



Harmony4D: A Video Dataset for In-The-Wild Close Human Interactions

NeurIPS 2024

Rawal Khirodkar* , Jyun-Ting Song* , Jinkun Cao, Zhengyi Luo, Kris Kitani

Human Mesh Recovery



PARE, ICCV 21'



Multi-HMR, ECCV 24'



BEV, CVPR 22'



BUDDI, CVPR 24'



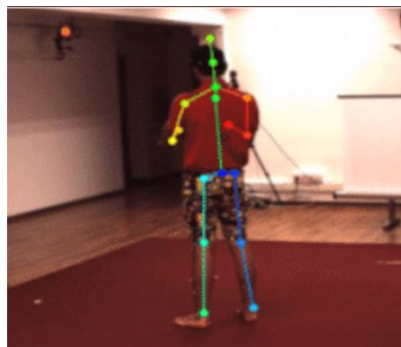
4DHumans, ICCV 23'

Existing 3D Human Datasets



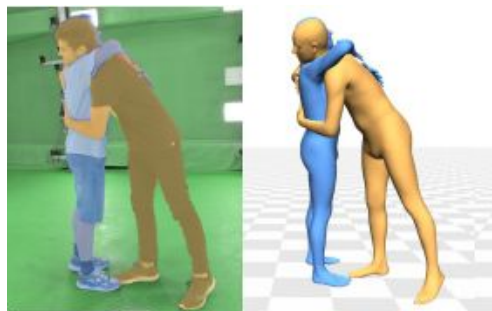
3DPW, ECCV 2018

lack human-human interaction

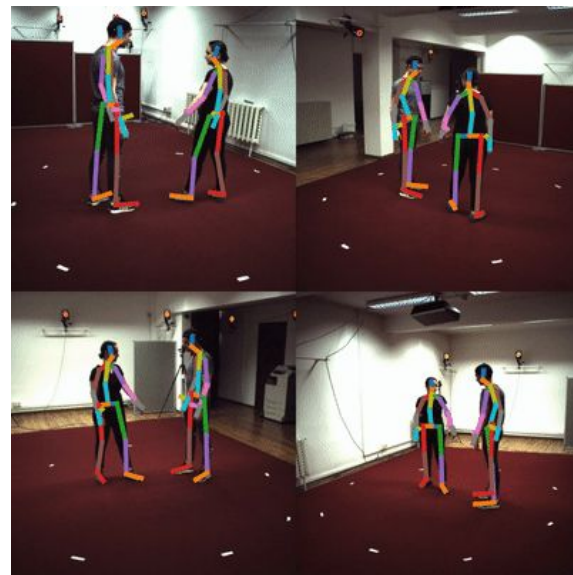


HUMAN3.6M, PAMI 2014

single human



Hi4D, CVPR 2023



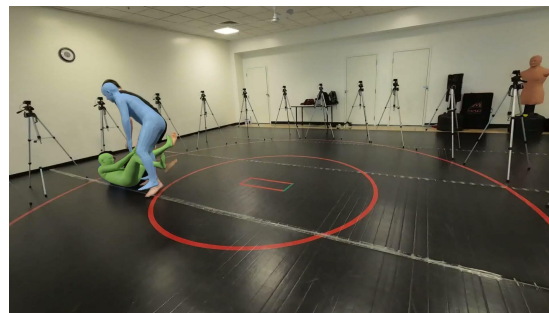
CHI3D, CVPR 2020

lab environment

Harmony4D Dataset



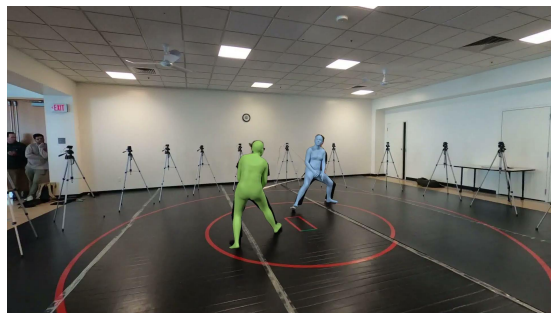
Hugging



Grappling



Karate



Fencing



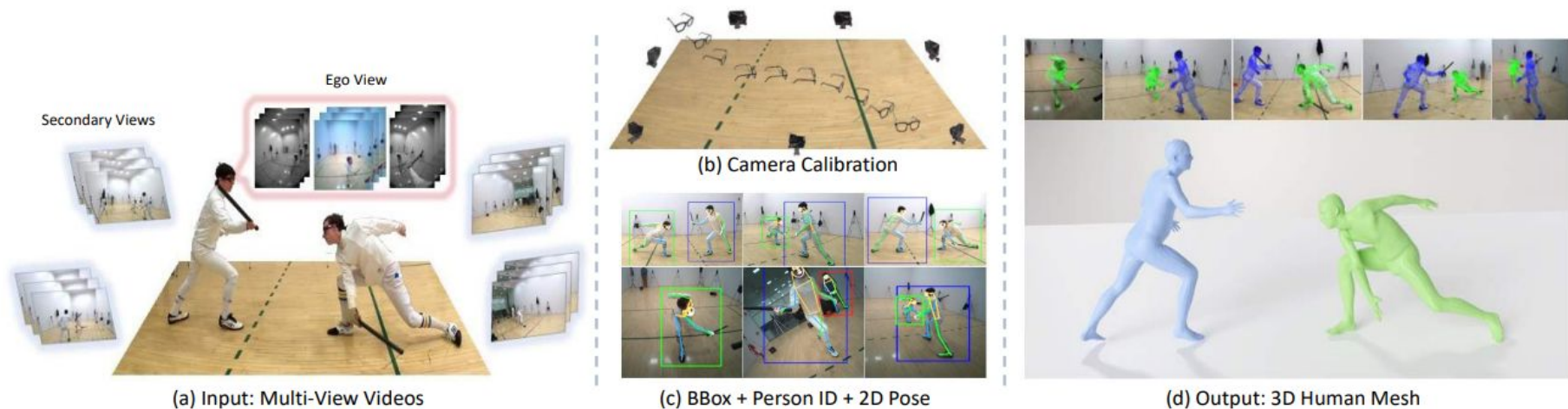
Dancing



MMA

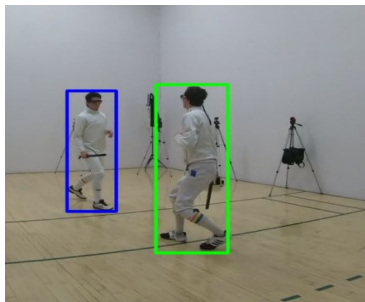
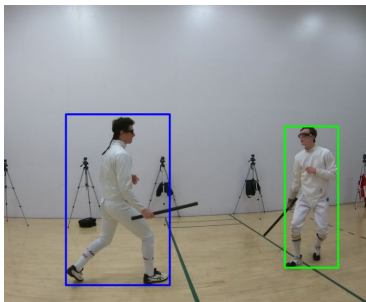
Preliminary - EgoHumans

- Markerless Multi-View Capture System

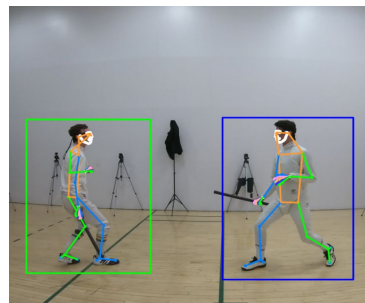
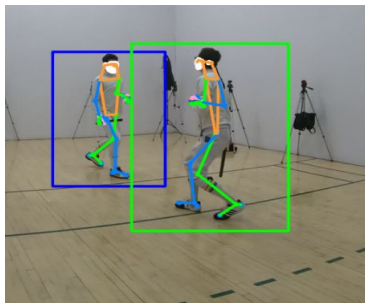
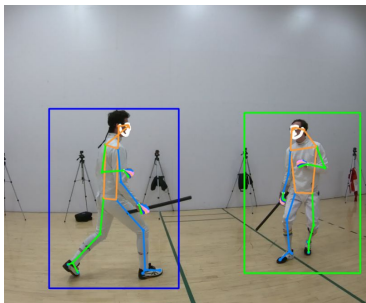


Preliminary - EgoHumans

BBoxes

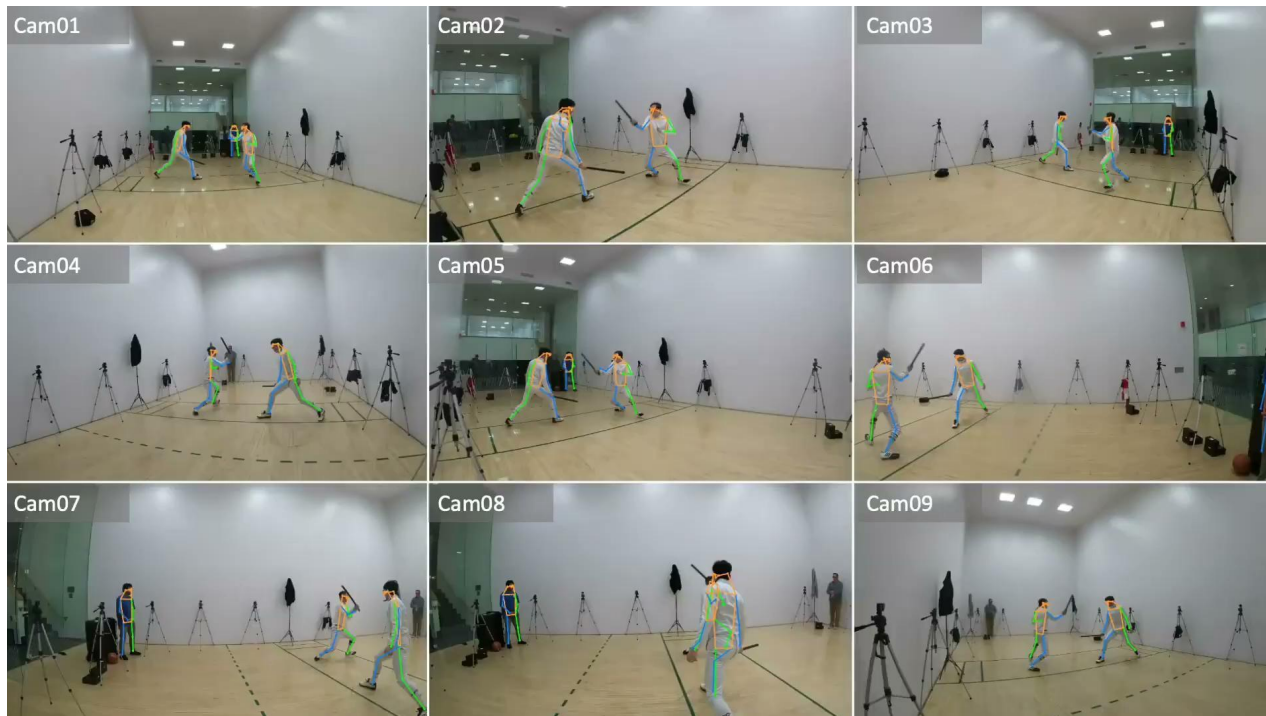


2D Poses



Triangulation

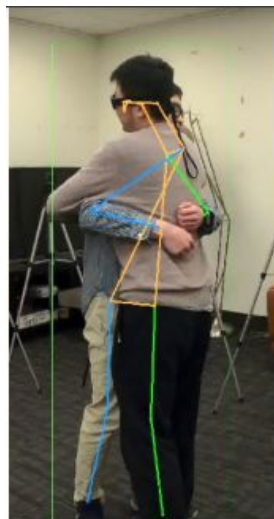
Preliminary - EgoHumans



Failure Case

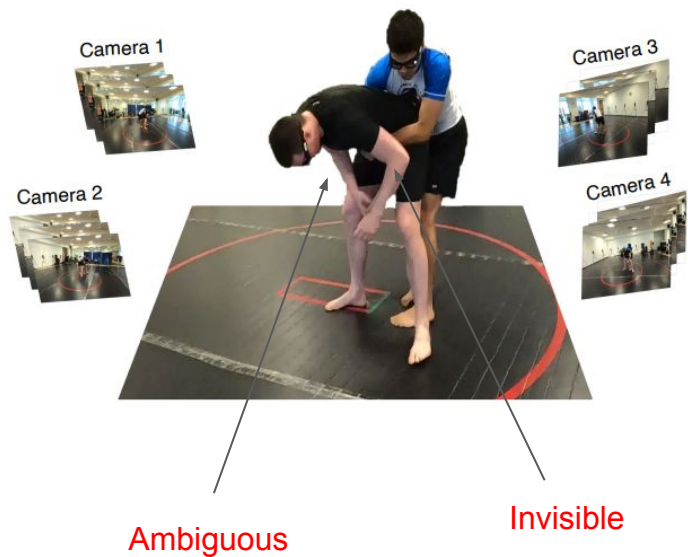


Image

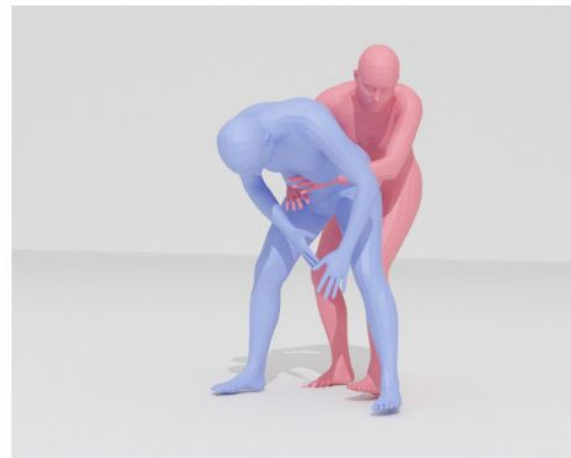


Wrong 3D Pose

Challenges

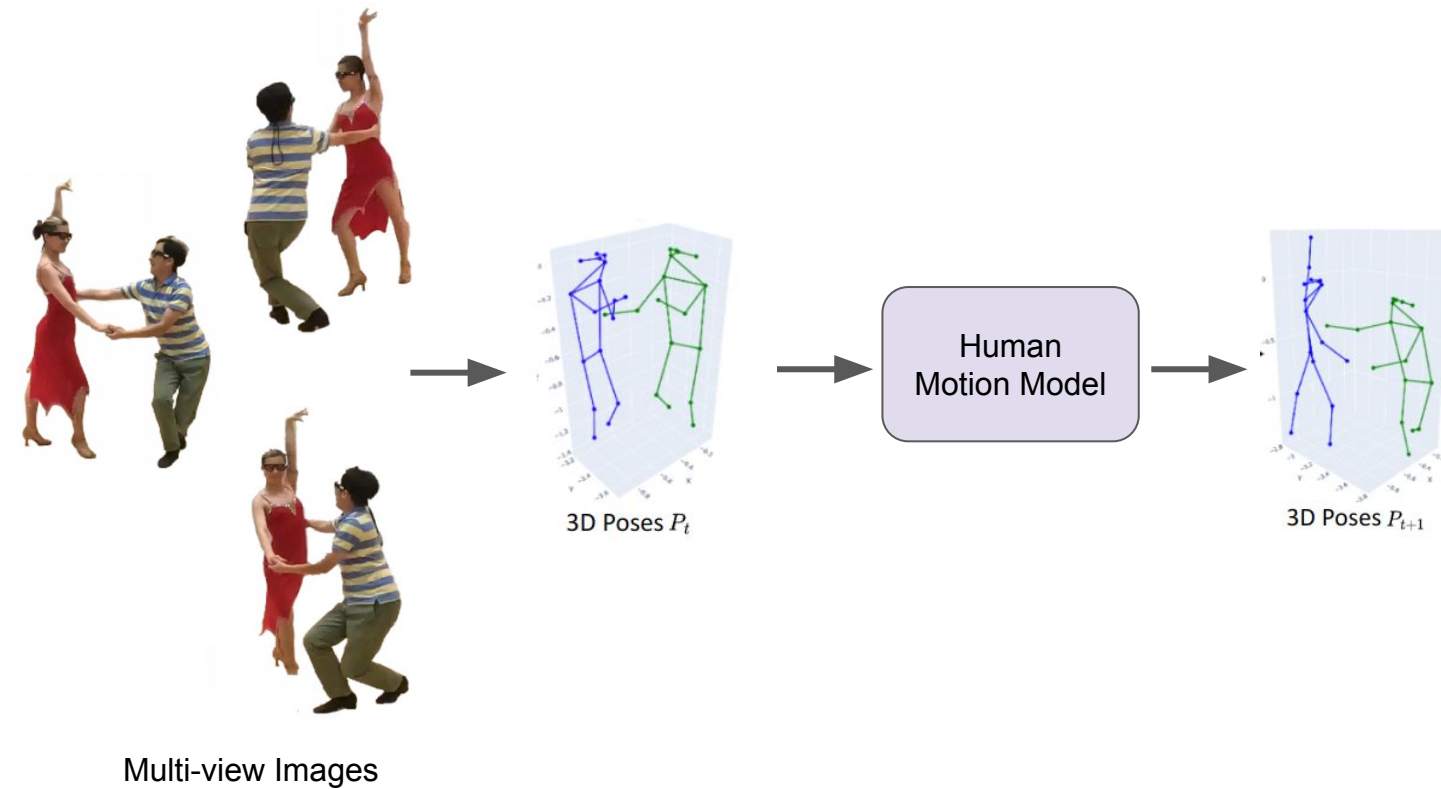


How do we get this?



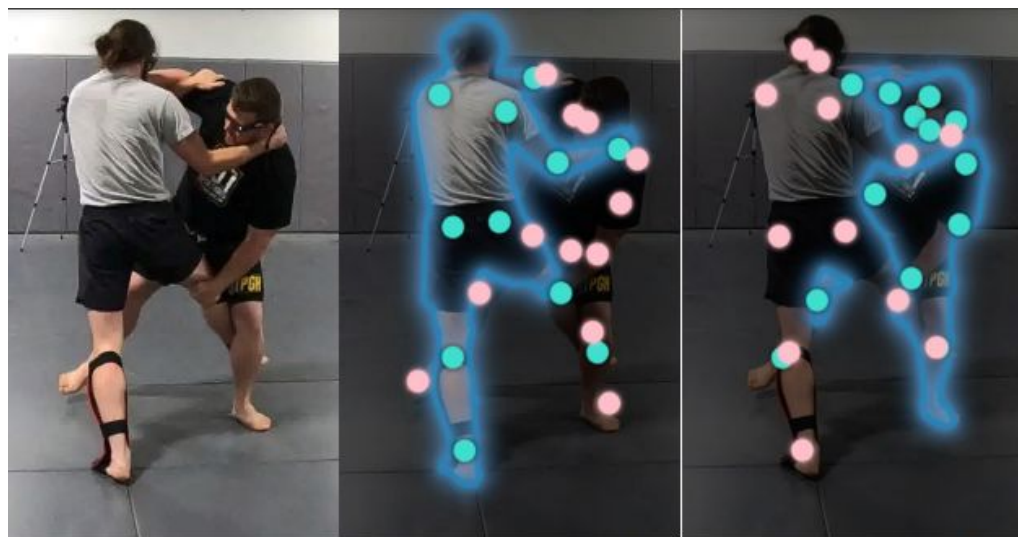
Reconstructed Humans

Human Motion Model



Mask-conditioned 2D Pose Estimator

- Segmentation Mask helps us disentangling occluded human poses

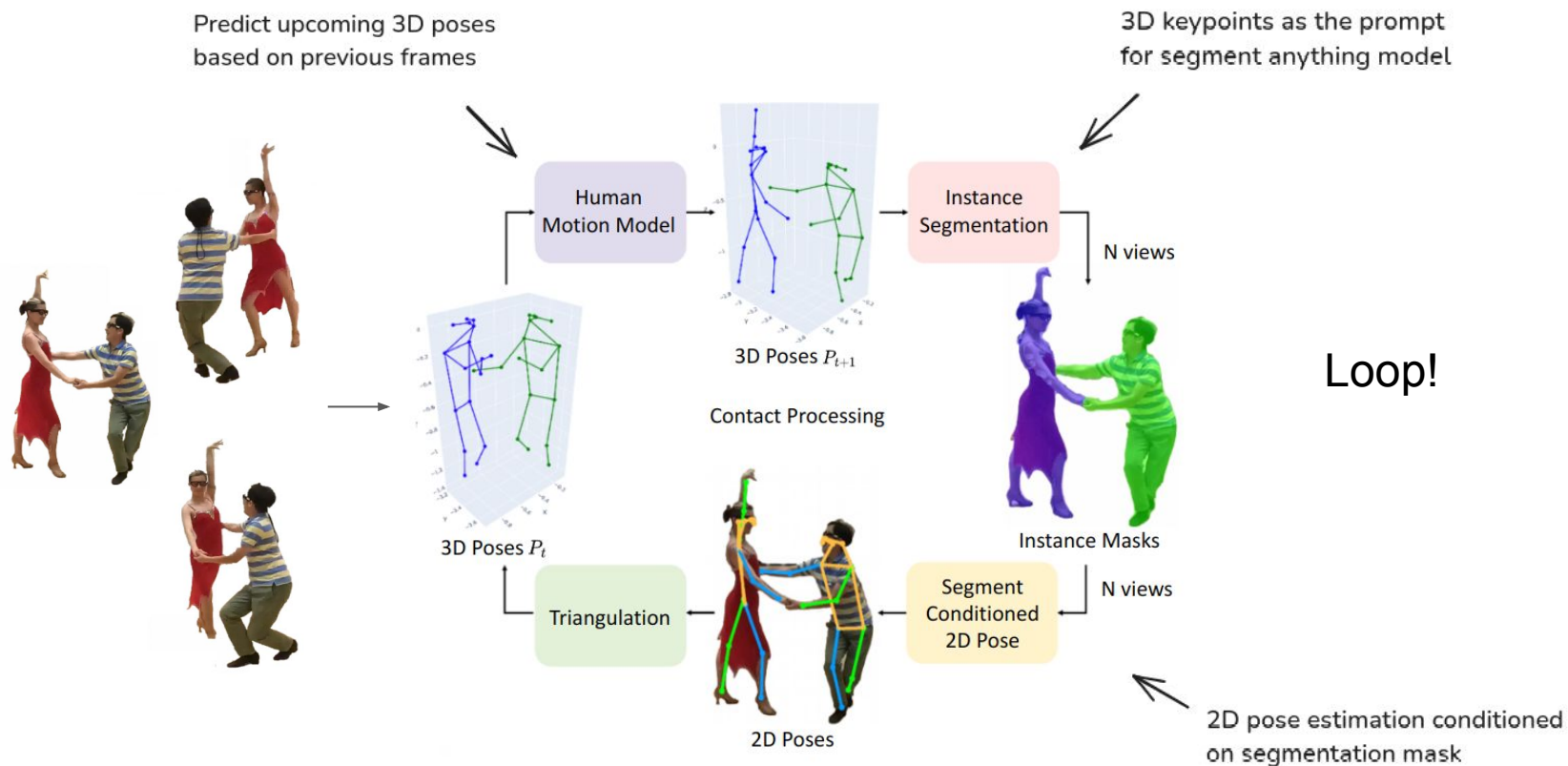


Image

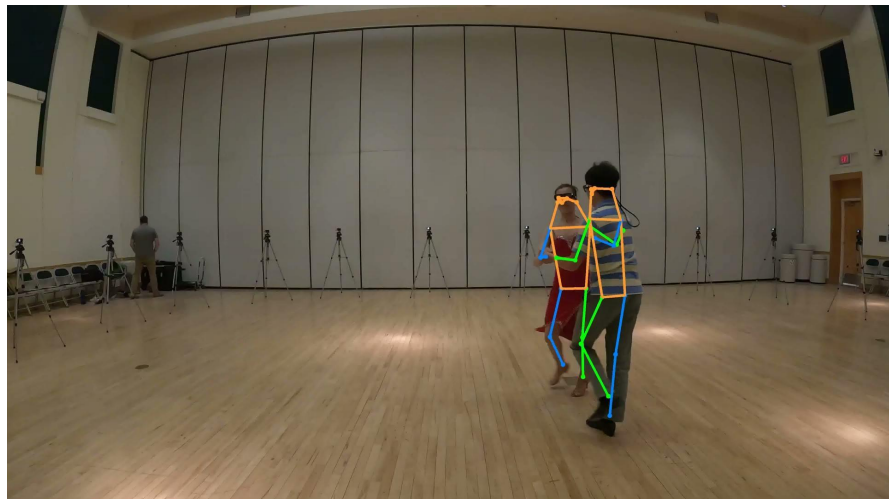
Person A

Person B

Processing Pipeline



Fitting SMPL

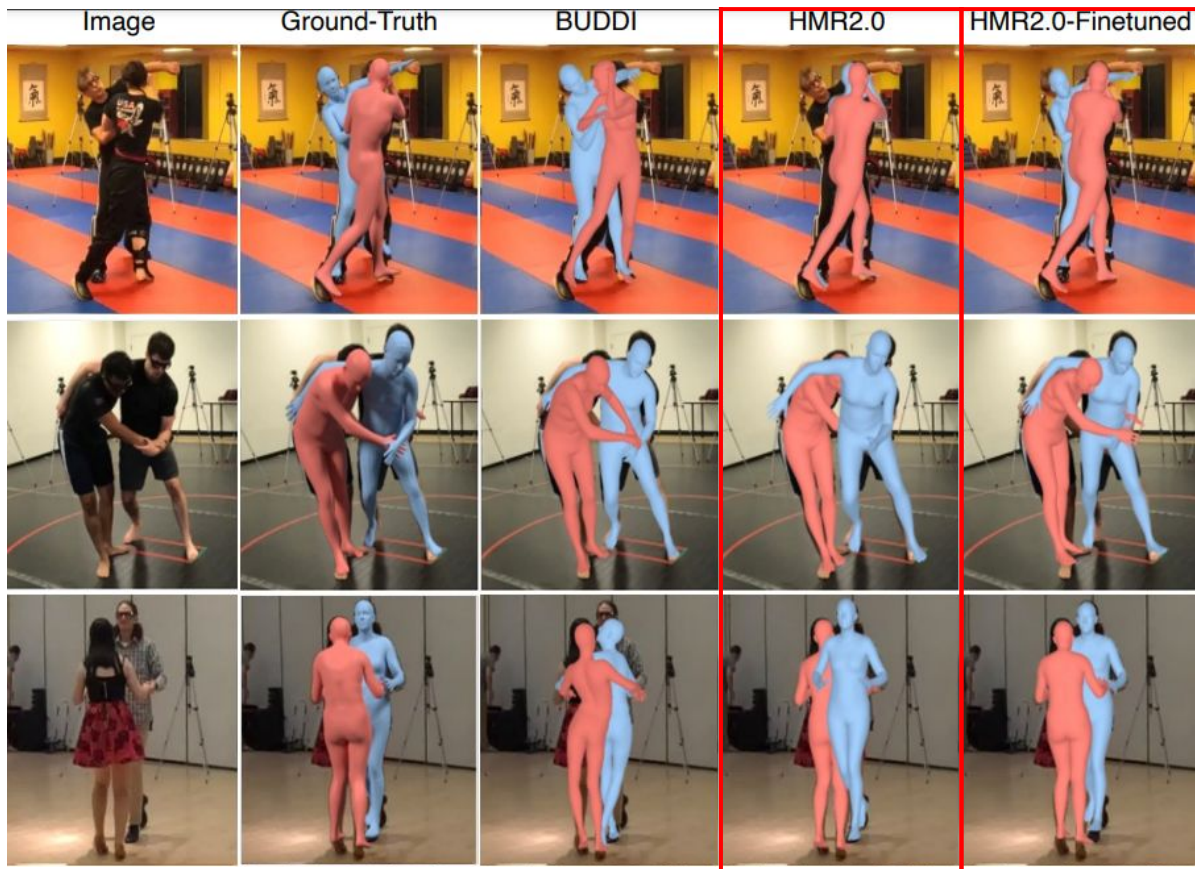


3D Poses



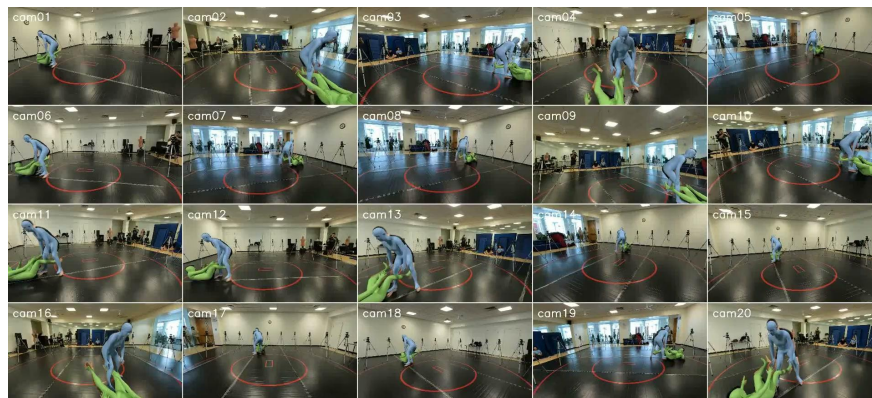
Human Meshes

Experiments



Harmony4D

- Multi-human Dataset
- 1.7M images
- 24 Subjects
- 6 Activities
- 5 Scenes
- Dynamic Activities



Grappling Sequence in Harmony4D Dataset

Dataset	In-The-Wild	Scenes	Subjects	Cameras	Images	Instances	Mesh
ShakeFive2 [74]	✗	1	6	1	34K	68K	✗
MuPoTs-3D [57]	✗	3	8	8	8K	22K	✗
MultiHuman [86]	✗	1	278	6	32K	69K	✓
ExPI [20]	✗	1	4	68	1.9M	3.8M	✗
CHI3D [15]	✗	1	10	4	315K	728K	✓
Hi4D [79]	✗	1	40	8	88K	176K	✓
Harmony4D (Ours)	✓	5	24	20	1.66M	3.32M	✓

Comparison with existing 3D multi-human datasets

Code & Data are released!

Project Page: <https://jyuntins.github.io/harmony4d>

Thank you!