Occlusion Handling in 3D Human Pose Estimation with Perturbed Positional Encoding
Supplementary Document

Niloofar Azizi[1], Mohsen Fayyaz[2], and Horst Bischof[1]

[1] Graz University of Technology, Graz, Austria {azizi, bischof}@tugraz.at
[2] Microsoft
mohsenfayyaz@microsoft.com

A Supplementary Material

Below we'll explore one particular example (missing two legs) thoroughly.

Two Legs are Missed The corresponding graph Laplacian matrix in this scenario when two edges (specifically legs; the ones in red) are missed is

$$\begin{bmatrix}
3 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 4 & -1 & -1 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 2 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1
\end{bmatrix}$$

Perturbed with Zero Edges To compute the perturbed eigenvectors when only zero edges are missed (which means addressing multiplicity)

- **Eigenvalue**
  - $[0.\ 0.\ 0.\ 0.10\ 0.27\ 0.36\ 1.\ 1.\ 1.46\ 1.68\ \ 2.00\ 2.35\ 3.08\ 3.15\ 4.32\ 5.21]$

In this scenario computing the perturbed eigenvectors contributes to resolving the issue of repeated eigenvalues for 0 and 1, thereby leading to an improvement.

Perturbed with One Edge In this scenario, we perturb the matrix with one randomly selected edge (*e.g.,* the blue one in the graph Laplacian matrix) κ-times and compute the average to extract the consistent part of the graph Laplacian eigenbasis.

Pertubed with Two Edges In this scenario, we consider perturbing the graph Laplacian matrix with two edges (*e.g.,* the orange ones in the graph Laplacian matrix) κ-times and compute the average to extract the consistent part of the graph Laplacian eigenbasis.

PerturbPE Efficacy on other GNNs Our proposed method, PerturbPE, is designed as a positional encoding approach for scenarios where some edges are missing. These features can be easily integrated into other GNN methods as well. We conducted experiments on SemGCN [3] using PerturbPE, and the results are in Table A.1. We made minor modifications to adapt and integrate PerturbPE to SemGCN. We reduced the features in the penultimate block to 16 and summed our positional features within this block.

	# Params	MPJPE
SemGCN	0.21M	43.1
SemGCN + PerturbPE	0.21M	**42.0**

Table A.1: PerturbPE Effect on SemGCN [3]

Recover 3D pose from partial 3D observation Similar to GFPose [1], we also consider the case where the partial 3D is observed. which is a common real-world scenario. Results are presented in Table A.2.

Similar to the findings in GFPose [1] the most challenging situation is observed when the left arm is missing. Nevertheless, under these circumstances, the introduction of perturbPE improves the accuracy, achieving a precision of 6.0 mm. In every other tested scenario, perturbPE consistently excels over GFPose.

Occ. Body Parts	Ours	GFPose [1]
Right Leg	**4.9**	5.2
Left Leg	**2.2**	5.8
Left Arm	**6.0**	9.4
Right Arm	**4.3**	8.9

Table A.2: Recover 3D pose from partial 3D observation: We train one model on Human3.6M dataset [2] under Protocol #1 (given 3D GT inputs). Lower is better.

Bibliography

[1] Hai Ci, Mingdong Wu, Wentao Zhu, Xiaoxuan Ma, Hao Dong, Fangwei Zhong, and Yizhou Wang. GFPose: Learning 3D Human Pose Prior With Gradient Fields. In *CVPR*, 2023. 2

[2] Catalin Ionescu, Dragos Papava, Vlad Olaru, and Cristian Sminchisescu. Human3.6M: Large Scale Datasets and Predictive Methods for 3D Human Sensing in Natural Environments. *IEEE TPAMI*, 36(7):1325–1339, 2014. 2

[3] Long Zhao, Xi Peng, Yu Tian, Mubbasir Kapadia, and Dimitris N. Metaxas. Semantic Graph Convolutional Networks for 3D Human Pose Regression. In *CVPR*, 2019. 2