ORAL PRESENTATION · INDOOR PEDESTRIAN NAVIGATION

# Seeing Around the Corner

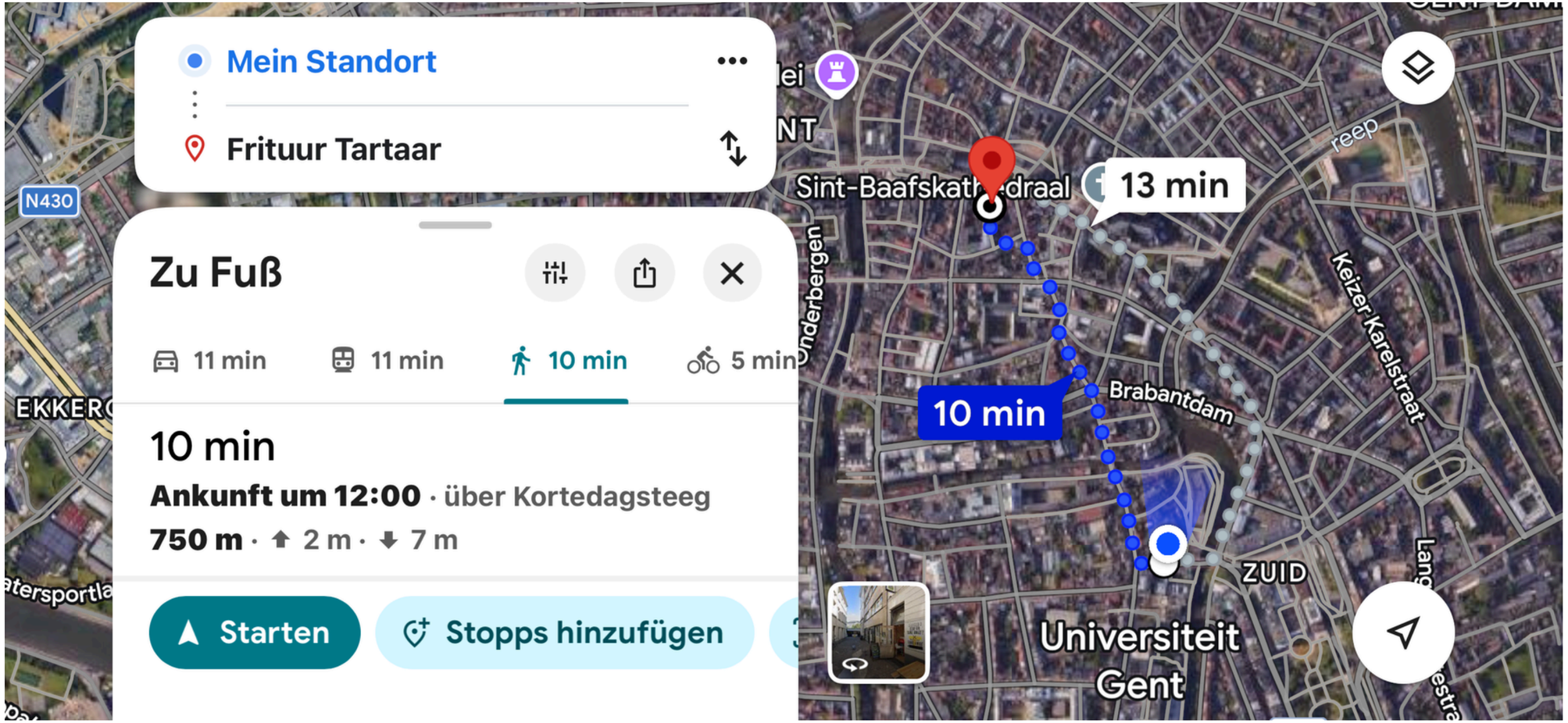
## Fusing Visual Flow and Inertial Sensors for Indoor Pedestrian Navigation.

**Noah Meißner**
Chair of Information Science · University of Regensburg

Tim Sieber
Chair of Information Science · University of Regensburg

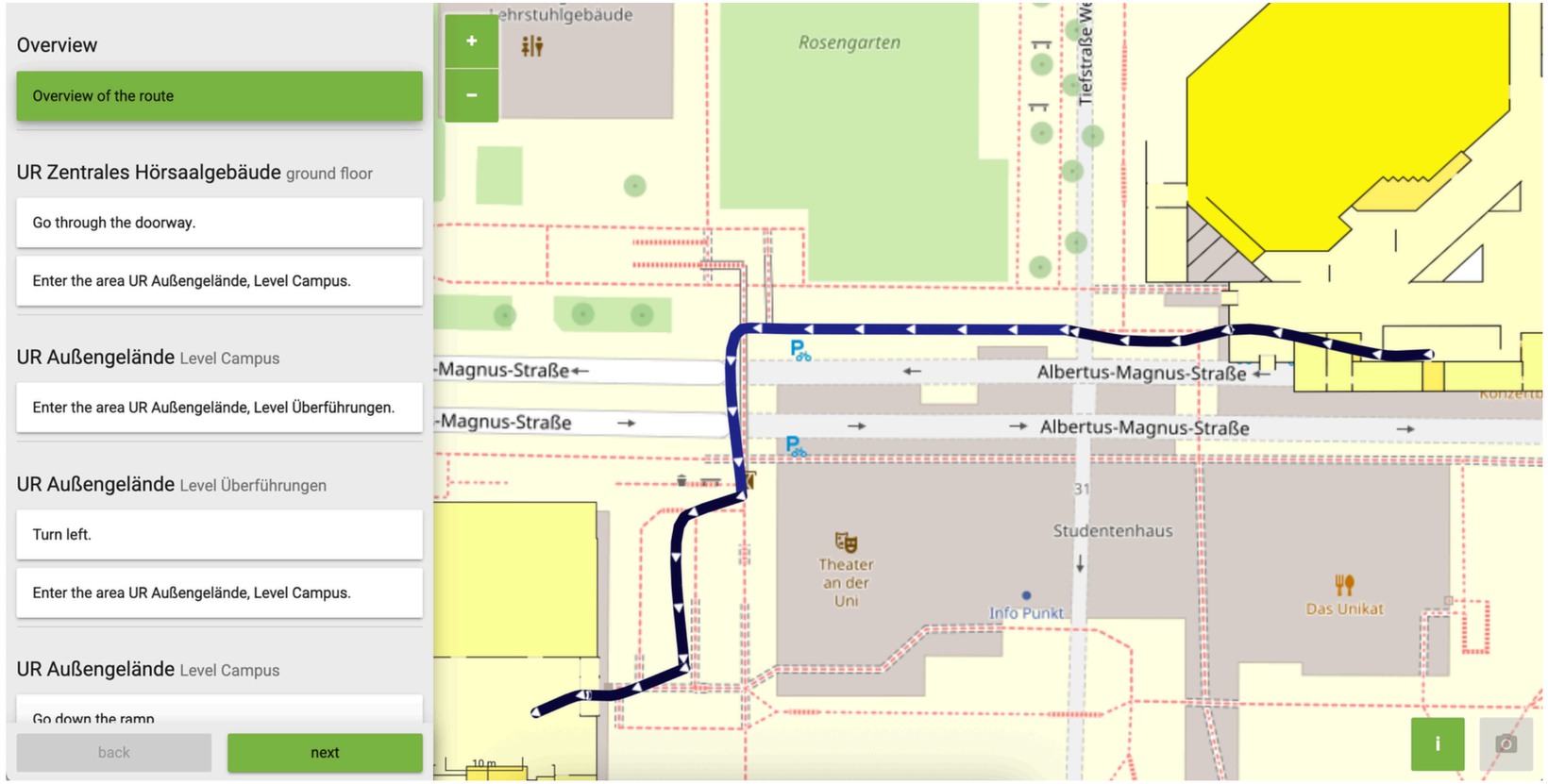
Bernd Ludwig
Chair of Information Science · University of Regensburg


# Where is the next Frituur shop?



# How can we find a space in the building?

# URWalking - Live System

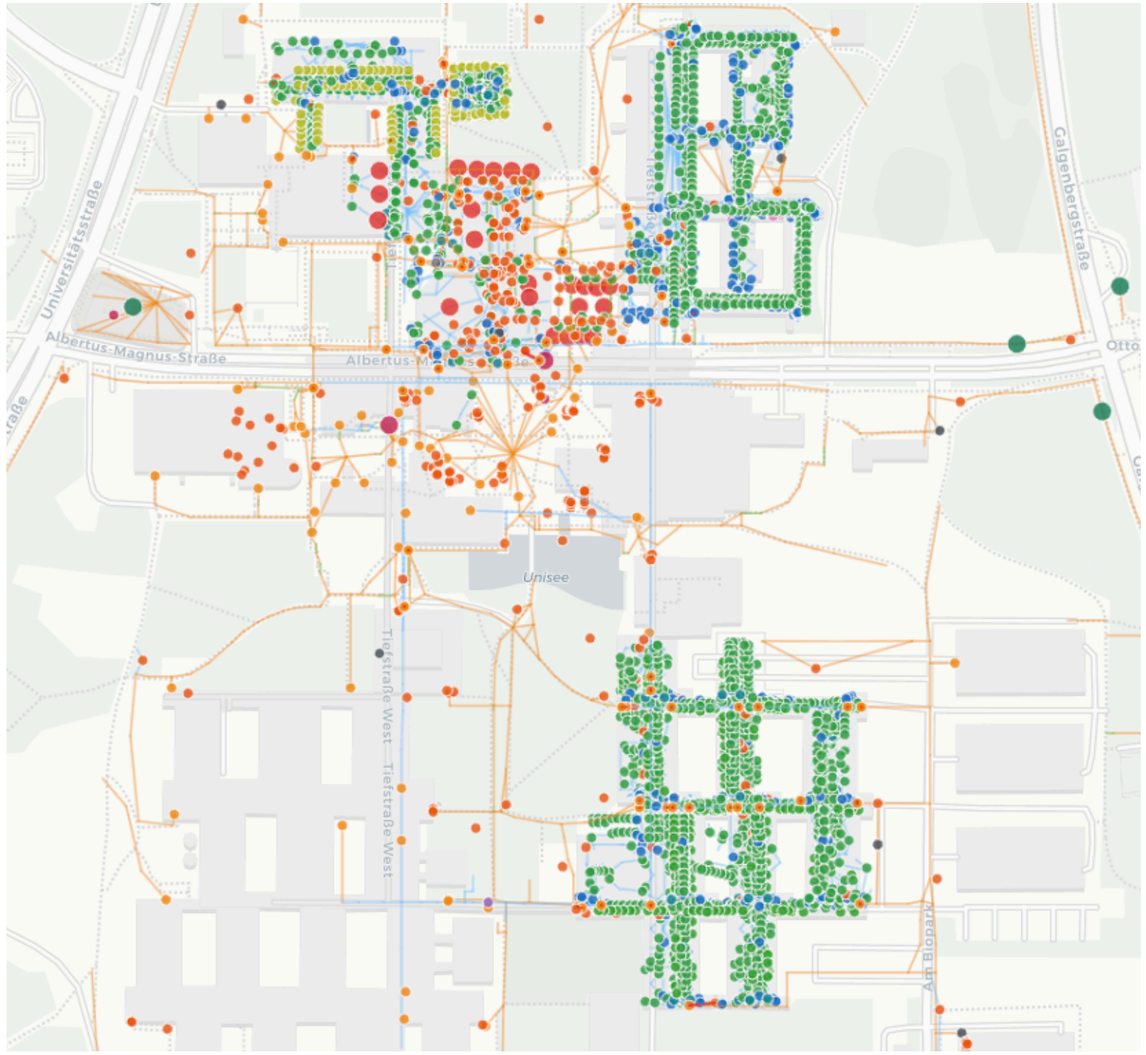


# Campusnavigation - UR Walking

01 **5000 rooms**
Located in 30+ Buildings

02 **50.000 search requests**
since 2018

03 **Live System**
urwalking.de




Top 3 Buildings (request history)

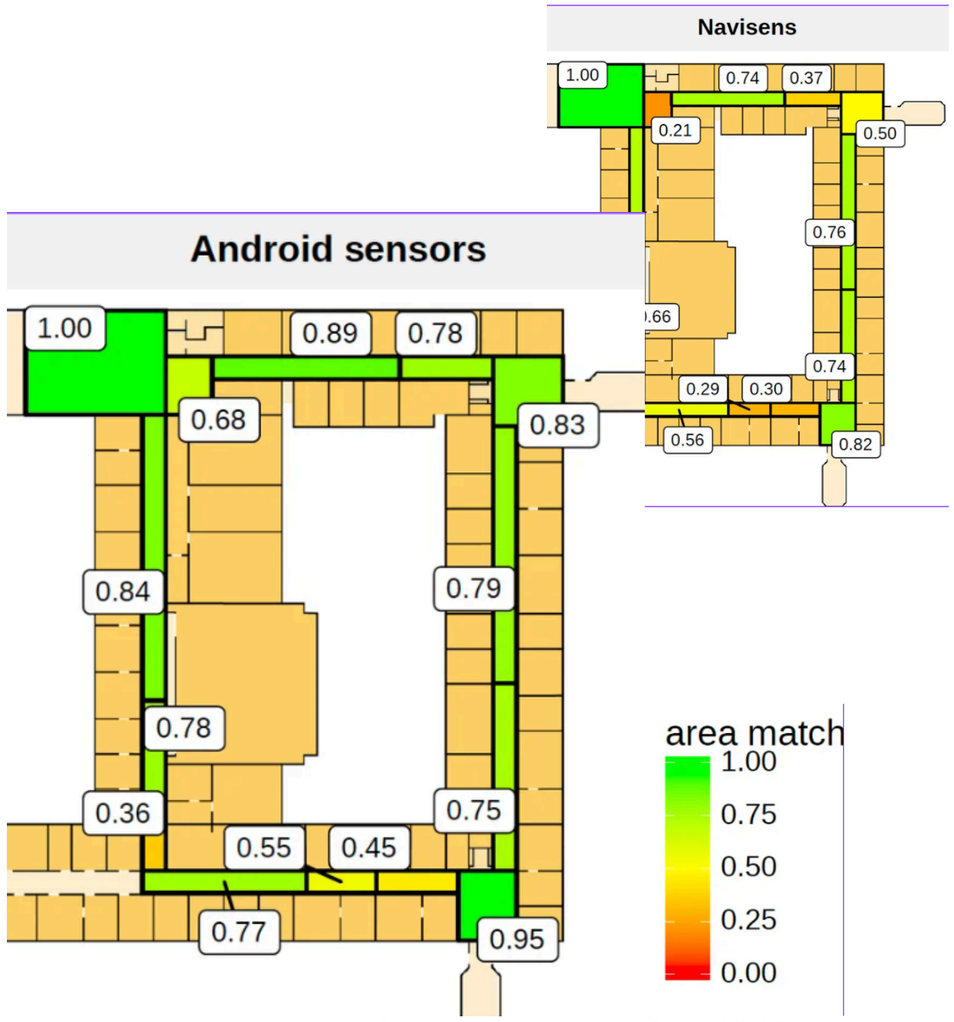
# Pedestrian Dead Reckoning (PDR)

01 **Turns increase localisation error** — Jackermeier et al.(2018)
in Area Matching

02 **Difficult to create a ground truth** — Sarlin et al. (2024)
LIDAR Sensors as Ground Truth

03 **URWalking needs scalability**
5000 rooms, 30+ buildings

_ **User base shows importance**
especially disabeld people




Jackermeier et al. (2018)

# Dataset - LaMAR

01 **45000m2** Sarlin et al. (2024)

around different buildings

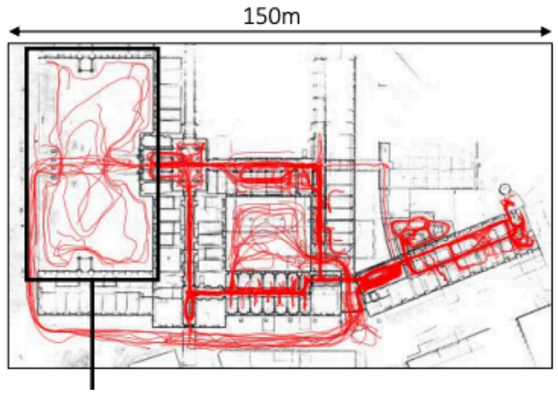
02 **But Time consuming** Sarlin et al. (2024)

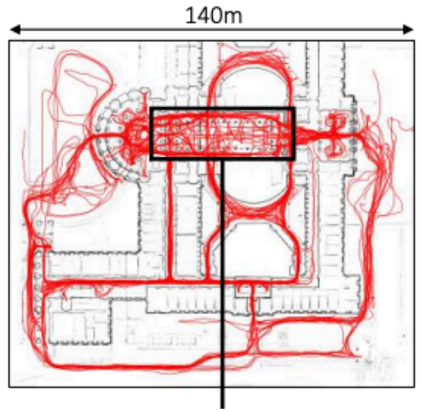
Use Laser Scans as ground truth

03 **URWalking needs scalability**

5000 rooms, 30+ buildings

_ **We do not need <1cm error**




## GT from laser scanners

- trolley motion
- 6x 113° 1080p RGB cameras every 1-3m
- lidar points
- WiFi, BT
- poses from lidar SLAM

- backpack motion
- 4x 90° 1080p RGB cameras every 1-3m
- lidar points
- WiFi, BT
- poses from lidar SLAM

Sarlin et al. (2024)

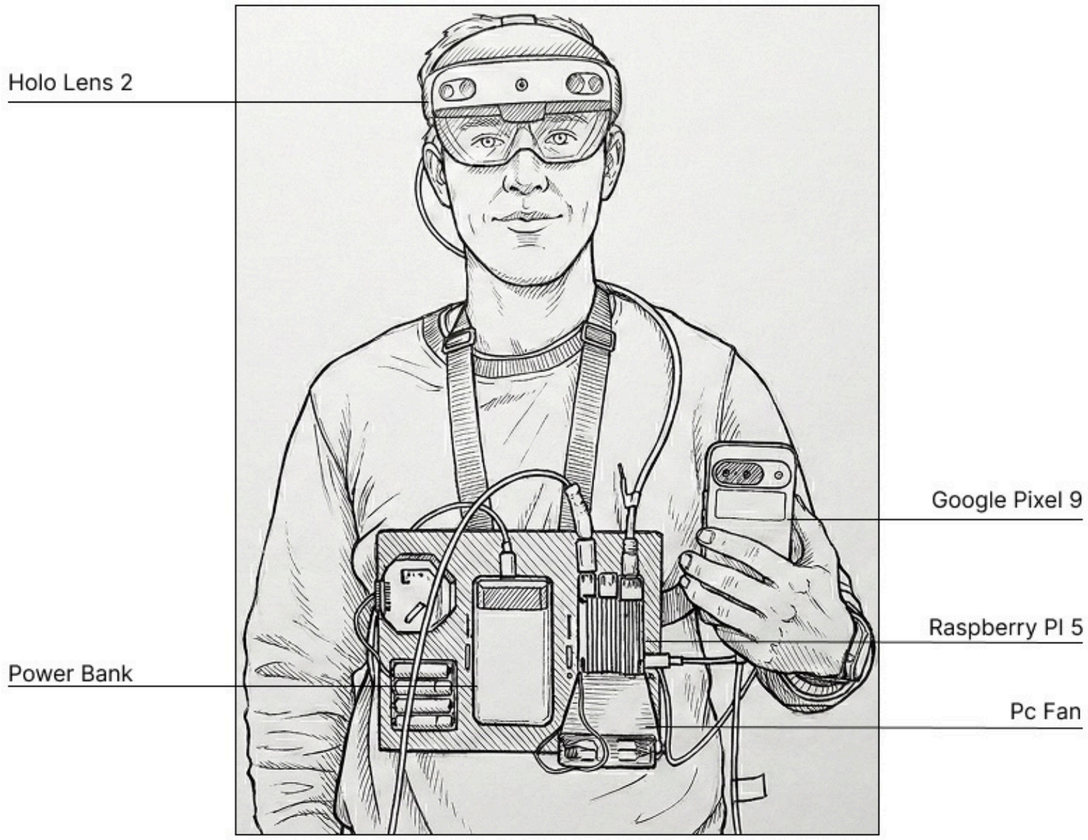
# PDR-Tracking System

01 **Microsoft HoloLens 2** — 6DOF POSE · GT
head-mounted · World-Locking SLAM

02 **Google Pixel 9** — TARGET DEVICE
chest-mounted upright · rear camera forward · 30–40 cm offset

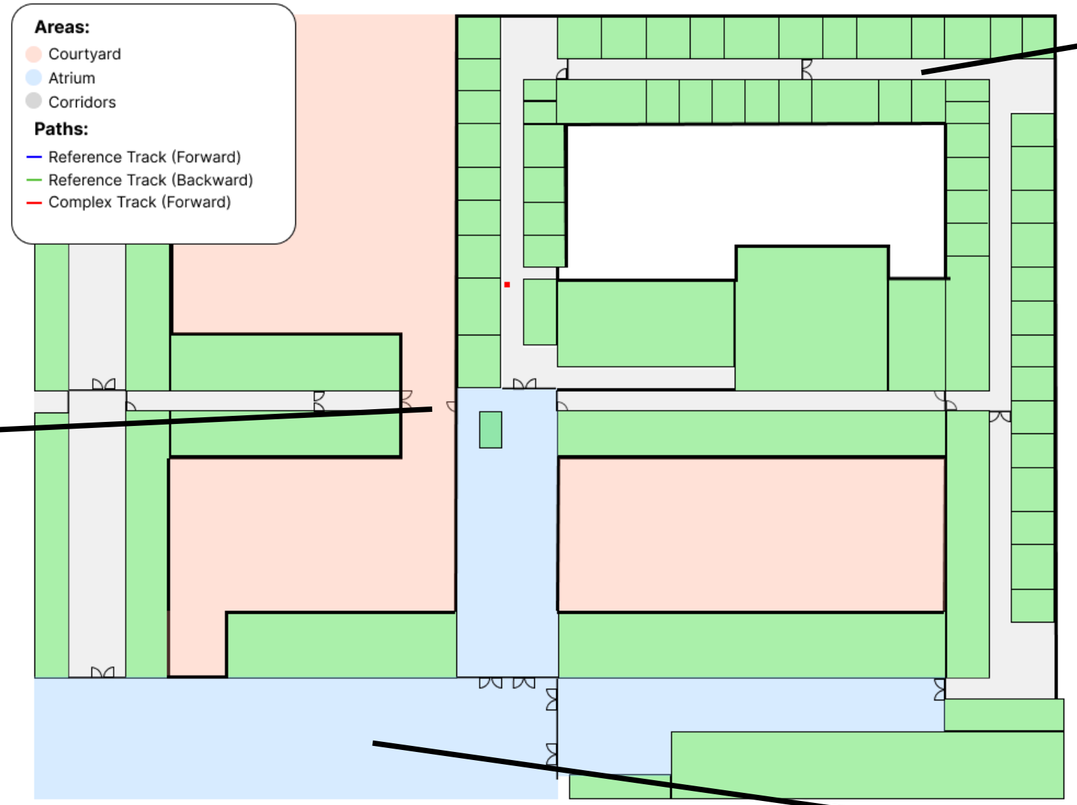
03 **Raspberry Pi 5** — CENTRAL TIME NODE
chest plate · UDP time-stamping at 30 Hz · power bank + PC fan

– **Recorded streams** — 7.95 ± 0.24 MS LATENCY
Video · IMU · 6DoF pose · eye tracking



Does this work?

# Where can we test the System?




# Dataset Creation

01 10 Participants
158cm to 185 cm height

02 27 Trajectories

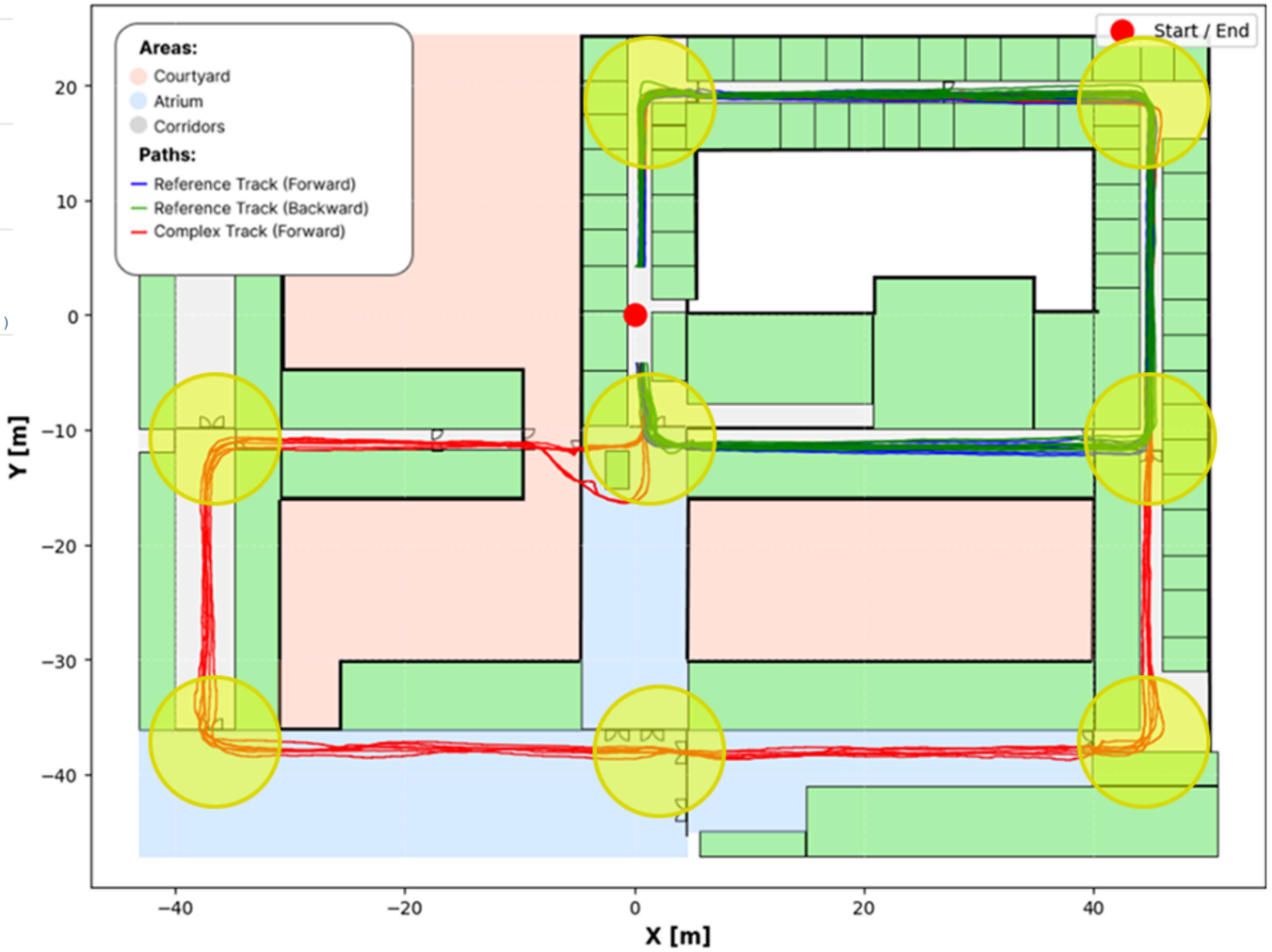
03 Reference track

Jackermeier et. al (2018)

0,13% (52 cm)
Mean Error
399 m

0.05% (8cm)
Mean Error
176 m

Building at University of Regensburg



# Optical Flow Calculation

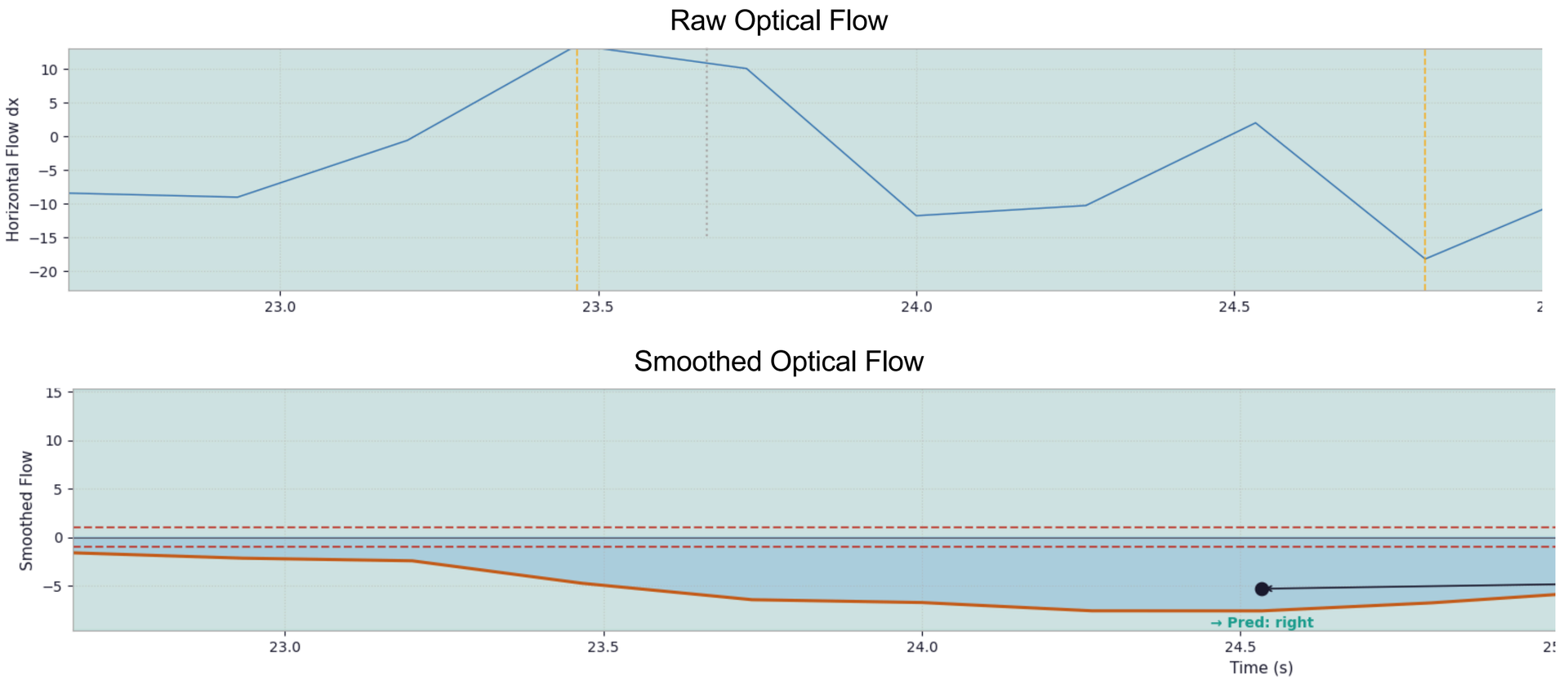
1. **Smooth optical Flow Signal**

$$\text{signal} = \frac{1}{|F_z|} \sum_{f \in F_z} o_f$$

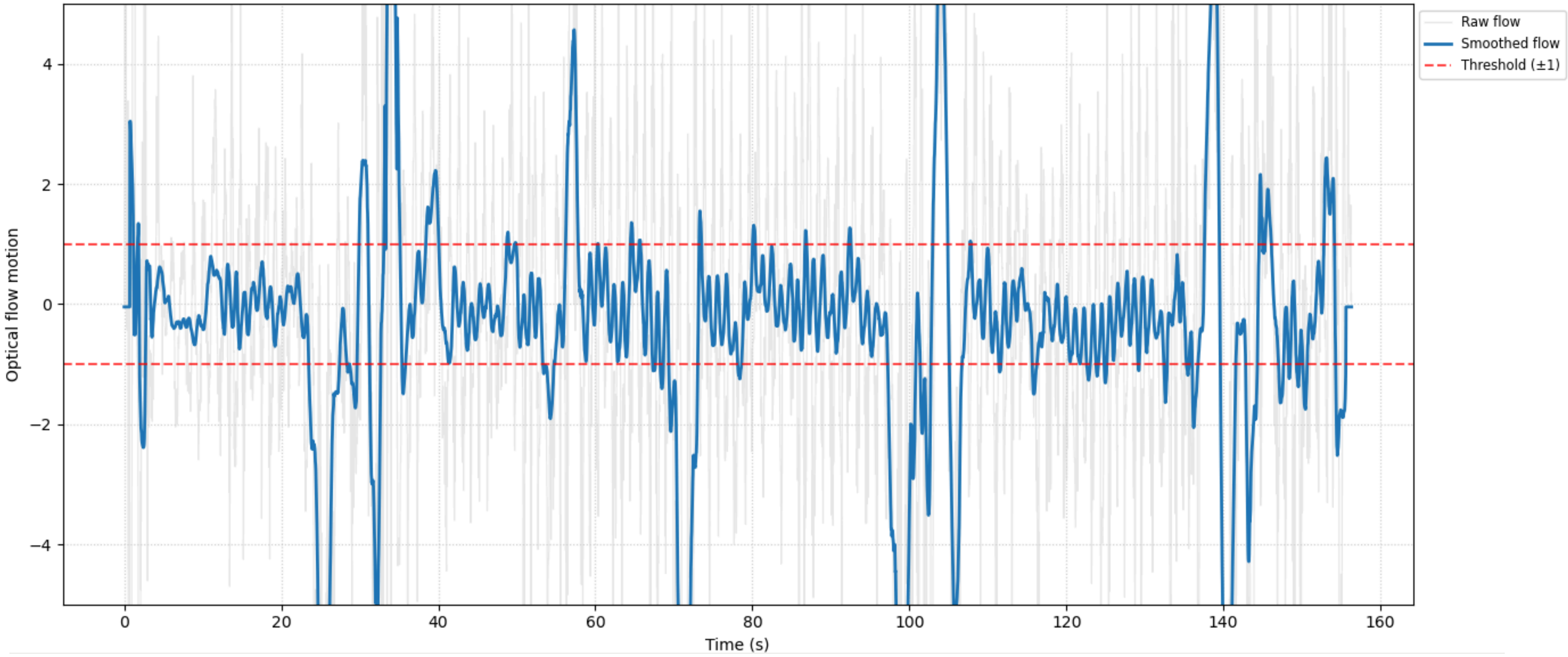
2. **Signal Translation**

signal < 0 Right
signal = 0 Straight
signal > 0 Left

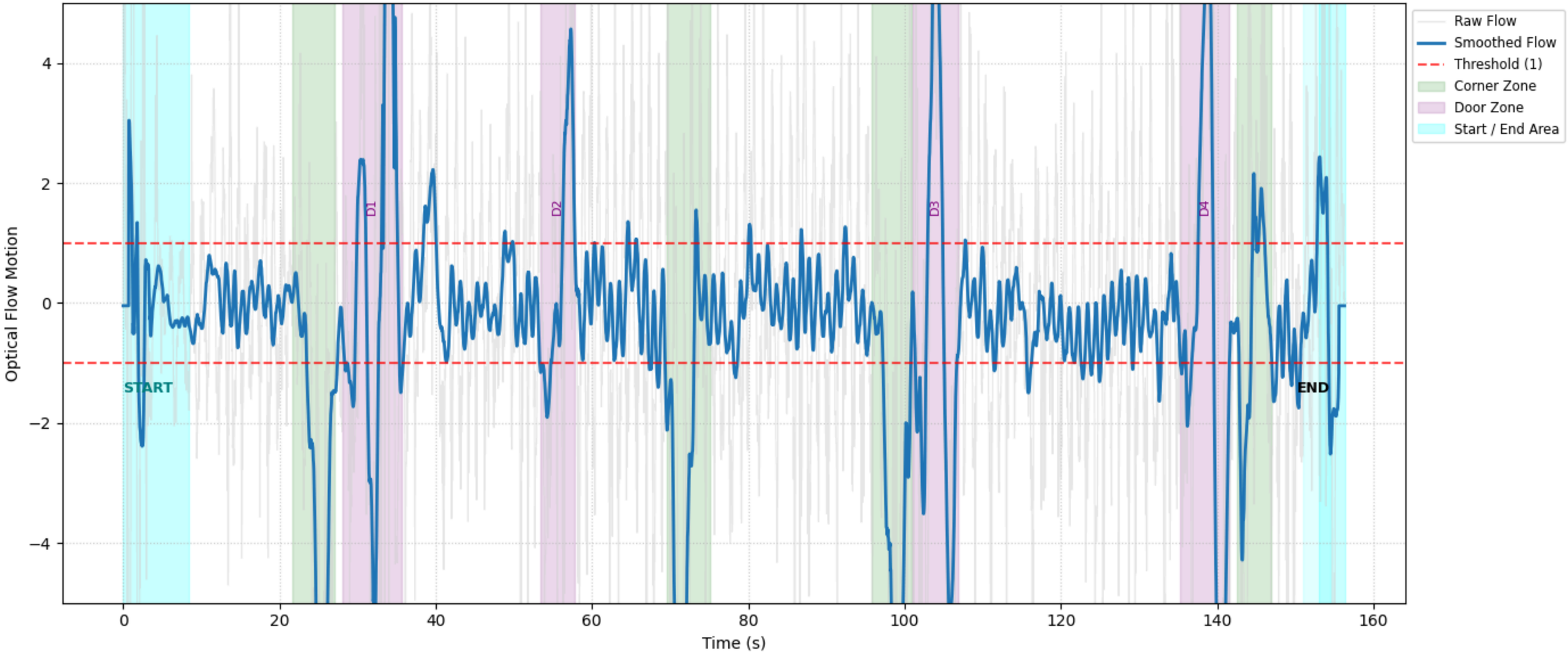
# Optical Flow Calculation



# Optical Flow - raw signal



# Optical Flow - motion



# Dataset Creation

01 10 Participants

158cm to 185 cm height

02 27 Trajectories

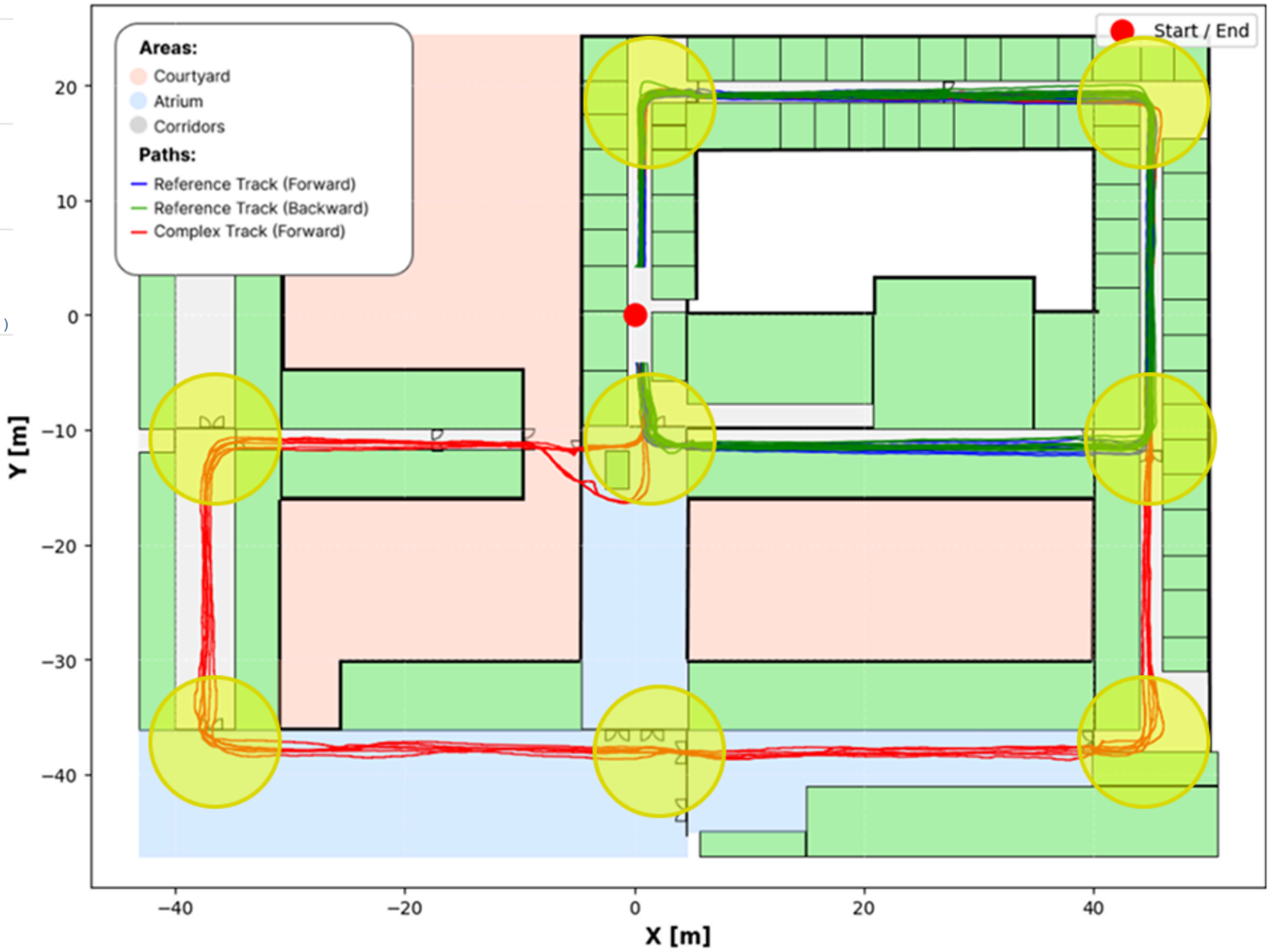
03 Reference track

Jackermeier et. al (2018)

0,13% (52 cm)
Mean Error
399 m

0.05% (8cm)
Mean Error
176 m

Building at University of Regensburg



# Classification Results

| Dataset | N | Accuracy | F1 Left | F1 straight | F1 right |
|---|---|---|---|---|---|
| Long | 56 | 0.929 | 0.77 | 0.87 | 0.99 |
| Forward | 40 | 0.9 | - | - | 0.95 |
| Back | 36 | 1 | 1 | - | - |

Configuration: Radius 450, Threshold 1.0

## Limitations

01 Unbalanced Dataset

Right 57%, Left 33% , Straight Ahead 11%

02 Limited Decision Zones

8 decision zones

03 10 participants

04 Incomplete class coverage per fold

04 Fixed spatial parameters

# Classification Results

| Dataset | N | Accuracy | F1 Left | F1 straight | F1 right |
|---|---|---|---|---|---|
| Long | 56 | 0.929 | 0.77 | 0.87 | 0.99 |
| Forward | 40 | 0.9 | - | - | 0.95 |
| Back | 36 | 1 | 1 | - | - |

Configuration: Radius 450, Threshold 1.0

## Limitations

01 Unbalanced Dataset

Right 57%, Left 33% , Straight Ahead 11%

02 Limited Decision Zones

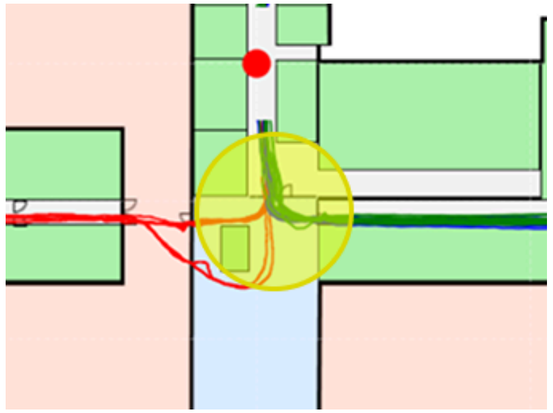
8 decision zones

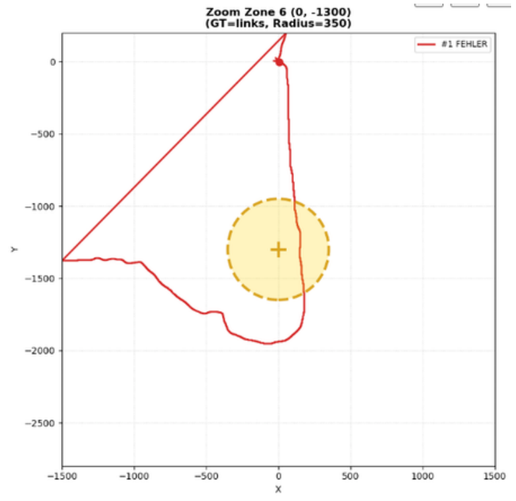
03 10 participants

04 Incomplete class coverage per fold

04 Fixed spatial parameters

# Classification Error Analysis






01 **Decision Zone - no turn**

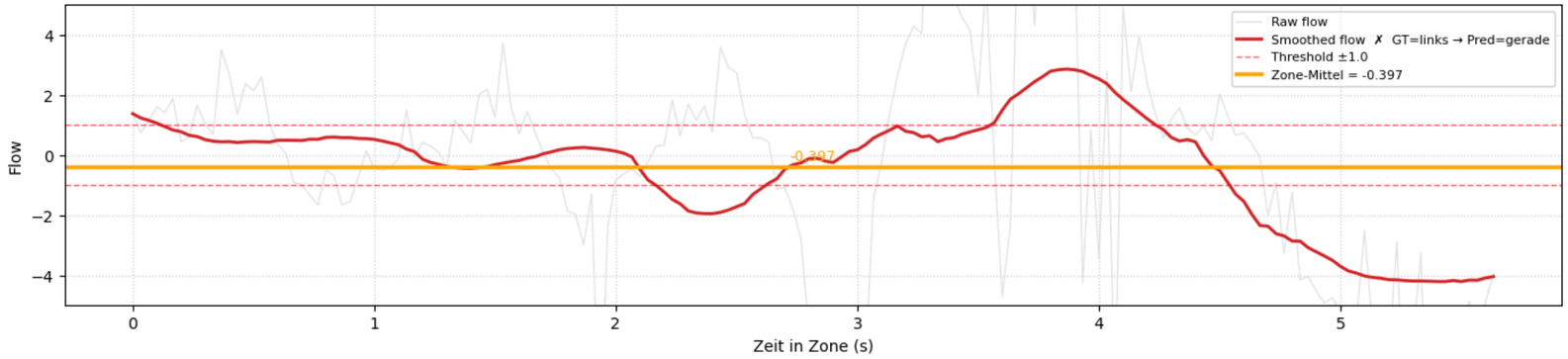
turn happens earlier

02 **Prediction straight**

classifies what happens

03 **Limitation in the method not the classification**

_ **shows need of classification in no decision areas**



# Conclusion & Future Work

- **Plug & Play Ground Truth**: Scalable 6DoF reference without offline scans
- **Spatial Grounding Works**: 90% accuracy with a minimal baseline, overcoming the typical PDR drop.

# Future Work

- **Real-time Integration**: Closing the loop in a live PDR pipeline.
- **Data Scaling & End-to-End Learning**: Multimodal sensor fusion and learned representations based on this setup.

**Linkedin**

**Github**

**Paper**