Ca3OH1

Contact Aware in 3D Object-Human Interaction

Thesis Proposal

Keywords: 3D human reconstruction, human-object interaction, contact aware geometric constraints, pose optimization,

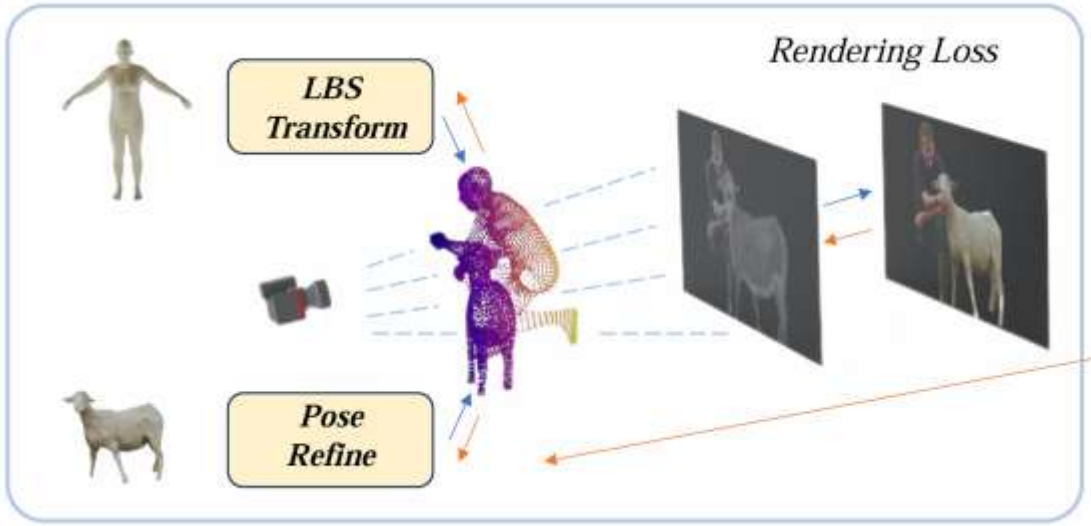
HCY WRT ZBH

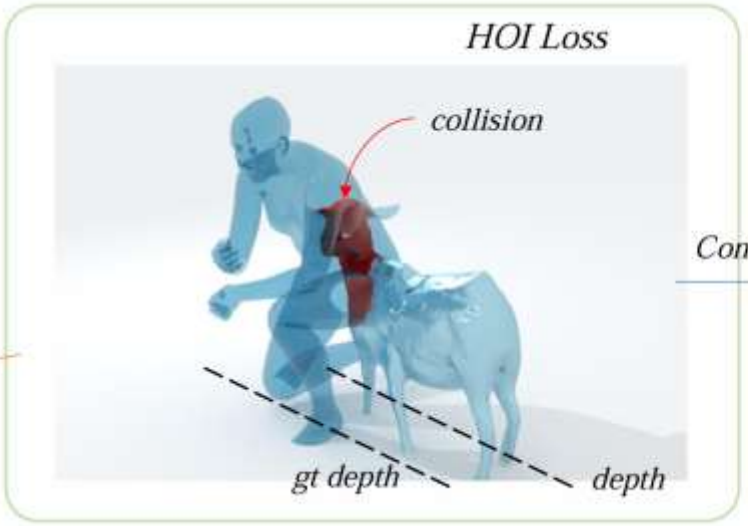
CONT
-ENTE

01

Backgrounds

- Base Research
- ‘Sandwich’ Problem





Loss Function

1. Rendering Loss → 2D alignment

- **L_1**: between Img_s and Img_{GT}
- **L_2**: between $Mask_s$ and $Mask_{GT}$
- **SSIM**: *Structural Similarity Index* Loss
- **LPIPS**: *Learned Perceptual Image Patch Similarity* Loss

$$L_r = w_h L_1 + w_o L_2 + w_{ho} L_{ho}$$

2. HOI Loss

$$L_{hoi} = L_{cont} + L_{colli} + L_{depth}$$

(a) Contact Loss L_{cont}

Chamfer distance between the *human contact area* and *object*.

(b) Collision Loss L_{colli}

Meshes do not intersect.

(c) Ordinal Depth Loss L_{depth}

Depth order consistency constraint.

3. Overall Loss

$$L = w_r L_r + w_{hoi} L_{hoi}$$

HOI-Gaussian Optimizer

Reconstructing In-the-Wild Open-Vocabulary Human-Object Interactions

Boran Wen[1,2] Dingbang Huang[1] Zichen Zhang[1] Jiahong Zhou[1] Jianbin Deng[1]
Jingyu Gong[2] Yulong Chen[3*] Lizhuang Ma[1*] Yong-Lu Li[1,2*]
[1]Shanghai Jiao Tong University, [2]Shanghai Innovation Institute, [3]East China Normal University
{wenboran, 2333970490, zhoujiahong, 1261021437, dengjb42}@sjtu.edu.cn
{llong.gm, lma, yonglu_li}@sjtu.edu.cn, yygong@cs.ecnu.edu.cn

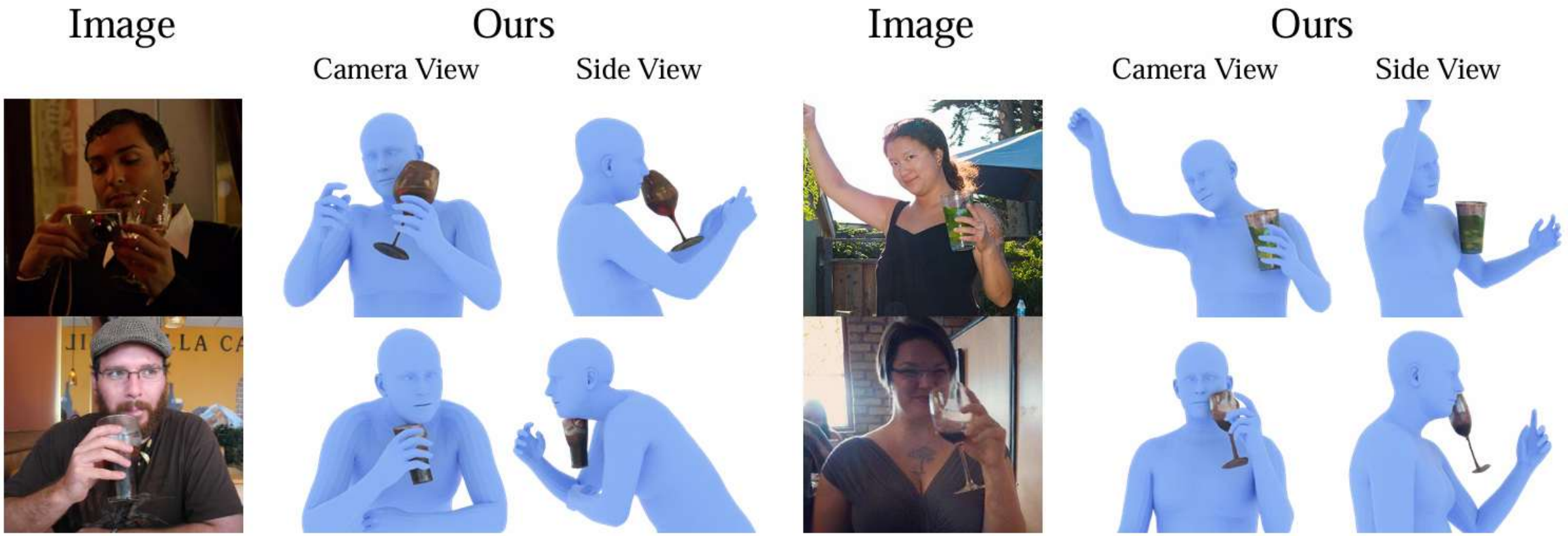


Figure 21. Failure cases of HOI-Gaussian.

C.2. Failure Cases

Fig. 21 shows some failure cases of our HOI-Gaussian optimizer. In these cases, human body parts occlude each other severely, and the object happens to be located between the occluded areas, which becomes challenging to determine which body part the object should contact with.

Sandwich

Contact in Gaussian Model

- We can easily identify areas where there's no interaction
- Occlude means lower opacity α
- Stable pos relationship between human and object

⇒Contact Interaction Score $$c = w_\alpha \cdot Norm(\alpha^h) + w_d \cdot d_C(p^h, p^o)^h$$

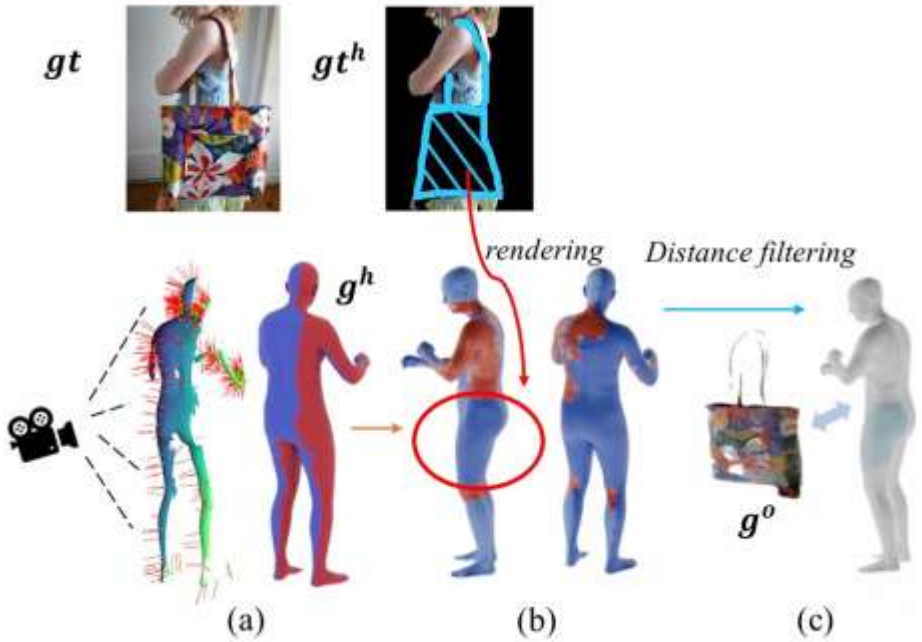
where Norm means the normalization of the vector to a range between 0 and 1,

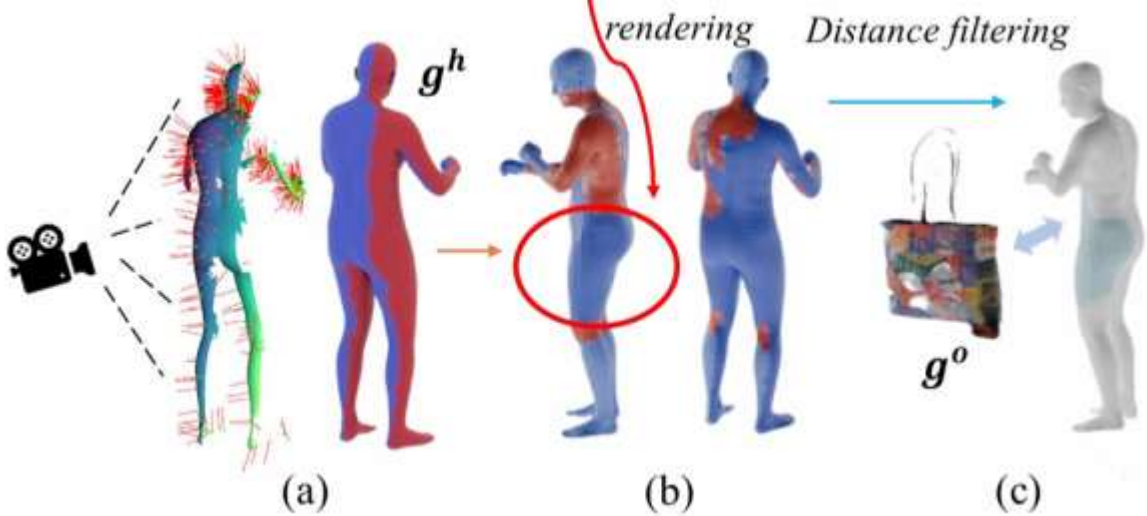
and d_C means the Chamfer distance.

(a) Opacity initialization using human normals.

(b) The distribution of human body point cloud opacity scores is visualized to identify the blue region as a potential inter action area.

(c) Based on the approximate distance between the human body and the object, the optimized contact region is further identified, shown in light blue.





Camera View
Side View
Camera View
Side View

02

Research Plan

- What to do

What TODO:

1. 条件分支扩展

在人体重建模型的**编码器或隐空间**中新增物体 mesh 分支，通过以下步骤优化重建：

- 提取物体 mesh 的几何特征（如表面法线、关键点、体积信息）；
- 将物体特征与人体特征进行融合（如注意力机制、特征拼接）；
- 通过融合特征调整人体姿态预测结果，确保人体与物体的交互合理性。

2. 交互几何约束

对人体重建结果施加**额外的交互几何约束**，优化人体与物体的空间关系：

- 接触约束：确保交互部位（如手部）与物体表面存在合理接触，避免悬浮或穿透；
- ~~穿插惩罚：对人体 mesh 与物体 mesh 的重叠区域施加惩罚项，减少几何冲突；~~
- 先验约束：结合少量人工标注或启发式规则（如"握持时手部应包裹物体"），定义交互先验，引导重建方向。

(a) Contact Loss L_{cont}

Chamfer distance between the *human contact area* and *object*.

(b) Collision Loss L_{colli}

Meshes do not intersect.

3. 基于数据增强/后处理的优化

通过数据集预处理或后处理模块，提升人体与物体的匹配精度：

- 数据增强：对 BEHAVE/InterCap 数据集进行人工标注补充（如交互关键点、接触区域），扩大约束样本；
- 后处理微调：在现有重建结果基础上，设计优化模块（如迭代最近点算法、姿态优化器），微调人体关节位置与物体相对关系；
- 误差修正：通过分析重建误差（如手部-物体距离），建立误差映射模型，实现针对性修正。

What TODO:

Foundation:

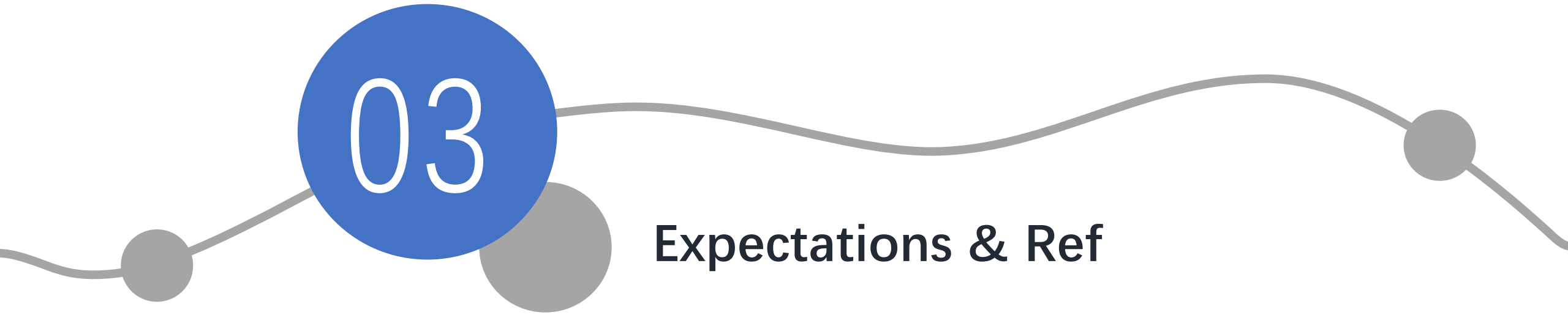
- Collect the ‘Sandwich’ cases in open datasets.
- Get the intermediate results to verify our conjecture.

Main:

- Improve **Contact in Gaussian Model**
- Add a **mesh branch** to the encoder or latent space of the human reconstruction model.
- **Prior constraints**: Combine a small number of manual annotations or heuristic rules to define interactive priors and guide the reconstruction direction.

Polishing:

- Dataset **Pre**processing & **Post**-processing module



Expectations & Ref

Expectation:

High:

- Fix the ‘Sandwich’ problem partly without adding prior constraints.
- With prior constraints automatically added, ensure a high accuracy.
- Adjust the dataset to provide knowledge of prior constraints.

Low:

- With prior constraints (auto/human) added, ensure a high accuracy.
- Improve the performance with better Loss and pipeline.
- Adjust the dataset to provide knowledge of prior constraints.

3DHOI

用途：人体-物体交互分析的基础工作，提供交互语义与几何关联的参考

Project Page: wenboran2002.github.io/3dhoi/

SMPL eXpressive Home Download SMPL-X Model License SMPL-X Body License

Expressive Body Capture:

3D Hands, Face, and Body from a Single Image

G. Pavlakos*, **V. Choutas***, **N. Ghorbani**, **T. Bolkart**, **A. A. A. Osman**, **D. Tzionas** and **M. J. Black** (*authors contributed equally)
Computer Vision and Pattern Recognition (CVPR) 2019, Long Beach, CA

Ca3OH1

Contact Aware in 3D Object-Human Interaction

Thesis Proposal

Keywords: 3D human reconstruction, human-object interaction, contact aware geometric constraints, pose optimization,

HCY WRT ZBH

01

Contact Area

Contact in Gaussian Model

- We can easily identify areas where there's no interaction
- Occlude means lower opacity α
- Stable pos relationship between human and object

⇒Contact Interaction Score $$c = w_\alpha \cdot Norm(\alpha^h) + w_d \cdot d_C(p^h, p^o)^h$$

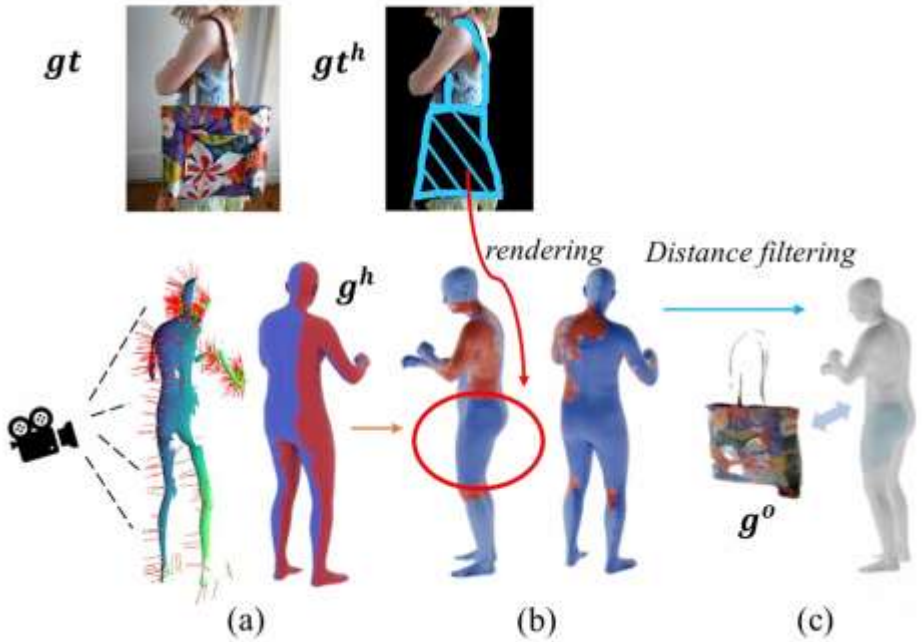
where Norm means the normalization of the vector to a range between 0 and 1,

and d_C means the Chamfer distance.

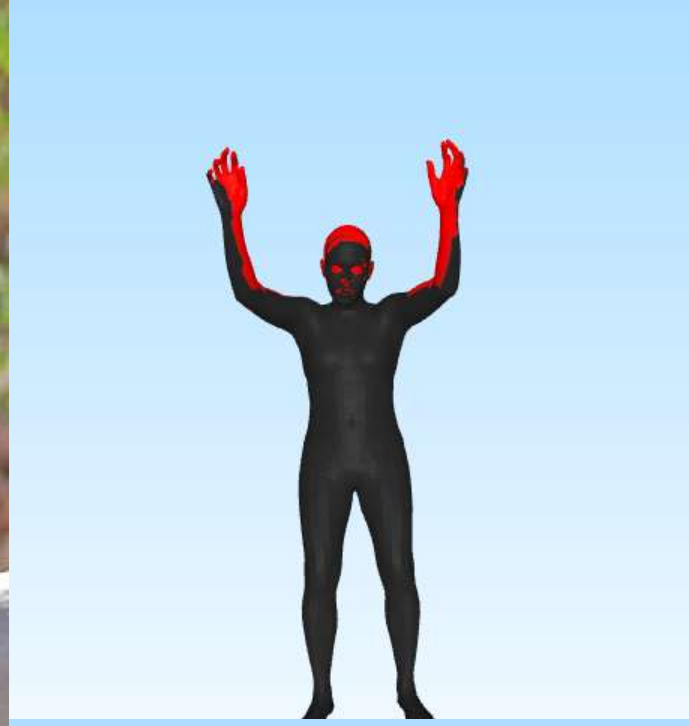
(a) Opacity initialization using human normals.

(b) The distribution of human body point cloud opacity scores is visualized to identify the blue region as a potential inter action area.

(c) Based on the approximate distance between the human body and the object, the optimized contact region is further identified, shown in light blue.




iStock
Credit: paresh3d
152496940



02

Only ‘Sandwich’ ?


vevo

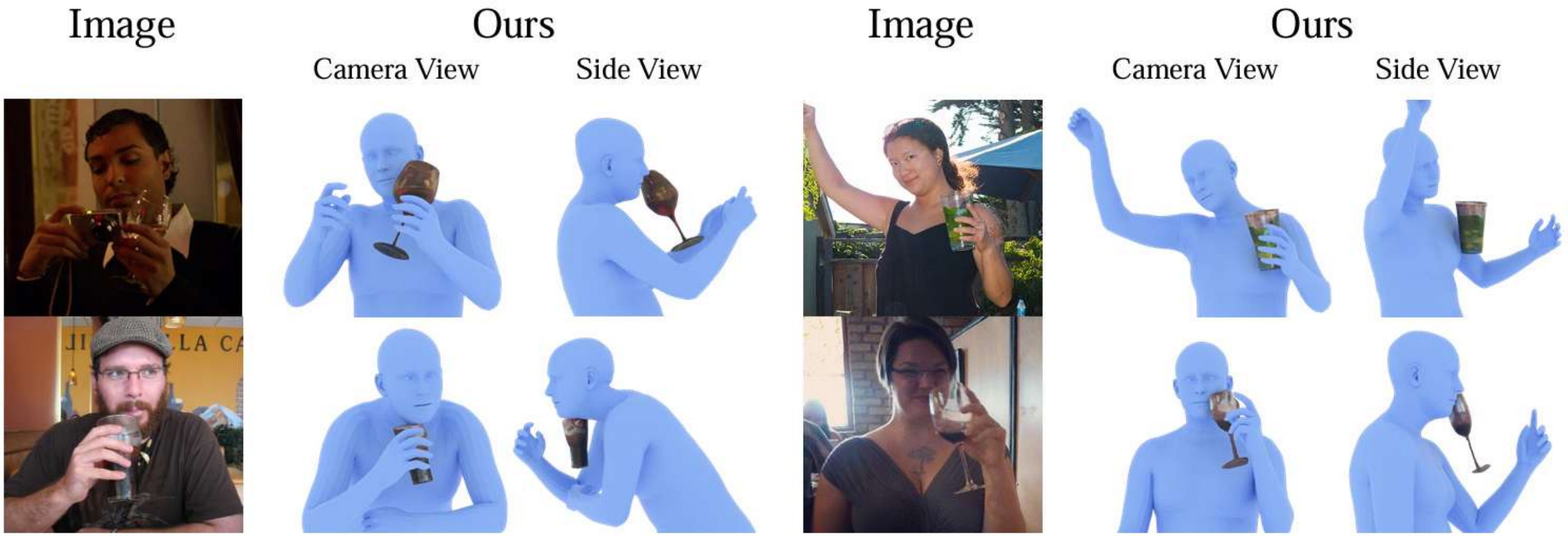


Figure 21. Failure cases of HOI-Gaussian.

C.2. Failure Cases

Fig. 21 shows some failure cases of our HOI-Gaussian optimizer. In these cases, human body parts occlude each other severely, and the object happens to be located between the occluded areas, which becomes challenging to determine which body part the object should contact with.

Sandwich

人体在优化过程中没有改变

sheep
- box_annotation.json
- calibration.json
- extrinsic.json
- h_mesh.obj
- image.jpg
- normals_smplx.npy
- obj_annot1.obj
- obj_pcd_h_align.obj
- object_mask.png
- person_mask.png
- smplx_parameters.json

Dataset Structure:

```
- motorcycle
  - HICO_train2015_00013672
    - smplx_parameters.json # SMPL-X paramet
    - h_mesh.obj # Human mesh
    - object_mesh.obj # Object mesh
    - image.jpg # Image
    - person_mask.png # Person mask save as
    - object_mask.png # Object mask save as
    - depth.png # Depth generated by ZoeDept
    - box_annotation.json # Bounding box ann
```

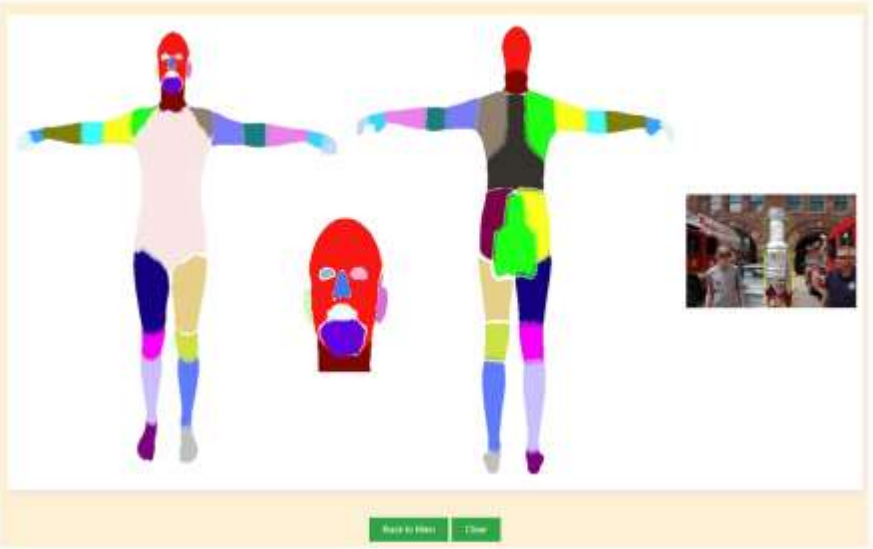
Expectation:

High:

- Adding mesh(.obj) & additional Loss to VLM model to fix the problem in human reconstruction before HOI reconstruction.
- Hopefully, this will solve the ‘Sandwich’ problem at the same time.

Low:

- Train a model to figure out the exact contact area. (‘Sandwich’ Only)
    1. (HxWxRGB -> the id of contact area)
    2. (HxWxRGB -> every vertex-10457)



Ca3OH1

Contact Aware in 3D Object-Human Interaction

Thesis Proposal

Keywords: 3D human reconstruction, human-object interaction, contact aware geometric constraints, pose optimization,

HCY WRT ZBH

Done this week:

Foundation:

- Collect the ‘Sandwich’ cases in open datasets.

  (And the data-selector script)

- Get the intermediate results to verify our conjecture.

  (And the test pipeline)

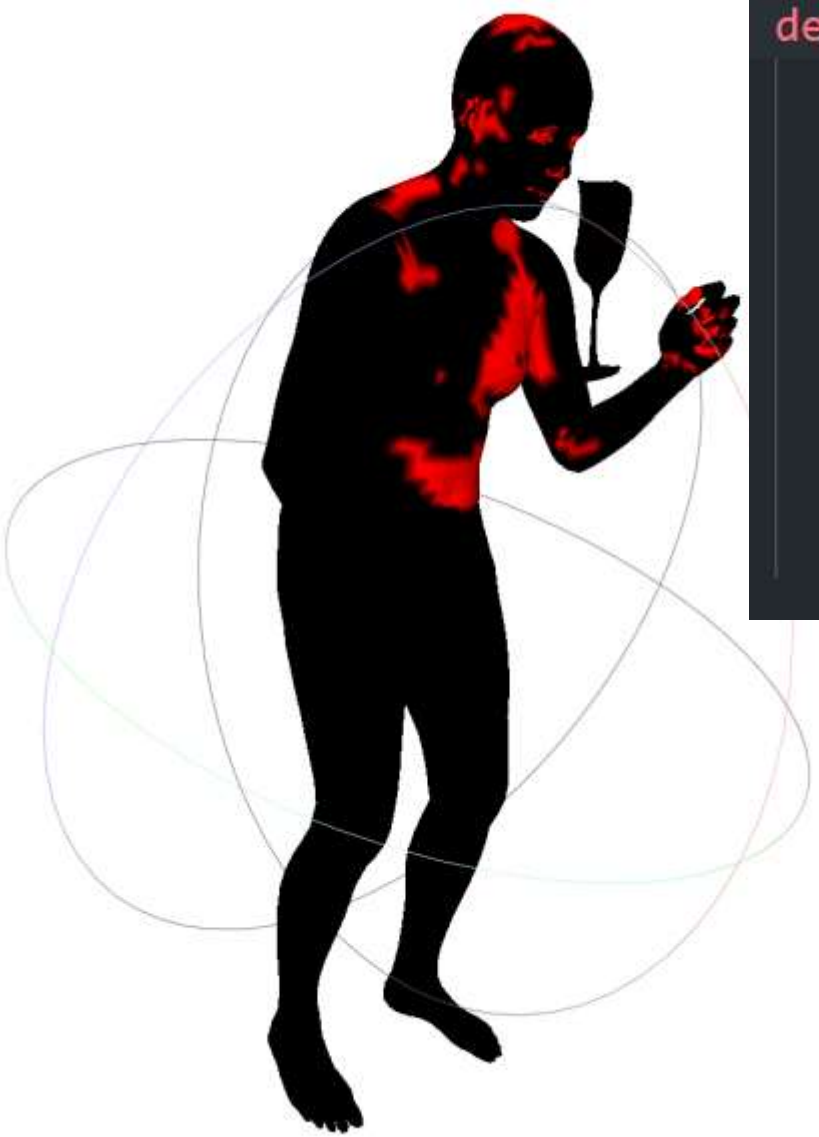
Main:

- A way to get **Contact Area GT**.
- Fix the problem that contact loss **Never** be optimized.
- The result with **Contact Area GT**.

01

Contact Area GT

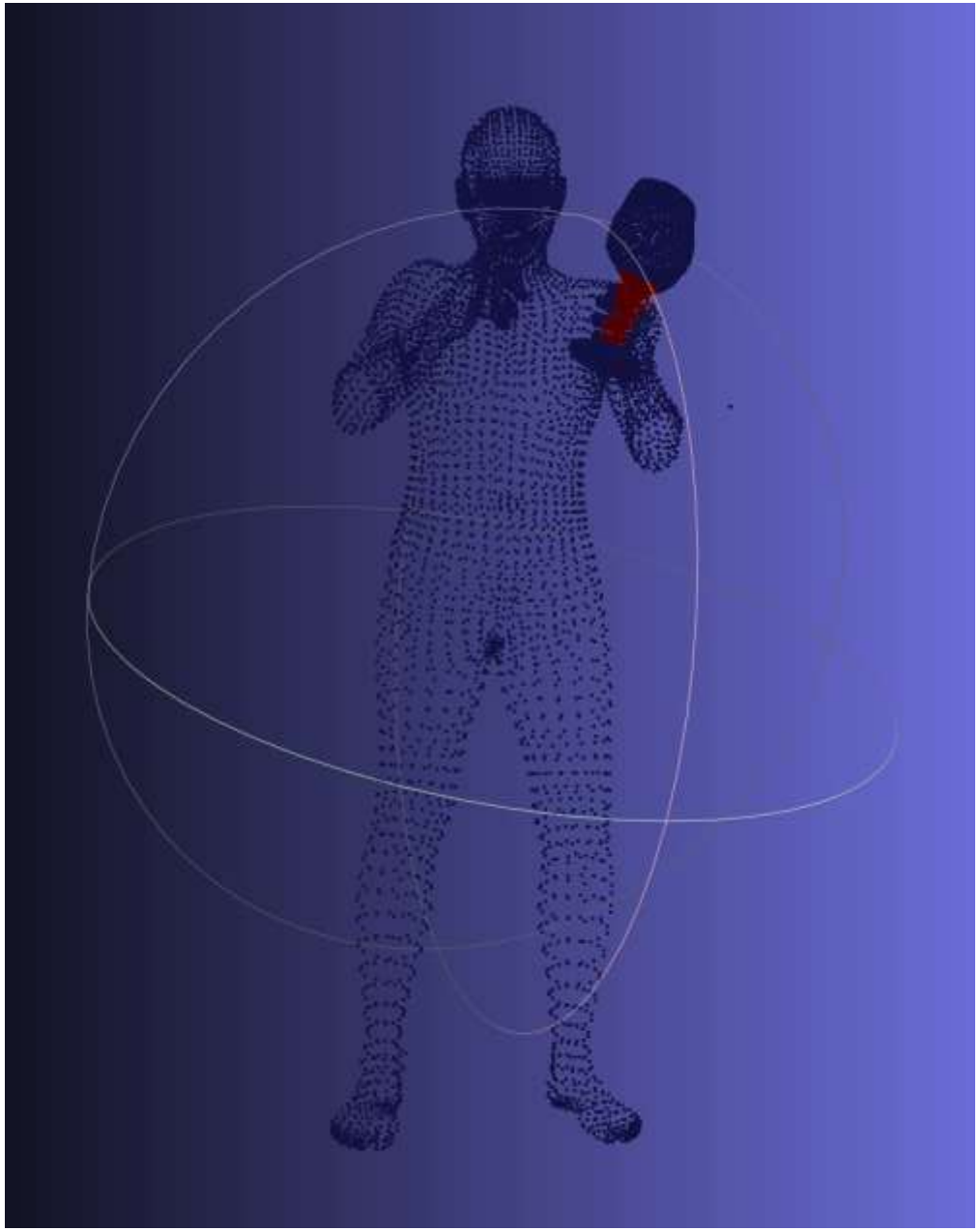
Contact Area on Object Matters too:



```python
def compute_contact_loss(self, hverts, overts,h_contact,o_contact):
    h_v_contact = hverts[h_contact]
    h_v_contact=h_v_contact.unsqueeze(0)
    o_v_contact = overts[o_contact]
    o_v_contact=o_v_contact.unsqueeze(0)
    hdist,odist=chamfer_distance(h_v_contact,o_v_contact)
    ho_distance = hdist.mean() + odist.mean()

    return {"loss_contact": ho_distance}
```

The way to get CAGT:



02

Experiments

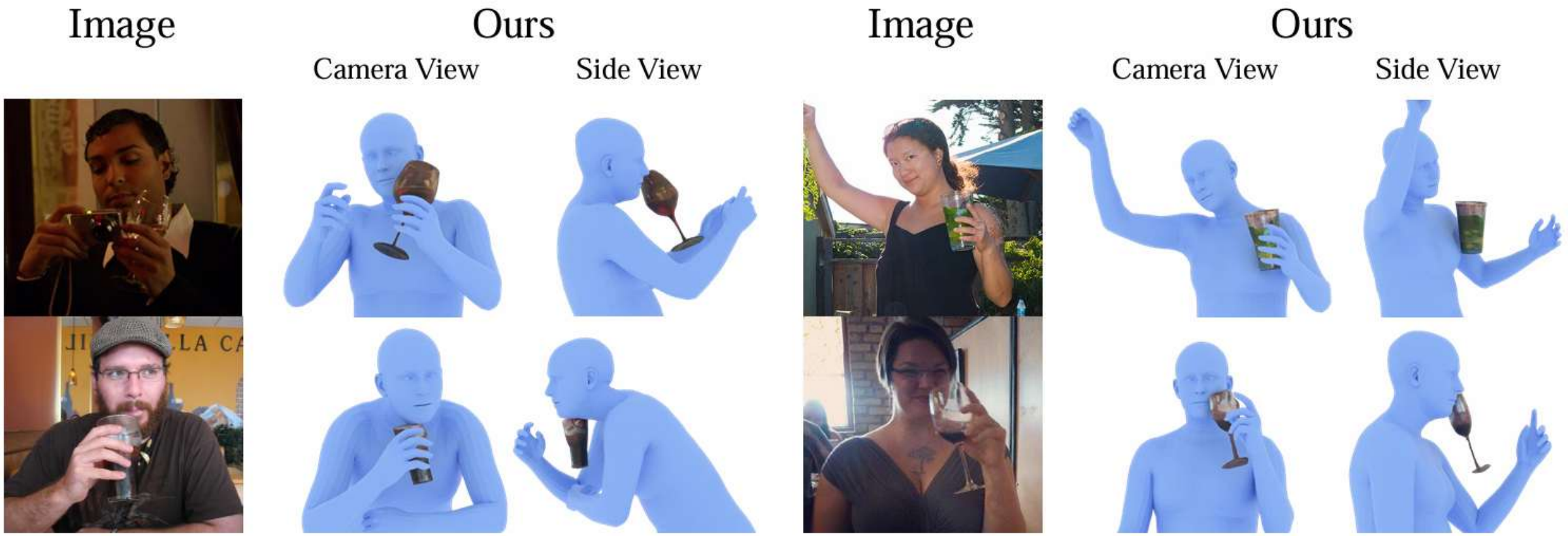


Figure 21. Failure cases of HOI-Gaussian.

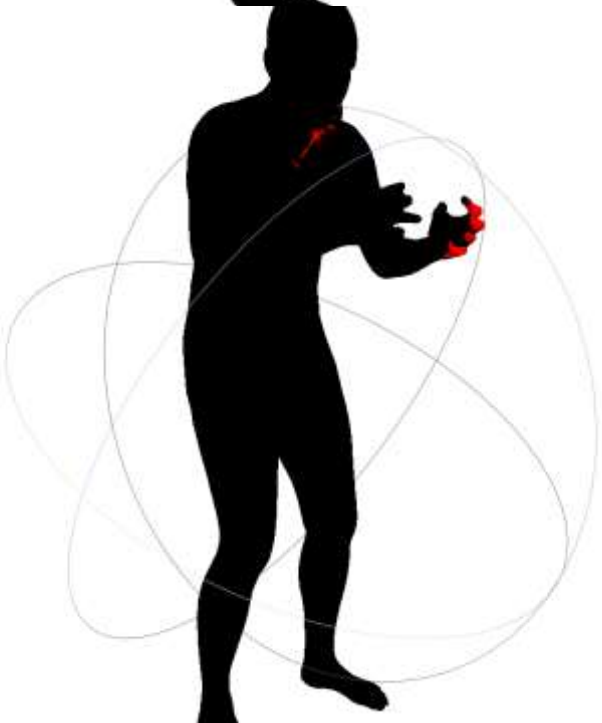
C.2. Failure Cases

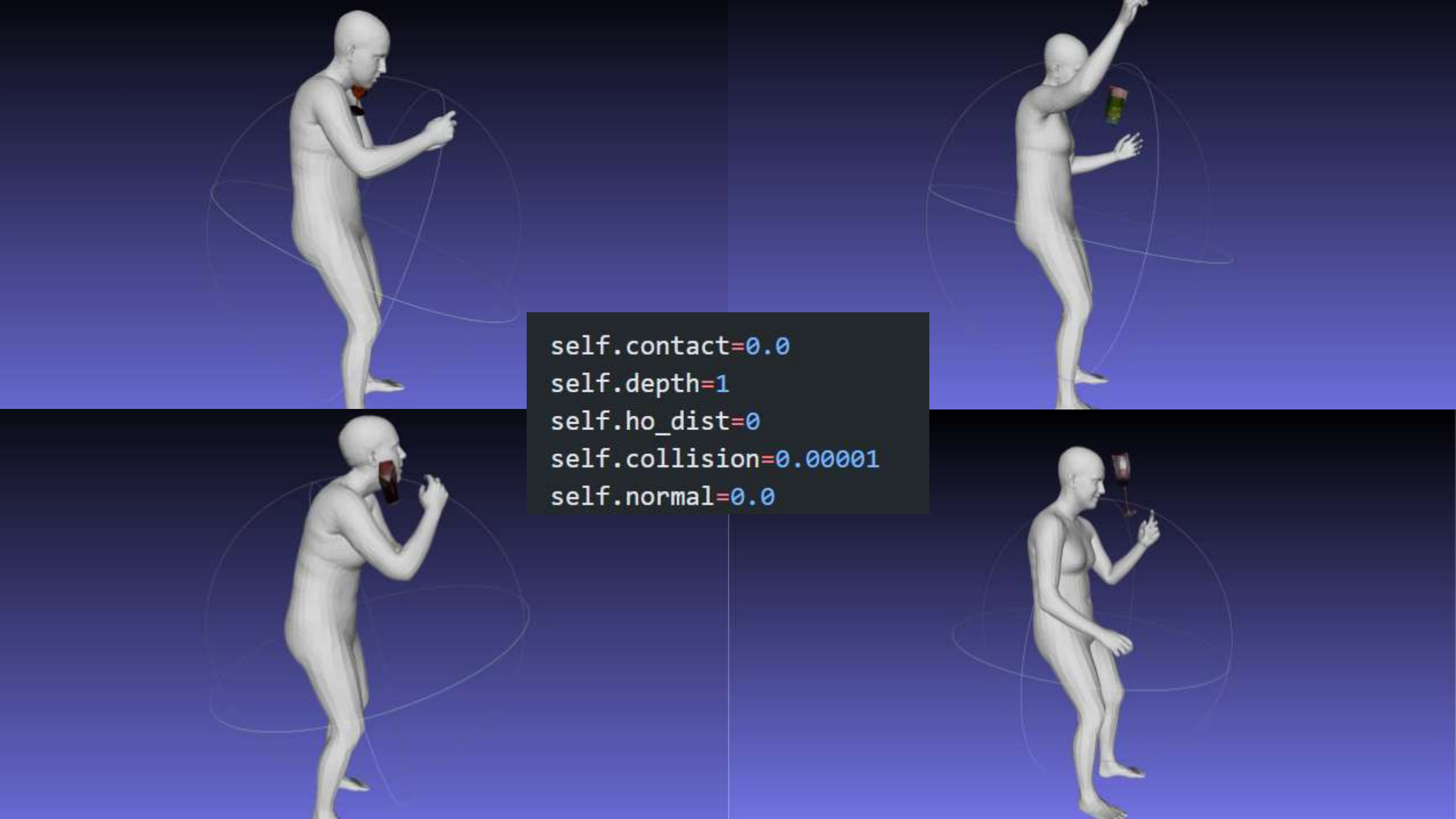
Fig. 21 shows some failure cases of our HOI-Gaussian optimizer. In these cases, human body parts occlude each other severely, and the object happens to be located between the occluded areas, which becomes challenging to determine which body part the object should contact with.

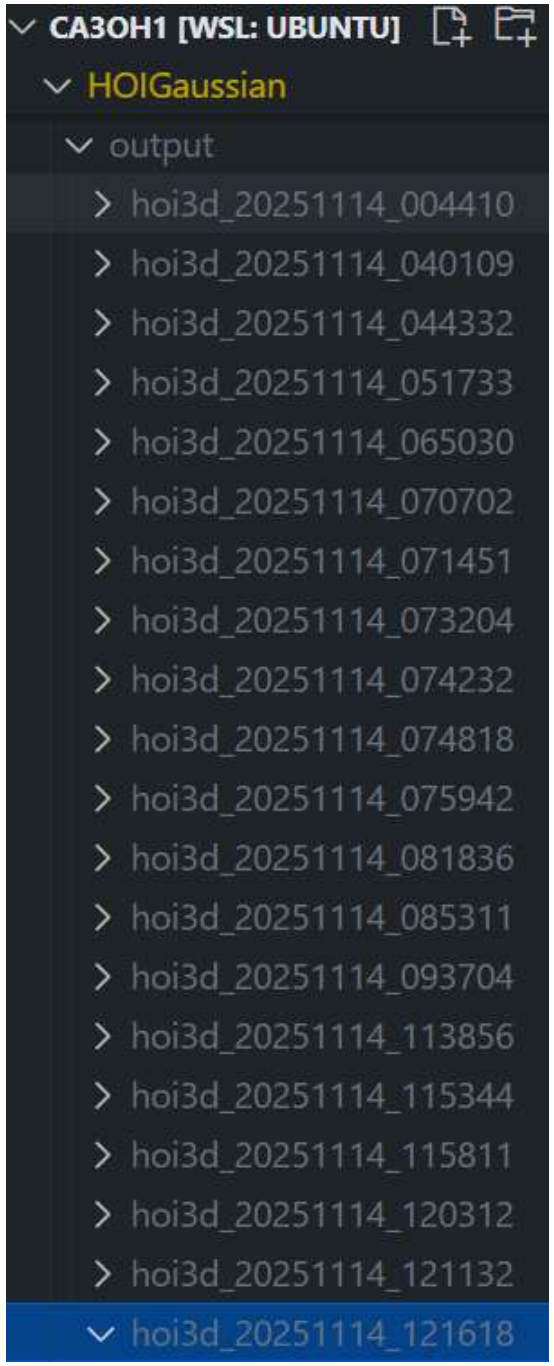
Sandwich

Paper

GT




self.contact=0.0
self.depth=1
self.ho_dist=0
self.collision=0.00001
self.normal=0.0




```
self.contact=100
self.depth=0
self.ho_dist=0
self.collision=0
self.normal=0.0
```

```
0, contact=1.21, depth=0, collision=0]
0, contact=1.21, depth=0, collision=0]
, contact=1.21, depth=0, collision=0]
, contact=1.21, depth=0, collision=0]
0, contact=1.21, depth=0, collision=0]
0, contact=1.21, depth=0, collision=0]
0, contact=1.21, depth=0, collision=0]
0, contact=1.21, depth=0, collision=0]
0, contact=1.21, depth=0, collision=0]
0, contact=1.21, depth=0, collision=0]
0, contact=1.21, depth=0, collision=0]
, contact=1.21, depth=0, collision=0]
, contact=1.21, depth=0, collision=0]
0, contact=1.21, depth=0, collision=0]
0, contact=1.21, depth=0, collision=0]
0, contact=1.21, depth=0, collision=0]
0, contact=1.21, depth=0, collision=0]
, contact=1.21, depth=0, collision=0]
, contact=1.21, depth=0, collision=0]
n will switch to that of iio.v3.imread.
```

```
337         if contact_loss != 0 and contact_loss is not None:
```

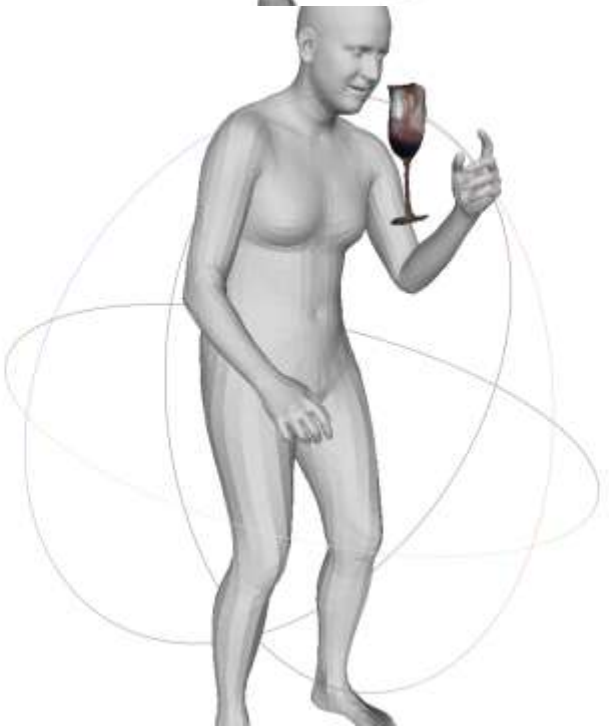
train.py Git 本地更改(工作树) - 第 7 个更改(共 13 个)

```
                else:
274 335             collision_loss = None
275 336
276             if contact_loss != 0 and contact_loss is None:
    337         if contact_loss != 0 and contact_loss is not None:
277 338             contact_loss.backward(retain_graph=True)
278 339
                if depth_loss != 0 and depth_loss != None:
338             contact_loss.backward(retain_graph=True)
```


```
contact=1.19, depth=0, collision=0]
 contact=1.19, depth=0, collision=0]
 contact=0.2, depth=0, collision=0]
contact=0.2, depth=0, collision=0]
contact=0.208, depth=0, collision=0]
 contact=0.208, depth=0, collision=0]
 contact=0.278, depth=0, collision=0]
 contact=0.278, depth=0, collision=0]
 contact=0.158, depth=0, collision=0]
contact=0.158, depth=0, collision=0]
contact=0.0845, depth=0, collision=0]
 contact=0.0845, depth=0, collision=0]
 contact=0.0977, depth=0, collision=0]
 contact=0.0977, depth=0, collision=0]
 contact=0.0857, depth=0, collision=0]
 contact=0.0857, depth=0, collision=0]
 contact=0.0654, depth=0, collision=0]
 contact=0.0654, depth=0, collision=0]
```

```
self.contact=100
self.depth=0
self.ho_dist=0
self.collision=0
self.normal=0.0
```

Paper




self.contact=100
self.depth=0
self.ho_dist=0
self.collision=0
self.normal=0.0
GT

TODO & Problems

TODO:

- Debug: Prevent the ‘not None’ like cases.

- Convert the current dataset to train Contact Area GT Model/Network!

- **Structure & Training!**

- Enable scaling rotating & **SMPLX Changing** in HOI-Optimizer.

-

Problems:

- **As an output, the size of contact score on Object is not fixed:**

  Current solution, down sampling to under 8000, and add 0 to 8000

- **Dataset Size might not enough:**

  .....

THANKS