CVPR Supplementary Material

A. Labeling Tool

The interface of our human body orientation labeling tool is illustrated in Fig. A1.

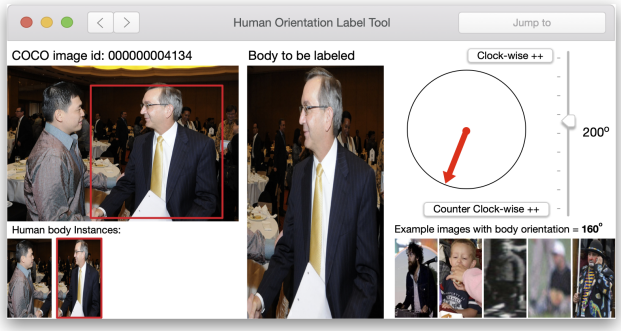

Figure A1. User interface of the labeling tool.

B. More Details on the HBOE Experiments

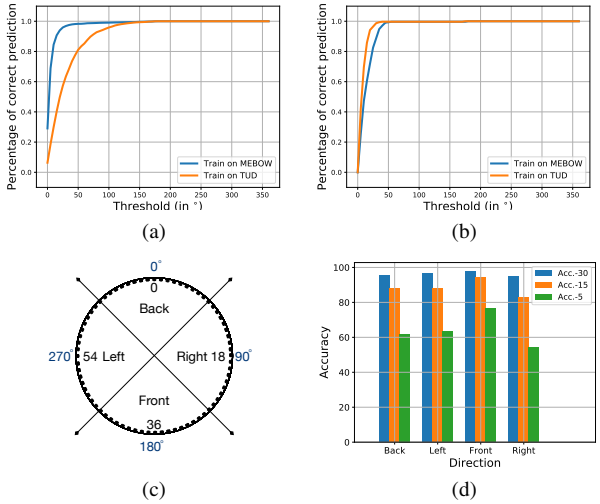

Figure A2. Breakdown analysis of the performance of our HBOE baseline model.

Breakdown analysis of the errors. *First*, we show the cumulative percentage of correct HBOE prediction with respect to the threshold of a correct prediction in Fig. A2 (a) and (b). Specifically, we compare the performance of our baseline model trained on the MEBOW dataset and that trained on the TUD dataset, respectively, using 1) the test set of the MEBOW dataset (Fig. A2 (a)) and 2) the test set of the TUD dataset (Fig. A2 (b)). Based on the same set of experiments in Table 3, these two sub-figures present a more detailed comparison, and they also support that the model trained on our MEBOW dataset has much better generalizability than it trained on TUD dataset. *Second*, we show how our baseline HBOE model performs when the camera point of view is towards the *Front*, *Back*, *Left*, and *Right* of the person in Fig. A2 (d). The association of the ground-truth orientations with the *Front*, *Back*, *Left*, and *Right* breakdown categories are shown in Fig. A2 (c). It is not surprising that our model performs best when the camera point of view is towards the *Front* of the person because a larger portion of MEBOW dataset falls into this category, as shown in Fig. 1 (a) in the main paper.

C. Additional 3-D Human Pose Estimation Evaluation on the Human3.6M Dataset

We also conducted 3-D human pose estimation experiments with Protocol I in [48]. The evaluation results are shown in Tabel A1.

Method	PA MPJPE
Chen *et al.* [11]	82.7
Moreno *et al.* [32]	76.5
Zhou *et al.* [56]	55.3
Sun *et al.* [47]	48.3
Sharma *et al.* [44]	40.9
Sun *et al.* [48]	40.6
Moon *et al.* [31]	34.0
Baseline*	34.7
Baseline 2**	34.3
ours	33.1

Table A1. 3-D human pose estimation evaluation on the Human3.6M dataset using Protocol I. *Our baseline is a re-implementation of Sun *et al.* [48], trained on Human3.6M + MPII, as in the original paper. **Our baseline 2 is a re-implementation of Sun *et al.* [48], trained on Human3.6M + MPII + COCO (2-D Pose).

D. More Qualitative Human Body Orientation Estimation Results

More qualitative human body orientation estimation examples are shown in Fig. A3 to supplement Fig. 4 in the main paper.

E. More Qualitative 3-D Pose Estimation Results on the Human3.6M Dataset

More example 3-D pose estimation results on the test set of the Human3.6M dataset are included in Fig. A4.

F. More Qualitative 3-D Pose Estimation Results on the COCO Dataset

More example 3-D pose estimation results on the test set of the COCO dataset are shown in Fig. A5.

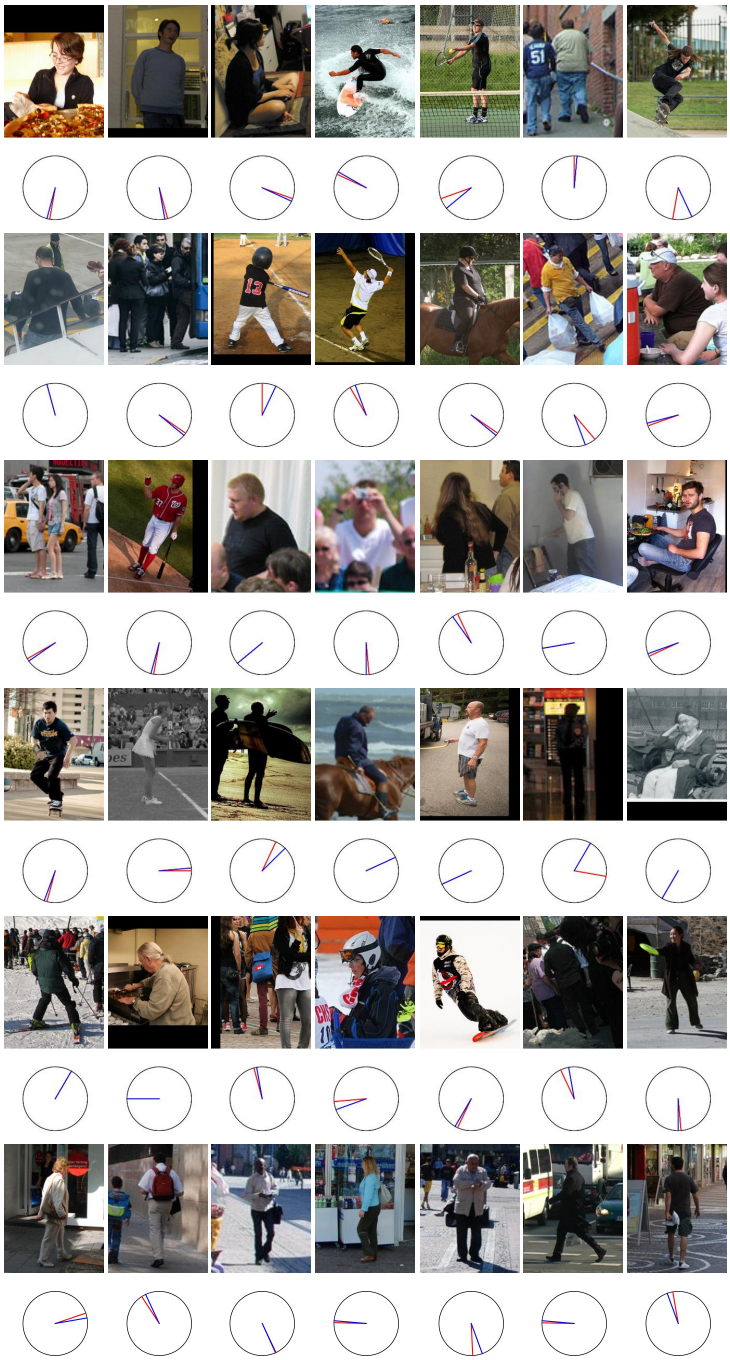

Figure A3. HBOE results generated by our baseline model (with HRNet as the backbone and $\sigma = 4.0$) on MEBOW (row 1 to row 5) and TUD dataset (row 6). Red arrow: ground truth; Blue arrow: prediction.

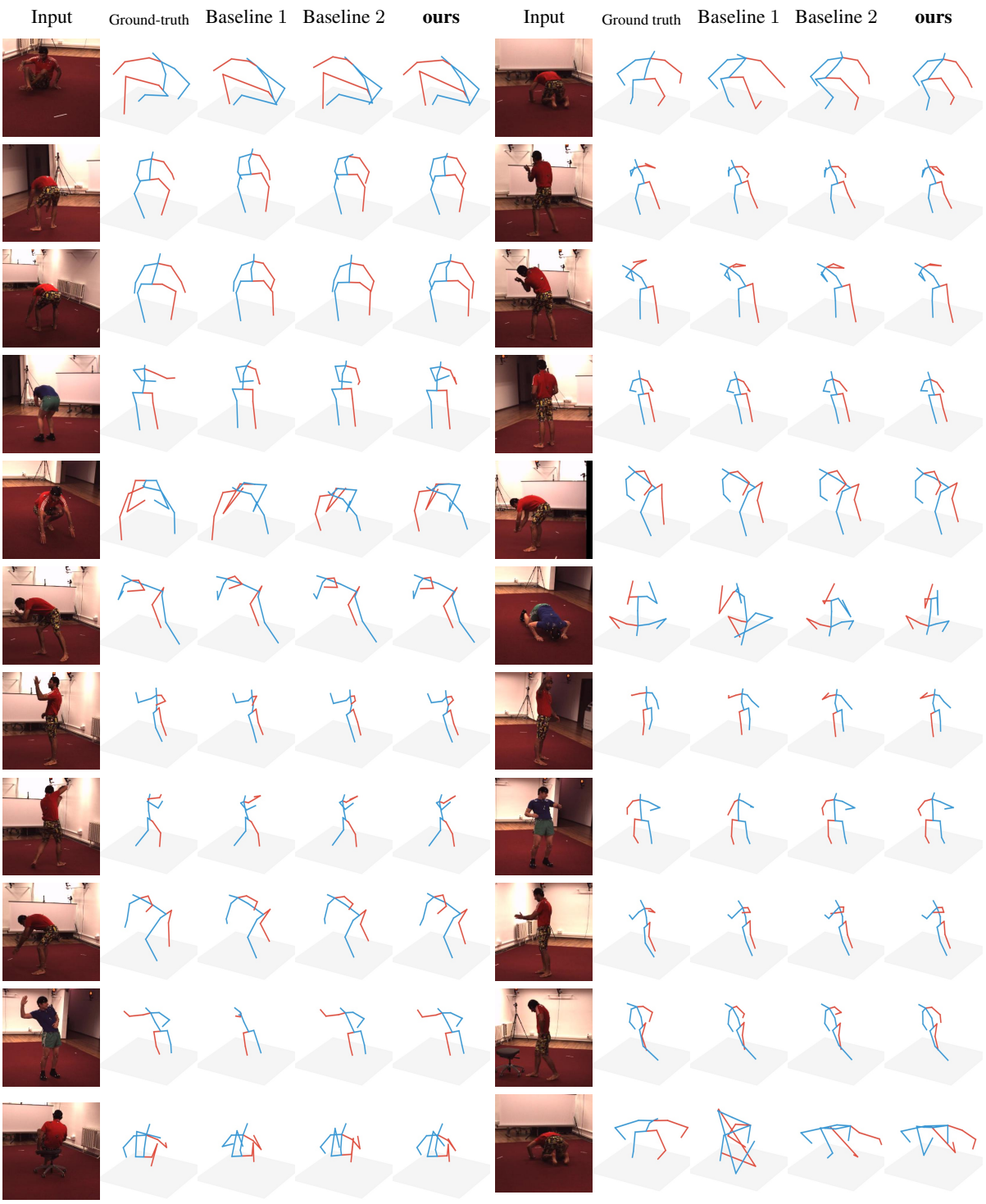

Figure A4. More example 3-D pose estimation results on the Human3.6M dataset.

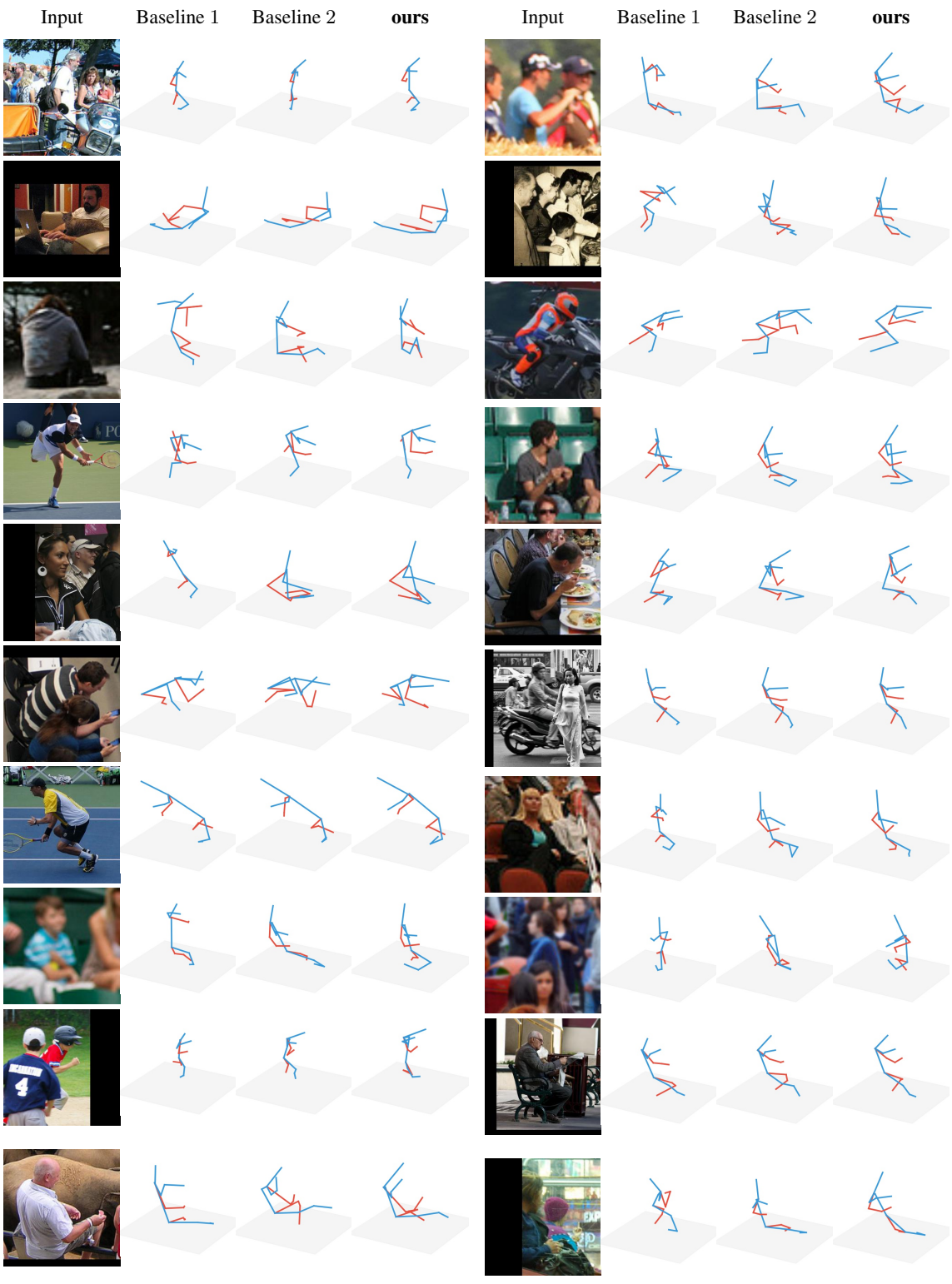

Figure A5. More example 3-D pose estimation results on the COCO dataset.