CLiFF-LHMP: Using Spatial Dynamics Patterns for Long-Term Human Motion Prediction

Authors:

Yufei Zhu[1], Andrey Rudenko[2], Tomasz P. Kucner[3],

Luigi Palmieri[3], Kai O. Arras[3], Achim J. Lilienthal[1,4], Martin Magnusson[1]

1.AASS MRO lab, Örebro University, Sweden 2.Bosch Corporate Research, Robert Bosch GmbH, Germany 3.Finnish Center for Artificial Intelligence, School of Electrical Engineering, Aalto University, Finland. 4.TU Munich, Germany.

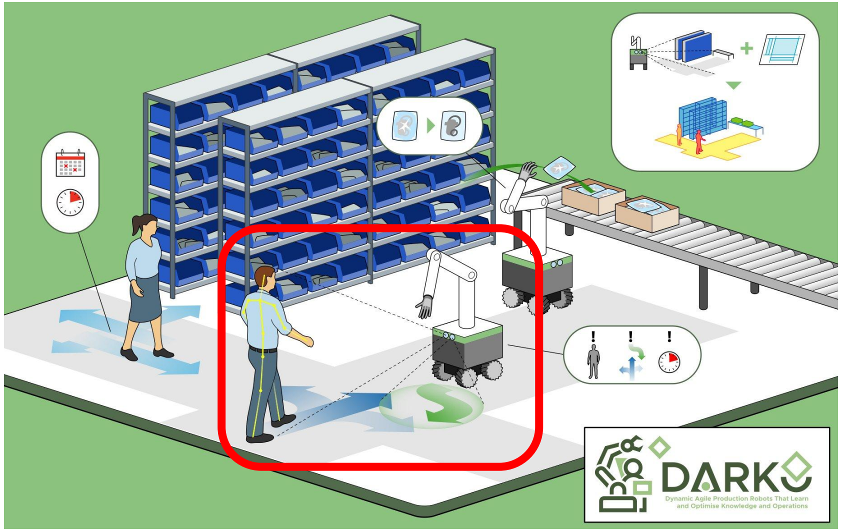
yufei.zhu@oru.se, andrey.rudenko@de.bosch.com, tomasz.kucner@aalto.fi, luigi.palmieri@de.bosch.com, kaioliver.arras@de.bosch.com, achim.lilienthal@oru.se, martin.magnusson@oru.se

Introduction

Long-term Human Motion Prediction

For robots, **accurate prediction** of trajectories of surrounding people over long time is a key skill to improve:

- motion planning,
- tracking,
- automated driving,
- human-robot interaction,
- surveillance



Human-robot co-production in a warehouse environment

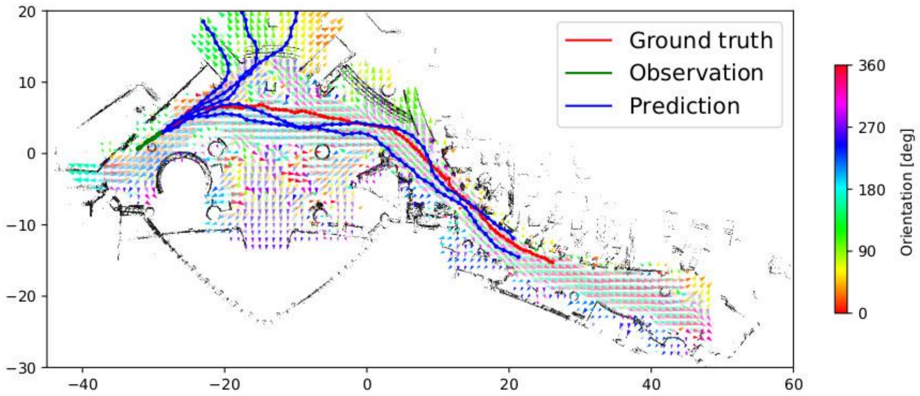
Long-term Human Motion Prediction

Task:

- very long-term human motion prediction (LHMP, up to 50 s)

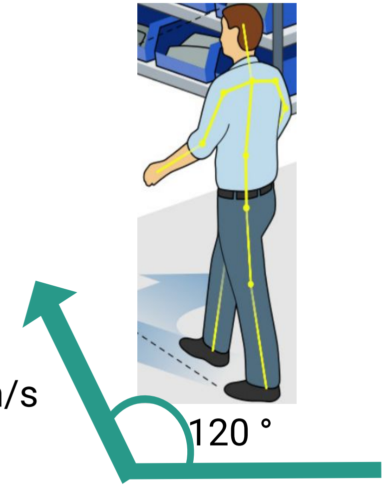
For prediction, we

- exploit *maps of dynamics* (MoDs)
- present a MoD-informed prediction approach, CLiFF-LHMP



A long-term (50s) prediction result in the ATC dataset, with a CLiFF MoD built from observations as colored arrows. 

Maps of Dynamics

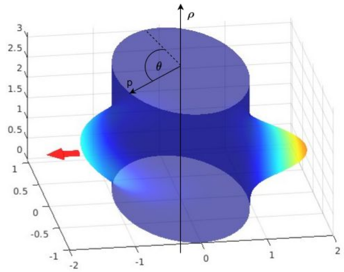


Motion: (speed, heading)

Location: (x, y)

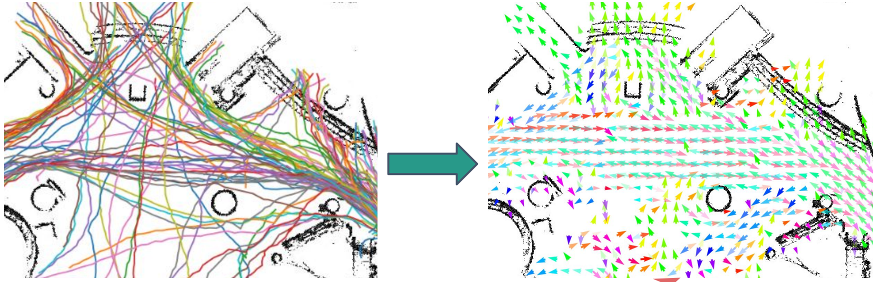
Provide dynamics information as a **map**:

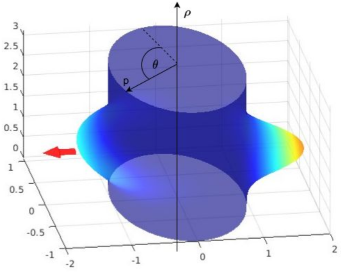
- encode motion as a feature of the environment

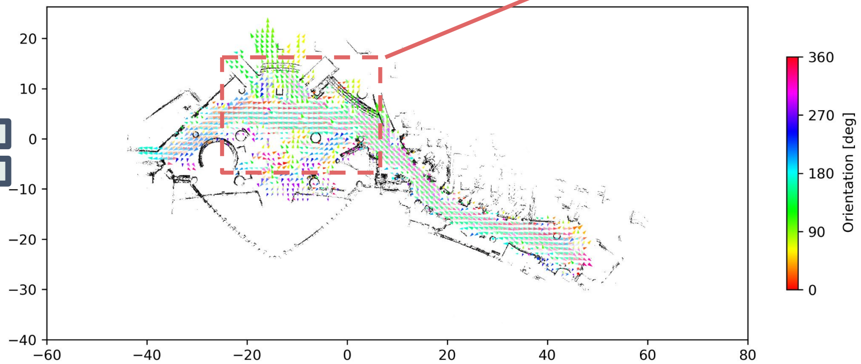


A distribution in CLiFF-map

Maps of Dynamics





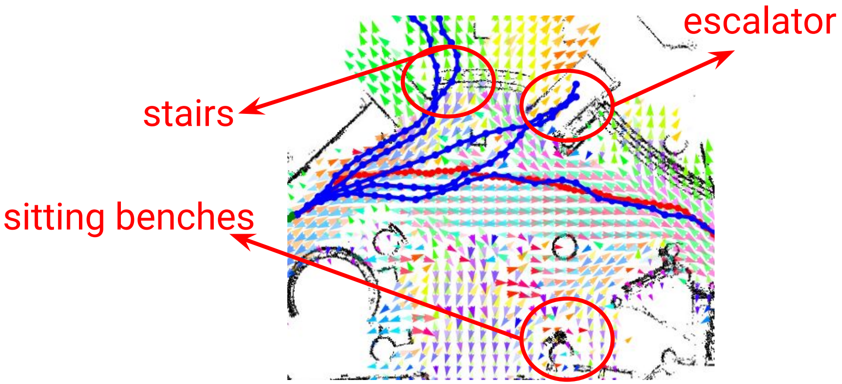


For each location

Maps of Dynamics

Advantage of using MoDs for LHMP

- data-efficiency
- explainability of motion prediction
- implicitly account for obstacle layouts and complex environment topology



[1] T. P. Kucner, M. Magnusson, E. Schaffernicht, V. H. Bennetts and A. J. Lilienthal, "Enabling Flow Awareness for Mobile Robots in Partially Observable Environments," in IEEE Robotics and Automation Letters, vol. 2, no. 2, pp. 1093-1100, April 2017, doi: 10.1109/LRA.2017.2660060.

Method

Method: CLiFF-LHMP

Algorithm 1: CLiFF-LHMP

Input: $\mathcal{H}$, x_{t_0}, y_{t_0}, Ξ
Output: $\mathcal{T}$

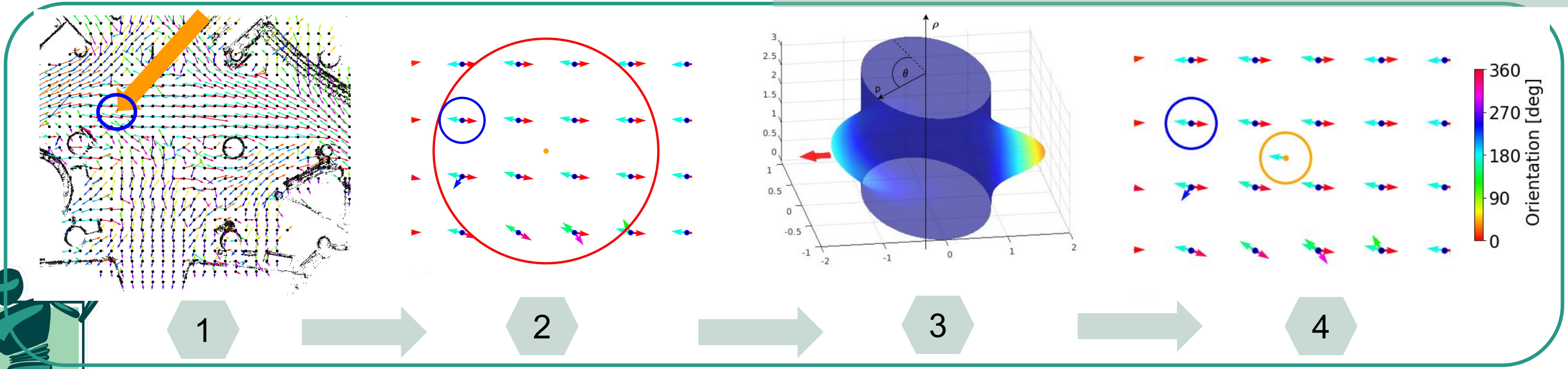
$\mathcal{T} = \{\}$
$\rho_{\text{obs}}, \theta_{\text{obs}} \leftarrow$ getObservedVelocity($\mathcal{H}$)
$s_{t_0} = (x_{t_0}, y_{t_0}, \rho_{\text{obs}}, \theta_{\text{obs}})$
for $t = t_0 + 1, \ldots, t_0 + T_p$ **do**
  $x_t, y_t \leftarrow$ getNewPosition(s_{t-1})
  $\theta_s \leftarrow$ sampleDirectionFromCLiFFmap(x_t, y_t, Ξ)
  $(\rho_t, \theta_t) \leftarrow$ predictVelocity(θ_s, ρ_{t-1}, θ_{t-1})
  $s_t \leftarrow (x_t, y_t, \rho_t, \theta_t)$
  $\mathcal{T} \leftarrow \mathcal{T} \cup s_t$
return $\mathcal{T}$

CLiFF-LHMP predict trajectories in two steps.

Step 1: sample a direction from CLiFF-map

CLiFF-map:

- use GMM to describe **multimodal** flow patterns in each location



Method: CLiFF-LHMP

Step2: use the sampled velocity to bias a CVM predictor

Algorithm 1: CLiFF-LHMP

Input: $\mathcal{H}$, x_{t_0}, y_{t_0}, Ξ

Output: $\mathcal{T}$

$\mathcal{T} = \{\}$

$\rho_{\text{obs}}, \theta_{\text{obs}} \leftarrow$ getObservedVelocity($\mathcal{H}$)

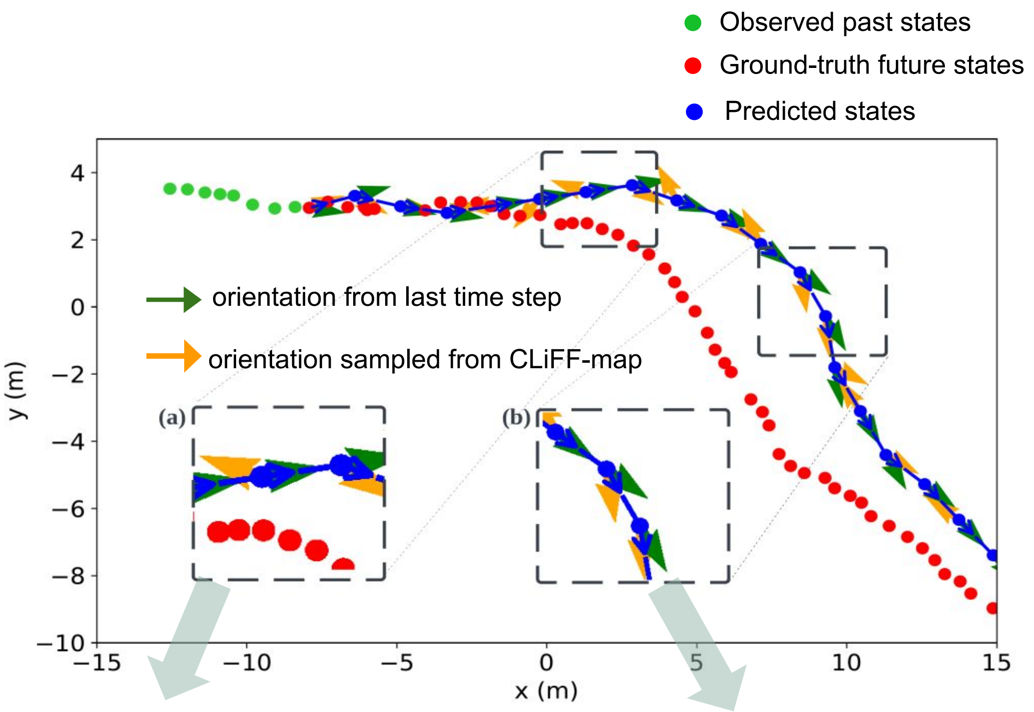
$s_{t_0} = (x_{t_0}, y_{t_0}, \rho_{\text{obs}}, \theta_{\text{obs}})$

for $t = t_0 + 1, \ldots, t_0 + T_p$ **do**

- $x_t, y_t \leftarrow$ getNewPosition(s_{t-1})
- $\theta_s \leftarrow$ sampleDirectionFromCLiFFmap(x_t, y_t, Ξ)
- $(\rho_t, \theta_t) \leftarrow$ predictVelocity(θ_s, ρ_{t-1}, θ_{t-1})
- $s_t \leftarrow (x_t, y_t, \rho_t, \theta_t)$
- $\mathcal{T} \leftarrow \mathcal{T} \cup s_t$

return $\mathcal{T}$

Output: a sequence of states that represent the predicted trajectory.



sampled direction **opposes** the CVM prediction
Trust more CVM

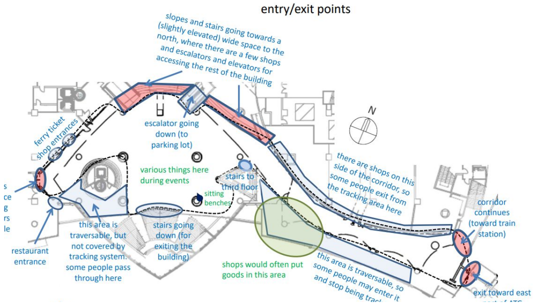
sampled direction **closes** CVM prediction
Trust more CLiFF-map

Evaluation

Experiments: datasets

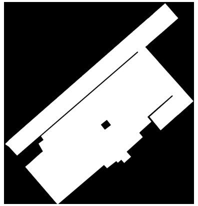
ATC [1]:

92 days of data from Japan shopping mall, 900 m²

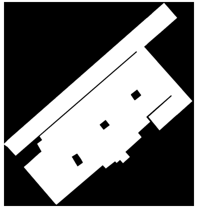


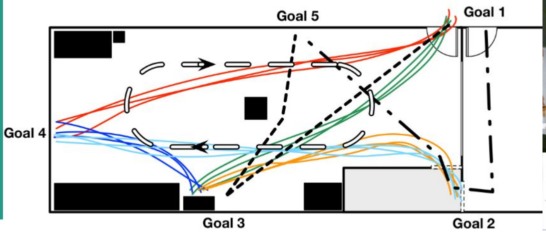
THÖR [2]: each have 4 rounds of data from Örebro robot lab, 150 m²

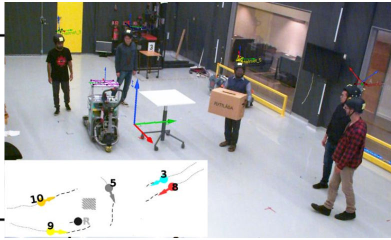
- THÖR1: one obstacle



- THÖR3: three obstacles







Both are publicly available real-world datasets.

[1] D. Brscic, T. Kanda, T. Ikeda, T. Miyashita, "Person position and body direction tracking in large public spaces using 3D range sensors", IEEE Transactions on Human-Machine Systems, Vol. 43, No. 6, pp. 522-534, 2013

[2] A. Rudenko, T. P. Kucner, C. S. Swaminathan, R. T. Chadalavada, K. O. Arras, A. J. Lilienthal, “THÖR: Human-Robot Navigation Data Collection and Accurate Motion Trajectories Dataset”, IEEE Robotics and Automation Letters, 5(2):676-682, 2020

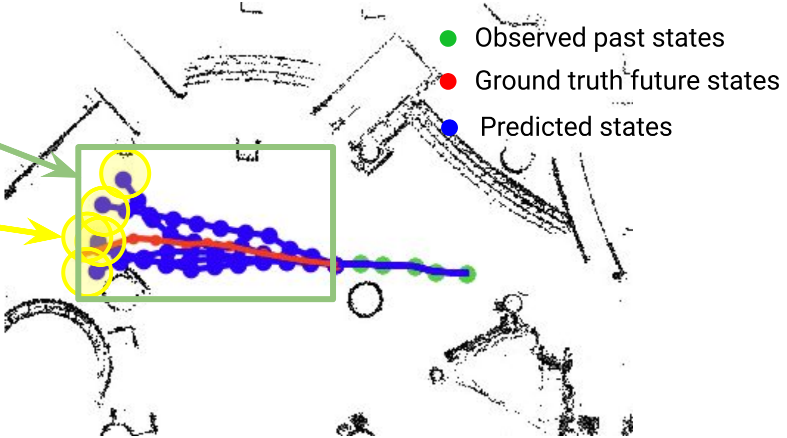
Experiments

Evaluation metrics

ADE: Average displacement error

FDE: Final displacement error

Top-k ADE/FDE: displacement error of the best predicted trajectory

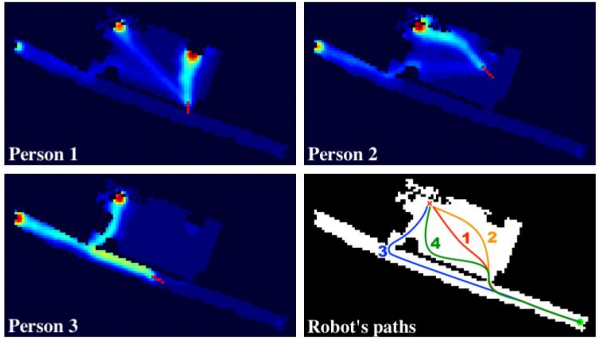


Baselines

IS-MDP: Individual-sampling MDP method, proposed by A. Rudenko [1]:

- environment-aware, planning-based long-term human motion prediction

CVM: constant velocity model



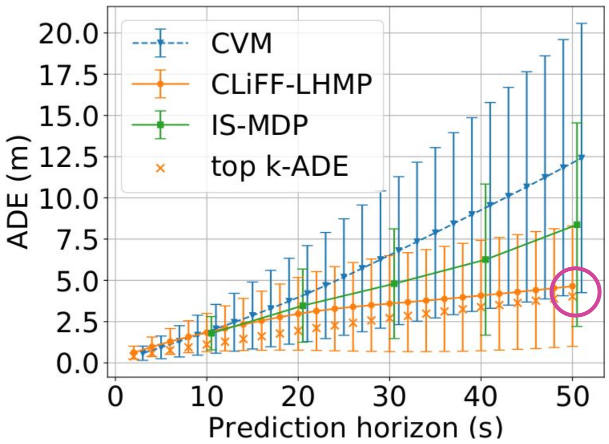
[1] A. Rudenko, L. Palmieri and K. O. Arras, "Joint Long-Term Prediction of Human Motion Using a Planning-Based Social Force Approach," 2018 IEEE International Conference on Robotics and Automation (ICRA), 2018, pp. 4571-4577, doi: 10.1109/ICRA.2018.8460527.

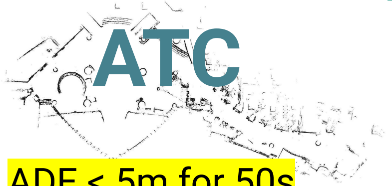
Results

Dataset	Horizon	ADE / FDE (m)		
		CLiFF-LHMP	IS-MDP	CVM
ATC	50 s	**4.6 / 9.6**	8.4 / 21.3	12.4 / 27.1
THÖR1	12 s	**1.5 / 2.6**	1.6 / 3.5	1.8 / 3.8
THÖR3	12 s	**1.3 / 2.6**	1.5 / 3.6	2.8 / 6.1

TABLE II

LONG-TERM PREDICTION HORIZON RESULTS ON DIFFERENT DATASETS. WITH $O_s = 3.2$ s, ERROR REPORTED ARE ADE/FDE IN METERS.

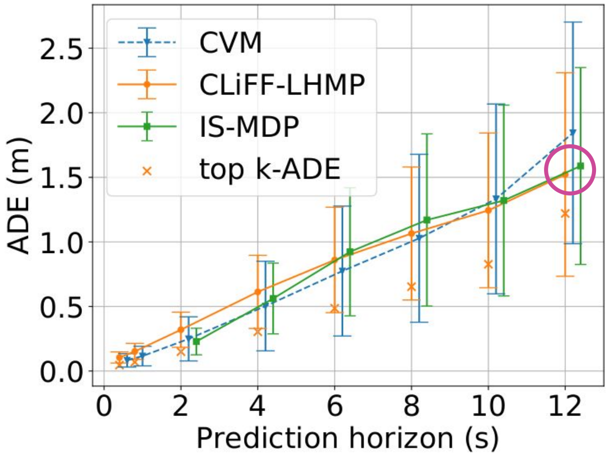


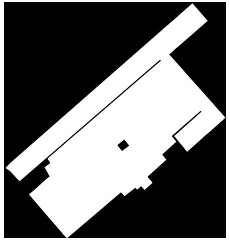


ATC

ADE < 5m for 50s

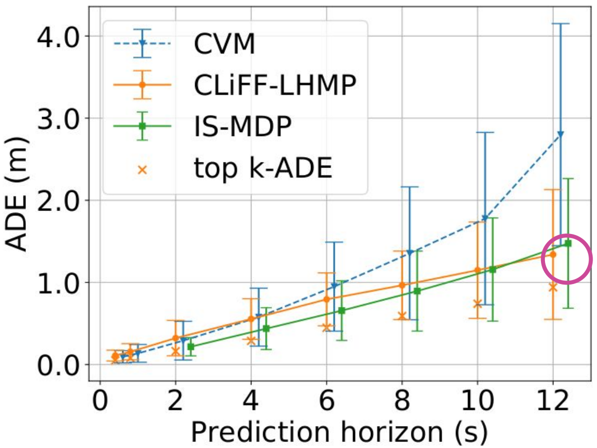
45% ADE and 55% FDE improvement compared to IS-MDP

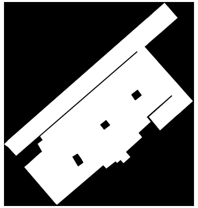




6.3% ADE and 25.7% FDE improvement compared to IS-MDP

THÖR1





13.3% ADE and 27.8% FDE improvement compared to IS-MDP

THÖR3

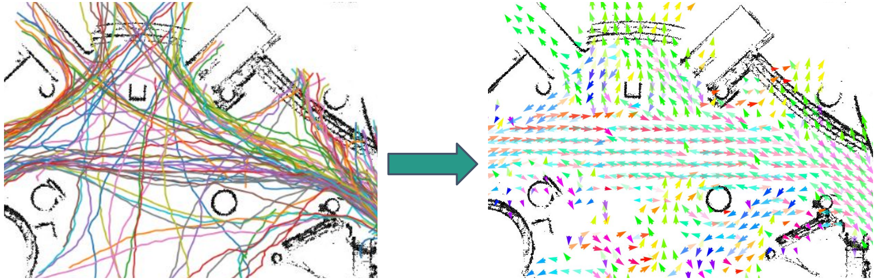
Results: IS-MDP comparison

IS-MDP (baseline):

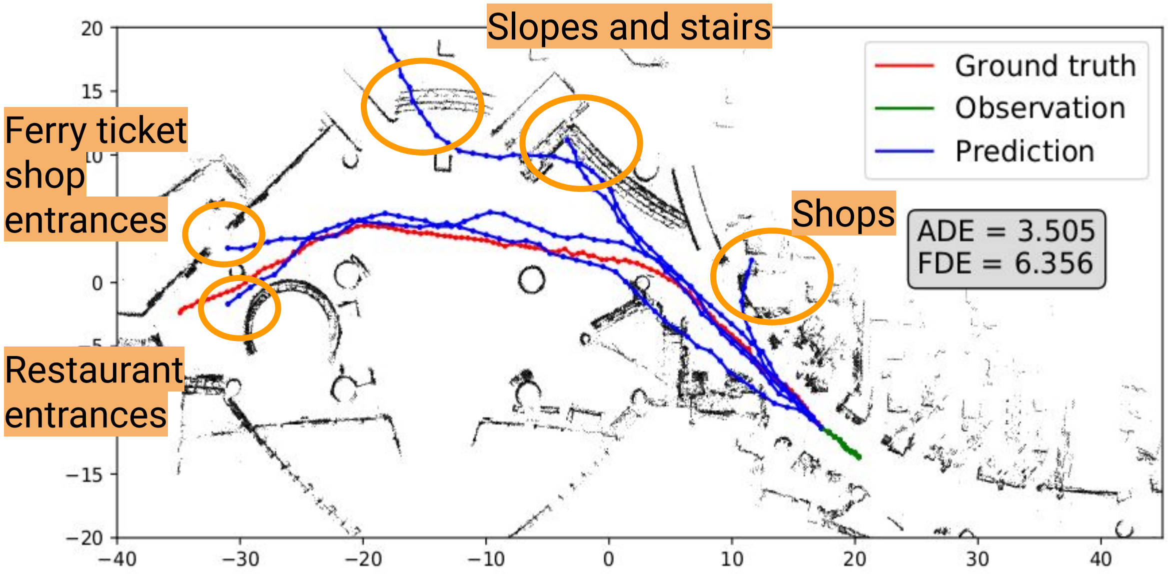
- requires additional input (goal points and the obstacle map)

Our method:

- implicitly exploits location-specific motion patterns
- no explicit knowledge of goals and the obstacle map



Results



Prediction in ATC with prediction horizon 50s

Capture the human motion pattern.

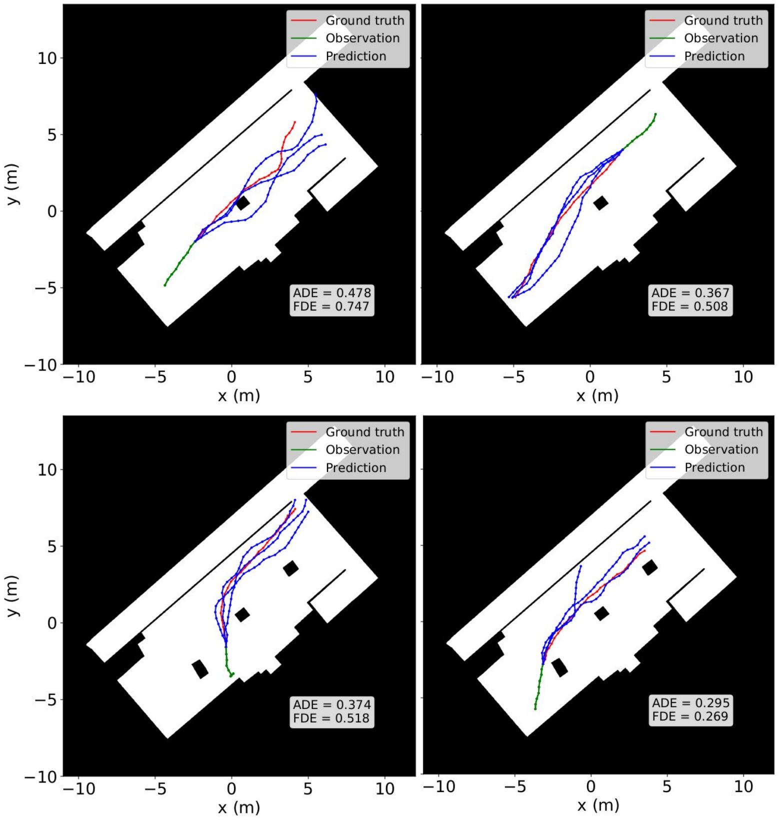
Predict **realistic** trajectories that follow the complex topology of the environment, e.g.

- navigating around corners or obstacles
- passing through narrow passages such as doors, stairs and exits

Results

THÖR1

THÖR3



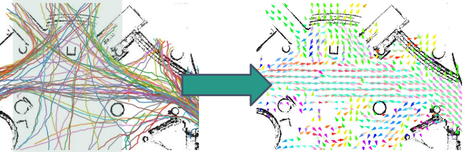
Prediction in THÖR with prediction horizon 12s

Predict trajectories avoid the **obstacles**.

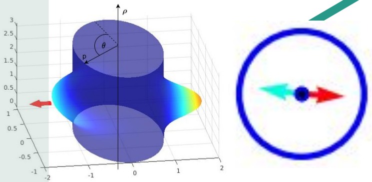
Keep predictions in more intensively used areas, avoiding semantically-insignificant and empty regions, e.g. corners of the room.

CLiFF-LHMP: Using Spatial Dynamics Patterns for Long-Term Human Motion Prediction

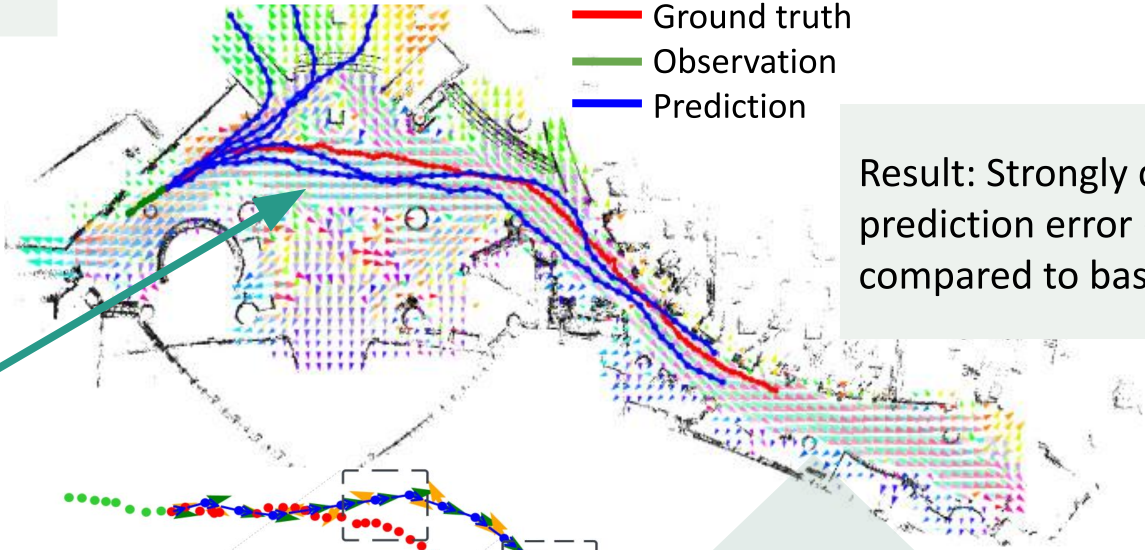
Our method: CLiFF-LHMP

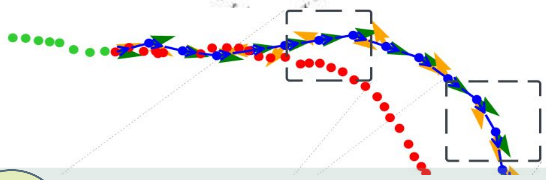


1. build CLiFF-map from observed motion



2. sample a direction from CLiFF-map





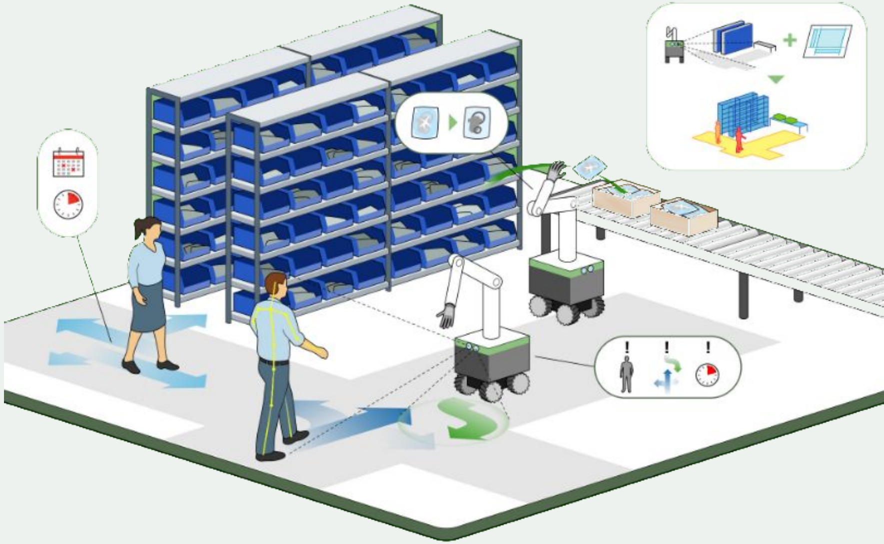
3. bias a constant velocity model predictor with the sampled direction

4. iterate and output a sequence of future states

Result: Strongly decreased prediction error (-45% at 50s) compared to baseline

DARKO

Dynamic Agile Production Robots That Learn and Optimise Knowledge and Operations



This work has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No.101017274 (DARKO).