Combined understanding of 3D plane articulation and partial human pose estimation

Siyi Chen[1], Shengyi Qian[1], David F. Fouhey[1]
[1]CSE, University of Michigan, Ann Arbor, Michigan.

Introduction

Watching the scene of a partial human, people will have a sense of the whole 3D body pose. Seeing an opening door or a sliding window, people will have a sense of the 3D planes' articulation. Further, we can understand the 3D relationship between the two. This project aims to understand the plane's articulation & partial human pose together in 3D.

Objectives

1. Independently predict 3D partial human poses as SMPL[6] meshes in pytorch3d following Rockwell's methods[3]. And use PointRend[7] to obtain 2D human masks.
2. Independently predict 2D plane masks as well as 3D articulation information.
3. (In progress) Use a differential render in pytorch3d similar to the NMR in PHOSA[1] for back-propagation, to build a similar system optimizing the position and size of the person considering 3d space relationship and interaction for each single image.
4. (In progress) Further improve the system, including building a video-level system, and optimizing the pose of the humans.

Dataset

This project used internet videos with plane articulations annotated. Scenes mainly include people moving objects.

Tools

1. GraphCMR pretrained following self-training with similar dataset[3].
2. Articulation prediction system based on planeRCNN pretrained with the same dataset.
3. PointRend on Detectron2 for 2D segmentation[7].
4. Pytorch3d's differential 3D render for building the optimization system.
5. SMPL for representing 3D human pose[6].

Structure

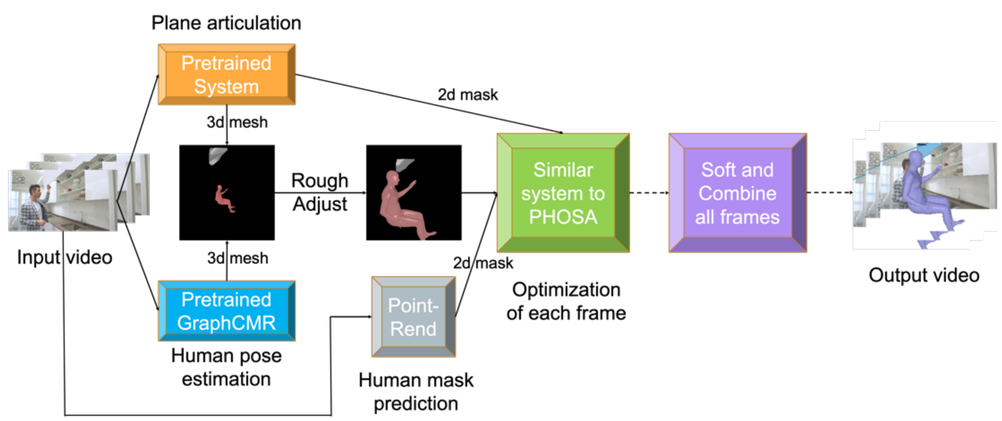

Figure 1: Overview of the pipeline.

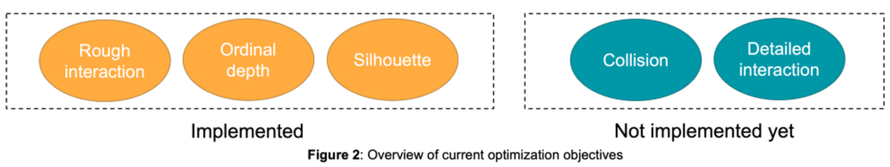

Figure 2: Overview of current optimization objectives

Discussion

1. Current implementation of the optimization component is not complete, lacking restrictions of specific parts of interaction[1]. For example, people tend to interact with the world with hands. Need improvements.
2. For frames containing only hands of the person or no person, predictions of whole body SMPL[6] pose do not make sense. Need a pre-judgement, or a choice focusing only on hands.
3. The total system is quite time consuming.
4. The masks from PointRend[7] can be very noisy, affecting the optimization based on 2D masks. Though when there are larger parts of human, the prediction is generally very accurate.

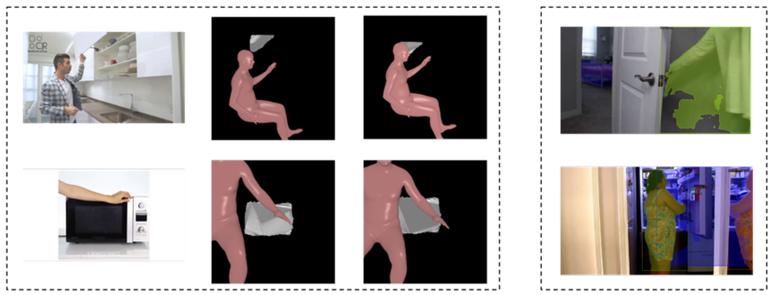

Figure 3: Left group: Two sample results; From left to right: input frame, rough optimized result, and detailed optimized result. Right group: Examples of noisy masks; Up: bad mask predictions; Down: multiple people.

Conclusions

This project tries to do a combined understanding of human action and plane articulation. In future, components to soften and combine all frames as well as optimizing human poses are intended to be designed.

Acknowledgement

Thanks to Professor David Fouhey and Shengyi Qian's kind help. Thanks to the inspiring talks with Dandan Shan, Alexander Raistrick, Linyi Jin, Richard Higgins, Max Hamilton, Sarah Jabbour, Chris Rockwell, and all the other kind lab members. That's really a cool summer.

References

[1] Jason Y. Zhang, Sam Pepose, Hanbyul Joo, Deva Ramanan, Jitendra Malik, Angjoo Kanazawa. Perceiving 3D Human-Object Spatial Arrangements from a Single Image in the Wild. In *ECCV*, 2020.
[3] Chris Rockwell, David F. Fouhey. Full-Body Awareness from Partial Observations. In *ECCV*, 2020.
[4] Chen Liu, Kihwan Kim, Jinwei Gu, Yasutaka Furukawa, and Jan Kautz. PlaneRCNN: 3D plane detection and reconstruction from a single image. In *CVPR*, 2019.
[5] Nikos Kolotouros, Georgios Pavlakos, Kostas Daniilidis. Convolutional Mesh Regression for Single-Image Human Shape Reconstruction. In *CVPR*, 2019.
[6] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J Black. SMPL: A skinned multi- person linear model. ACM Transactions on Graphics (TOG), 34(6):248, 2015.
[7] Alexander Kirillov, Yuxin Wu, Kaiming He, Ross Girshick. PointRend: Image Segmentation as Rendering. In *CVPR*, 2020.

SURE SUMMER UNDERGRADUATE RESEARCH IN ENGINEERING UNIVERSITY OF MICHIGAN