On Self-Contact and Human Pose

Lea Müller[1], Ahmed A. A. Osman[1], Siyu Tang[2], Chun-Hao P. Huang[1], Michael J. Black[1]

[1]Max Planck Institute for Intelligent Systems [2]ETH Zürich

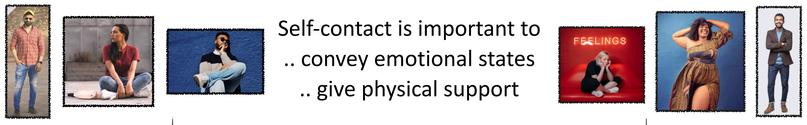
Code & Data
tuch.is.tue.mpg.de

CVPR VIRTUAL JUNE 19-25

PERCEIVING SYSTEMS
MAX PLANCK INSTITUTE FOR INTELLIGENT SYSTEMS

ETH zürich

MOTIVATION



Self-contact is important to
.. convey emotional states
.. give physical support

Why not studied? -> No suitable datasets

- Common labels (e.g. keypoints): No contact information
- Scans avoid self-contact (Table 1)
- Mocap. and Scans: No in-the-wild images
- Recently: discrete contact annotations (rare & no 3D mesh)

Dataset	Hasler et al. [1]	3D BodyTex [2]	SCAPE [3]	FAUST [4]
Meshes	520	400	70	500
Self-contact meshes	0	3	0	60

Table 1. 3D scan datasets with mesh count (total and self-contact).

CONTRIBUTION

- Three new datasets: 3DCP, MTP, and DSC
- "Mimic The Pose" & SMPLify-XMC (images in-the-wild & 3D reference poses)
- TUCH: the first HPS regressor for self-contact poses trained end-to-end
- Code and data available

REFERENCES

[1] Hasler et al., A statistical model of human pose and body shape, CGF 2009
[2] Ahmed et al., A survey on deep learning advances on different 3d data representations, CoRR 2018
[3] Anguelov et al., SCAPE: Shape Completion and Animation of PEople, TOG 2005
[4] Bogo et al., FAUST: Dataset and evaluation for 3D mesh registration, CVPR 2014
[5] Mahmood et al., AMASS: Archive of motion capture as surface shapes, ICCV 2019
[6] Pavlakos et al., Expressive body capture: 3d hands, face, and body from a single image, CVPR 2019
[7] Kolotouros et al., Learning to Reconstruct 3D Human Pose and Shape via Model-fitting in the Loop, ICCV 2019

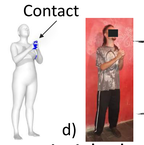
NEW DATASETS

3DCP - 3D meshes with self-contact in SMPL-X topology
a) scan six subjects in self-contact poses (3DCP Scan)

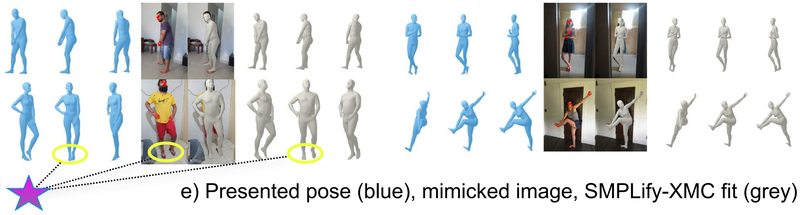
188 scans with self-contact

b) refine near-contact poses select from AMASS [5] (3DCP Mocap)

MTP - images in the wild with 3D reference pose
c) render meshes from 3DCP (presented pose)
d) ask Workers on Mechanical Turk to mimic poses
e) apply **SMPLify-XMC**
.. adapts SMPLify-X [6] optimisation
.. corrects small differences between presented and mimicked pose
.. uses presented pose and contact and subject measurements

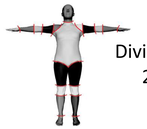


3.7K MTP images



e) Presented pose (blue), mimicked image, SMPLify-XMC fit (grey)

DSC - discrete self-contact annotations

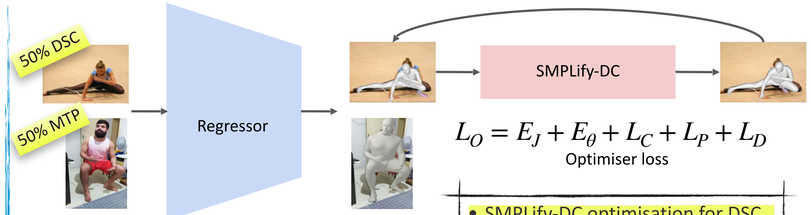


Divide mesh into 24 regions

Label pairwise contact

30K labels for LSP & DeepFashion

REGRESSOR



$$L_R = E_J + E_\theta + E_\beta + L_C + L_P$$
Regressor loss

$$L_O = E_J + E_\theta + L_C + L_P + L_D$$
Optimiser loss

- SMPLify-DC optimisation for DSC data (SPIN [7] design)
- MTP data treated as ground truth

TUCH highlights:

- SMPLify-DC for DSC data
  - Minimise minimum distance between regions in contact, $L_D =$

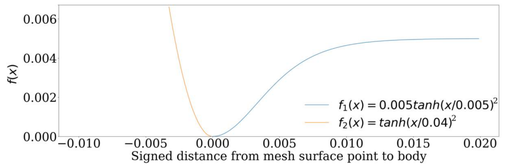
$$\sum_{i=1}^{K} \sum_{j=i+1}^{K} S_{ij} \min_{v \in R_i,\, u \in R_j} (||v - u||)^2$$

- Contact energy terms
  - L_C encourages self-contact for vertices that are outside and near each other

$$\mathcal{L}_C = \sum_{p_i \in M_{D_I^0}} \alpha_1 \tanh\left(\frac{f_g(p_i)}{\alpha_2}\right)^2,$$

  - L_P pushes inside vertices to the surface

$$\mathcal{L}_P = \sum_{p_i \in M_{D_I}} \beta_1 \tanh\left(\frac{f_g(p_i)}{\beta_2}\right)^2,$$



RESULTS

Largest improvement

	MPJPE				PA-MPJPE			
	contact	no contact	unclear	total	contact	no contact	unclear	total
SPIN	100.2	95.5	96.7	96.9	59.1	61.7	55.7	59.2
TUCH	**85.1**	**86.6**	**81.9**	**84.9**	**54.1**	**58.6**	**51.2**	**55.5**

Table 2. Quantitative results on 3DPW.