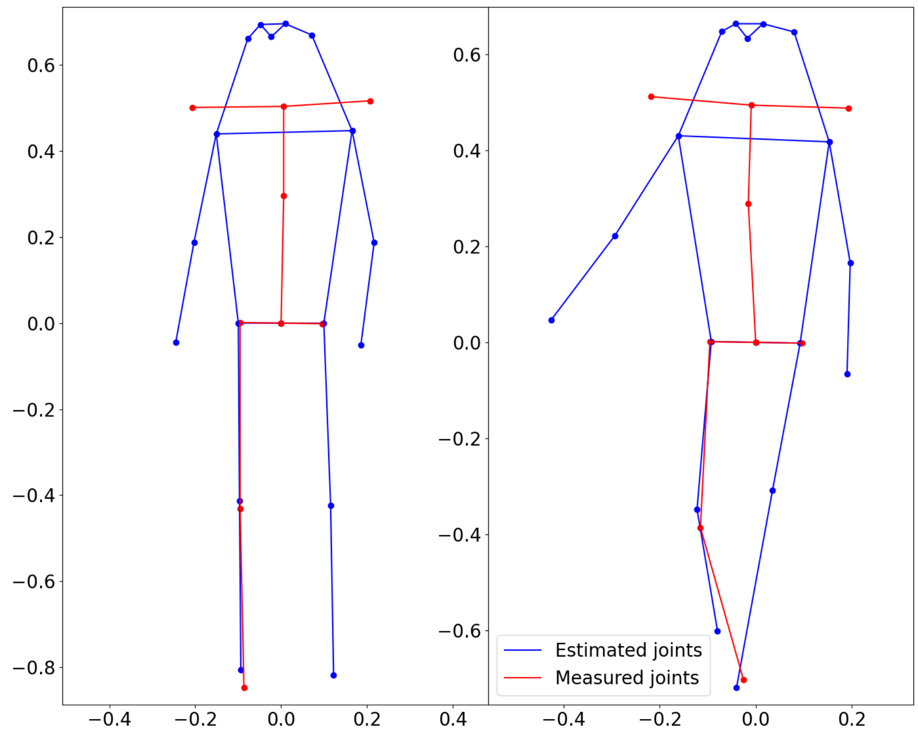

Figure 2: Estimated body joints together with measured joints for two frames in the video.

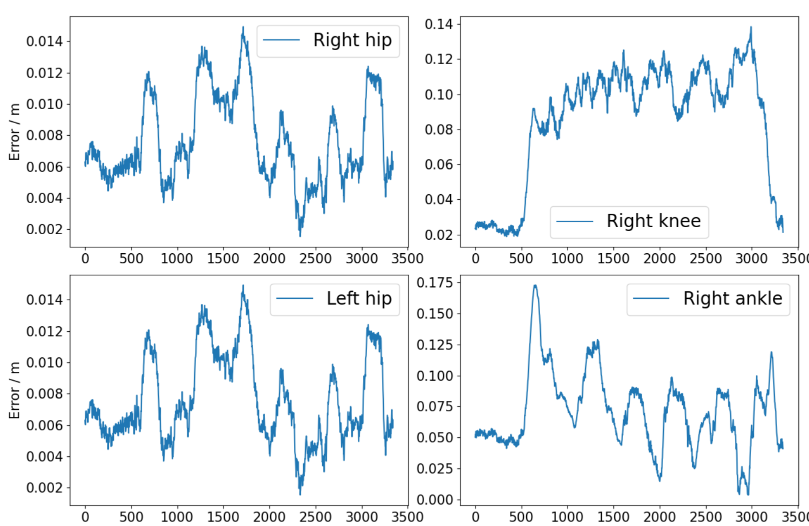

Figure 3: Euclidean distance between the measured and estimated joint positions.

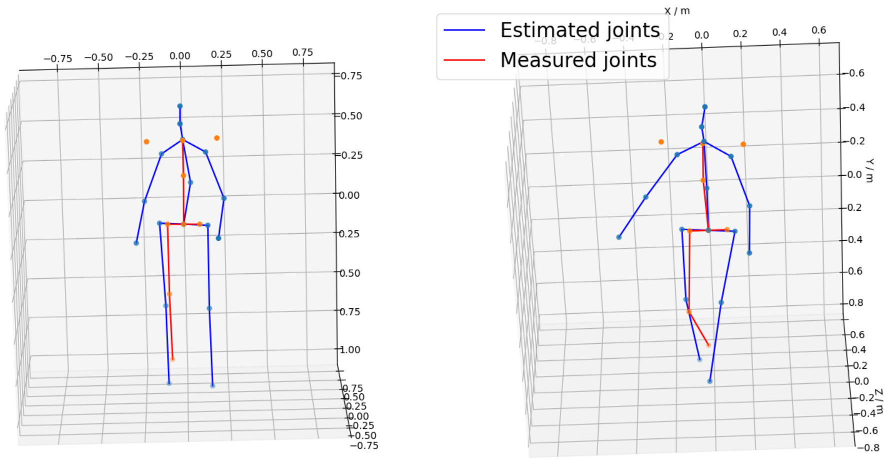

Figure 4: Estimated 3D body joints together with measured joints for the same two frames as in Figure 2.

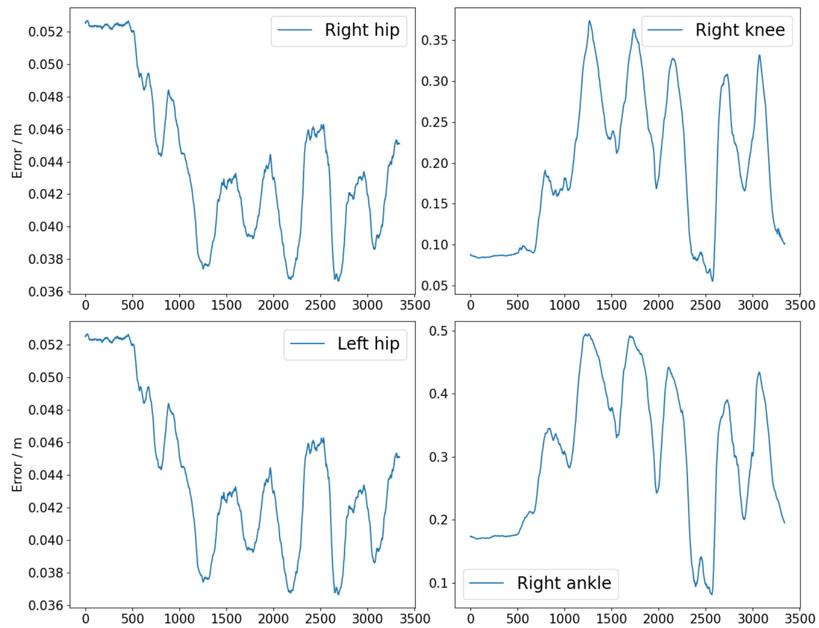

Figure 5: Euclidean distance between the measured and estimated 3D joint positions.

References

[1] MMPose - OpenMMLab Pose Estimation Toolbox and Benchmark. Available at `https://github.com/open-mmlab/mmpose`.

[2] Tsung Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. Microsoft COCO: Common objects in context. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 8693 LNCS, pages 740–755. Springer Verlag, may 2014.

[3] Dario Pavllo, Christoph Feichtenhofer, David Grangier, and Michael Auli. 3D human pose estimation in video with temporal convolutions and semi-supervised training. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2019-June, pages 7745–7754. IEEE Computer Society, jun 2019.

[4] Feng Zhang, Xiatian Zhu, Hanbin Dai, Mao Ye, and Ce Zhu. Distribution-Aware Coordinate Representation for Human Pose Estimation. pages 7091–7100. Institute of Electrical and Electronics Engineers (IEEE), aug 2020.