Robust Human Registration with Body Part Segmentation on Noisy Point Clouds

Kai Lascheit[1], Daniel Barath[1,4,5], Marc Pollefeys[1,2], Leonidas Guibas[3], Francis Engelmann[3]

[1]ETH Zürich, [2]Microsoft, [3]Stanford University, [4]Google, [5]HUN-REN SZTAKI

Introduction

Input *Ground Truth* *Baseline: ArtEq* *Baseline: NICP* *Ours*

- Extracting human poses from point clouds is essential in augmented reality and human-robot interaction
- Current methods focus on cleaned input data, disregarding noise and obstructions
- We show significant improvements on real-world datasets by combining body part segmentation with SMPL-X fitting

Experiments

Method	BEHAVE V2V ↓ in mm	BEHAVE MPJPE ↓ in mm	BEHAVE Time ↓ in s	EgoBody V2V ↓ in mm	EgoBody MPJPE ↓ in mm	EgoBody Time ↓ in s	InterCap V2V ↓ in mm	InterCap MPJPE ↓ in mm	InterCap Time ↓ in s
ArtEq	140.6	162.4	**0.105**	538.7	605.7	**0.098**	422.0	515.0	**0.103**
NICP	59.9	–	33.14	232.1	–	17.2	257.7	–	22.2
SegFit (Ours)	**37.0**	**30.7**	1.86	**47.9**	**42.2**	1.79	**147.2**	**140.7**	1.37

Table: Pose and shape scores across different datasets. Metrics are vertex-to-vertex (V2V) distance, mean-per-joint-position-error (MPJPE) and average time.

- We evaluate pose fitting results on real-world datasets
- We demonstrate how finetuning with re-assigned part labels improves segmentation accuracy

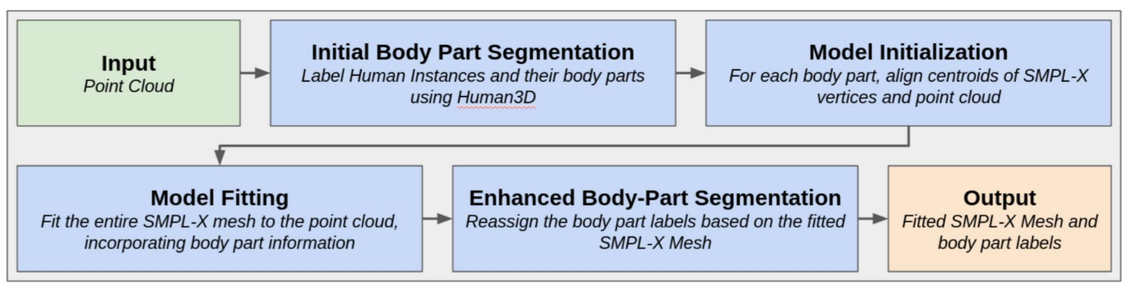
Method	Metric	BEHAVE	EgoBody	Intercap
Human3D	Acc	74.61%	77.06%	55.85%
+ SegFit		**91.22%**	**84.71%**	**67.86%**
Human3D	IoU	57.38%	62.89%	45.37%
+ SegFit		**77.26%**	**69.48%**	**55.13%**
Human3D	mAP	73.42%	73.98%	59.54%
+ SegFit		**85.74%**	**77.20%**	**67.69%**

Table: Segmentation accuracy when re-assigning body parts using our method

Human3D Model	Metric	BEHAVE	InterCap
w/o Fine-Tuning	Accuracy	74.61%	55.85%
w/ Fine-Tuning		**89.96%**	**67.98%**
w/o Fine-Tuning	IoU	57.38%	45.37%
w/ Fine-Tuning		**75.36%**	**56.71%**
w/o Fine-Tuning	AP	73.42%	59.54%
w/ Fine-Tuning		**84.56%**	**68.18%**

Table: Segmentation accuracy after self-supervised finetuning of Human3D

Method



- Human3D is applied to identify human instances and segment body parts
- Corresponding body part centroids of SMPL-X and identified human instances are aligned to initialize the pose fitting
- The complete SMPL-X model is fitted to the point cloud to refine the pose fitting
- Body part labels are recomputed based on nearest neighbors to SMPL-X model
- Enhanced body part labels can be used to finetune Human3D on given datasets

Conclusion

- We propose a new method to enhance accuracy of SMPL-X fitting on real-world data
- We achieve this by combining a pre-trained segmentation network with a two-step mesh fitting approach.
- We additionally show that through re-assigning part labels after fitting the mesh, self-supervised fine-tuning on diverse datasets can be achieved.

Input *SMPL-X Fitting* *Segmentation*