Divide and Fuse: Body Part Mesh Recovery from Partially Visible Human Images (Supplementary)

This supplementary material includes the following contents:

A. HPPM template design details.
B. HPPM part-joint correspondence.
C. Evaluation metrics.
D. Adaptiveness on full-body visible inputs.

A HPPM Template Design Details

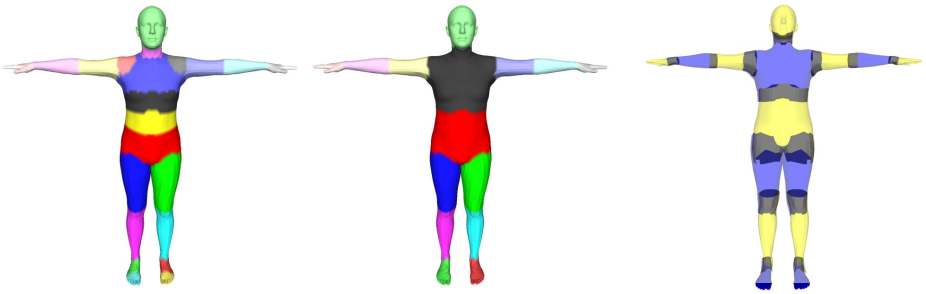
As shown in Fig. S1, we use the SMPL template to design our HPPM template according to the following three steps. First, we initiate with a segmentation of the mesh according to Eq. (1) from the main paper, showcased in Fig. S1a. This segmentation creates 23 distinct segments, corresponding to the number of bones in the SMPL skeleton, to ensure each segment is aligned with a specific bone. Second, adjacent segments that exhibit nearly-rigid behavior are grouped together to form larger, coherent sections of the mesh, leading to the grouped segmentation shown in Fig. S1b. Finally, to ensure smooth transitions and coverage around joint areas, we apply graph dilation. This process expands each segment to slightly overlap with its neighbors, enhancing the mesh's flexibility and realism around body joints. For a given body part p, its n-nearest-neighbor dilated mesh mask is determined by:

$$m_{pn} = A^n m_p, \tag{1}$$

where $m_p \in \{0,1\}^N$ is the initial vertex mask for part p, indicating whether each of the N vertices in the SMPL template belongs to part p ($m_p = 1$ for inclusion, and $m_p = 0$ otherwise). The matrix A represents the adjacency matrix of the SMPL mesh graph, and $m_{pn} \in \{0,1\}^{N_p}$ denotes the vertex mask of part p after applying n-nearest-neighbor dilation, with N_p being the dilated number of vertices. In this paper, the dilation parameter is empirically set to $n = 5$, resulting in the dilated mesh depicted in Fig. S1c. This step ensures each body part mesh extends to overlap with adjacent parts, facilitating a more integrated and natural representation of the human body parts in the final HPPM template.

B HPPM Part-joint Correspondence

The correspondence between HPPM body parts and their associated joints is outlined in Tab. S1. We utilize a "checkmark" (✓) to denote which joints correspond to each HPPM body part. Specifically, for every HPPM body part, only those joints marked with a checkmark are included in its joint regressor, establishing a clear linkage between body segments and their respective joints for accurate pose representation.



(a) (b) (c)

Fig. S1: HPPM template design details.

Part Names \ Joints	Pelvis	Right Hip	Right Knee	Right Ankle	Left Hip	Left Knee	Left Ankle	Torso	Neck	Nose	Head	Left Shoulder	Left Elbow	Left Wrist	Right Shoulder	Right Elbow	Right Wrist
Abdomen	✓	✓			✓			✓									
Left Thigh					✓	✓											
Right Thigh		✓	✓														
Left Calf						✓	✓										
Right Calf			✓	✓													
Chest								✓	✓			✓			✓		
Left Foot				✓													
Right Foot							✓										
Head									✓	✓	✓						
Left Upper Arm												✓	✓				
Right Upper Arm															✓	✓	
Left Forearm													✓	✓			
Right Forearm																✓	✓
Left Hand														✓			
Right Hand																	✓

Table S1: Part-joint Correspondence Table.

C Evaluation Metrics

MPVE is defined as:

$$MPVE = \frac{1}{M}\sum(\|\hat{v} - v\|), \tag{2}$$

where M is the number of data cases, $\hat{v}$ corresponds to the estimated vertex locations, and v represents the ground-truth vertex locations. MPJPE is defined as:

$$MPJPE = \frac{1}{M}\sum(\|\hat{J} - J\|), \tag{3}$$

where M is the number of data cases, $\hat{J}$ is the estimated joint locations, and J is the ground-truth joint locations.

D Adaptiveness on Full-body Visible Inputs

Primarily designed for scenarios with partially visible humans, our method demonstrates notable adaptability to scenarios featuring fully visible body inputs. This capability is showcased without requiring any modifications to the original approach. As indicated in Tab. S2, the performance of our method in full-body

Table S2: Adaptiveness on full-body visible inputs. As shown in the table, our method can also get reasonable results on full-body visible input.

Model	3DPW [4]		Human3.6M [4]
	MPJPE/mm↓	MPVE/mm↓	MPJPE/mm↓
HMR [1]	130	-	88.0
HMMR [2]	116.5	139.3	83.7
BMP [5]	104.1	119.8	-
VIBE [3]	93.5	113.4	65.6
D&F(Ours)	100.5	123.7	64.6

visibility images remains competitive with that of previous methods, highlighting the flexibility and efficacy of our approach.

References

1. Kanazawa, A., Black, M.J., Jacobs, D.W., Malik, J.: End-to-end recovery of human shape and pose. In: CVPR (2018)
2. Kanazawa, A., Zhang, J.Y., Felsen, P., Malik, J.: Learning 3d human dynamics from video. In: CVPR. pp. 5614–5623 (2019)
3. Kocabas, M., Athanasiou, N., Black, M.J.: Vibe: Video inference for human body pose and shape estimation. In: CVPR. pp. 5253–5263 (2020)
4. Von Marcard, T., Henschel, R., Black, M.J., Rosenhahn, B., Pons-Moll, G.: Recovering accurate 3d human pose in the wild using imus and a moving camera. In: ECCV. pp. 601–617 (2018)
5. Zhang, J., Yu, D., Liew, J.H., Nie, X., Feng, J.: Body meshes as points. In: CVPR. pp. 546–556 (2021)