Supplementary Material for Hierarchical Graph Networks for 3D Human Pose Estimation

Han Li[1]
qingshi9974@sjtu.edu.cn
Bowen Shi[1]
sjtu_shibowen@sjtu.edu.cn
Wenrui Dai*[1]
daiwenrui@sjtu.edu.cn
Yabo Chen[1]
chenyabo@sjtu.edu.cn
Botao Wang[2]
botaow@qti.qualcomm.com
Yu Sun[2]
sunyu@qti.qualcomm.com
Min Guo[2]
mguo@qti.qualcomm.com
Chenlin Li[1]
lcl1985@sjtu.edu.cn
Junni Zou[1]
zoujunni@sjtu.edu.cn
Hongkai Xiong[1]
xionghongkai@sjtu.edu.cn

[1] Shanghai Jiao Tong University
Shanghai, China

[2] Qualcomm AI Research
Shanghai, China

1 Sensitivity Analysis

We evaluate the sensitivity of our method to the input 2D poses. In this experiment, we train all the models using ground truth 2D and 3D pose pairs. In test, for each input ground truth 2D pose, we randomly add Gaussian noises with different standard deviations to one of the 2D joints (*e.g.* head). Then we calculate the mean 3D error increment for the rest joints. As shown in Table 1, our HGN and SemGCN [4] is affected more by noise compared with the fully connected network (FCN) [1]. This is mainly because the misleading information from input noise can be massively propagated between nodes in GCN. However, our model still has a slight advantage over baseline in robustness. It is worth noting that if we remove the semantic relationship matrix of adjacent nodes learned by SemGConv, the error increment will be evidently reduced.

Table 1: Mean 3D errors increment for the rest of the joints when one joint is corrupted by different levels of additive Gaussian noises, e.g., 5, 10, 15, 20, 25.

	$\mathcal{N}(0,5^2)$	$\mathcal{N}(0,10^2)$	$\mathcal{N}(0,15^2)$	$\mathcal{N}(0,20^2)$	$\mathcal{N}(0,25^2)$
Martinez *et al*. [1]	1.63	5.35	10.15	15.57	21.32
SemGCN [2]	4.63	12.08	20.24	28.93	38.41
Ours (HGN)	4.48	11.52	19.49	28.42	36.54
Ours w/o SemGConv	3.36	9.29	15.95	23.18	29.52

References

[1] Julieta Martinez, Rayat Hossain, Javier Romero, and James J Little. A simple yet effective baseline for 3d human pose estimation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2640–2649, 2017.

[2] Long Zhao, Xi Peng, Yu Tian, Mubbasir Kapadia, and Dimitris N. Metaxas. Semantic graph convolutional networks for 3D human pose regression. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3425–3435, 2019.