

# Event-based Egocentric Human Pose Estimation in Dynamic Environment



<sup>†</sup>Wataru Ikeda, <sup>†</sup>Masashi Hatano, <sup>†</sup>Ryosei Hara, <sup>†§</sup>Mariko Isogawa

<sup>†</sup>Keio University, <sup>§</sup>JST Presto

Accepted at ICIP 2025



## Overview

### Problem Statement

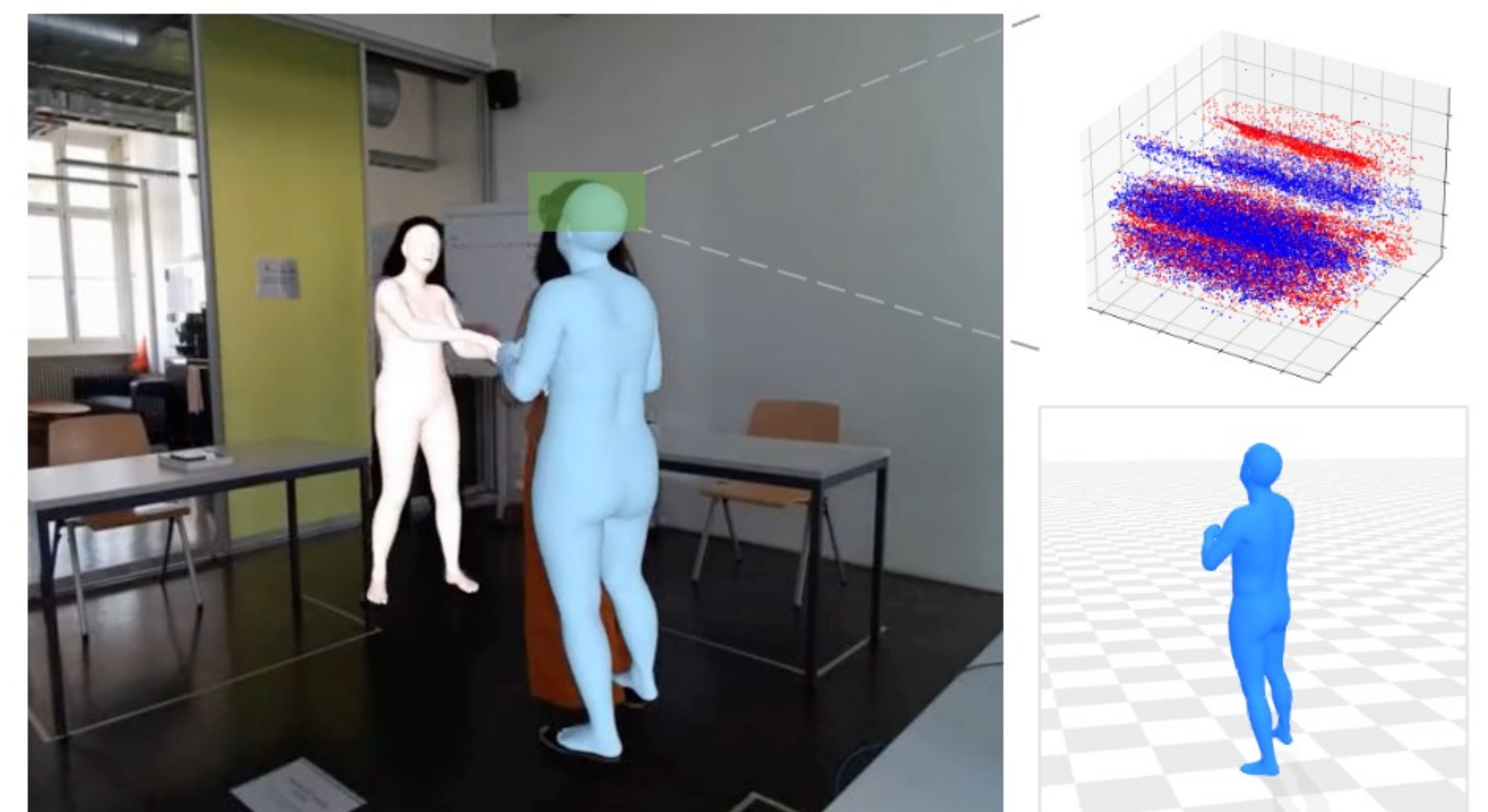
We aim to estimate 3D human pose with event-based egocentric camera in dynamic environment.

### Challenges

- RGB methods: Difficult to capture in **low-light conditions**, and fast motion causes **motion blur**.
- Handling Dynamic Environments: Existing methods do not work well in **dynamic environments**. They assume a static scene.

### Contributions

Proposed a novel task and a framework / Developed the Motion Segmentation Module / Created a new dataset



## Method

### 1. Voxelization

Convert the event point cloud into a **voxel grid**. This keeps the time information and reduces the amount of computation.

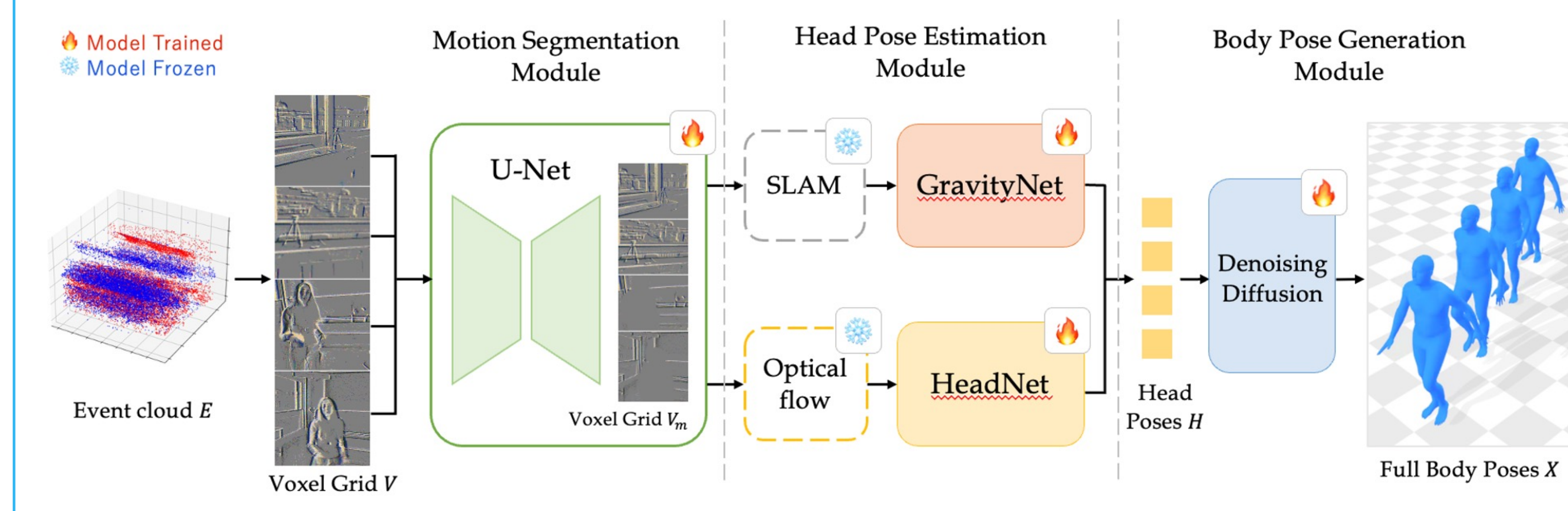
### 2. Motion Segmentation

**Remove moving objects** and extract background information. This makes the method work in dynamic environments.

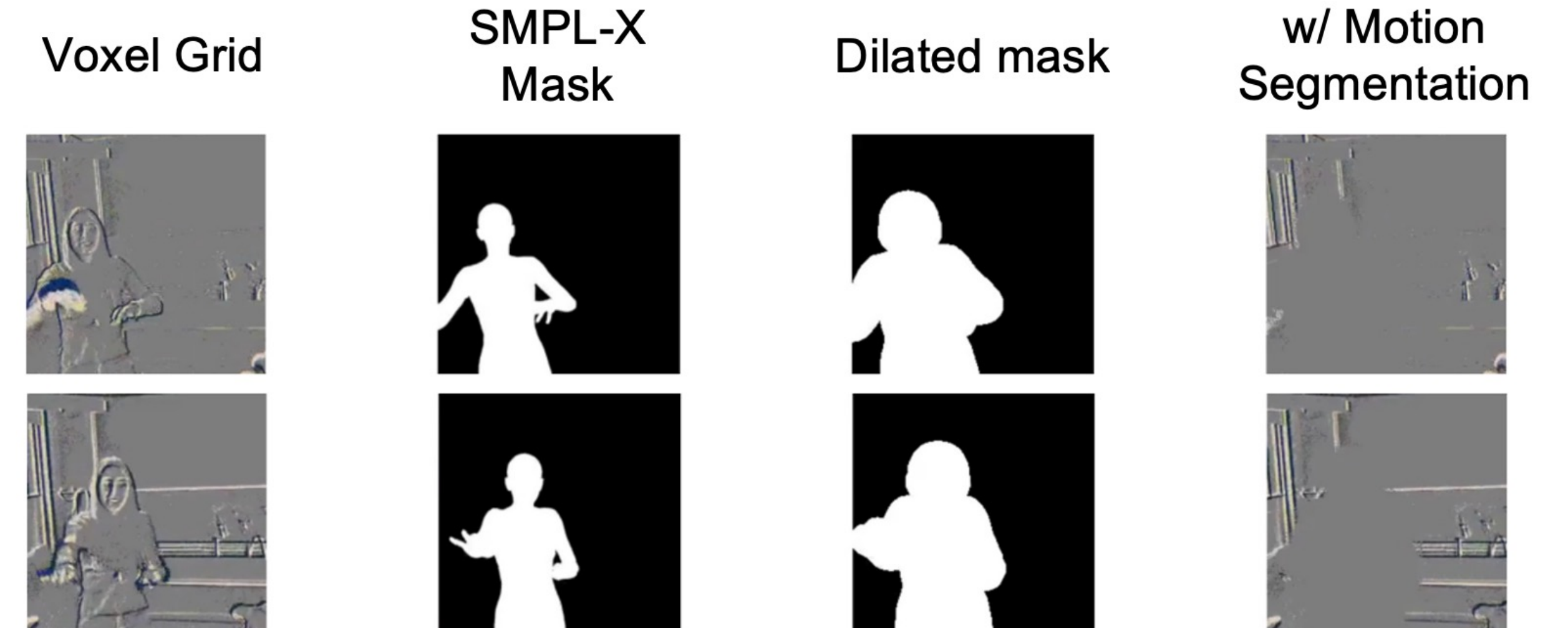
### 3. Head and Body Pose Estimation

- Use the processed voxel grid from steps 1 and 2 to **estimate the head pose**
- Conditioned on the head pose, generate the body pose using a diffusion model

## Pipeline Overview



## Motion Segmentation Result



## Experiment

### Dataset

- **Synthetic events** are generated from EgoBody
- Dynamic masks are created from SMPL-X

### Quantitative Evaluation

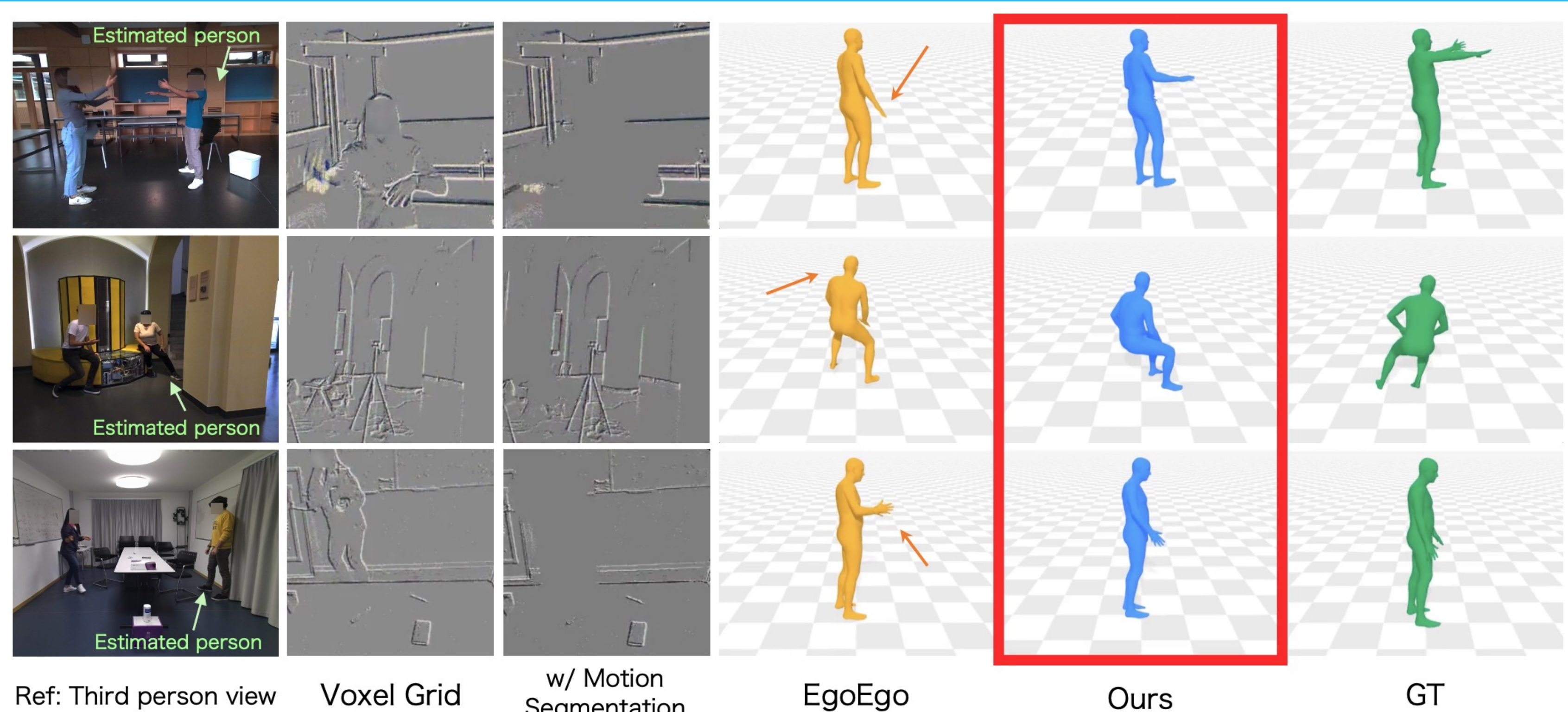
- Our method outperformed in 4 out of 5 metrics
- For head pose, it performed better in both of the 2 metrics

### Qualitative Evaluation

Our method gives more accurate results, especially for **hand** and **head positions**.

### Future Work

Test on real data / Combine with other modalities / Interactions with people and the environment



	$O_{head} \downarrow$	$T_{head} \downarrow$	MPJPE $\downarrow$	Accel $\downarrow$	FS $\downarrow$
EgoEgo*	0.293	126.6	<b>119.5</b>	2.87	0.79
Ours	<b>0.282</b>	<b>121.8</b>	121.5	<b>2.69</b>	<b>0.64</b>
EgoEgo w/ GT mask	0.281	121.8	121.4	2.60	0.55