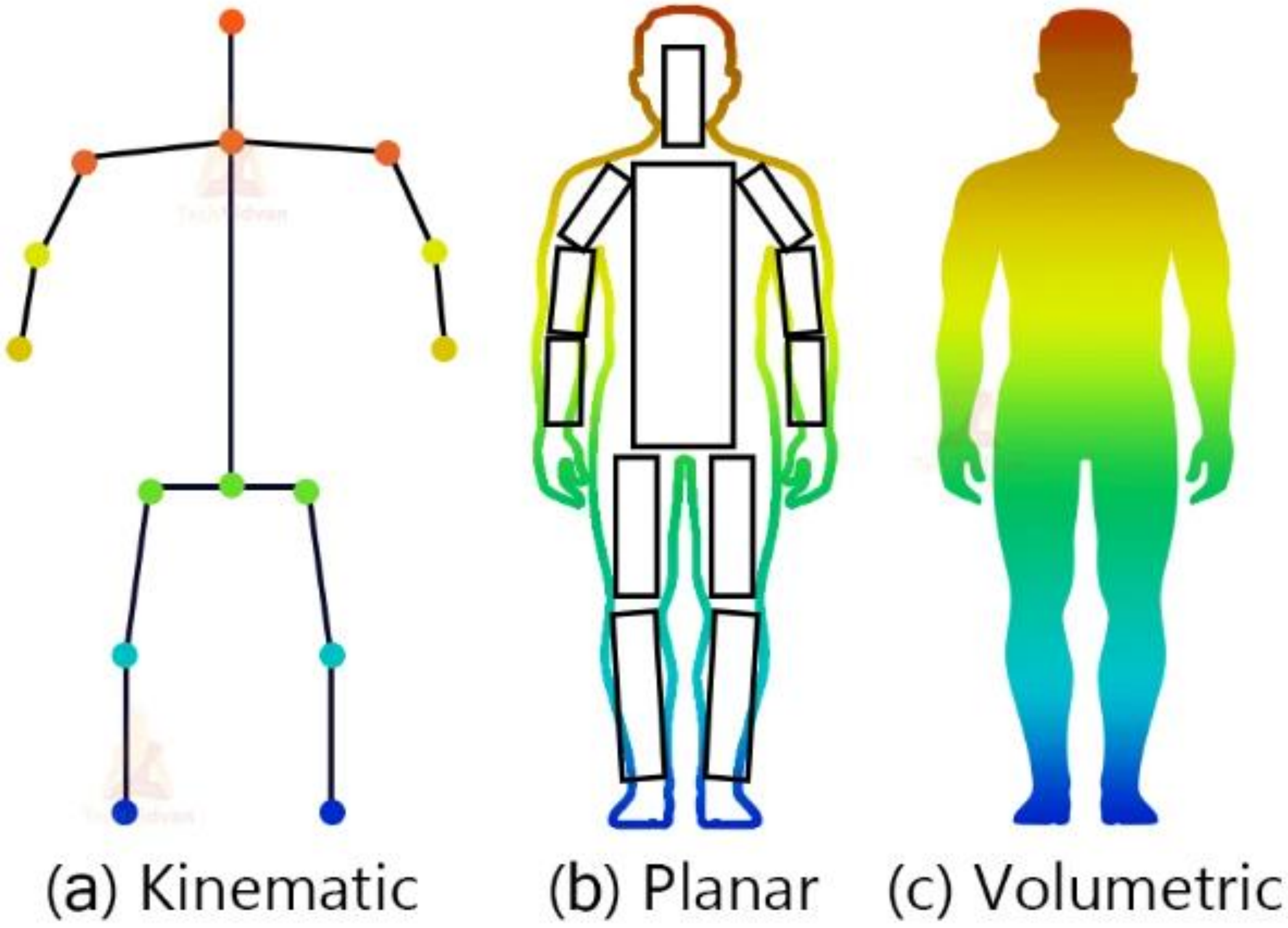
Improving Real-Time Omnidirectional 3D Multi-Person Human Pose Estimation With People Matching and Unsupervised 2D-3D Lifting

Pawel Knap[1], Peter Hardy[1], Alberto Tamajo[1], Hwasup Lim[2], Hansung Kim[1]

1. University of Southampton, School of Electronics & Computer Science
2. Korea Institute of Science and Technology

Human Pose Estimation

- Computer Vision tasks to identifies and classifies the poses of human body parts and joints in images or videos.
- Applications in healthcare, entertainment, surveillance, sports, education, and beyond



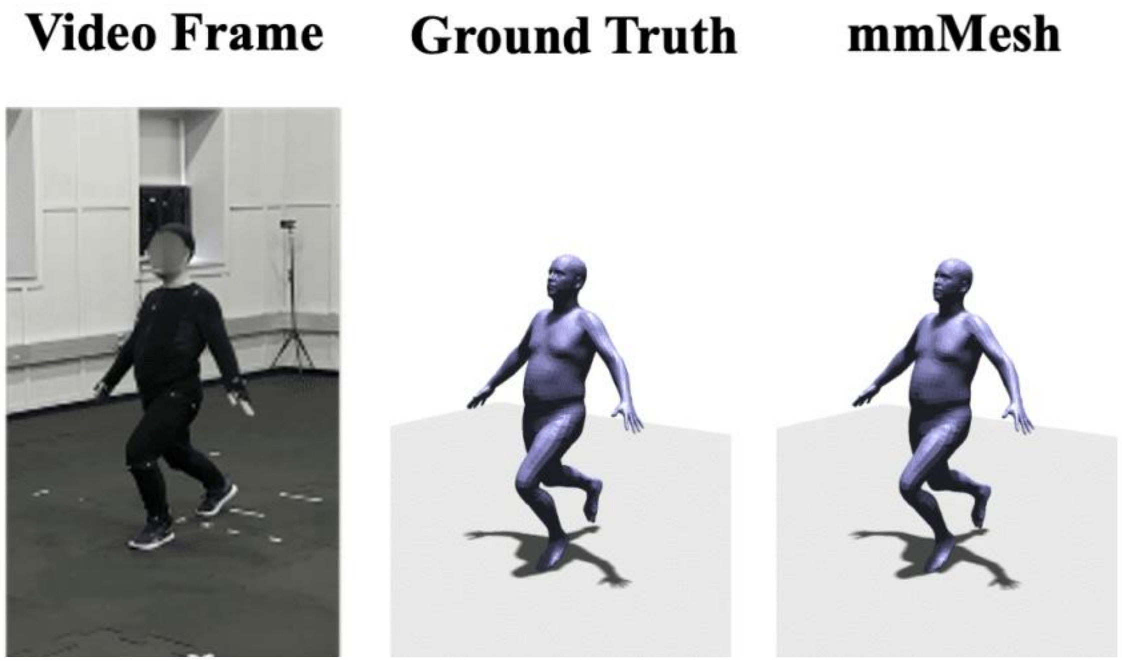
2D Human Pose Estimation

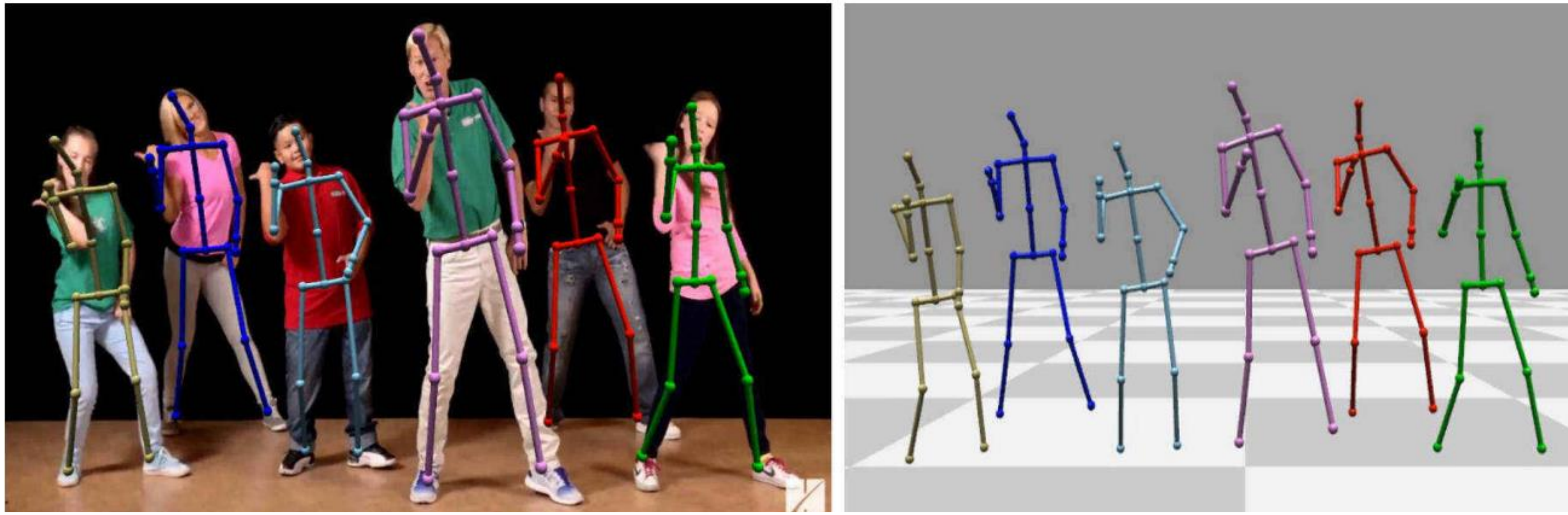
- Locations of the body joints in 2D space
- Bottom-up, and top-down methods
- Classic approaches OpenPose and DeepPose

OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields, Cao et al., 2019
DeepPose: Human Pose Estimation via Deep Neural Networks, Toshev & Szegedy, 2014

3D Human Pose Estimation

- Locations of the body joints in 3D space
- Classic approaches: mmMesh, XNect





mmMesh: Towards 3D Real-Time Dynamic Human Mesh Construction Using Millimeter-Wave, Xue et al., 2021

XNect: Real-time Multi-Person 3D Motion Capture with a Single RGB Camera, Mehta et al., 2020

Sensors for Human Pose Estimation

- Monocular camera: Cheap, yet with limitations like occlusion and depth ambiguity.
- Set of cameras: Addresses monocular camera limitations but is costly and lacks transferability, restricting applications.
- RADARs: Effective in occlusion handling but produces sparse data.
- LIDARs: High-resolution output, but expensive with scarce data.
- IR-based sensors (Kinect): Faces challenges outdoors.
- Motion Capture Sensor Systems: Limited in applications

Human Pose Estimation – major problems

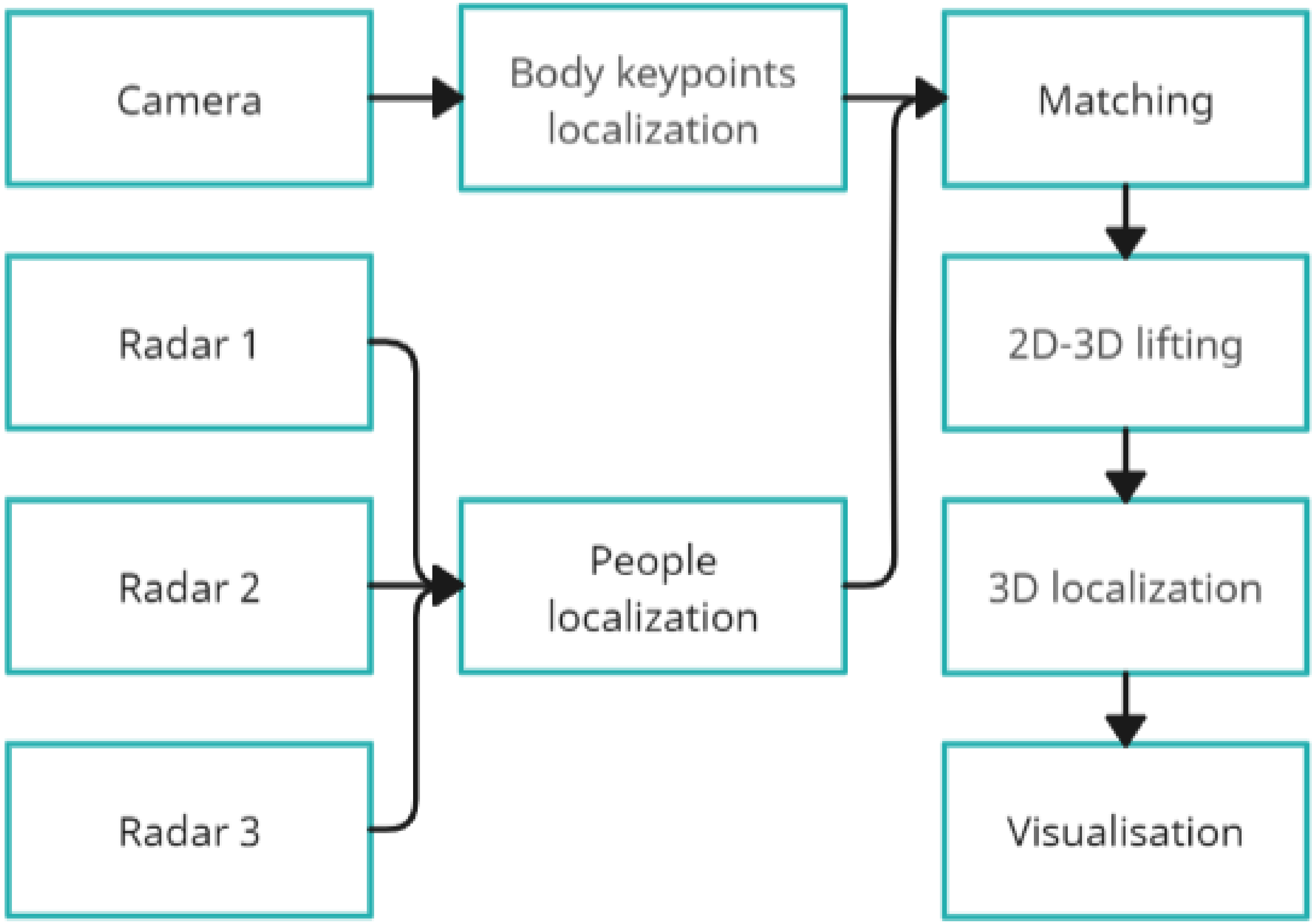
- Majority of solutions focus on single-person 3D HPE
- Multi-person solutions have limited range and limited occlusion handling capabilities [Mehta 2020, Carraro 2019]

Mehta, Dushyant, et al. "XNect: Real-time multi-person 3D motion capture with a single RGB camera", 2020
Carraro, M. Munaro et al. “Real-time marker-less multi-person 3d pose estimation in rgb-depth camera networks”, 2019

Contributions

- Multi-person 3D pose estimation
- Omnidirectional real-time HPE systems for real-world localization – applications in robotics
- New design for systems using blend of sensors
- Cheap and robust real-time, multi-person system

System Overview



Keypoints detection

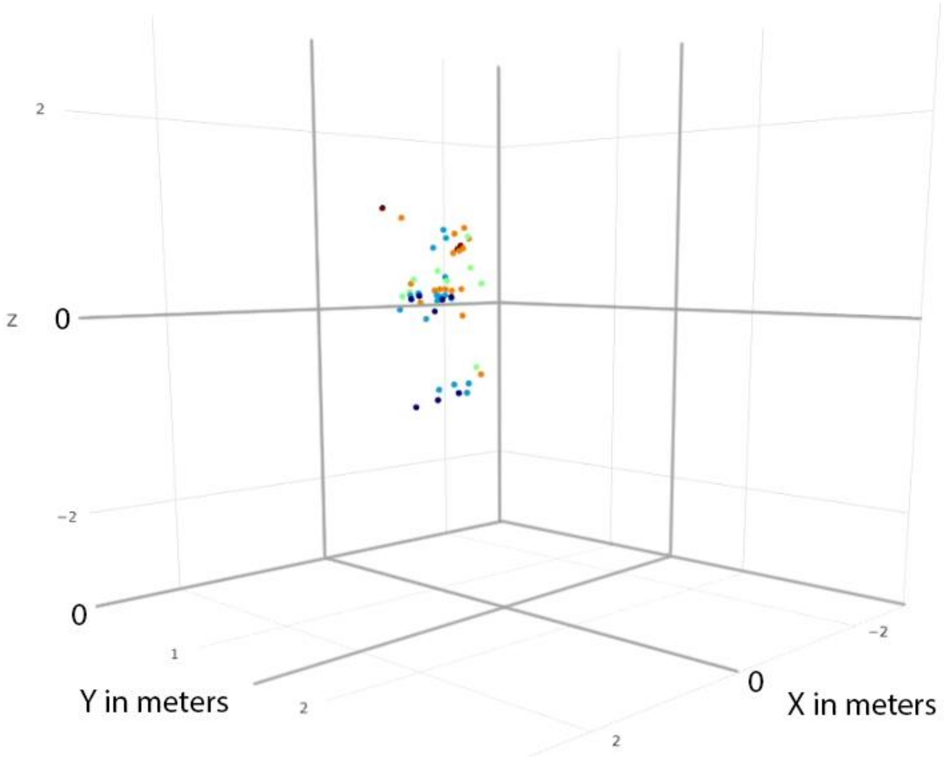
- Openpose for 2D keypoints detection
- It uses Part Affinity Fields
- Handle multiple-people
- No tracking. Detection happens in each frame

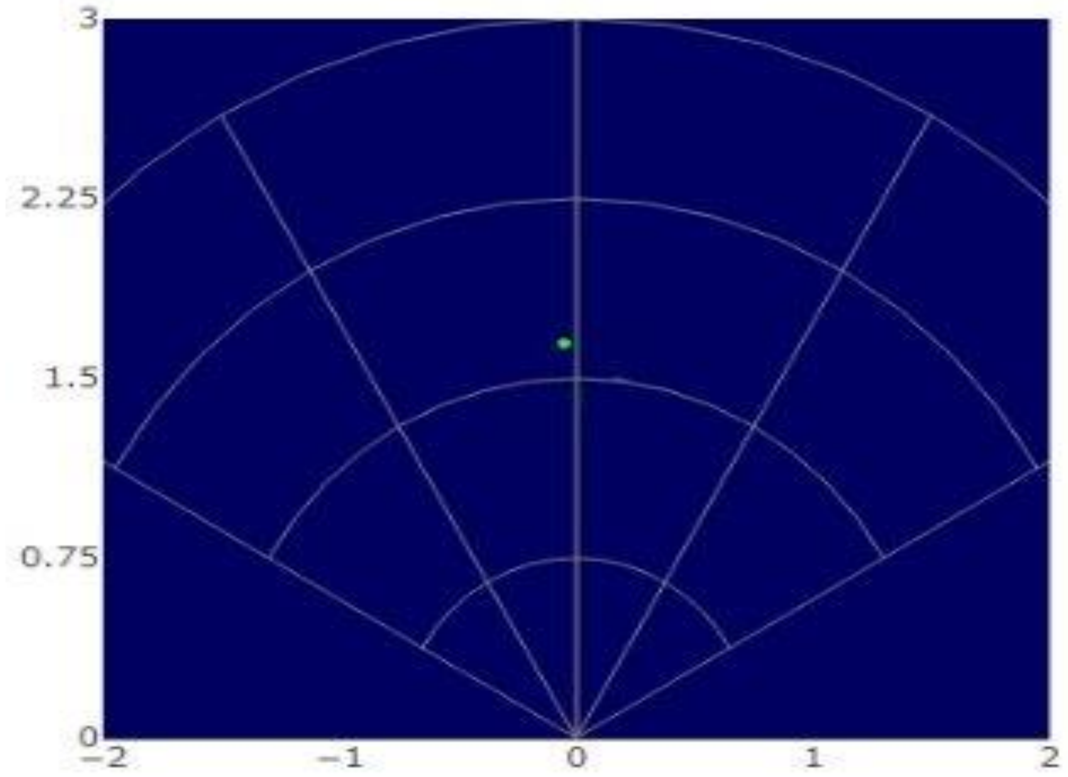



Z. Cao etal. “Openpose: Realtime multi-person 2d pose estimation using part affinity fields”, 2019.

People counting algorithm

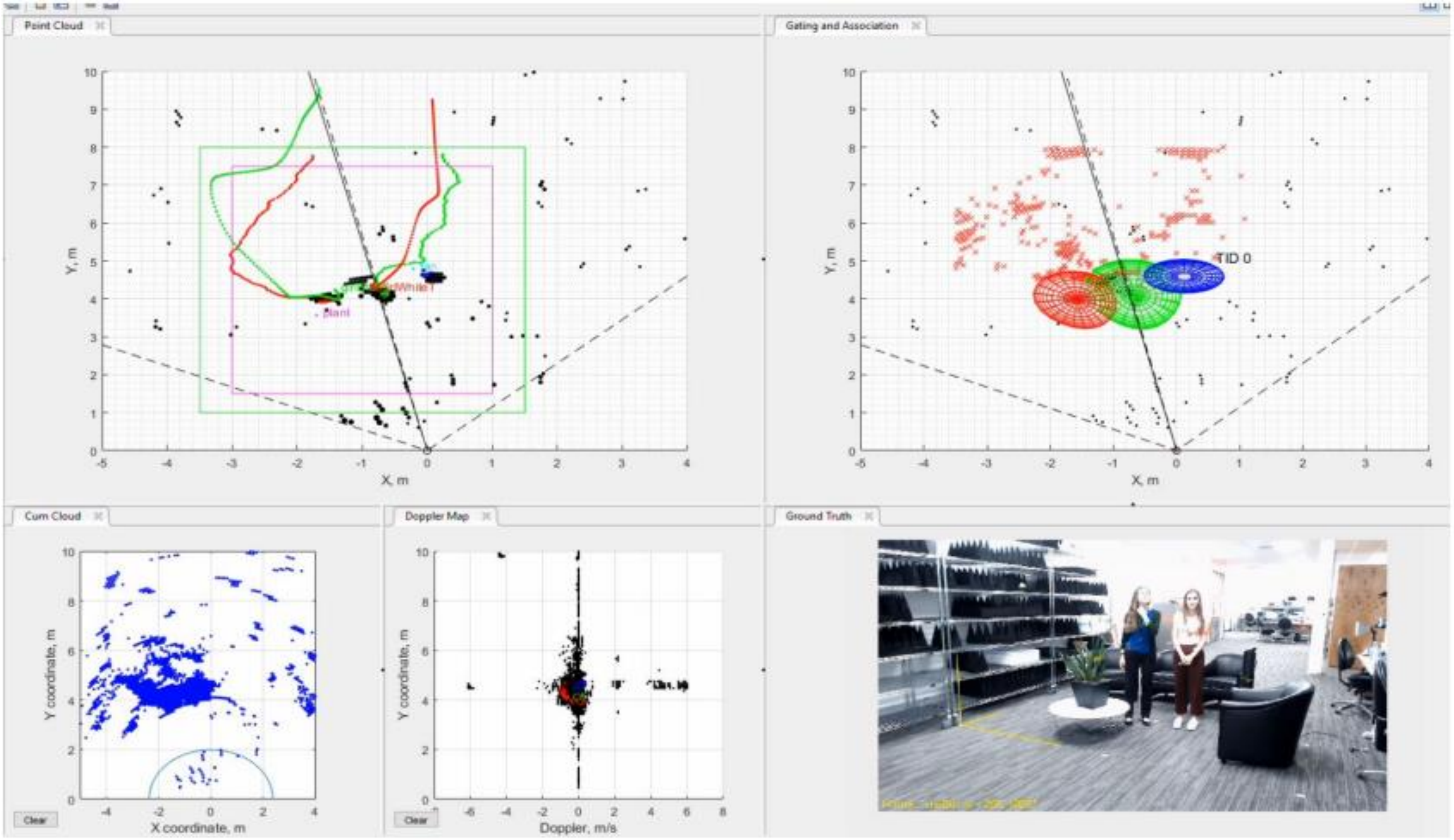
- People localisation done with people tracking algorithm from Texas Instruments (Garcia 2019)





Keegan Garcia, Bringing intelligent autonomy to fine motion detection and people counting with TI mmWave sensors, 2019

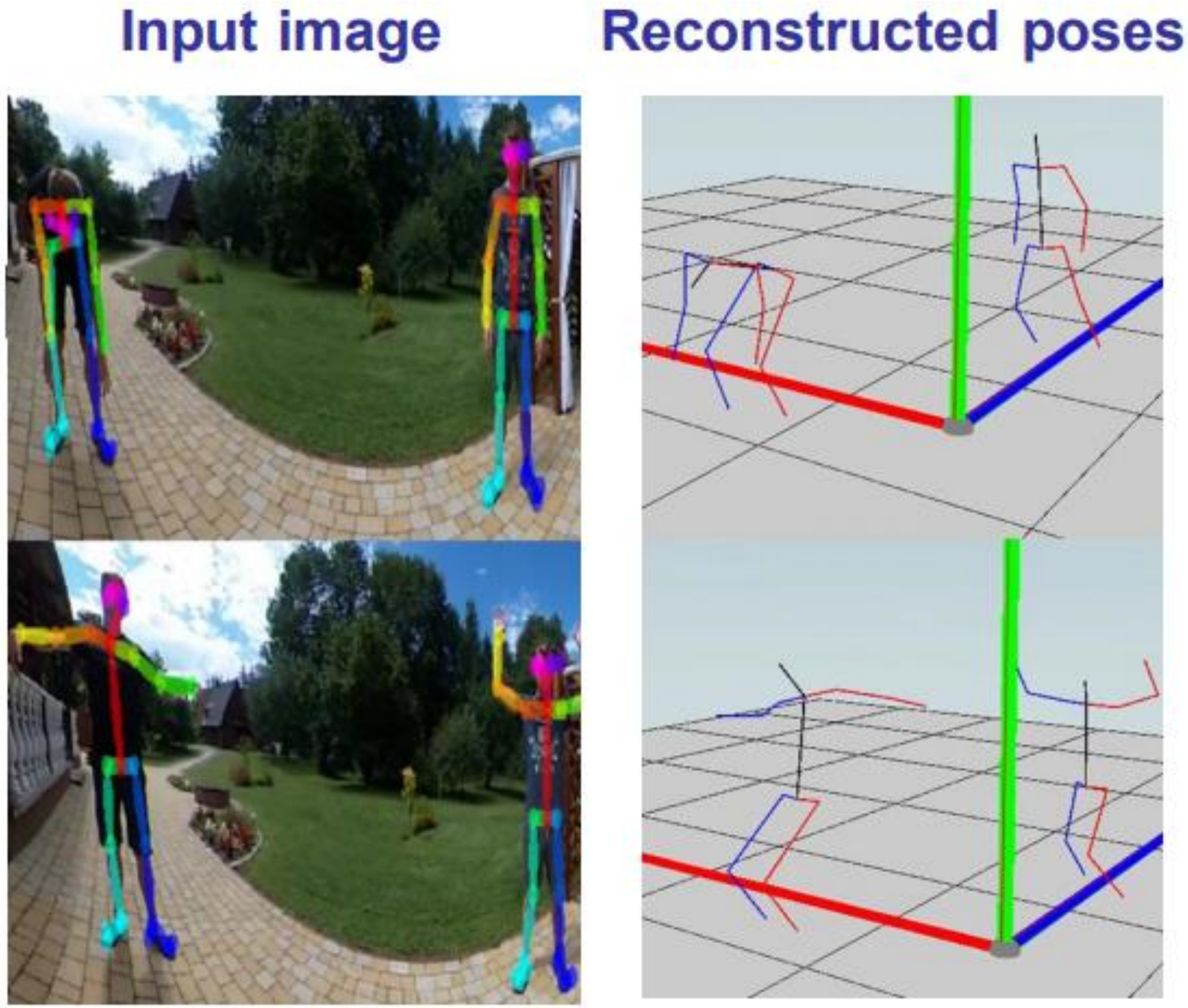
People counting algorithm



People Tracking and Counting Reference Design Using mmWave Radar Sensor – Texas Instruments

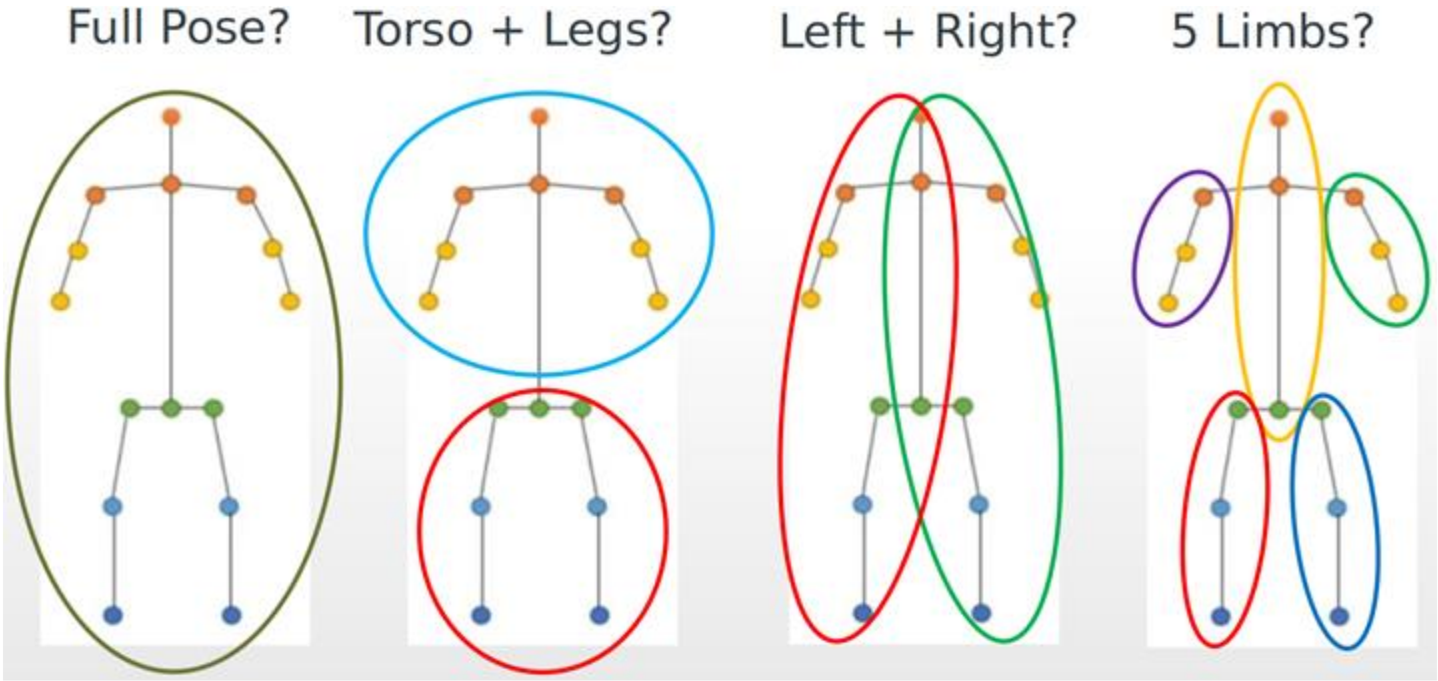
Matching

- Binary search tree
- Based on disparity between the average image x coordinate, and the radars coordinates transformed into the image coordinate space through a learned transform
- Transform based on Oh et al. - learned by Levenberg–Marquardt algorithm
- Radar data used to move poses to correct position in 3D space



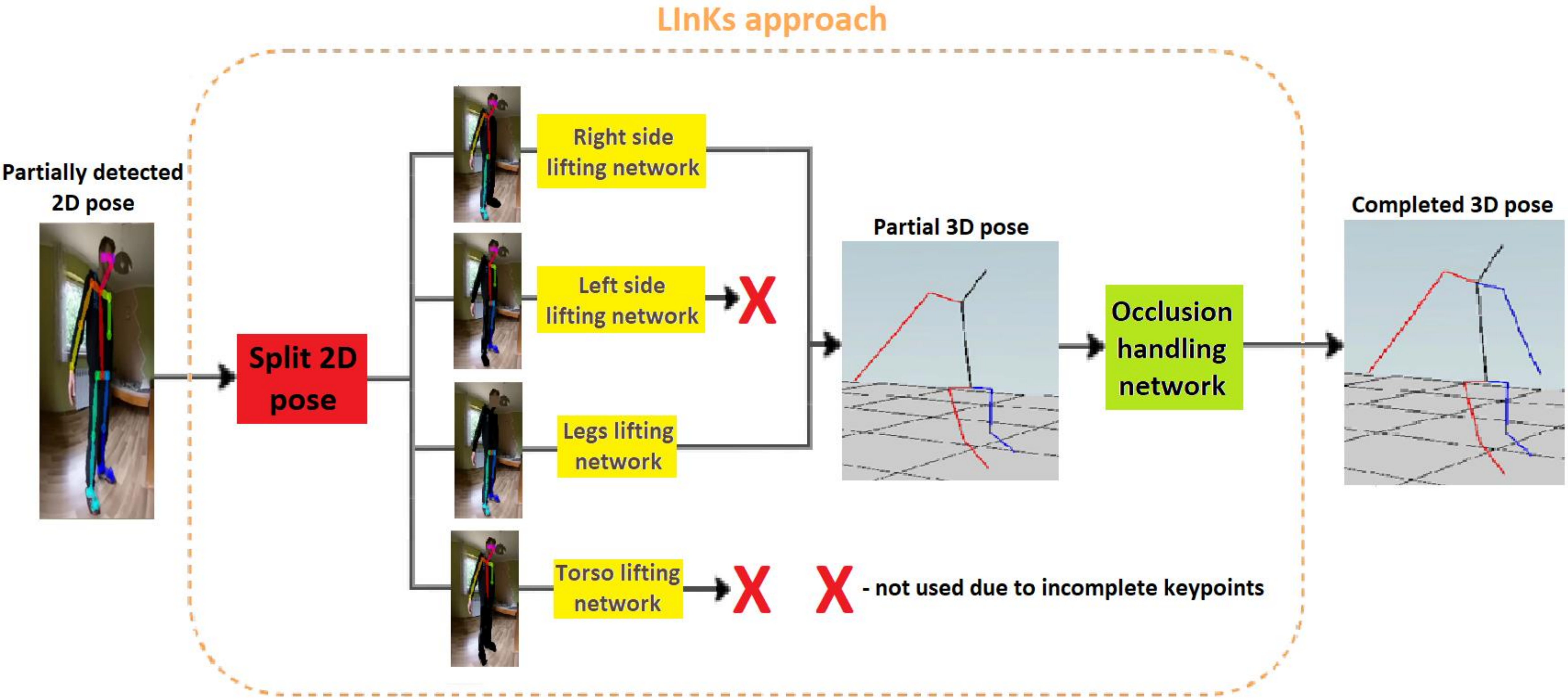
J. Oh, K.-S. Kim, M. Park, and S. Kim, “A comparative study on cameraradar calibration methods”, 2018

Overview of the lifting algorithm

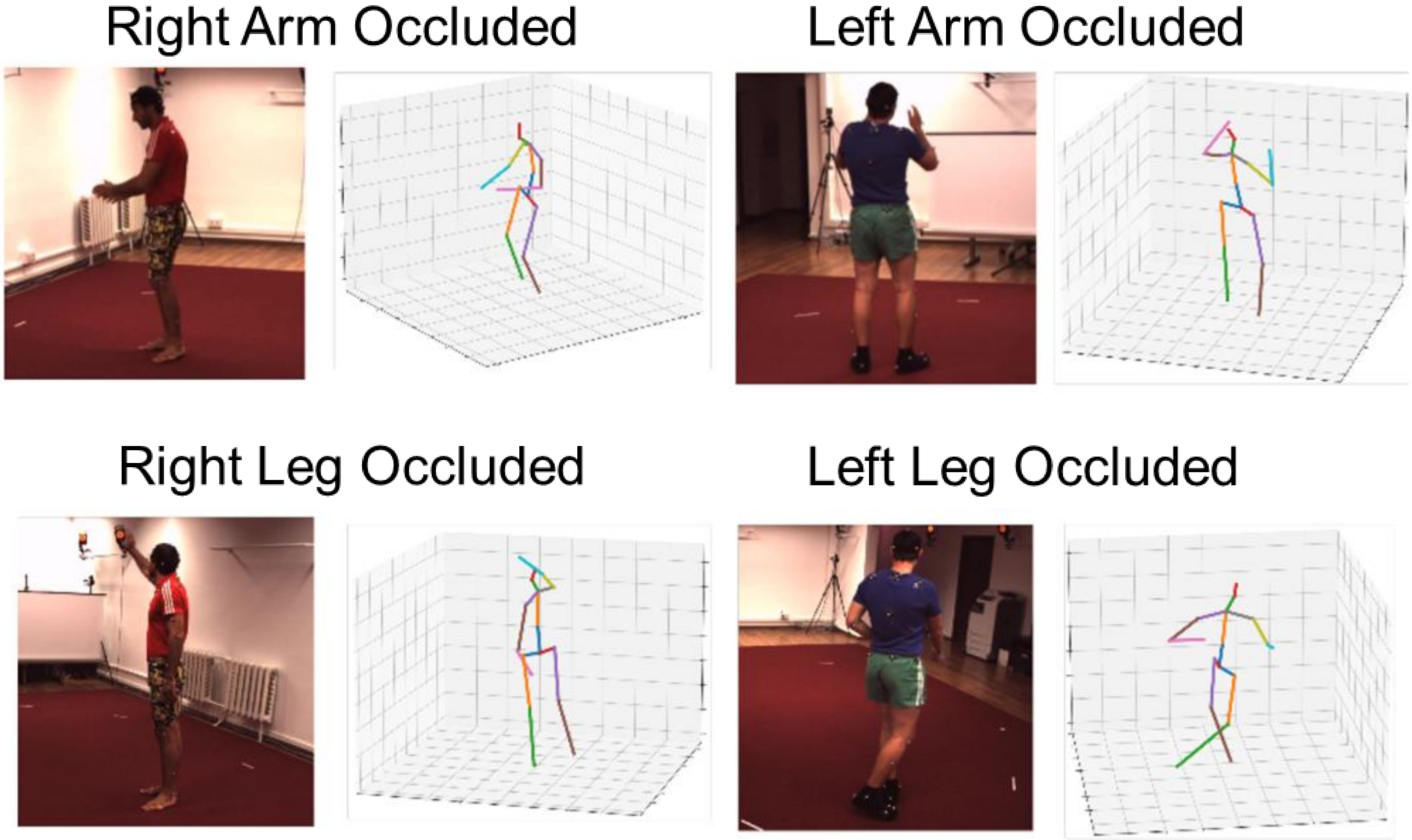


Pose Detected	Front Cameras	Rear Cameras	Avg.
Full 2D Pose	54.8%	35.5%	45.1%
Partial 2D Pose	92.4%	84.8%	88.6%

Overview of the lifting algorithm



Qualitative Results of lifting algorithm

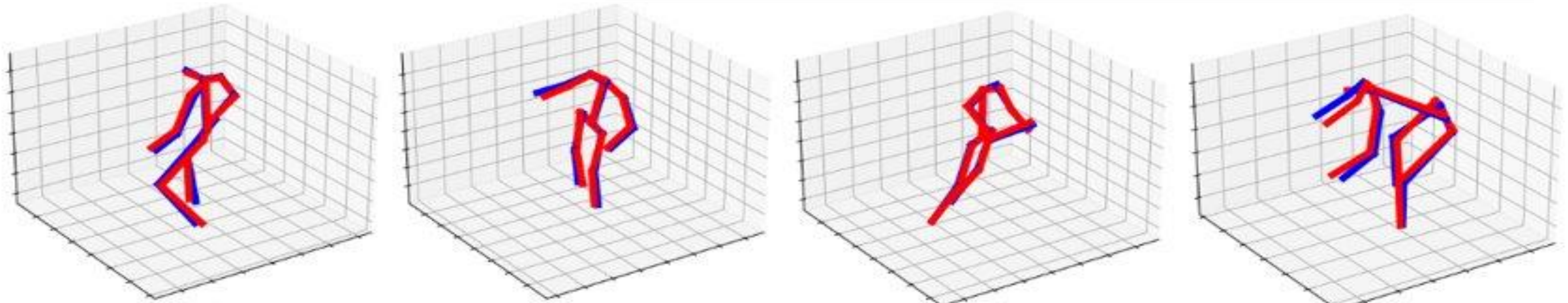


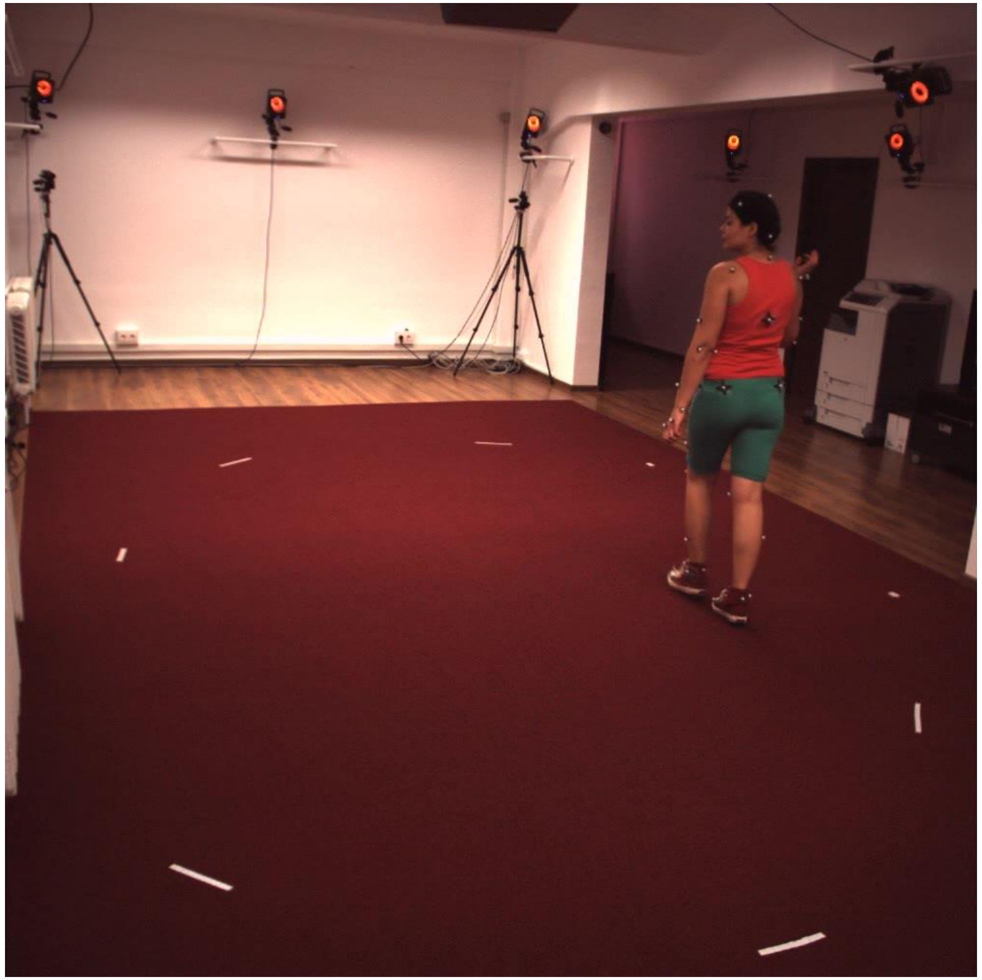
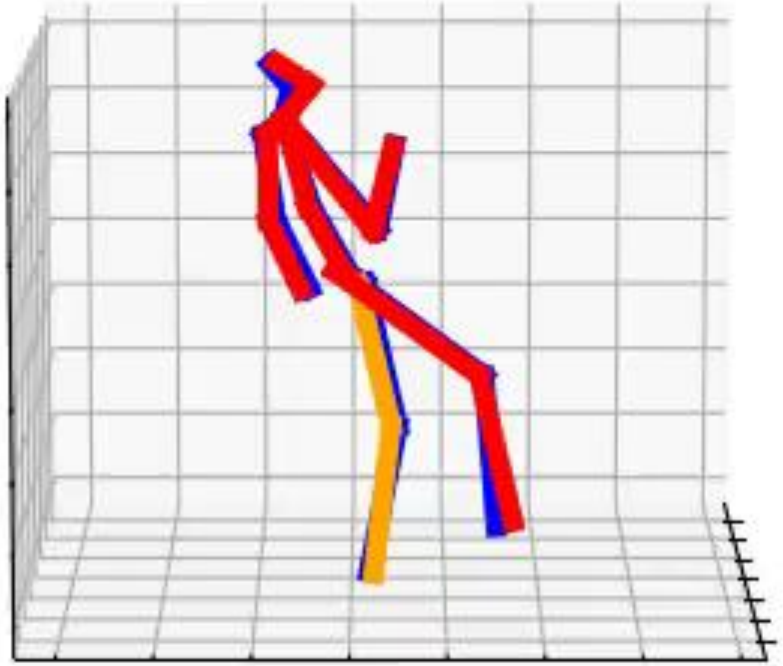
Pose reconstruction errors on the Human3.6M

- The Procrustes aligned mean per-joint position error (PA-MPJPE)
- Normalized mean per joint position error (N-MPJPE)
- The smaller, the better for both

TABLE III

Method	Occlusion	PA-MPJPE	N-MPJPE
LInKs [2]	None	33.8	61.6
Ours (Recreation)	None	37.2	61.7
Ours (Recreation)	Left Arm	52.1	78.1
Ours (Recreation)	Left Leg	46.0	73.2
Ours (Recreation)	Right Arm	49.8	75.7
Ours (Recreation)	Right Leg	44.5	71.6
Ours (Recreation)	Left Arm & Leg	62.0	86.0
Ours (Recreation)	Right Arm & Leg	60.2	83.7
Ours (Recreation)	Both Legs	69.3	99.8
Ours (Recreation)	Torso	88.4	122.0



Matching algorithm results

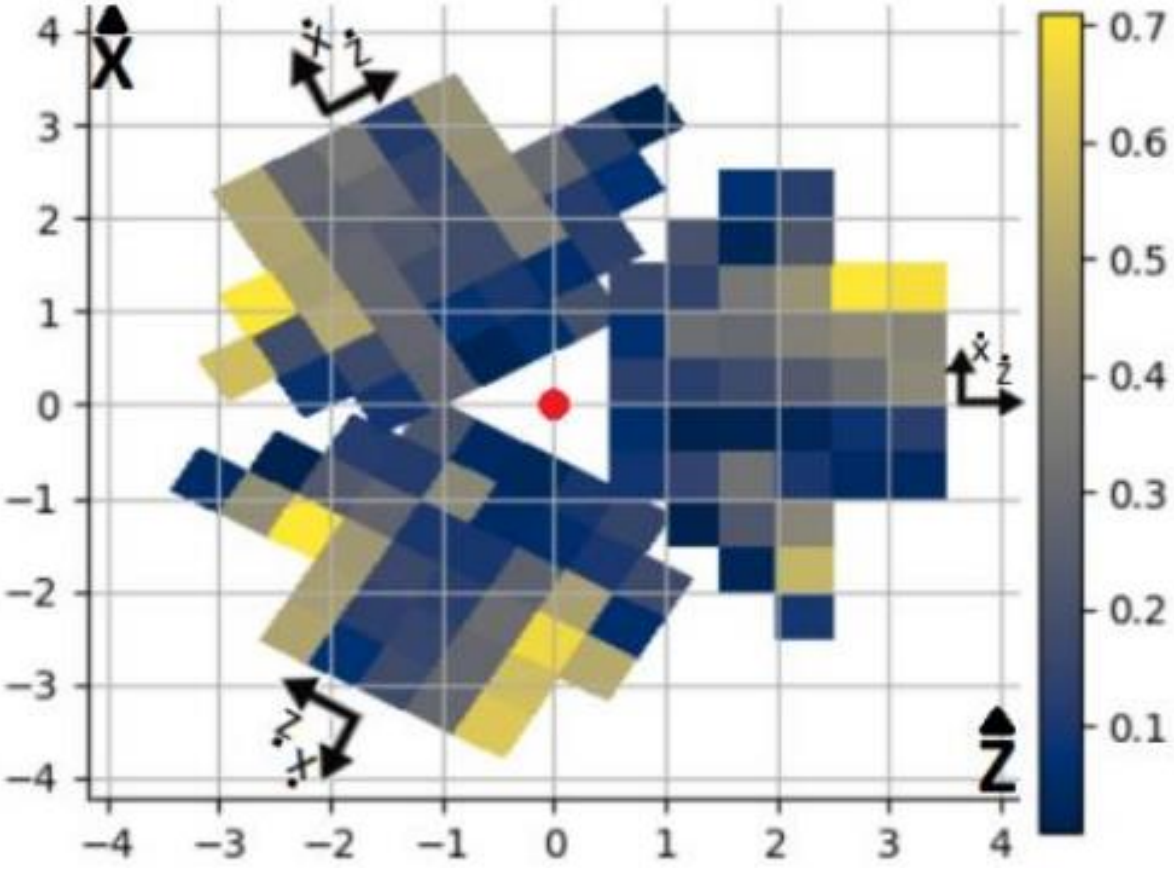
TABLE II

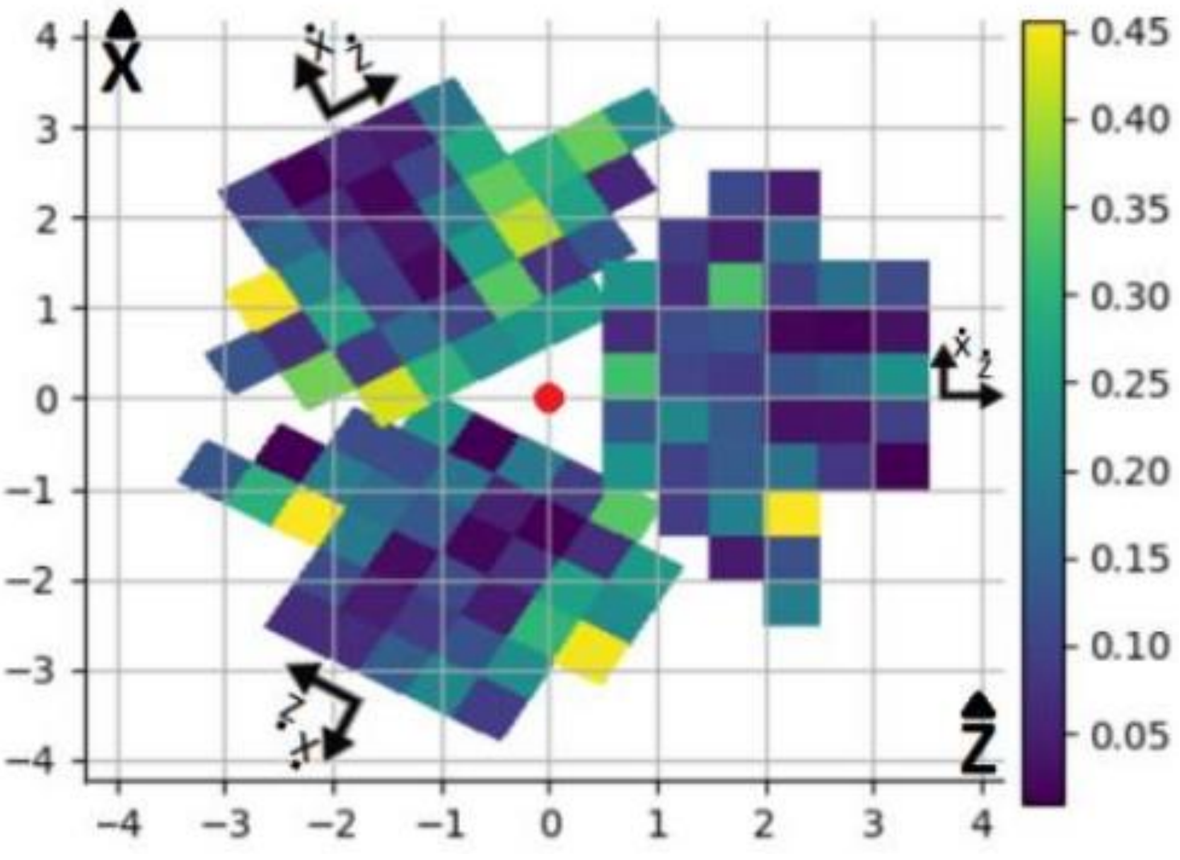
	Radar 1 ↓	Radar 2 ↓	Radar 3 ↓
Preliminary Work [3]	23.89% ± 6.57%	33.57% ± 50.55	66.89% ± 263.89
Ours	2.52% ± 2.51	9.44% ± 13.27	1.94% ± 1.52

- Low average matching error of 4.63%
- Error represents the absolute difference between the radar and camera matching values of an individual, divided by the camera values

Radar detection errors

- Localisation error in meters.
- The errors were evaluated in each radar’s xˆ (left) and zˆ (right) directions.
- The figures represent these errors in the (Xˆ , Zˆ) 2D global coordinate system.





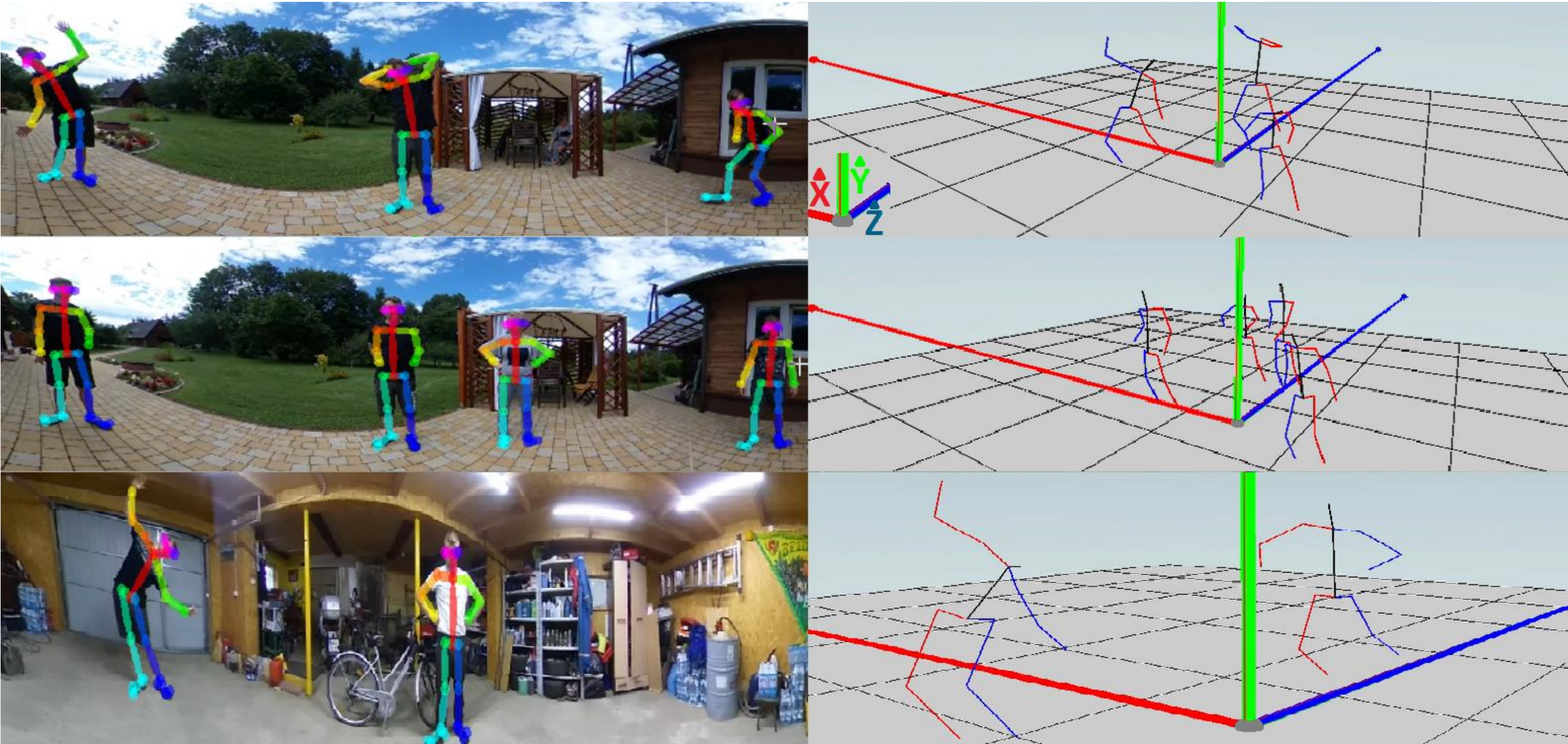
Radar detection errors

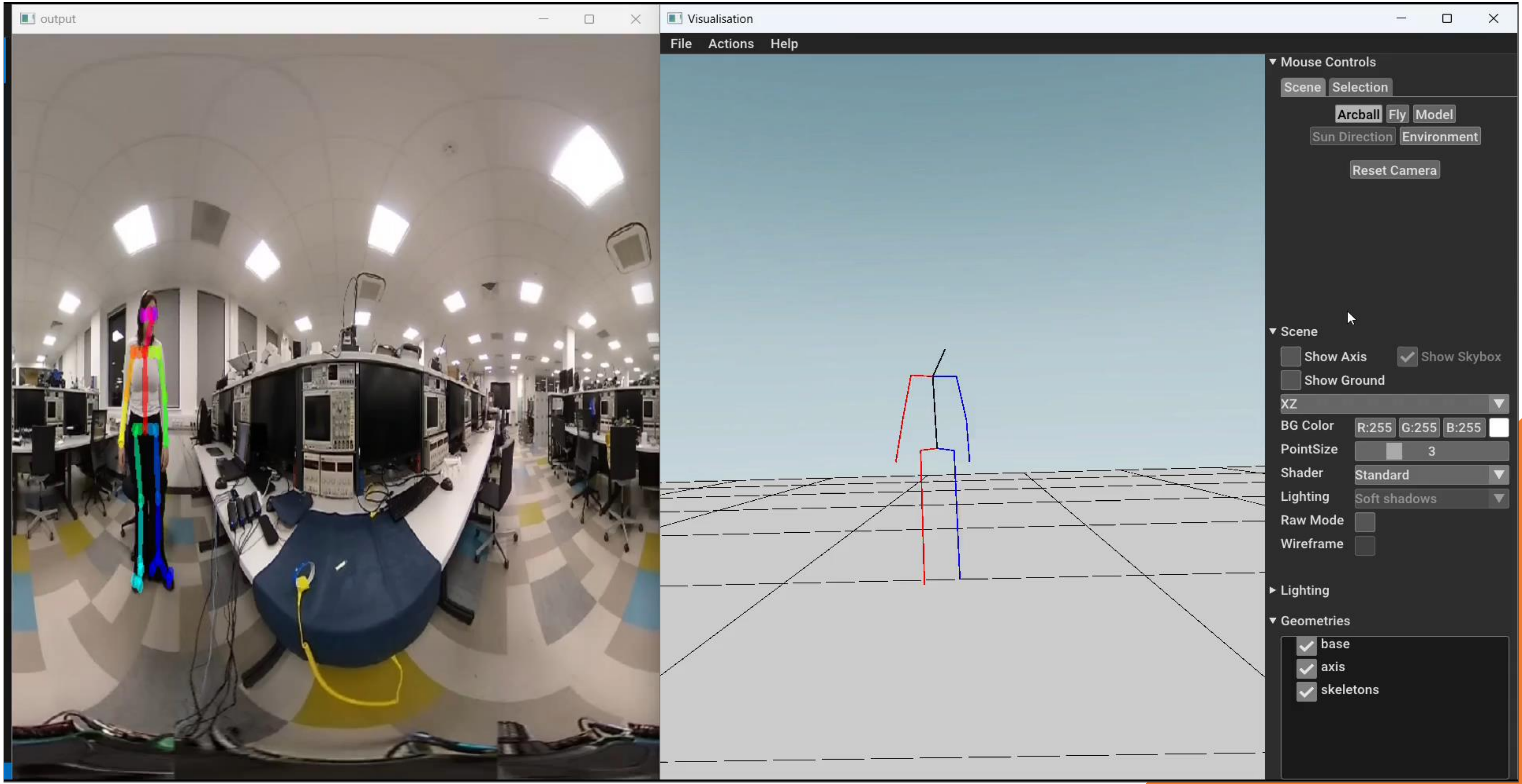
- Localization errors in centimeters were reduced with radar and camera calibrations

TABLE I

Radar	Direction	Preliminary [3]	Ours
1	x	20.65	16.45
	z	11.41	11.45
2	x	26.19	24.86
	z	15.39	10.77
3	x	16.88	15.94
	z	13.83	13.46

Qualitative system's results





https://youtu.be/FAFVYWSzu7Q

Conclusions

- Robust detection system
- Performs consistently regardless of the number of individuals.
- Theoretically can handle any number of detected people.
- Another limitation is the system's inability to accurately detect when a person is facing away from it.

[1] Z. Cao etal. “Openpose: Realtime multi-person 2d pose estimation using part affinity fields”, 2019.

[2] Peter Hardy and Hansung Kim. “LInKs - Lifting Independent Keypoints - Partial Pose Lifting for Occlusion Handling with Improved Accuracy In 2D-3D Human Pose Estimation”, 2023.

[3] Aarti Amin et al. “Real-time 3d multi-person pose estimation using an omnidirectional camera and mmwave radars”, 2023