



ZMO.AI



Code

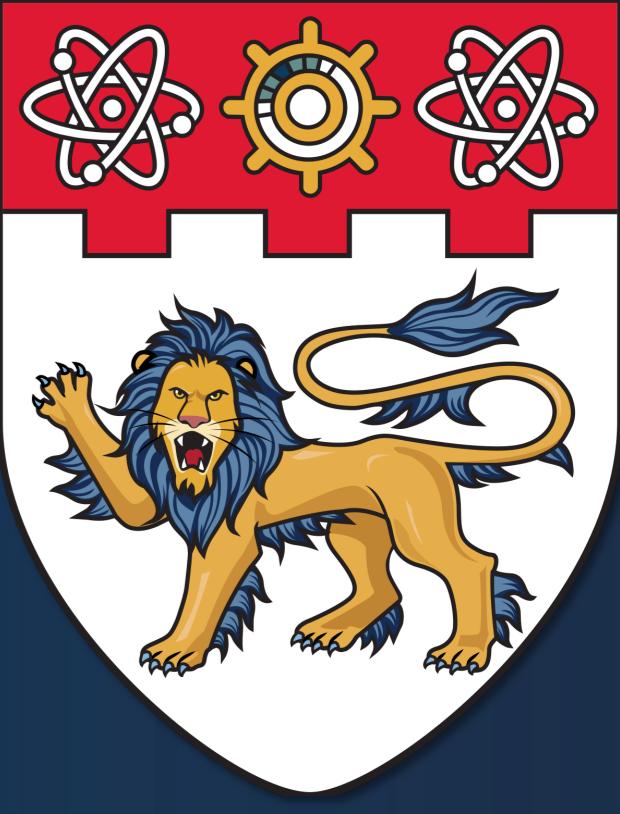
UNIF: United Neural Implicit Functions for Clothed Human Reconstruction and Animation

Shenhan Qian^{1,2} Jiale Xu² Ziwei Liu³ Liqian Ma^{1†} Shenghua Gao^{2,4,5}

¹ZMO AI Inc. ²ShanghaiTech University ³S-Lab, Nanyang Technological University

⁴Shanghai Engineering Research Center of Intelligent Vision and Imaging ⁵Shanghai Engineering Research Center of Energy Efficient and Custom AI IC

ECCV
TEL AVIV 2022



Introduction

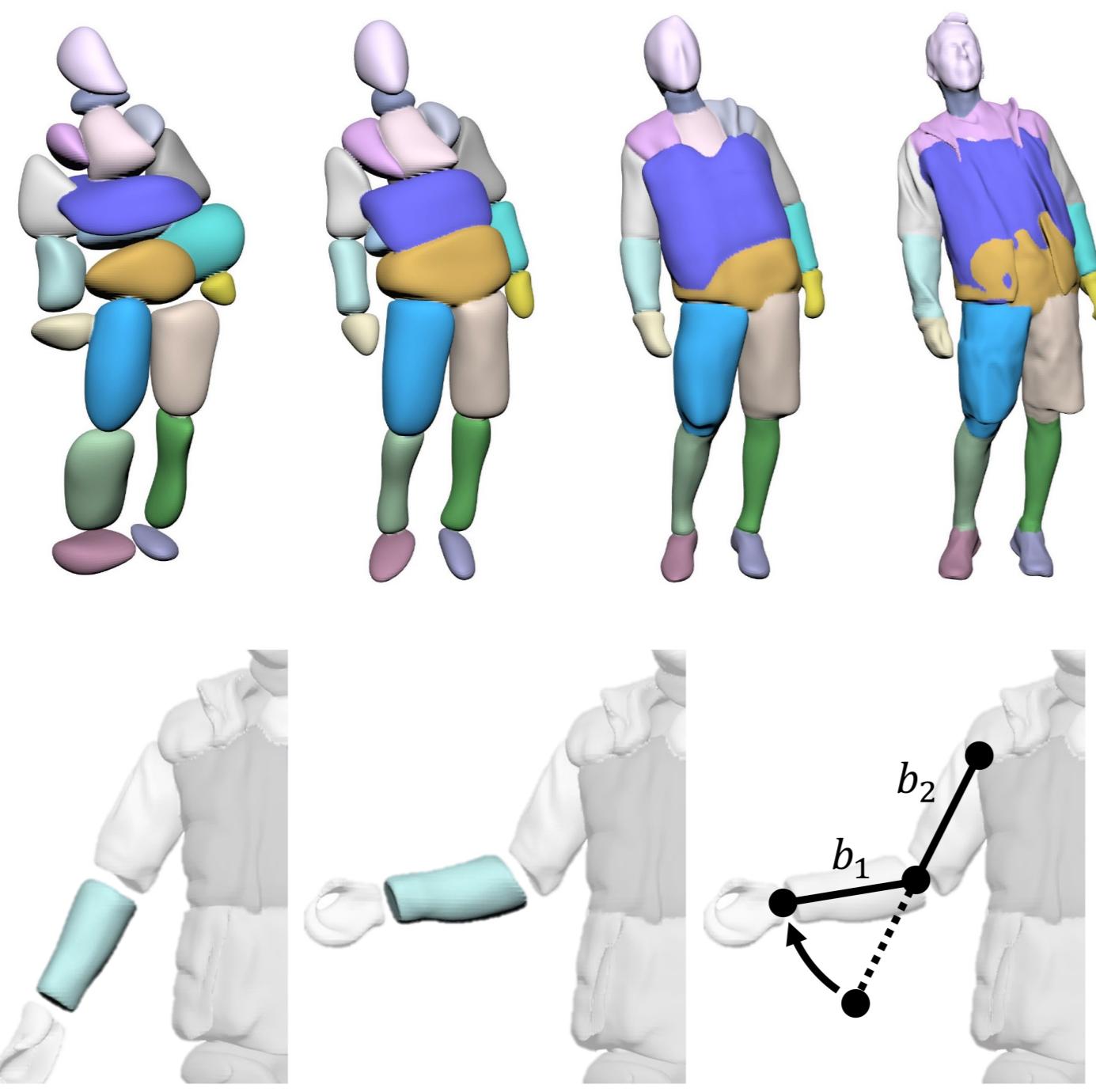
Challenges of Articulated Shape Representation

LBS-based:

- The neural skinning network for neural implicit functions barely generalize to novel poses.

Part-based:

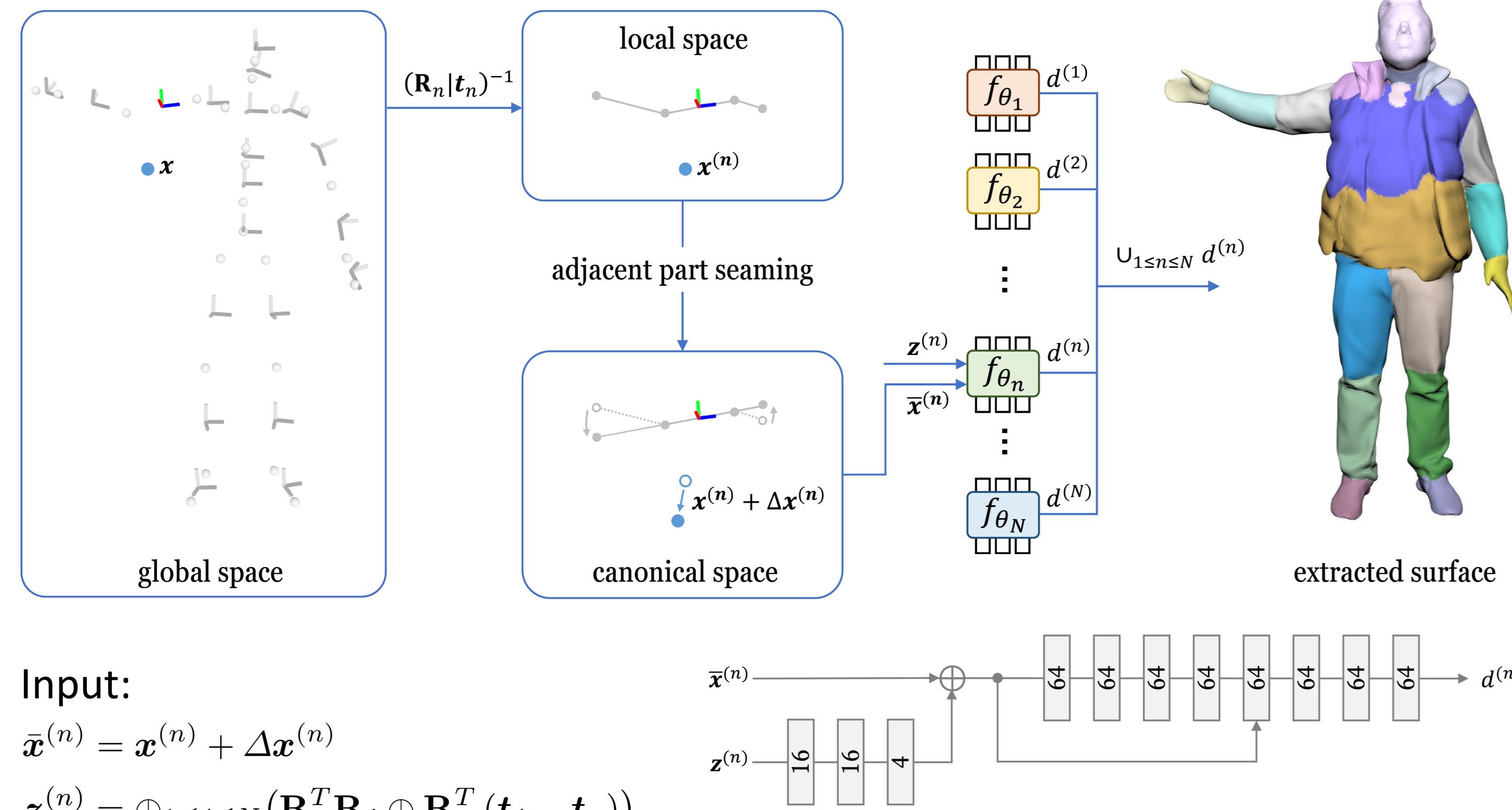
- Relying on ground-truth part labels.
- Lack of explicit modeling of non-rigid deformations.



Method

Pipeline

The data we use is a sequence of point clouds, which captures the shapes of a person in varying poses. For each frame, we first fit the body skeleton to the point cloud. Then, we set up local coordinate systems based on the skeleton and define a neural implicit function in each local space. Finally, we optimize the union of the functions to reconstruct the surface.



Input:

$$\bar{x}^{(n)} = x^{(n)} + \Delta x^{(n)}$$

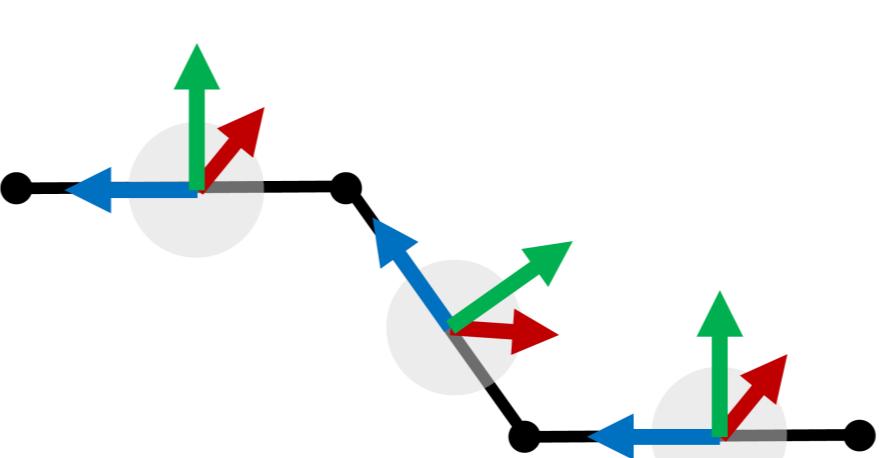
$$z^{(n)} = \oplus_{1 \leq j \leq N} (\mathbf{R}_n^T \mathbf{R}_j \oplus \mathbf{R}_n^T (t_j - t_n))$$

Supervision:

$$\mathcal{L}_{\text{recon}} = \frac{1}{|I|} \sum_{i \in I} (|d| + \lambda_{\text{normal}} \|\nabla_x d - n_i\|_2)$$

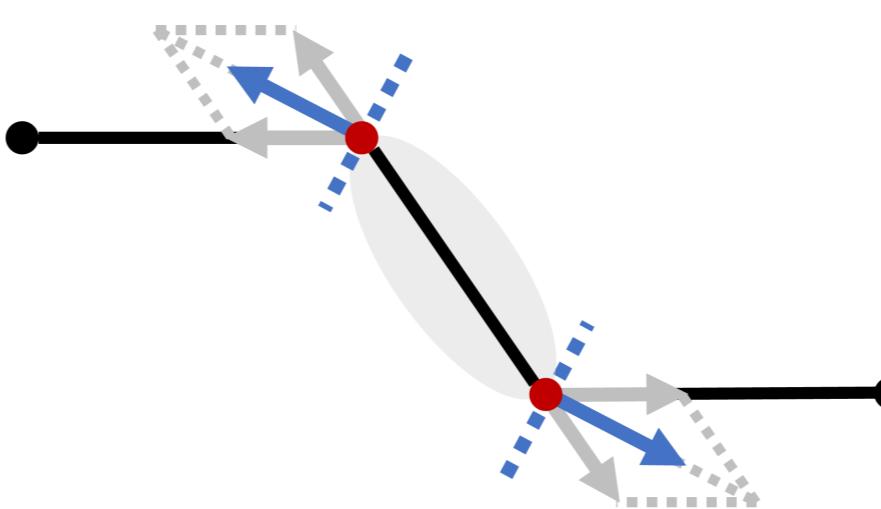
Partition-from-Motion

Bone-centered initialization: We initialize each part into a small sphere ($r = 0.01$) at the bone center. Then, parts are not intersected, and the SDF of a part approximately equals the distance to the bone center. This ensures that most points are assigned to the right part when training begins.



Bone limit loss and section normal loss: When two parts barely have relative motions in the training set, they are at high risk of overlapping. This leads to artifacts when the model is animated under novel poses. Therefore, we propose a bone limit loss and a section normal loss for regularization:

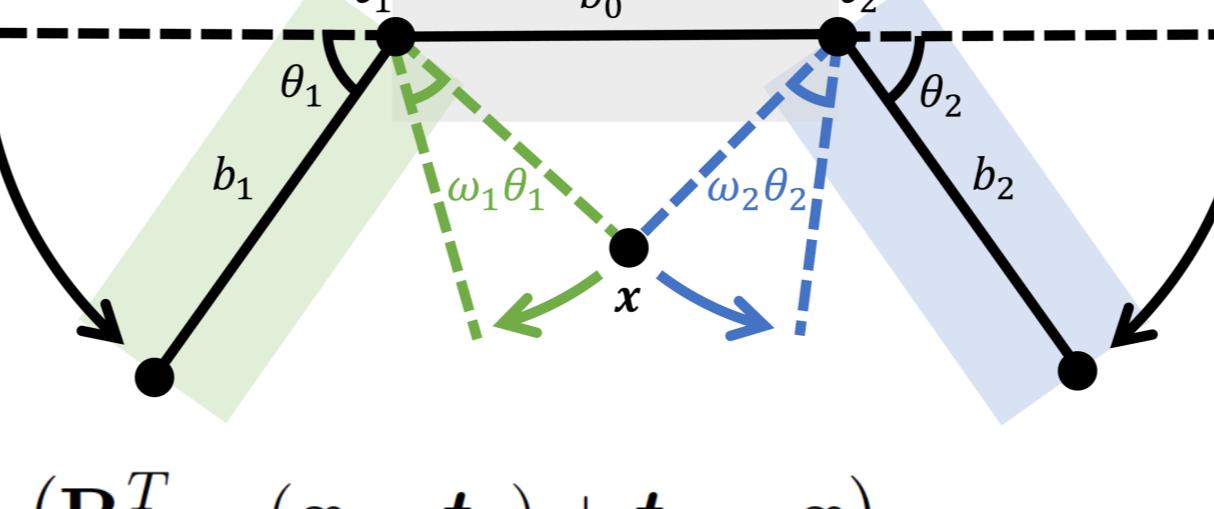
$$\mathcal{L}_{\text{lim}} = \frac{1}{N \cdot |J^{(n)}|} \sum_{n=1}^N \sum_{j \in J^{(n)}} |d_j^{(n)}|, \quad \mathcal{L}_{\text{sec}} = \frac{1}{N \cdot |J^{(n)}|} \sum_{n=1}^N \sum_{j \in J^{(n)}} \|\nabla_x d_j^{(n)} - n_j^{(n)}\|_2$$



Adjacent Part Seaming (APS)

Adjacent part seaming by local rotations:

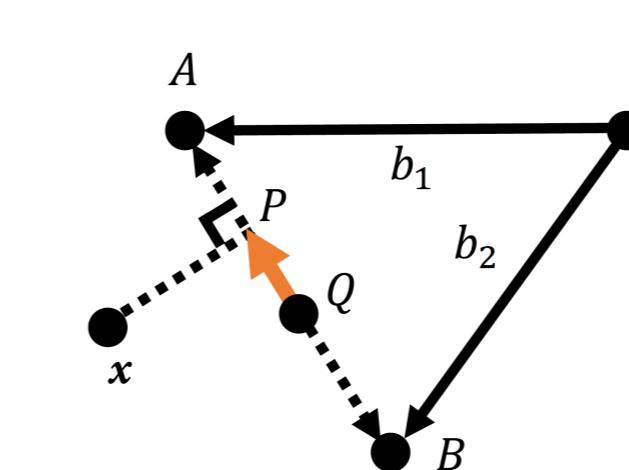
When trying to recover a point x on the bone b_0 to its original position, the point x is expected to go through two different rotations. We blend the offset vectors:



$$\Delta x = (\mathbf{R}_{w_1 \theta_1}^T (x - t_1) + t_1 - x) + (\mathbf{R}_{w_2 \theta_2}^T (x - t_2) + t_2 - x)$$

Blending weights from “Competing Parts”: We define the tendency of a point x to move with either bone b_1 or bone b_2 as

$$r_1 = \exp(\alpha_1 \frac{\vec{QP} \cdot \vec{QA}}{\|\vec{QA}\|^2} + \beta_1), \quad r_2 = \exp(\alpha_2 \frac{\vec{QP} \cdot \vec{QB}}{\|\vec{QB}\|^2} + \beta_2)$$



Then we define the blending weights of the point x with respect to bone b_1 and b_2 as

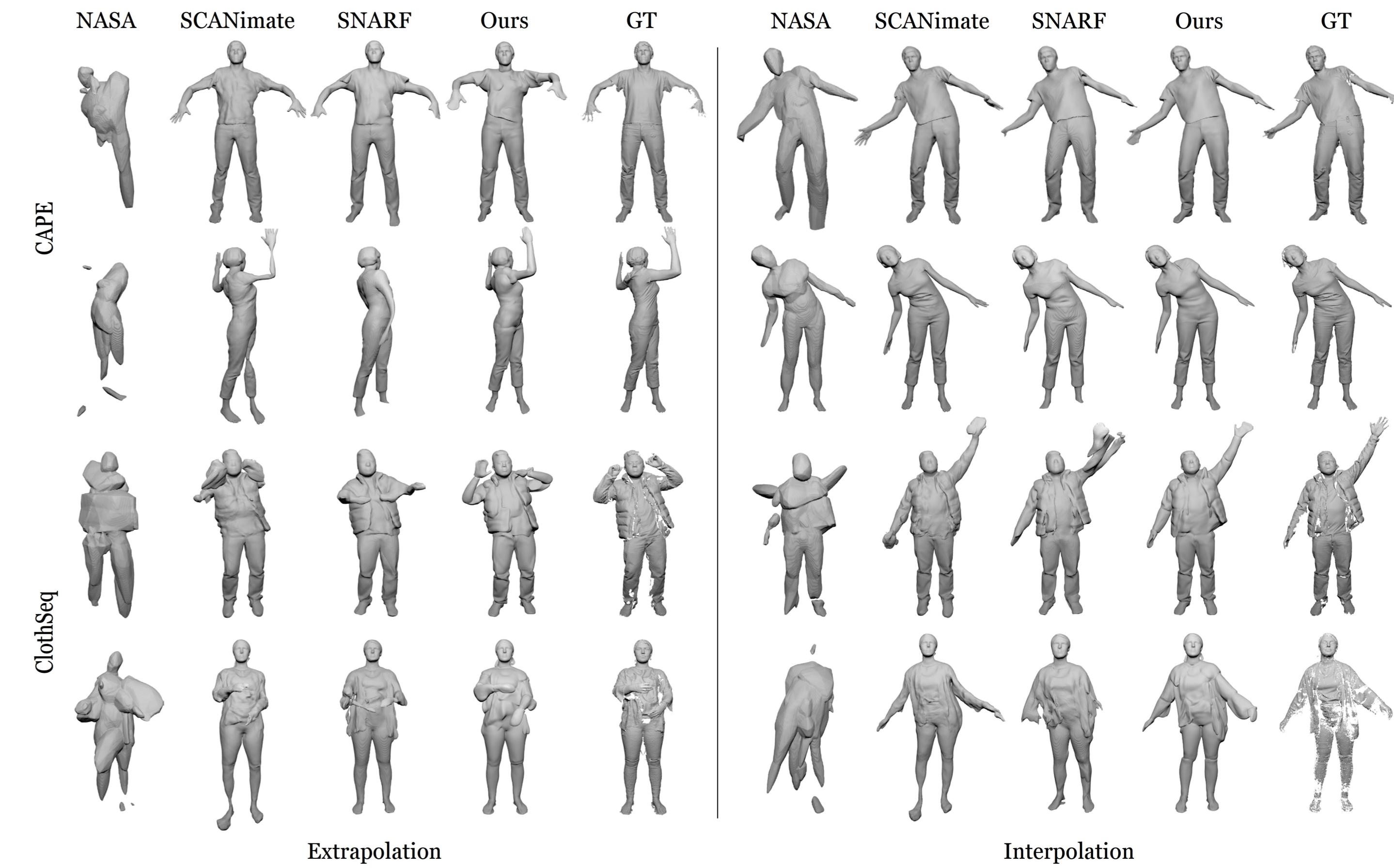
$$w_1 = \frac{r_1}{r_1 + r_2}, \quad w_2 = \frac{r_2}{r_1 + r_2}$$

Experiments

Comparisons on CAPE and ClothSeq datasets

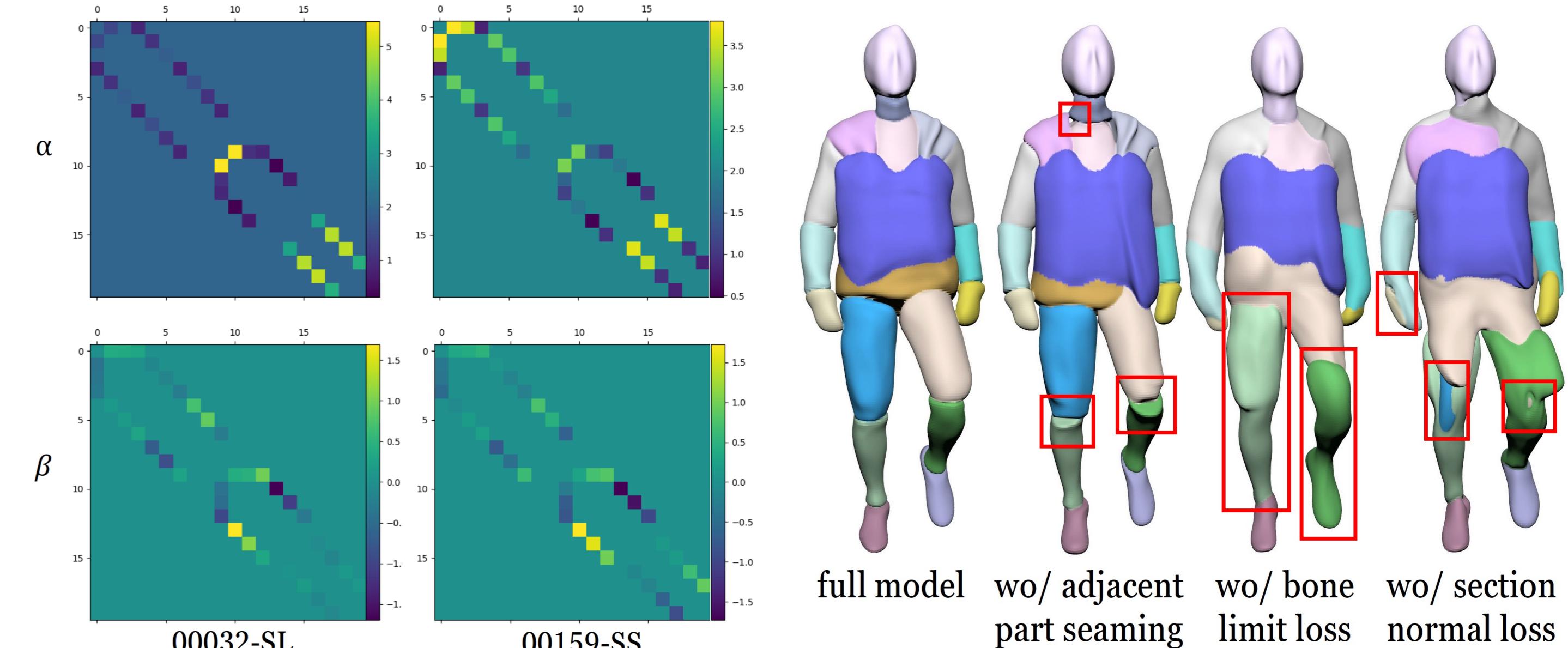
seq.	SCANimate			SNARF			NASA			Ours		
	CD↓	F1↑	p2s↓	Rec.↑	CD↓	F1↑	p2s↓	Rec.↑	CD↓	F1↑	p2s↓	Rec.↑
0032-SL	10.19	66.16	10.19	67.31	10.71	65.40	10.68	66.96	98.23	15.78	103.35	15.16
0032-SS	9.89	65.56	9.45	66.54	15.49	49.58	15.55	48.85	131.79	9.01	74.52	11.81
0096-SL	14.53	56.97	16.89	57.25	12.19	63.70	14.35	63.93	92.74	10.69	93.72	10.53
0096-SS	11.25	64.89	11.51	65.50	23.57	72.47	24.68	73.42	101.51	14.37	86.82	14.77
0159-SL	7.93	75.24	7.49	76.96	29.34	68.65	33.26	67.22	118.10	8.08	153.04	7.47
0159-SS	6.52	84.34	6.15	85.71	20.76	78.39	26.82	77.42	85.46	11.73	81.09	12.37
3223-SL	8.12	77.95	8.60	78.28	25.29	68.29	30.17	67.20	66.91	21.31	73.49	20.06
3223-SS	9.45	75.10	10.93	74.24	13.90	83.83	16.32	84.41	70.15	22.78	67.47	23.04
	6.80	85.81	6.80	88.76	4.93	95.51	5.06	97.93	10.00	74.16	10.01	75.38
	5.70	90.45	5.23	93.39	4.07	96.79	3.99	98.23	10.28	68.45	10.38	69.26
	8.48	89.50	10.69	91.94	6.48	96.93	8.92	98.07	15.47	61.22	18.34	62.08
	7.08	82.76	6.47	85.05	4.05	96.41	3.84	97.84	12.73	67.40	11.29	69.12
	5.18	91.79	4.35	96.01	3.77	96.35	3.18	99.27	11.82	66.37	11.39	69.81
	4.77	93.75	4.20	96.71	3.42	97.74	3.18	99.19	12.28	65.86	12.04	67.05
	5.31	93.40	5.26	96.81	5.06	95.70	5.86	95.84	8.17	84.47	7.92	85.61
	4.89	94.09	4.74	97.15	3.76	97.68	3.88	99.24	7.80	86.61	6.95	87.93
	14.33	56.25	14.29	58.02	17.72	58.49	21.58	59.16	69.76	16.88	68.24	16.59
	1.10	61.26	10.85	62.58	13.40	57.48	13.91	57.72	116.14	8.51	97.25	9.61
	14.32	54.06	14.19	54.80	15.06	60.22	16.47	60.19	65.73	19.93	39.26	21.29
	10.05	71.78	7.38	79.40	8.43	80.82	8.67	83.92	22.37	44.95	21.97	45.42
	8.84	74.84	7.77	78.33	8.81	80.20	7.86	81.80	33.66	31.61	34.35	31.34
	13.20	57.74	12.52	59.28	11.21	73.04	11.08	74.40	48.69	38.02	33.54	40.06

* E: pose extrapolation, I: pose interpolation



Visualizations

Optimized rigidness coefficients: the matrices of the scaling factor α are symmetric, while the matrices of the bias factor β are skew-symmetric.



Ablation study: We run experiments with each main component disabled and visualize parts in an unseen pose at the early stage of training.

