Supplementary Materials: The Best of Both Worlds: Combining Model-based and Nonparametric Approaches for 3D Human Body Estimation

Zhe Wang University of California, Irvine zwang15@ics.uci.edu Jimei Yang Adobe Research jimyang@adobe.com Charless Fowlkes University of California, Irvine fowlkes@ics.uci.edu

1. Qualitative Results

We show more qualitative results on COCO [3] in Fig 1, and 3DOH [8] in Fig 2.

2. Part Segmentation in UV Space

We first use the reference T-pose mesh and the LSP joint regressor provided by [1,7] to get the T-pose 14 joint location. Then we calculate the joint-vertex euclidean distance and assign vertex to joint based on the smallest distance. After that, we use the barycentric interpolation (mapping between vertex triangle and UV space triangles) to get the UV space assignment probability ($128 \times 128 \times 14$). Following these operations, we use argmax to get the final assignment for each UV grid to the joint location.

3. Implementation Details

For AMASS [5] data, we only get SMPL-H [6] fitting instead of SMPL fitting data, however, SMPL-H does not included hands rotations as in SMPL. We sample random rotations from SPIN [1] fitting or the predictions from our *DMP* stages for its training data. For AIST++ [2], it does not included β parameters, we sample β from SPIN [1] fitting or the predictions from our *DMP* stages for its training data. We use the original rotation representation from SMPL [4] (axis-angle representation) for the fast training purpose.

References

- Nikos Kolotouros, Georgios Pavlakos, Michael J Black, and Kostas Daniilidis. Learning to reconstruct 3d human pose and shape via model-fitting in the loop. In *ICCV*, 2019.
- [2] Ruilong Li, Shan Yang, David A. Ross, and Angjoo Kanazawa. Ai choreographer: Music conditioned 3d dance generation with aist++. In *ICCV*, 2021. 1
- [3] 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 ECCV, 2014. 1

- [4] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J. Black. Smpl: A skinned multiperson linear model. In *siggraph*, 2015. 1
- [5] Naureen Mahmood, Nima Ghorbani, Nikolaus F. Troje, Gerard Pons-Moll, and Michael J. Black. AMASS: Archive of motion capture as surface shapes. In *ICCV*, 2019. 1
- [6] Georgios Pavlakos, Vasileios Choutas, Nima Ghorbani, Timo Bolkart, Ahmed A. A. Osman, Dimitrios Tzionas, and Michael J. Black. Expressive body capture: 3d hands, face, and body from a single image. In CVPR, 2019. 1
- [7] Wang Zeng, Wanli Ouyang, Ping Luo, Wentao Liu, and Xiaogang Wang. 3d human mesh regression with dense correspondence. In *CVPR*, 2020. 1
- [8] Tianshu Zhang, Buzhen Huang, and Yangang Wang. Objectoccluded human shape and pose estimation from a single color image. In CVPR, 2020. 1



Figure 1. More qualitative results on COCO dataset. (Best viewed in Color)



Figure 2. More qualitative results on 3DOH dataset. (Best viewed in Color)