# PICO: Reconstructing 3D People in Contact with Objects


Alpár Cseke[1,2*] Shashank Tripathi[1*†] Sai Kumar Dwivedi[1] Arjun S. Lakshmipathy[3] Agniv Chatterjee[4] Michael J. Black[1] Dimitrios Tzionas[5]

*equal contribution †project lead

[1]Max Planck Institute for Intelligent Systems, Tübingen [2]Meshcapade [3]Carnegie Mellon University [4]UT Austin [5]University of Amsterdam


## Key Insight

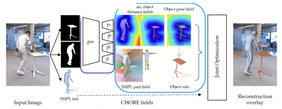
**3D contact correspondences** are crucial for reconstructing human-object interactions (HOI) in the wild

Color-coded 3D contact correspondences between human and object

## Problem

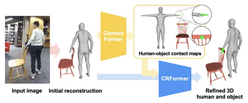
### Previous work

PHOSA, ECCV 2020 [1]



CHORE, CVPR 2023 [2]

Procigen, CVPR 2024 [3]



CONTHO, CVPR 2024 [4]

### Common limitations

- ✕ Assumes known object mesh
- ✕ Uses SMPL (flat hands)
- ✕ Relies on manual heuristics
- ✕ Restricted to a small number of object categories
- ✕ Does not generalize well to natural images (in-lab, except PHOSA)

### How to improve?

1 Get dense 3D contact correspondences for in-the-wild human-object interactions. **PICO-db**

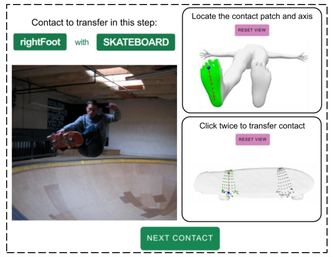
2 Leverage such data for reconstructing interactions in 3D. **PICO-fit**

## PICO-db: Dataset of Vertex-level HOI Contact Correspondences

### Annotation Process

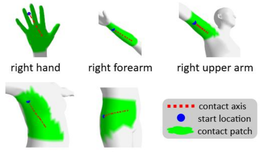
**Step 1:** Pick best object from top 3 *OpenShape* [6] similarity matches from *Objaverse-LVIS* [7] ; Query object size from a VLM

**Step 2:** Simple two-click contact transfer on our web UI using the *"contact axis"* [9]




### 3D Contact Parametrization


Break the human contact into patches and parametrize them via *"contact axes"* [9]

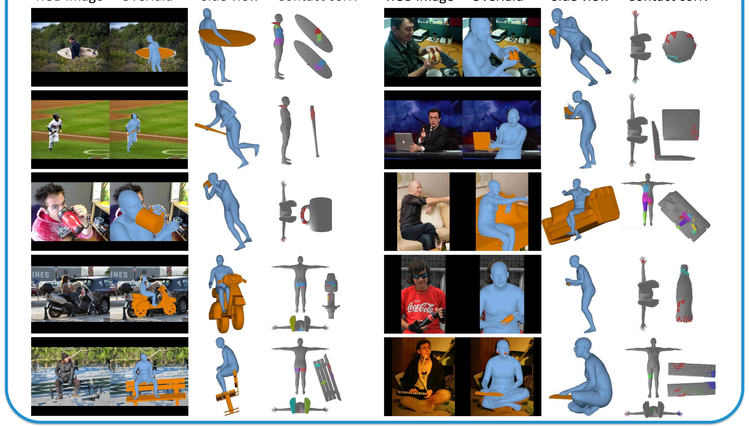



✓ Establishes 1:1 vertex-level correspondence on transfer

### The Dataset


**4000+ images**
**44 object categories**
**600+ unique object instances**

## PICO-fit Reconstructions

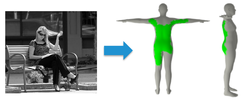



## PICO-fit: 3D HOI Reconstruction from Natural Images

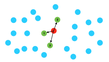
### I. Contact Pair Initialization

**Finding the best proxy contact from PICO-db:**

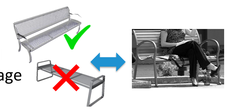
1. Predict human contact using DECO [5]



2. Nearest neighbor search: find K closest human contacts in PICO-db within the object category



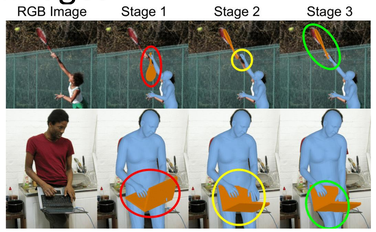
3. Select the best one, considering OpenShape [6] object match with the image

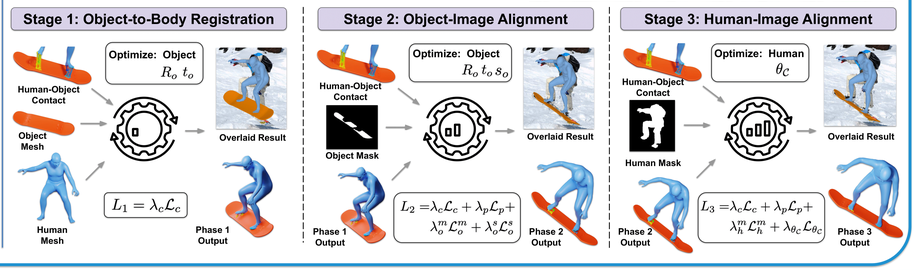


4. Take object mesh, object contact, and corresponding human contact from that sample

### II. Reconstruction Pipeline

1. Get object and contact pair
2. Predict SMPL-X human pose and shape from OSX [8]
3. Multi-stage optimization pipeline:
   - **Stage 1:** Minimize contact distance between human and object
   - **Stage 2:** Align 6DoF object pose using image cues
   - **Stage 3:** Fine-tune human limbs in contact using image cues

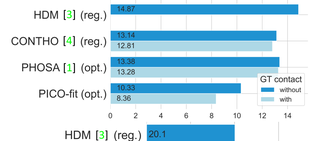



## Evaluations

**Reconstruction on Intercap**

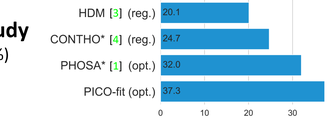
*PA-CD*$_{h+o}$ (cm) ↓



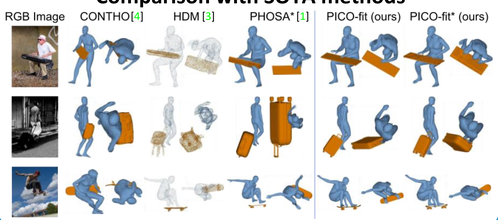

**Perceptual study**

preference rate (%)

X vs PICO-fit*




**Comparison with SOTA methods**




*Note:* methods marked by * use ground truth contact for reconstruction


[1] Zhang et al., Perceiving 3D human-object spatial arrangements from a single image in the wild., ECCV 2020
[2] Xie et al., CHORE: Contact, human and object reconstruction from a single RGB image., ECCV 2022
[3] Xie et al., Template free reconstruction of human object interaction with procedural interaction generation., CVPR 2024
[4] Nam, Jung et al., Joint reconstruction of 3D human and object via contact-based refinement transformer., CVPR 2024
[5] Tripathi, Chatterjee et al., DECO: Dense estimation of 3D human-scene contact in the wild., ICCV 2023
[6] Liu, Shi, Kuang et al., OpenShape: Scaling Up 3D Shape Representation Towards Open-World Understanding., NeurIPS 2023
[7] Deitke et al., Objaverse: A Universe of Annotated 3D Objects., CVPR 2023
[8] Lin et al., One-stage 3D whole-body mesh recovery with component aware transformer., CVPR 2023
[9] Lakshmipathy et al., Contact Edit: Artist Tools for Intuitive Modeling of Hand-Object Interactions., SIGGRAPH 2023


https://pico.is.tue.mpg.de