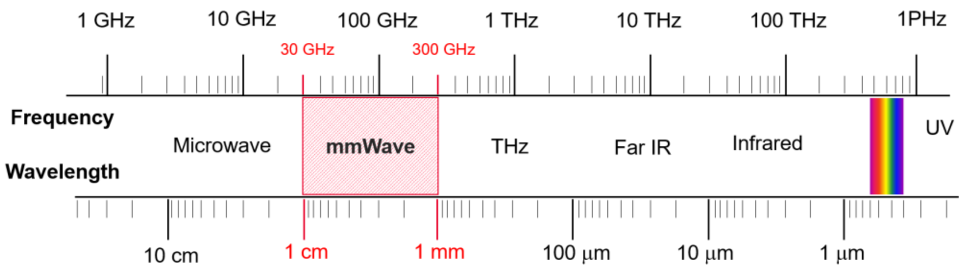
milliFlow: Scene Flow Estimation on mmWave Radar Point Cloud for Human Motion Sensing

Paper #3470

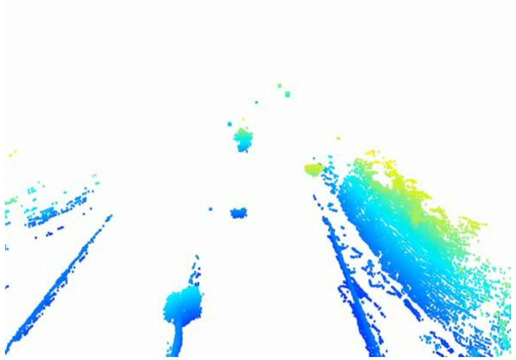
Fangqiang Ding[1], Zhen Luo[1], Peijun Zhao[2], Chris Xiaoxuan Lu[3]
[1]University of Edinburgh, [2]MIT, [3]UCL

Advantages of mmWave radar for human sensing

- Robust to poor lighting conditions (e.g., low light and glare) and airborne particles (e.g., smoke, fog, rain)



RGB camera



LiDAR point cloud

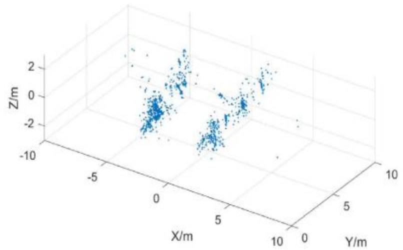
K-RADAR DATASET

Optical sensors (i.e., camera, LiDAR) can not see through airborne particles.

THE UNIVERSITY of EDINBURGH

Advantages of mmWave radar for human sensing

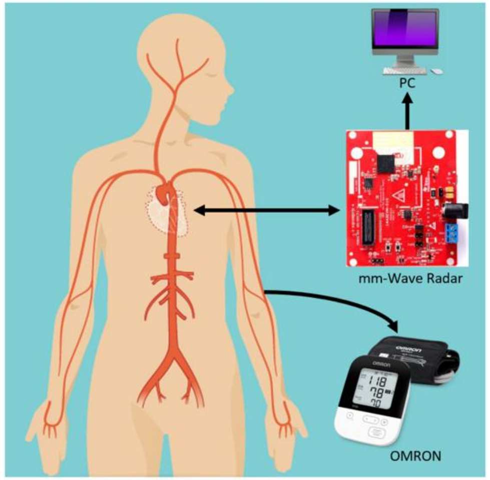
- Privacy-preserving sensing



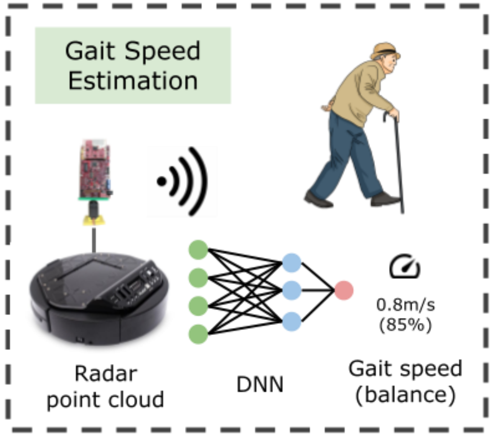
- Psychologically non-intrusive data

Applications of mmWave radar for human sensing

Vital sign monitoring

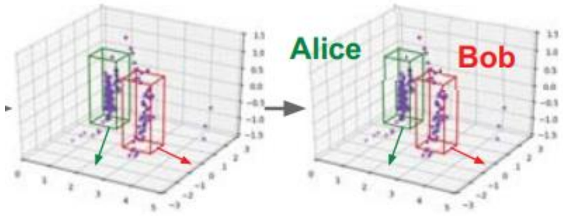


Gait analysis

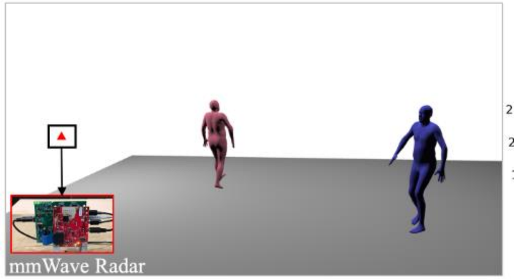


Gesture/activity recognition

Human tracking and identification

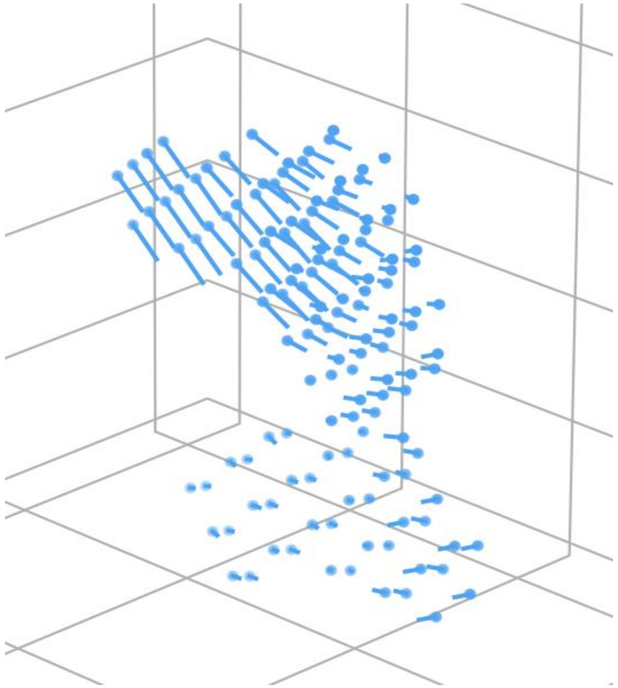


Pose/mesh estimation



THE UNIVERSITY of EDINBURGH

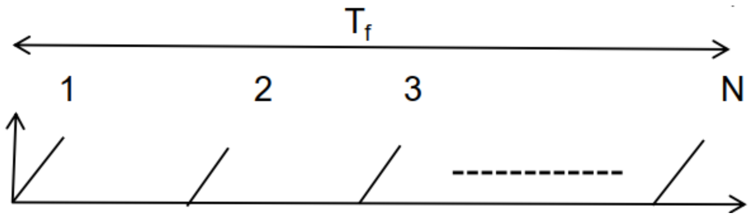
mmWave-based human motion sensing



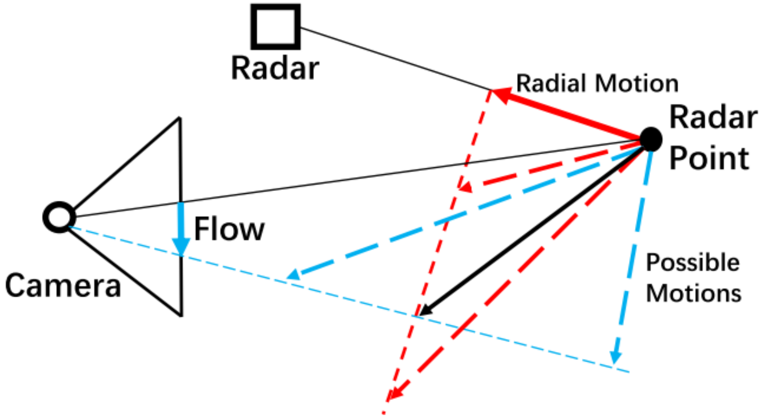
- Observation: pointwise velocity per radar frame is intuitively a strong cue for improving the motion estimation robustness for human sensing.

Fine-grained velocity is difficult to obtain when it comes to mmWave radars.

Why not doppler velocity measurement

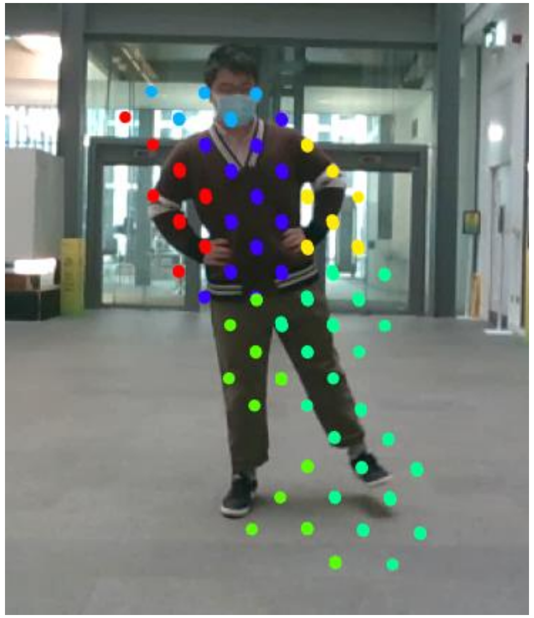


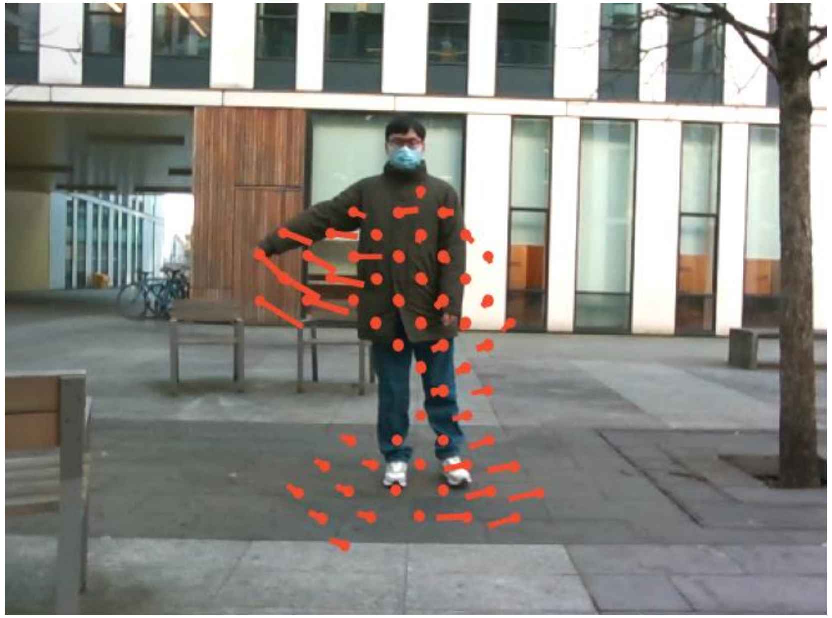
Doppler velocity $v > v_{res} = \frac{\lambda}{2T_f}$

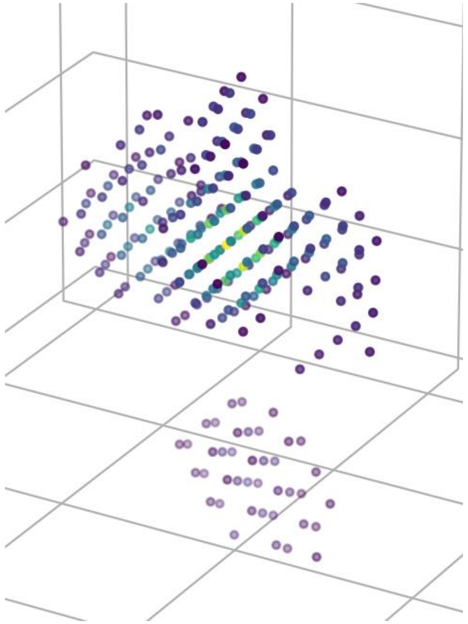


- Low-resolution (>0.15m/s)
- Ambiguity in tangential direction
- Our case: absent velocity measurement

Why not conventional tracking methods

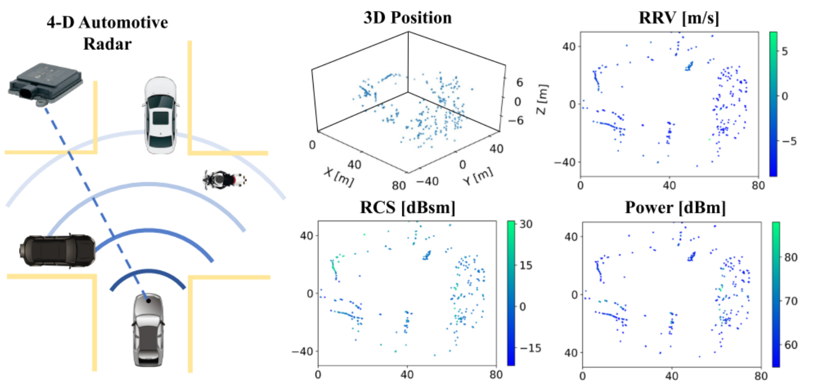


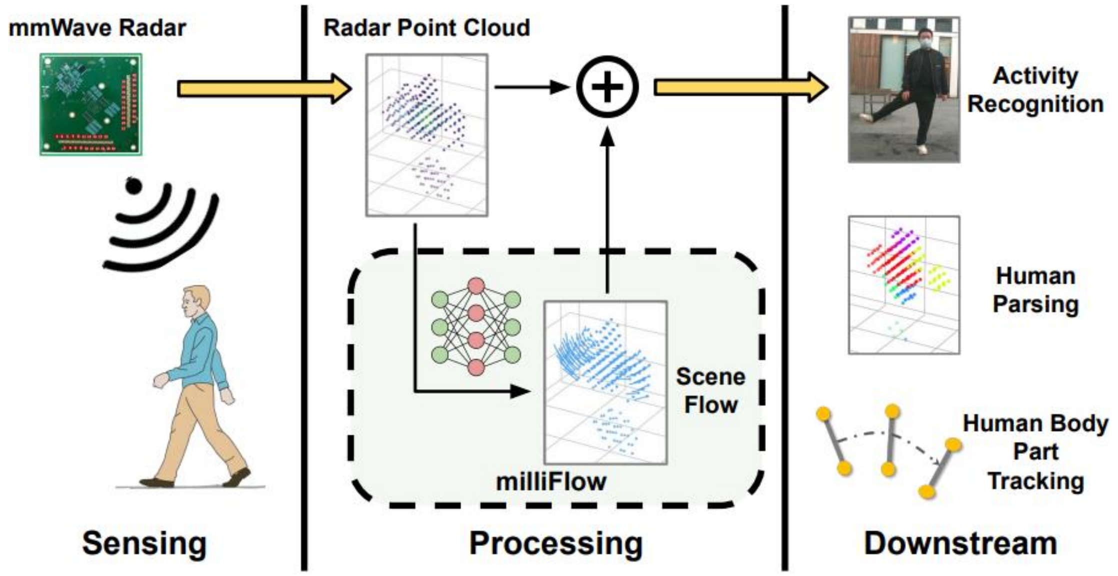




- Highly error-prone: extreme sparse, ghost points and missing body parts

Scene flow estimation with mmWave radar

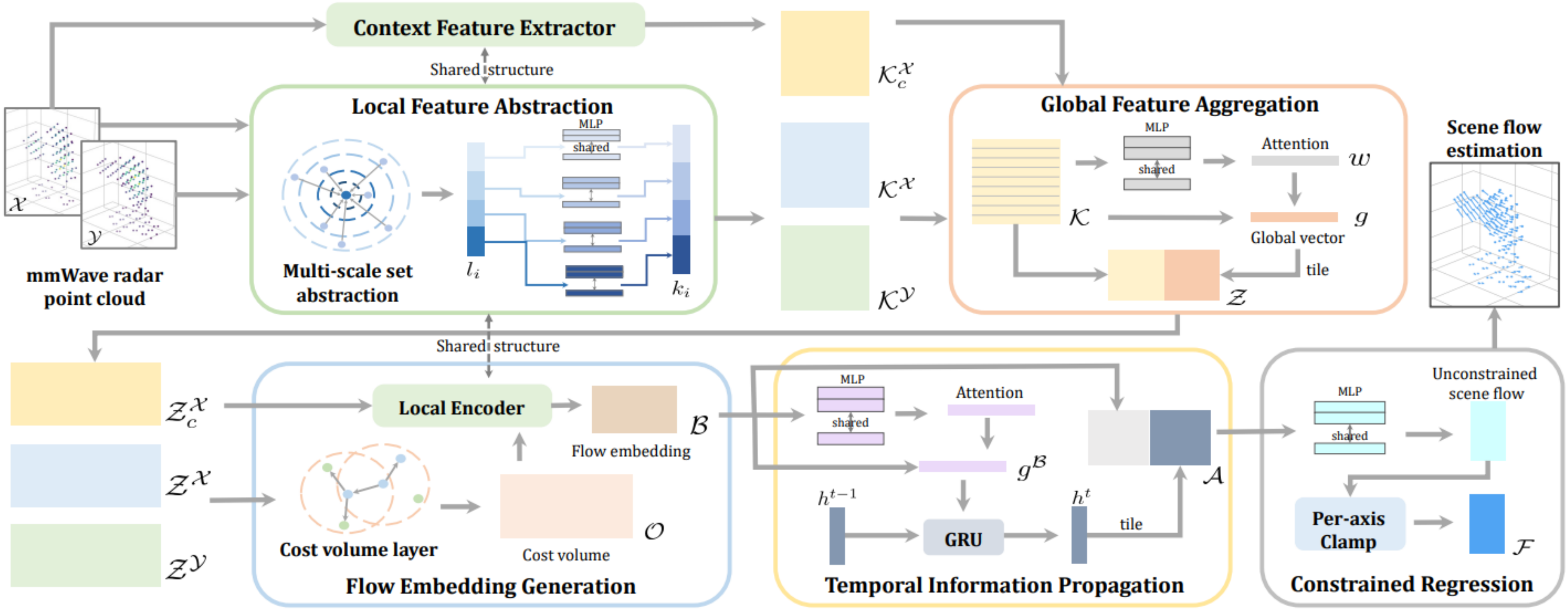




- Inspiration: 4D radar scene flow estimation in autonomous driving
- Insight: estimate and use scene flow as intermediate features to support human motion sensing
- Hard to transfer: different radar hardware; non-rigid human body motion

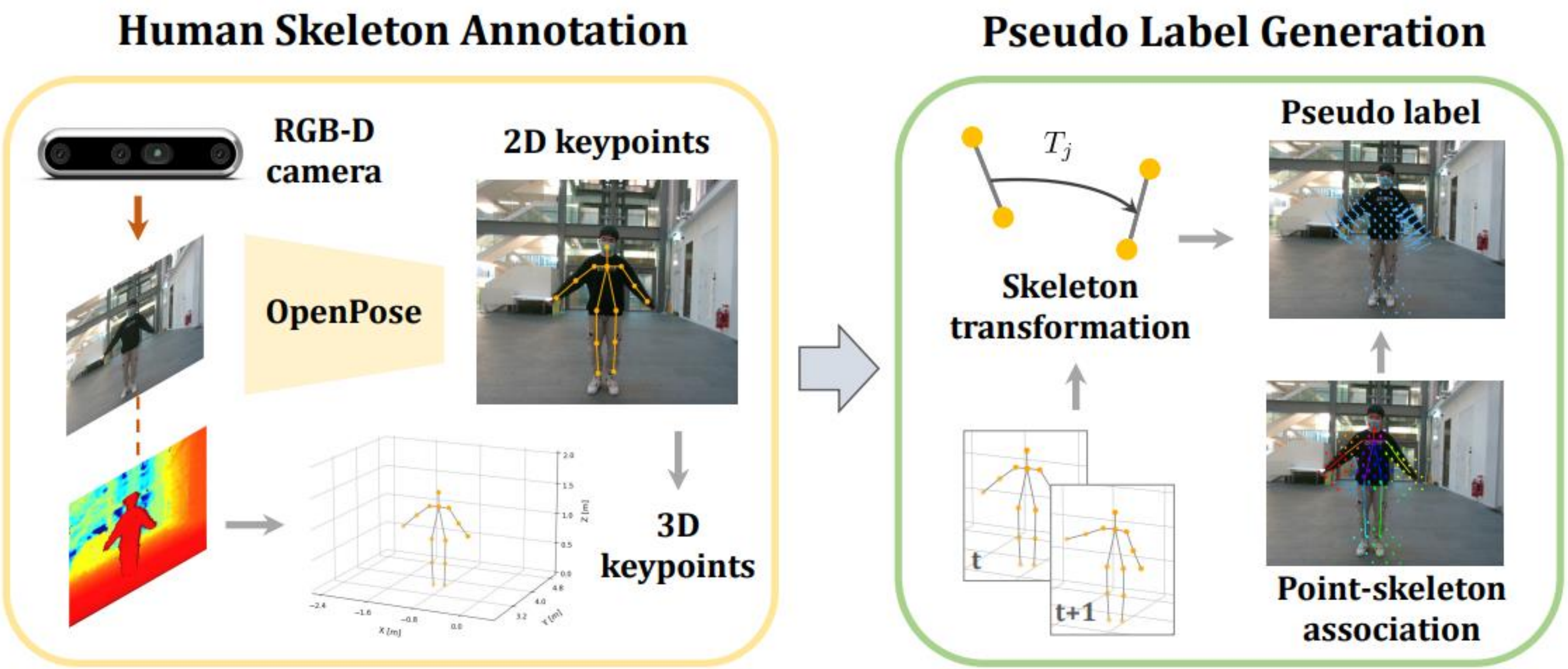
THE UNIVERSITY *of* EDINBURGH

milliFlow's overall network



- Global feature aggregation - sparsity and noise
- Temporal information propagation - lack of temporal cues
- Constrained regression – refrain from non-viable results

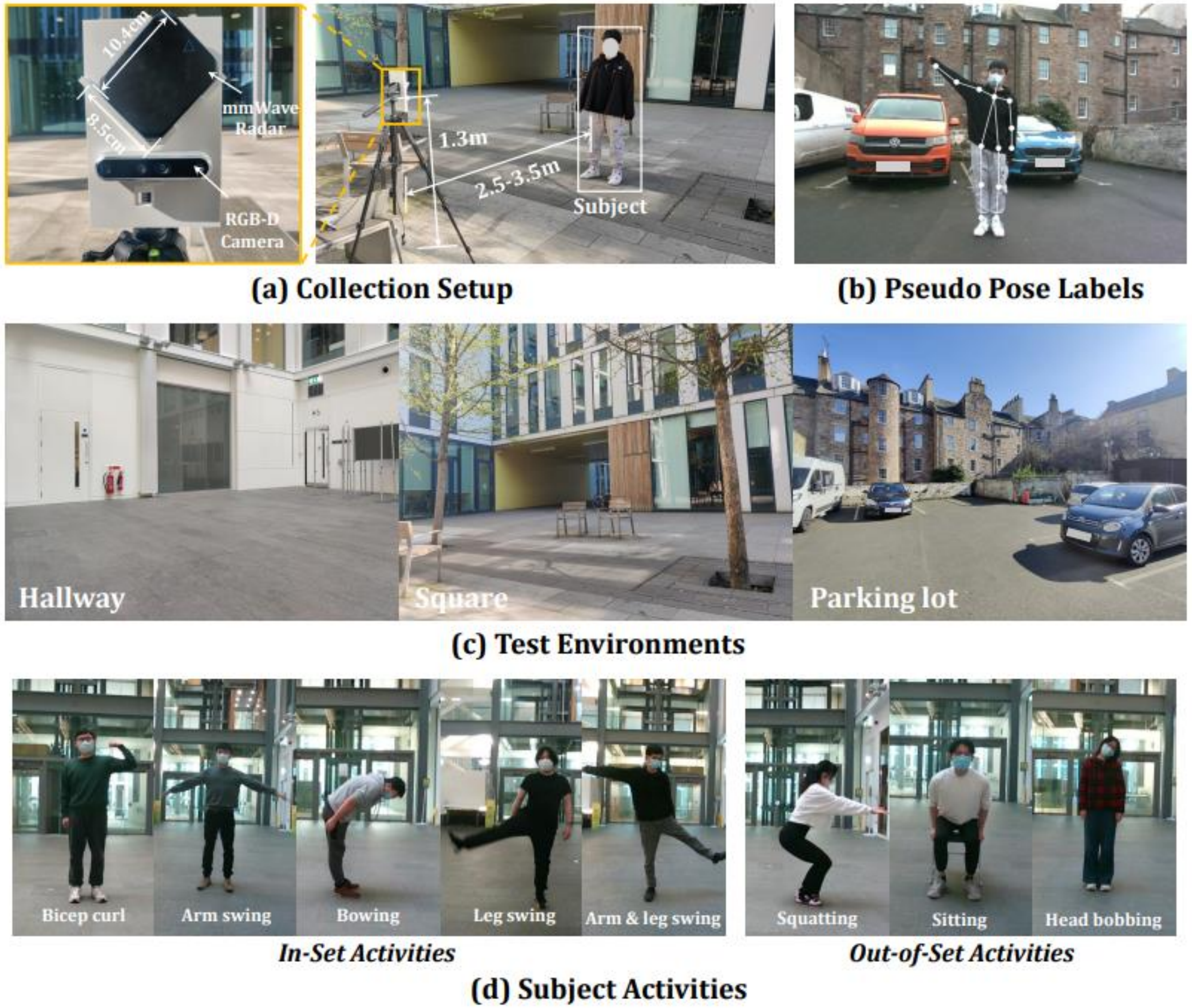
Automatic scene flow annotation for training



- Non-rigid human body movement
- Skeleton-based rigid-motion assumption

THE UNIVERSITY of EDINBURGH

Large-scale human motion sensing dataset



THE UNIVERSITY of EDINBURGH

mmWave scene flow results

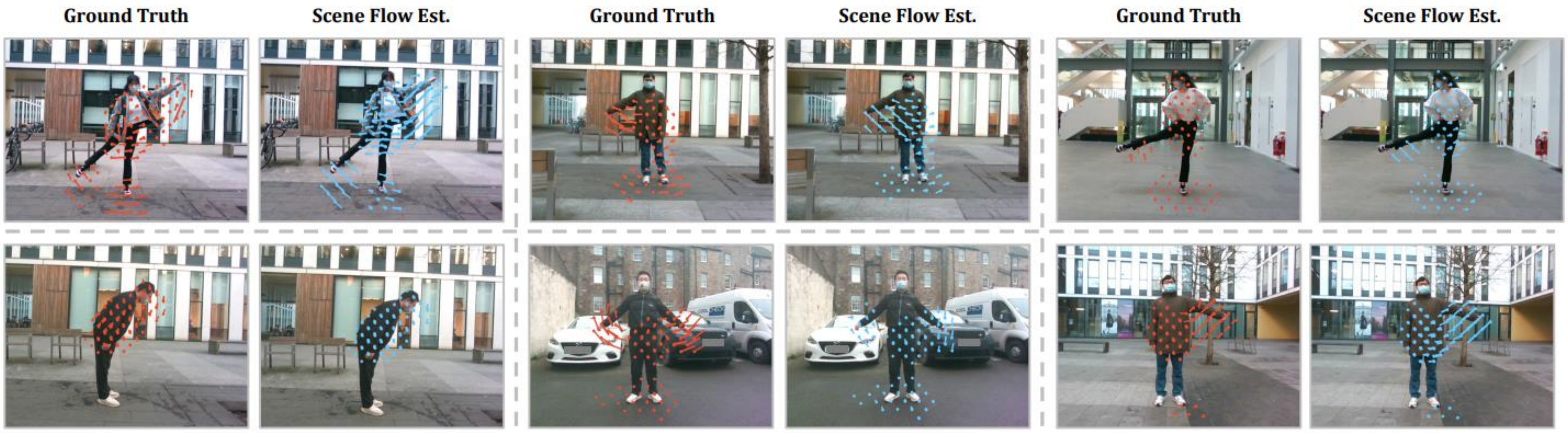
Method	EPE3D (m) ↓ All	EPE3D (m) ↓ Moving	EPE3D (m) ↓ Static	Acc3D ↑ Strict	Acc3D ↑ Relax
FlowNet3D [43]	0.293	0.290	0.259	0.016	0.095
PPWC-Net [71]	0.171	0.181	0.128	0.138	0.179
Graph Prior [49]	0.315	0.322	0.283	0.007	0.011
FLOT [50]	0.299	0.307	0.265	0.015	0.094
FlowStep3D [35]	0.243	0.251	0.216	0.062	0.109
PV-RAFT [69]	0.161	0.170	0.107	0.179	0.292
RaFlow [21]	0.107	0.115	0.094	0.271	0.427
Bi-PFNet [15]	0.159	0.168	0.111	0.153	0.264
milliFlow (ours)	**0.046**	**0.051**	**0.009**	**0.406**	**0.703**

- State-of-the-art comparison

	Method	EPE3D (m) ↓ All	EPE3D (m) ↓ Moving	EPE3D (m) ↓ Static	Acc3D ↑ Strict	Acc3D ↑ Relax
(a)	Full version	**0.046**	**0.051**	**0.009**	**0.406**	**0.703**
(b)	(a) w/o TP	0.053	0.062	0.018	0.382	0.676
(c)	(b) w/o GA	0.061	0.068	0.025	0.361	0.628
(d)	(c) w/o CF	0.071	0.077	0.028	0.315	0.536
(e)	(d) w/o CR	0.083	0.090	0.034	0.286	0.490

- Network ablation study

mmWave scene flow visualization



Showing radar points and scene flow vectors on the 2D image via perspective projection

Benefit downstream human sensing tasks

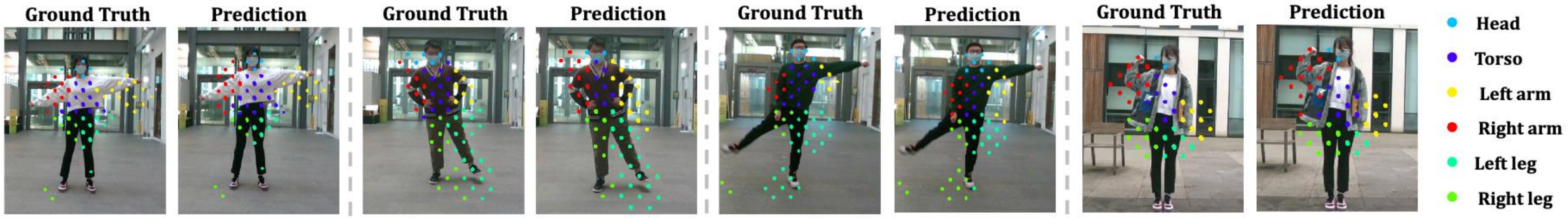
- Results on human action recognition and human parsing

Method	Raw	w. S1	Gain	w. S2	Gain
Ours	47.32	57.88	+10.56	57.78	+10.46
MMPointGNN [29]	52.46	60.16	+7.70	59.94	+7.48
RadHAR [67]	44.65	49.98	+5.33	50.53	+5.88
Average	48.14	56.01	+7.87	56.08	+ 7.94

Method	mIoU (%)	Gain (%)	oA (%)	Gain (%)
Raw	49.09	-	65.75	-
w. S1	52.72	+3.63	69.27	+3.52
w. S2	51.04	+1.95	68.21	+2.46

Strategy 1 (S1): take scene flow as point-level feature;
Strategy 2 (S2): use the latent representation from scene flow network

- Human parsing visualization



- Results on human body part tracking

Activity	Tracking length - mJE (m) ↓ 1	2	3	4
Arm swing	0.028	0.076	0.097	0.124
Leg swing	0.016	0.071	0.105	0.130
Arm & leg swing	0.030	0.108	0.146	0.178
Average	0.025	0.085	0.116	0.144

Thanks for watching the presentation!

milliFlow: Scene Flow Estimation on mmWave Radar Point Cloud for Human Motion Sensing

Paper #3470

Fangqiang Ding[1], Zhen Luo[1], Peijun Zhao[2], Chris Xiaoxuan Lu[3]
[1]University of Edinburgh, [2]MIT, [3]UCL