LONG-TERM HUMAN MOTION PREDICTION WORKSHOP
ICRA 2019

HUMAN MOTION PREDICTION: METHODS AND CHALLENGES

Andrey Rudenko, Luigi Palmieri, Kai O. Arras, Achim J. Lilienthal et al.

AASS www.mrolab.eu MR&O MOBILE ROBOTICS & OLFACTION
ÖREBRO UNIVERSITY
BOSCH

Human Motion Prediction: Methods and Challenges

Agenda

1. Motion Trajectory Prediction: A Survey
2. Highlights from our research

PART 1:
A SURVEY ON HUMAN MOTION TRAJECTORY PREDICTION

Motion Trajectory Prediction: A Survey
Introduction

- **Motion Trajectory Prediction: A Survey**
- Authors:
  - Andrey Rudenko, Bosch Corporate Research / Örebro University
  - Luigi Palmieri, Bosch Corporate Research
  - Michael Herman, Bosch Center for AI
  - Kris M. Kitani, CMU
  - Dariu M. Gavrila, TU Delft
  - Kai O. Arras, Bosch Corporate Research

Human Motion Trajectory Prediction: A Survey

Journal Title
XX(X):1–33
©The Author(s) 2019
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

Andrey Rudenko[1,2], Luigi Palmieri[1], Michael Herman[3], Kris M. Kitani[4], Dariu M. Gavrila[5] and Kai O. Arras[1]

Abstract

With growing numbers of intelligent systems in human environments, the ability of such systems to perceive, understand and anticipate human behavior becomes increasingly important. Specifically, predicting future positions of dynamic agents and planning considering such predictions are key tasks for self-driving vehicles, service robots and advanced surveillance systems.

This paper provides a survey of human motion trajectory prediction. We review, analyze and structure a large selection of work from different communities and propose a taxonomy that categorizes existing approaches based on the motion modeling approach and level of contextual information used. We provide an overview of the existing datasets and performance metrics. We discuss limitations of the state of the art and outline directions for further research.

Keywords

Survey, motion prediction, robotics, video surveillance, autonomous driving

Motion Trajectory Prediction: A Survey
Introduction

- **Motivation**:
  - Bring structure to the growing body of work on predicting trajectories and better understand the key aspects of the problem and outline the current state of the area
- **Scope**:
  - Over 170 methods for predicting 2D motion trajectories of humans and vehicles
  - Cross-disciplinary review of literature, datasets and benchmarking
    - Mobile robotics
    - Video surveillance
    - Autonomous driving
  - A new taxonomy of prediction methods
  - Discussion of the state of the art along several research hypotheses
- **Available on arXiv**, soon to be submitted to a journal

Motion Trajectory Prediction: A Survey
General formulation of the prediction problem

- Future trajectory (or distribution over states) is a **function** of the **observed environment**

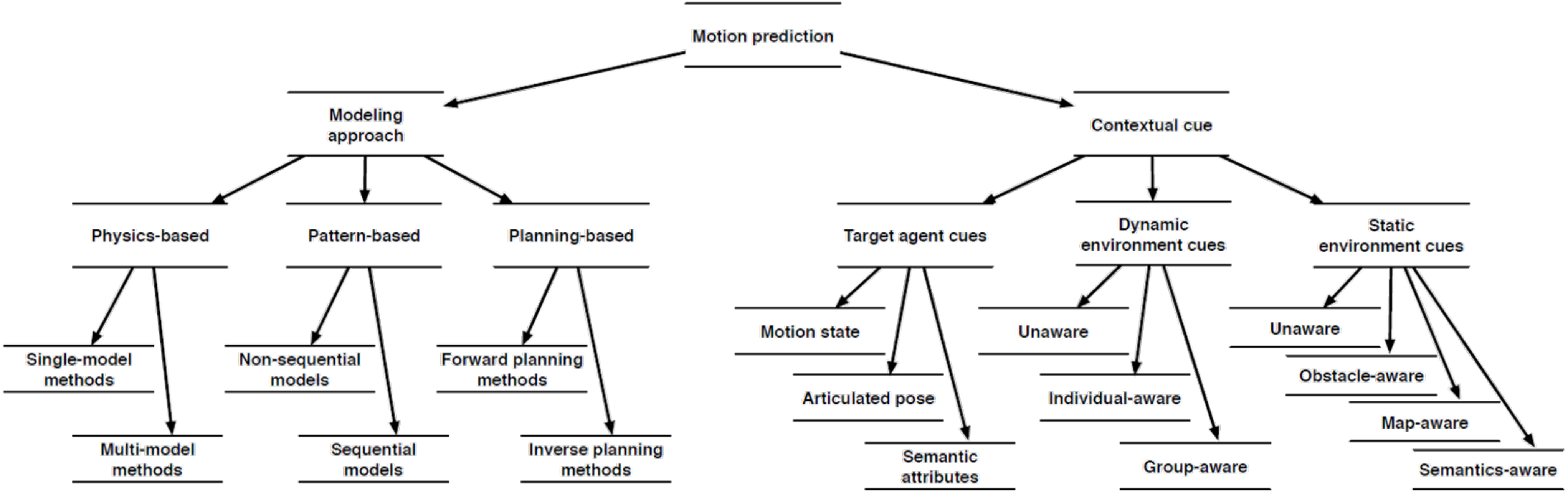
$$\mathcal{T}_{\text{future}} = f(\text{environment})$$

- Dynamics-based transition function
- Learning-based motion patterns
- Planning-based optimal motion

- Current position of the agent
- History of observations
- Map of static obstacles
- Positions of other dynamic agents
- Various semantic cues of the environment

Motion Trajectory Prediction: A Survey
Our taxonomy

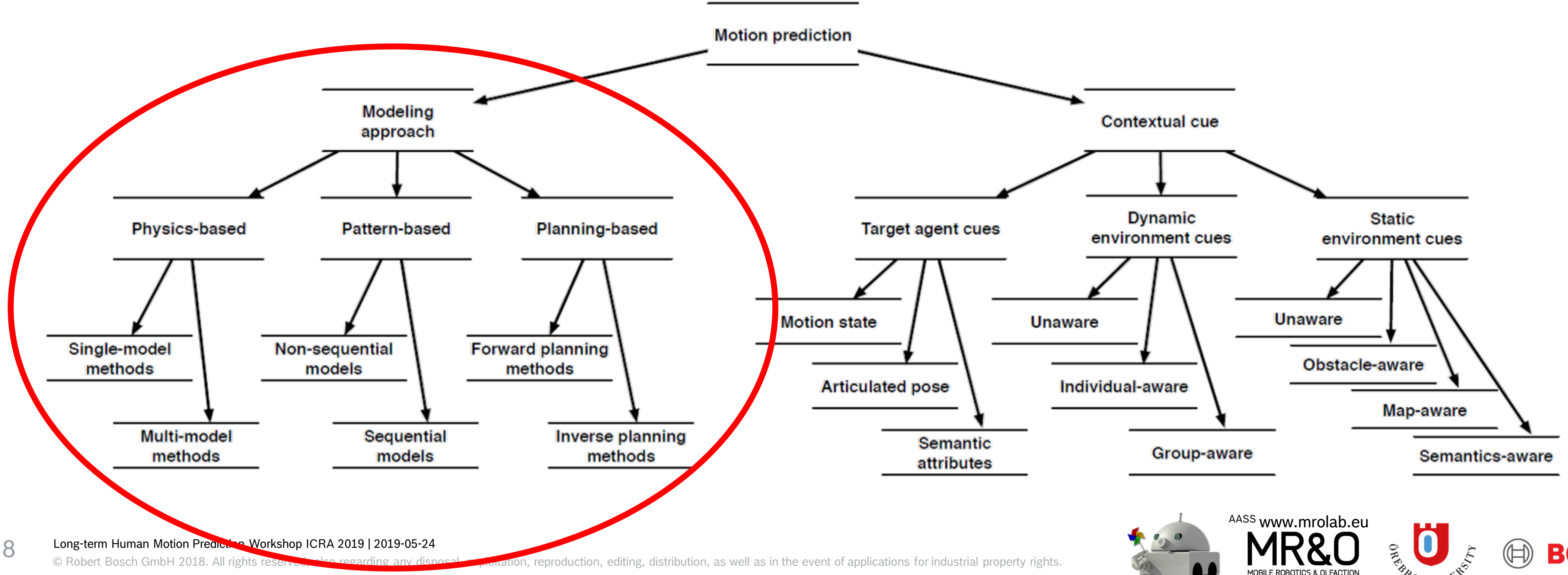
- Every prediction method is discussed in two dimensions:
  - What is the nature of the prediction function?
  - Which contextual cues does it exploit?



AASS www.mrolab.eu

Motion Trajectory Prediction: A Survey
Our taxonomy

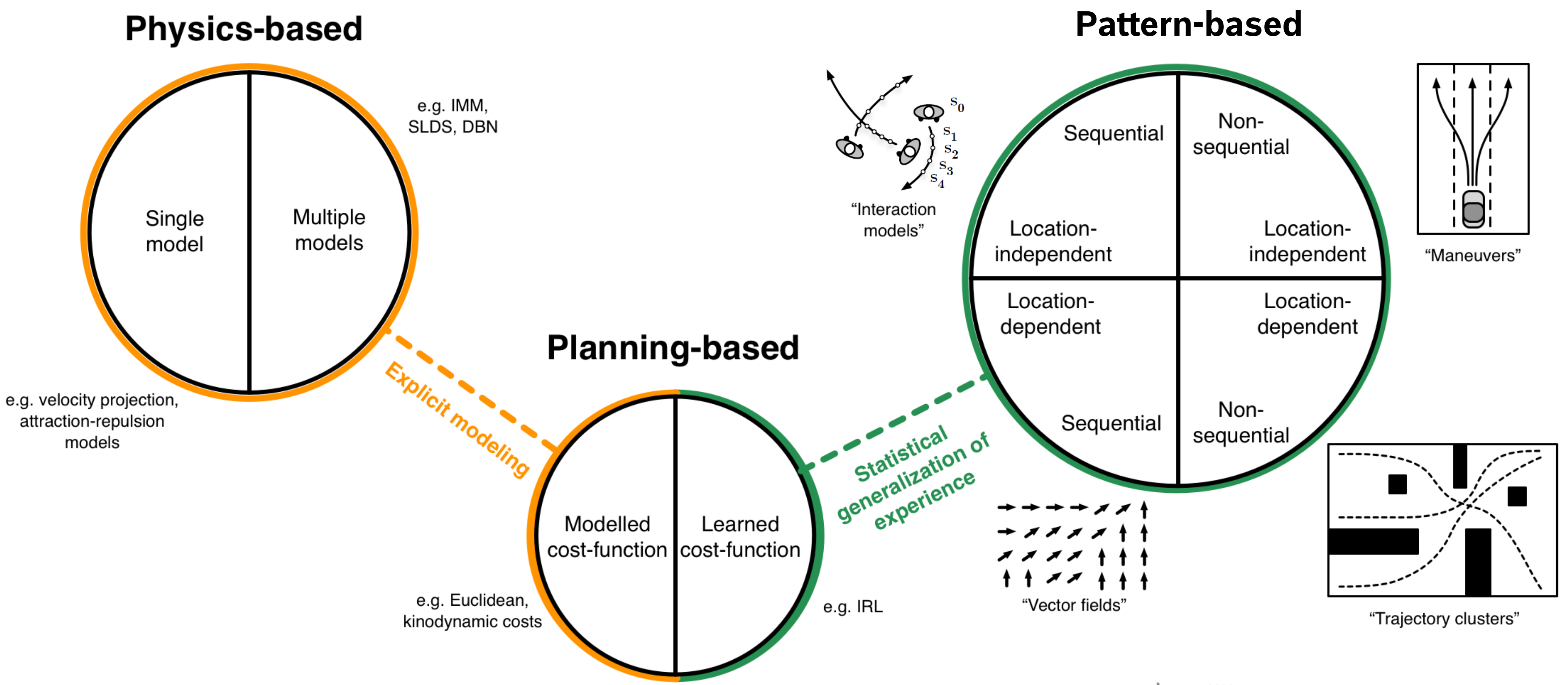
- Every prediction method is discussed in two dimensions:
  - What is the nature of the prediction function?
  - Which contextual cues does it exploit?



Motion Trajectory Prediction: A Survey
1st classification criteria (motion modelling approach)

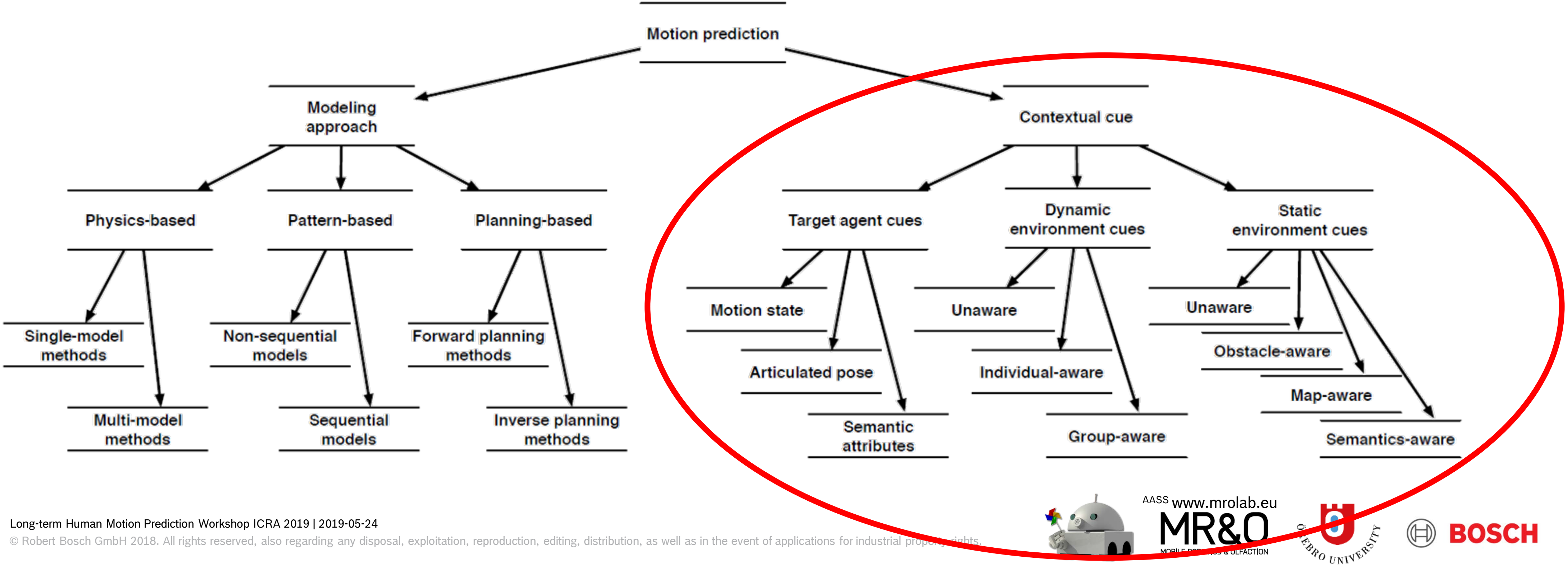
- **Physics-based** methods use a **dynamics transition function** to project the current state of the agent
  - Constant velocity, constant acceleration, constant curvature turn
  - Attraction-repulsion approaches
  - Reachability-based approaches
- **Pattern-based** approaches learn **generalized transitions and trajectories** from data
  - Vector fields
  - Clustering
  - Distribution over full trajectories
  - CNNs, RNNs, LSTMs
- **Planning-based** approaches assume a motion optimality criteria and predict **optimal paths** towards estimated goal states
  - MDPs, RRTs, A*, PRMs
  - Inverse Reinforcement Learning

Motion Trajectory Prediction: A Survey
1st classification criteria (motion modelling approach)



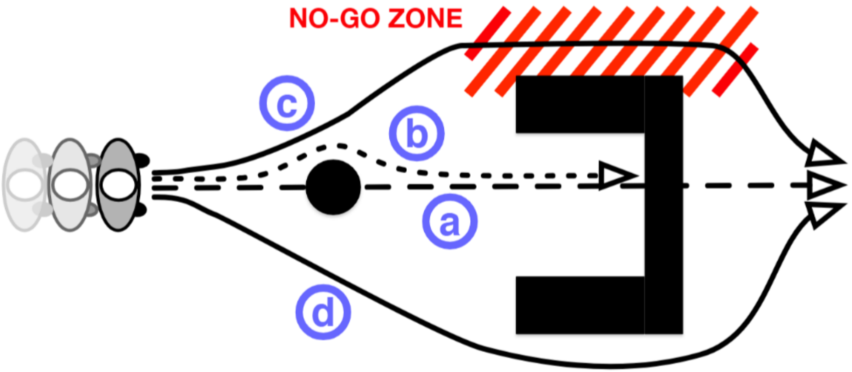
Motion Trajectory Prediction: A Survey
Our taxonomy

- Every prediction method is discussed in two dimensions:
  - What is the nature of the prediction function?
  - Which contextual cues does it exploit?

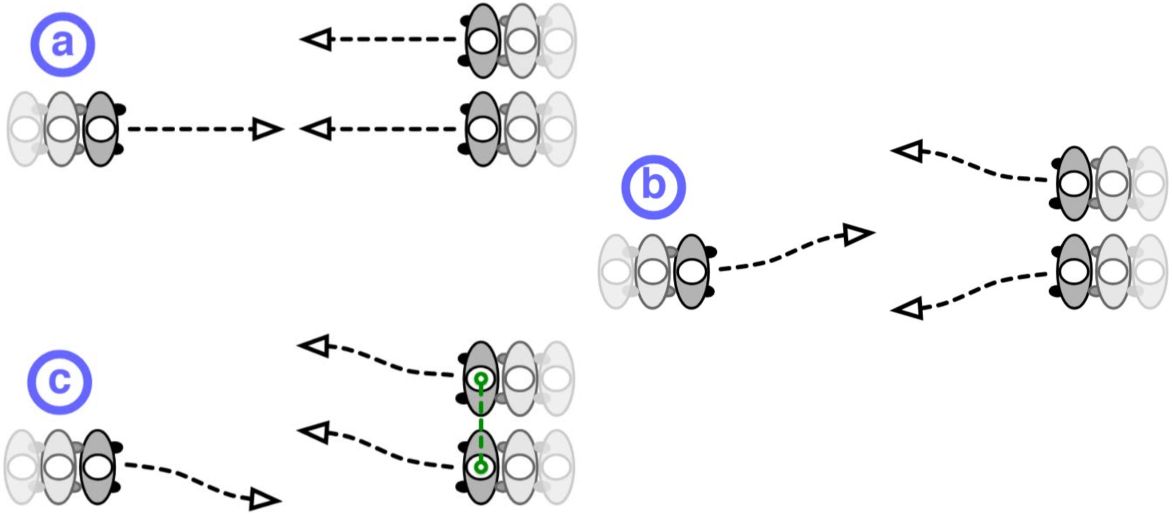


Motion Trajectory Prediction: A Survey
2nd classification criteria (contextual cues of the environment)

- Static environment
  - Unaware methods (a)
  - Obstacle-aware methods (b)
  - Map-aware methods (c)
  - Semantics-aware methods (d)



- Dynamic environment
  - Unaware methods (a)
  - Individual-aware methods (b)
  - Group-aware methods (c)



Motion Trajectory Prediction: A Survey
Discussion

- Discussion is centered around 3 questions:
  - **Question 1:** are all motion models equally good, in particular regarding contextual cues?
  - **Question 2:** is motion prediction solved?
  - **Question 3:** is benchmarking in good state?

Motion Trajectory Prediction: A Survey
Benefits and drawbacks

- **Physics-based methods:**
  - Good for short-term predictions
  - Simple to deploy
  - Fast and efficient inference
  - No or small amount of training data needed
  - Natural handling of prediction uncertainty
  - Generalization to new environments
  - Hard to non-homogeneous, complex motion dynamics

- **Pattern-based methods**
  - No or little modeling required
  - Can capture complex, non-homogeneous dynamics
  - Capture all contextual cues that are present in data
  - Fast inference
  - Require large amount of training data
  - Limited generalization to new environments
  - Limited social-awareness *

- **Planning-based methods**
  - Naturally handle maps and avoid local minima
  - No or small amount of training data needed
  - Reasoning over actions and goals
  - Generalization to new environments
  - Planning could be time-consuming

* Historically until LSTM-based methods

Motion Trajectory Prediction: A Survey
Discussion

- Discussion is centered around 3 questions:
  - **Question 1:** are all motion models equally good, in particular regarding contextual cues?
    - There are powerful, context-aware predictors from each modeling approach
    - It is possible to extend all method classes with all both static and dynamic contextual cues
    - Method classes have their inherent strengths and drawbacks
  - **Question 2:** is motion prediction solved?
  - **Question 3:** is benchmarking in good state?

Motion Trajectory Prediction: A Survey
Discussion

- Discussion is centered around 3 questions:
  - **Question 1:** are all motion models equally good, in particular regarding contextual cues?
    - There are powerful, context-aware predictors from each modeling approach
    - It is possible to extend all method classes with all both static and dynamic contextual cues
    - Method classes have their inherent strengths and drawbacks
  - **Question 2:** is motion prediction solved?
    - What are the requirements from the domains? (emergency breaking requirements: ISO 15622:2018 standard)
    - Are considered semantic cues enough? (see Rasouli and Tsotsos, *IEEE Transactions on Intelligent Transportation Systems* 2019)
    - How good is the transfer to new environments? (see Ballan et al., *ECCV* 2016, Srikanth et al., *arXiv* 2019)
    - What the methods are really capable of? (see Schöller et al., *arXiv* 2019)
  - **Question 3:** is benchmarking in good state?

Motion Trajectory Prediction: A Survey
Discussion

- Discussion is centered around 3 questions:
  - **Question 1:** are all motion models equally good, in particular regarding contextual cues?
    - There are powerful, context-aware predictors from each modeling approach
    - It is possible to extend all method classes with all both static and dynamic contextual cues
    - Method classes have their inherent strengths and drawbacks
  - **Question 2:** is motion prediction solved?
    - What are the requirements from the domains? (emergency breaking requirements: ISO 15622:2018 standard)
    - Are considered semantic cues enough? (see Rasouli and Tsotsos, *IEEE Transactions on Intelligent Transportation Systems* 2019)
    - How good is the transfer to new environments? (see Ballan et al., *ECCV* 2016, Srikanth et al., *arXiv* 2019)
    - What the methods are really capable of? (see Schöller et al., *arXiv* 2019)
  - **Question 3:** is benchmarking in good state?
    - Condition precision on the observation period, prediction horizon and the complexity of the situation
    - Robustness experiments
    - Need for a prediction benchmark (see TrajNet: http://trajnet.stanford.edu/)

PART 2:
HIGHLIGHTS FROM OUR RESEARCH

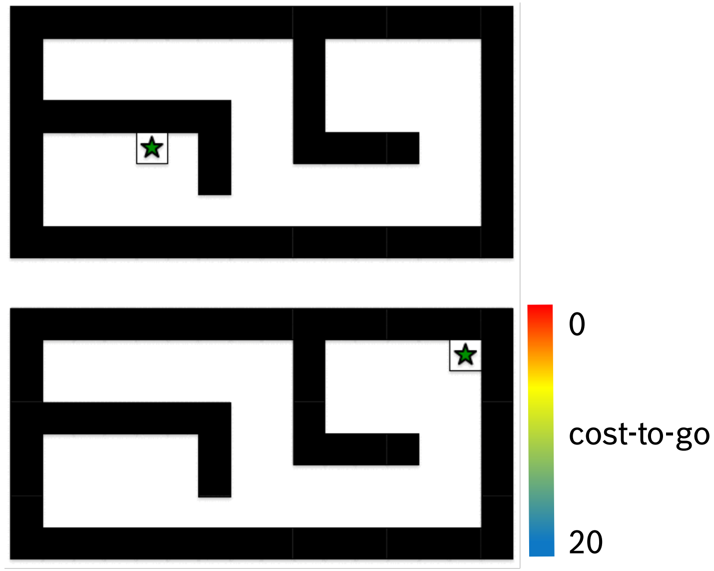
Highlights from our research
Overview

- Group-aware MDP-based predictor
  - A. Rudenko, L. Palmieri, A. Lilienthal, K.O. Arras, *ICRA Workshop* 2017, *ICRA* 2018, *IROS* 2019
  - Combination of
    - Global map-aware MDP-based predictor
    - Group-aware local interaction model by Moussaïd et al. 2010
- Semantic occupancy priors in urban environments
- A new dataset of indoor human trajectories

Human Motion Prediction Under Social Grouping Constraints
Method

- **Global optimal policy: MDPs**
- Stochastic policy sampling
- Human interactions: group social force

Value function $V_g^*(s)$



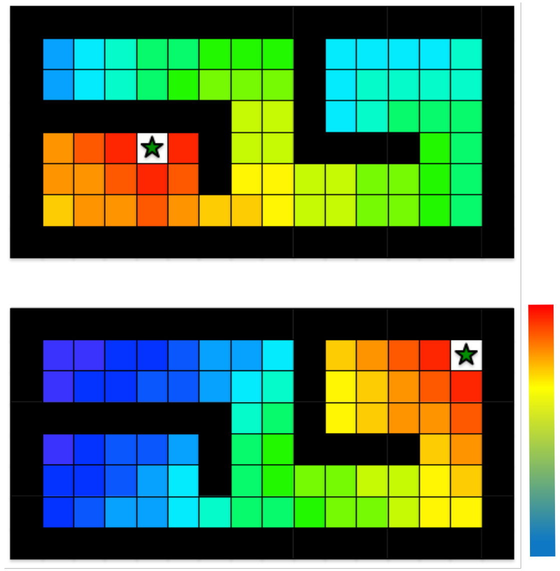
$$\begin{cases} Q^*(s,a) = \mathcal{R}(s,a) + \gamma \sum_{s'} \mathcal{P}(s,s',a) V^*(s') \\ V^*(s) = \max_a Q^*(s,a) \end{cases}$$

AASS www.mrolab.eu

Human Motion Prediction Under Social Grouping Constraints
Method

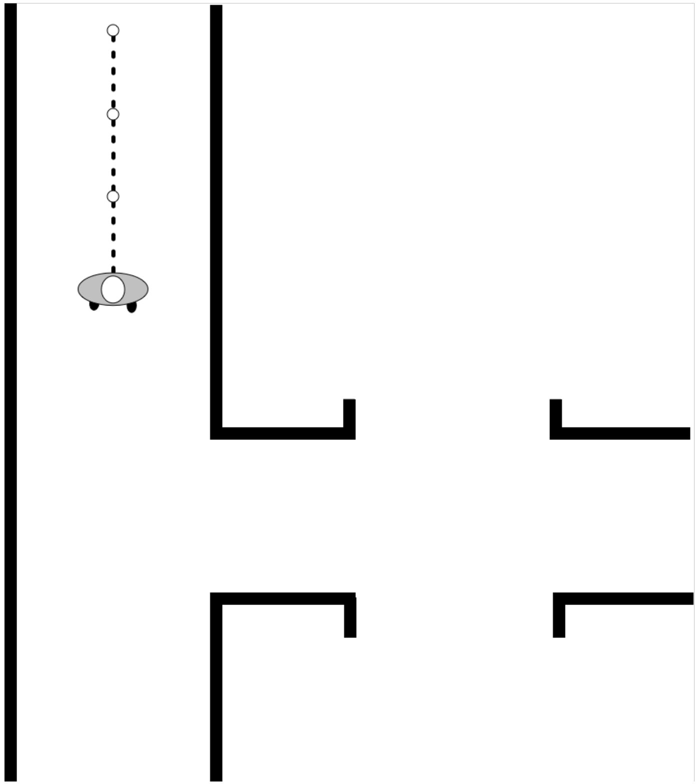
- **Global optimal policy: MDPs**

Value function $V_g^*(s)$

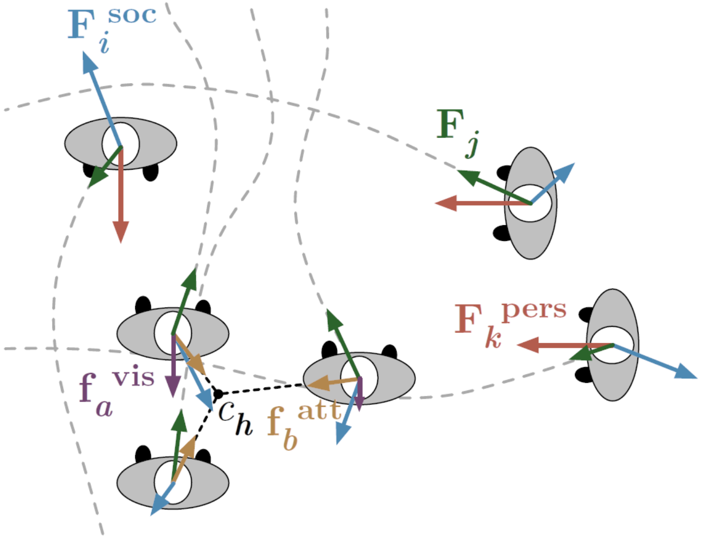


$$\begin{cases} Q^*(s,a) = \mathcal{R}(s,a) + \gamma \sum_{s'} \mathcal{P}(s,s',a) V^*(s') \\ V^*(s) = \max_a Q^*(s,a) \end{cases}$$

- **Stochastic policy sampling**



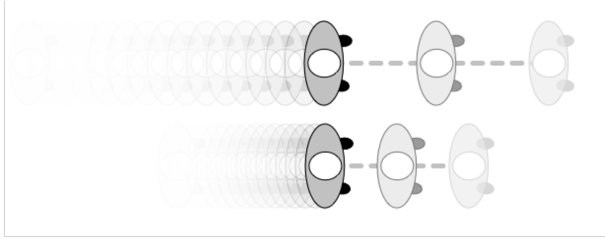
- **Interactions: group social force** (Moussaïd et al. 2010)



$$\mathbf{F}_i = \mathbf{F}_i^{pers} + \sum_{k \neq i}^{N} \mathbf{f}_{i,k}^{soc} + \mathbf{f}_i^{vis} + \mathbf{f}_i^{att}$$

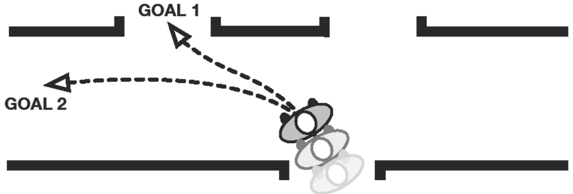
Human Motion Prediction Under Social Grouping Constraints
Method properties

- **Correct speed handling**



$$p(a) \text{ in } \hat{\pi}^i_g \;\propto\; \begin{cases} p(\langle \theta, v \rangle) \text{ in } \pi_g, & \text{if } v \leq v^i_{\text{obs}}, \\ p(\langle \theta, 2v^i_{\text{obs}} - v \rangle) \text{ in } \pi_g, & \text{if } v > v^i_{\text{obs}} \end{cases}$$

- **Reasoning about goals**

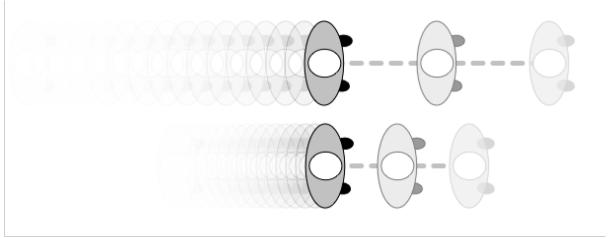


$$p(g) \;\propto\; \exp\left(\beta\left(V^*_g(s^i_{l(i)}) - V^*_g(s^i_1)\right)\right)$$

- **Multimodal predictions**
- **Semantic map input**

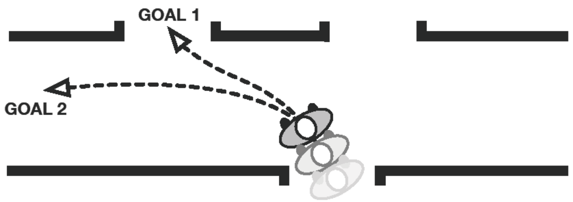
Human Motion Prediction Under Social Grouping Constraints
Method properties

- **Correct speed handling**



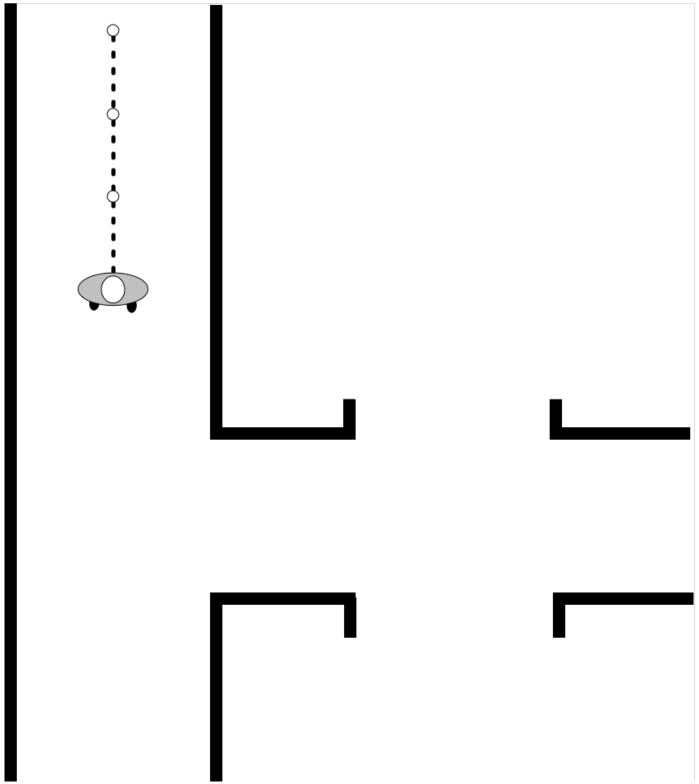
$$p(a) \text{ in } \hat{\pi}_g^i \propto \begin{cases} p(\langle\theta, v\rangle) \text{ in } \pi_g, \text{ if } v \leq v_{\text{obs}}^i, \\ p(\langle\theta, 2v_{\text{obs}}^i - v\rangle) \text{ in } \pi_g, \text{ if } v > v_{\text{obs}}^i \end{cases}$$

- **Reasoning about goals**

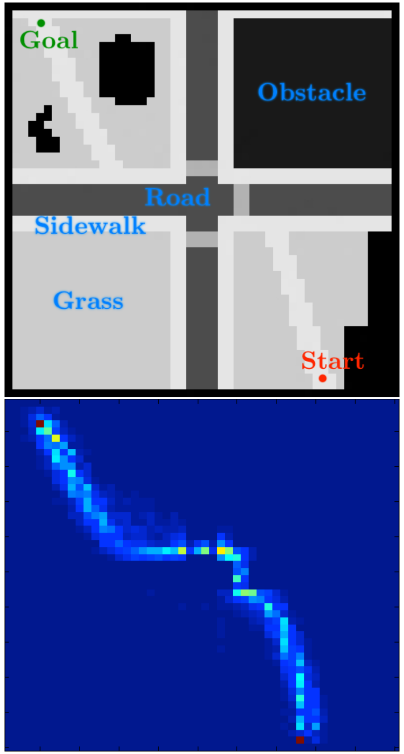


$$p(g) \propto \exp\left(\beta\left(V_g^*(s_{l(i)}^i) - V_g^*(s_1^i)\right)\right)$$

- **Multimodal predictions**

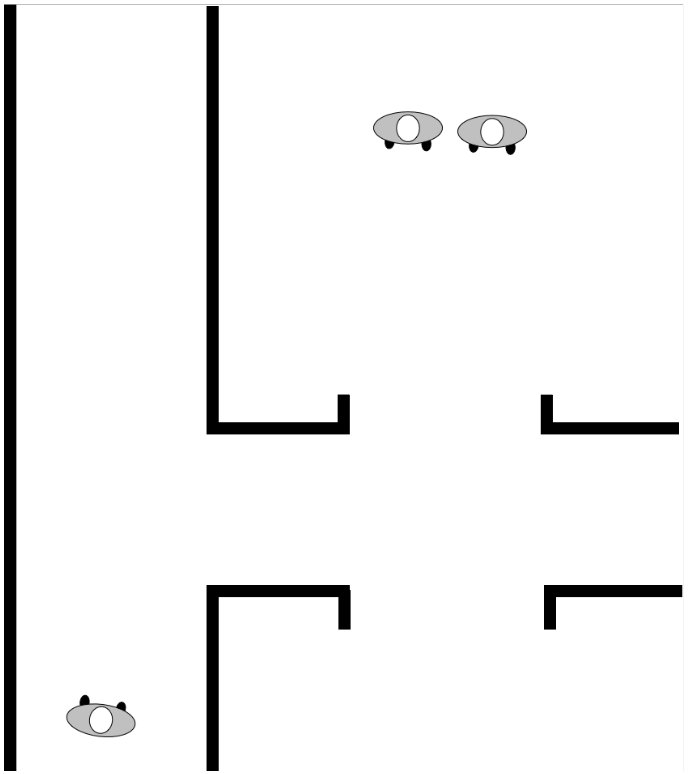


- **Semantic map input**



Human Motion Prediction Under Social Grouping Constraints
Method demonstration

- **Accounting for social interactions which include groups**

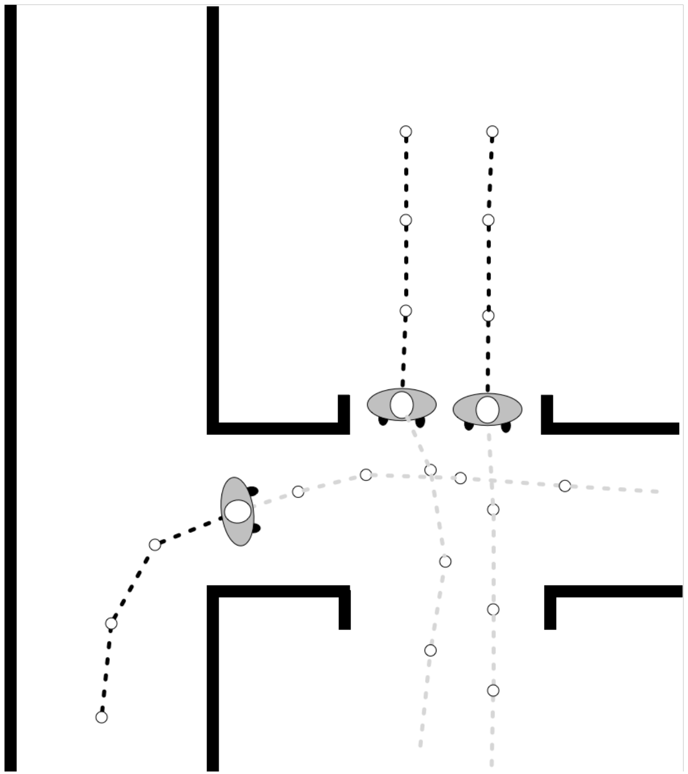


Scenario

Simulation

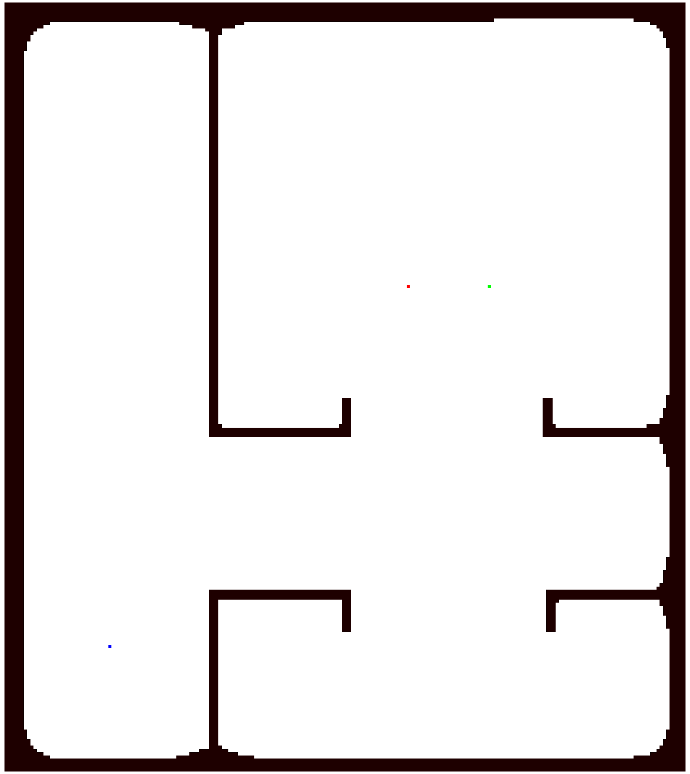
Human Motion Prediction Under Social Grouping Constraints
Method demonstration

- **Accounting for social interactions which include groups**



Scenario

Simulation



SEMANTIC OCCUPANCY PRIORS IN URBAN ENVIRONMENTS

AASS www.mrolab.eu

MR&O

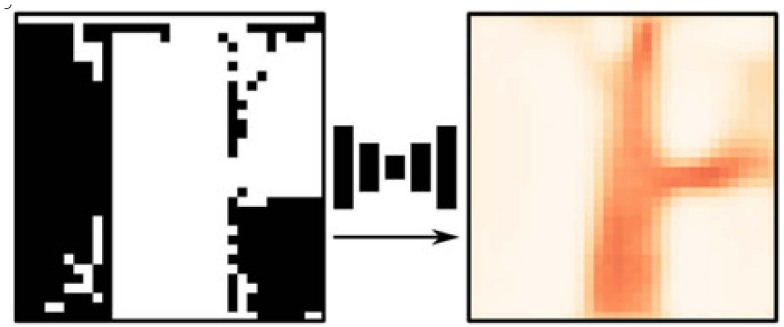
MOBILE ROBOTICS & OLFACTION

ÖREBRO UNIVERSITY

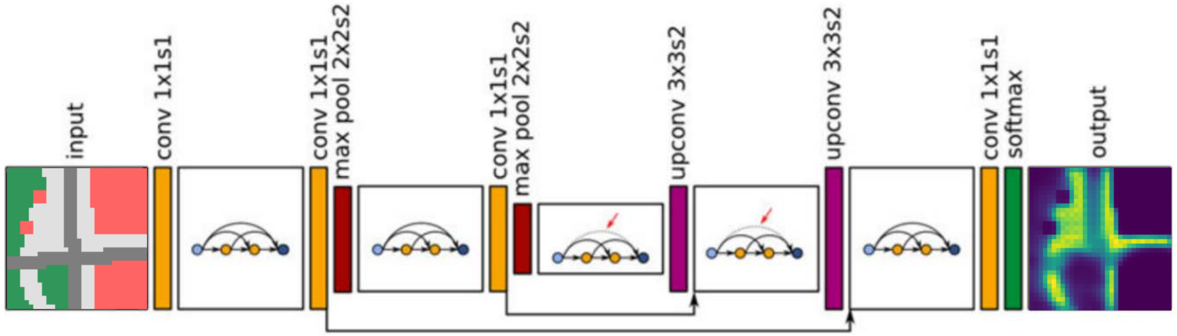
BOSCH

Semantic occupancy priors in urban environments
Overview

- Contributors: A. Rudenko, J. Döllinger et al.
- Long-term reasoning about dynamic agents in urban environment
- Estimating prior probability of observing a human in each cell of the semantic map
- Based on the occupancy priors estimation method by Doellinger et al. RA-L 2018
- Reasonably shallow CNN which generalizes well with small amounts of training examples

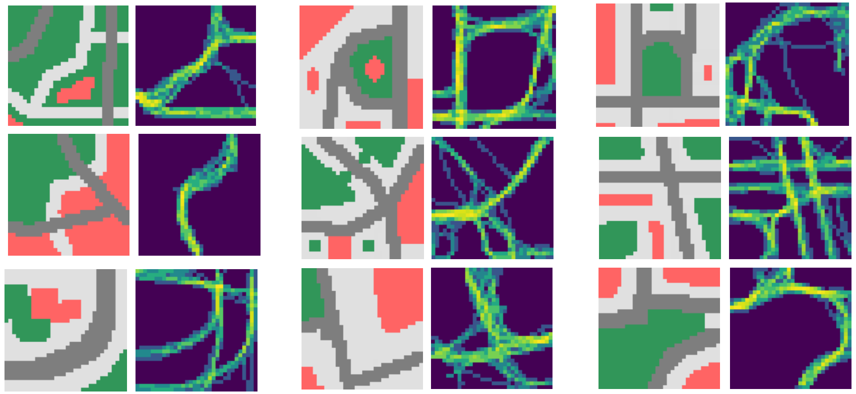


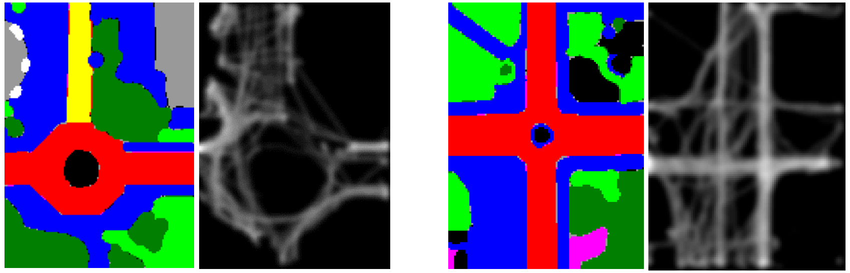
Döllinger et al., “Predicting Occupancy Distributions of Walking Humans With CNNs”, RA-L 2018



Semantic occupancy priors in urban environments
Training data

- **Simulated**: 80 urban scenes with 4 semantic classes, trajectories generated manually
  - Building
  - Sidewalk
  - Road
  - Grass
- **Real**: Stanford Drone Dataset, 25 scenes with sufficient coverage, 9 semantic classes
  - Road
  - Bicycle road
  - Pedestrian zone
  - Grass
  - Tree foliage
  - Building
  - Entrance
  - Obstacle
  - Bicycle parking



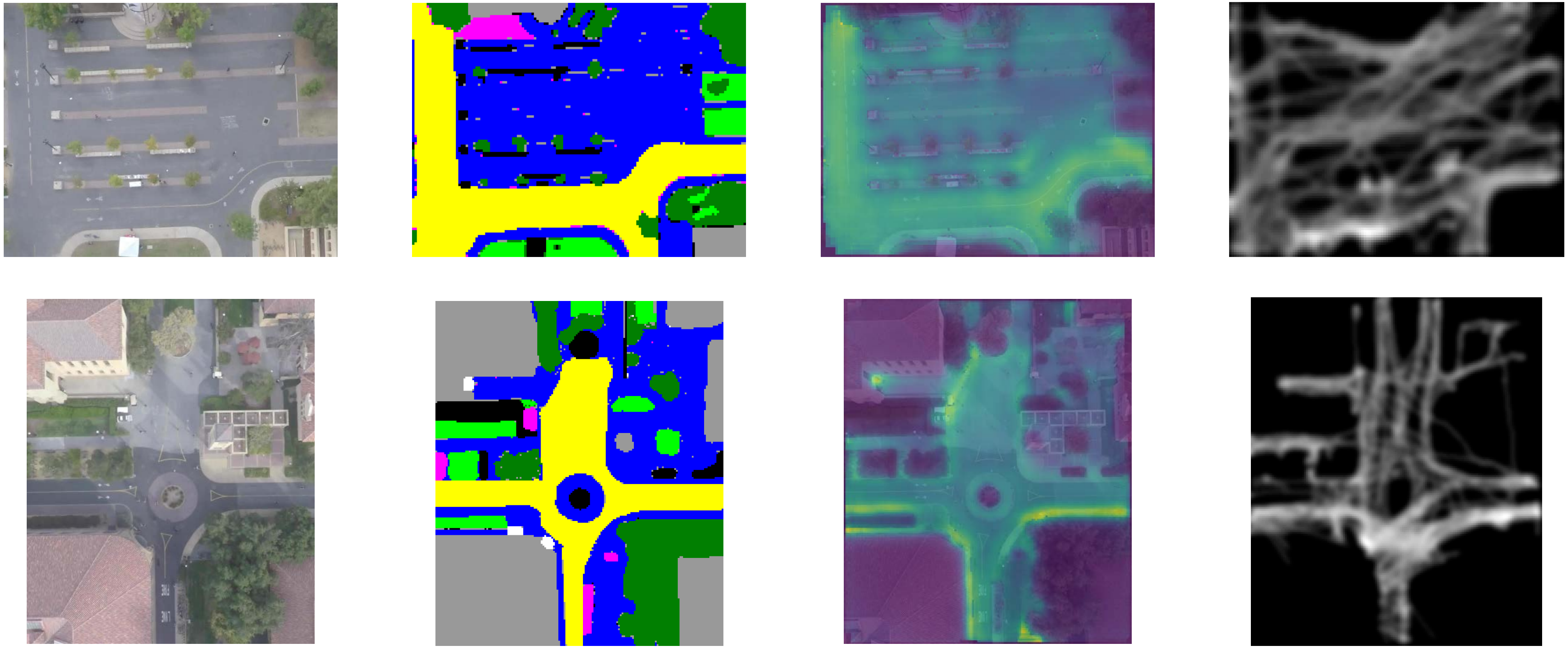


Scenes from the Stanford Drone Dataset, www.cvgl.stanford.edu/projects/uav_data/

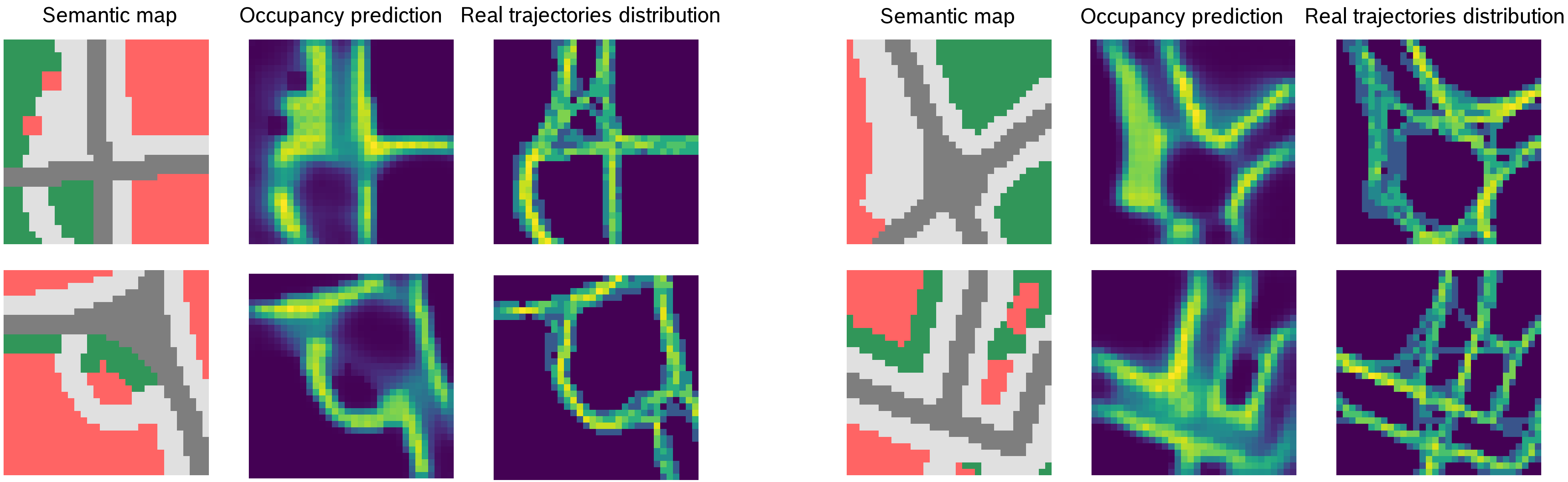
AASS www.mrolab.eu MR&O MOBILE ROBOTICS & OLFACTION
ÖREBRO UNIVERSITY
BOSCH

Semantic occupancy priors in urban environments
Results: Stanford Drone Dataset

Scene | Semantic map | Occupancy prediction | Real trajectories distribution



Semantic occupancy priors in urban environments
Results: simulated maps



- **Non-uniform probabilities** for states of the same semantic class
- Prediction for each cell based on the **global topology of the environment**
- Use-case: predicting “illegal crossroads” – places where people might cross the road

A NEW DATASET OF INDOOR HUMAN TRAJECTORIES

AASS www.mrolab.eu

MR&O

MOBILE ROBOTICS & OLFACTION

ÖREBRO UNIVERSITY

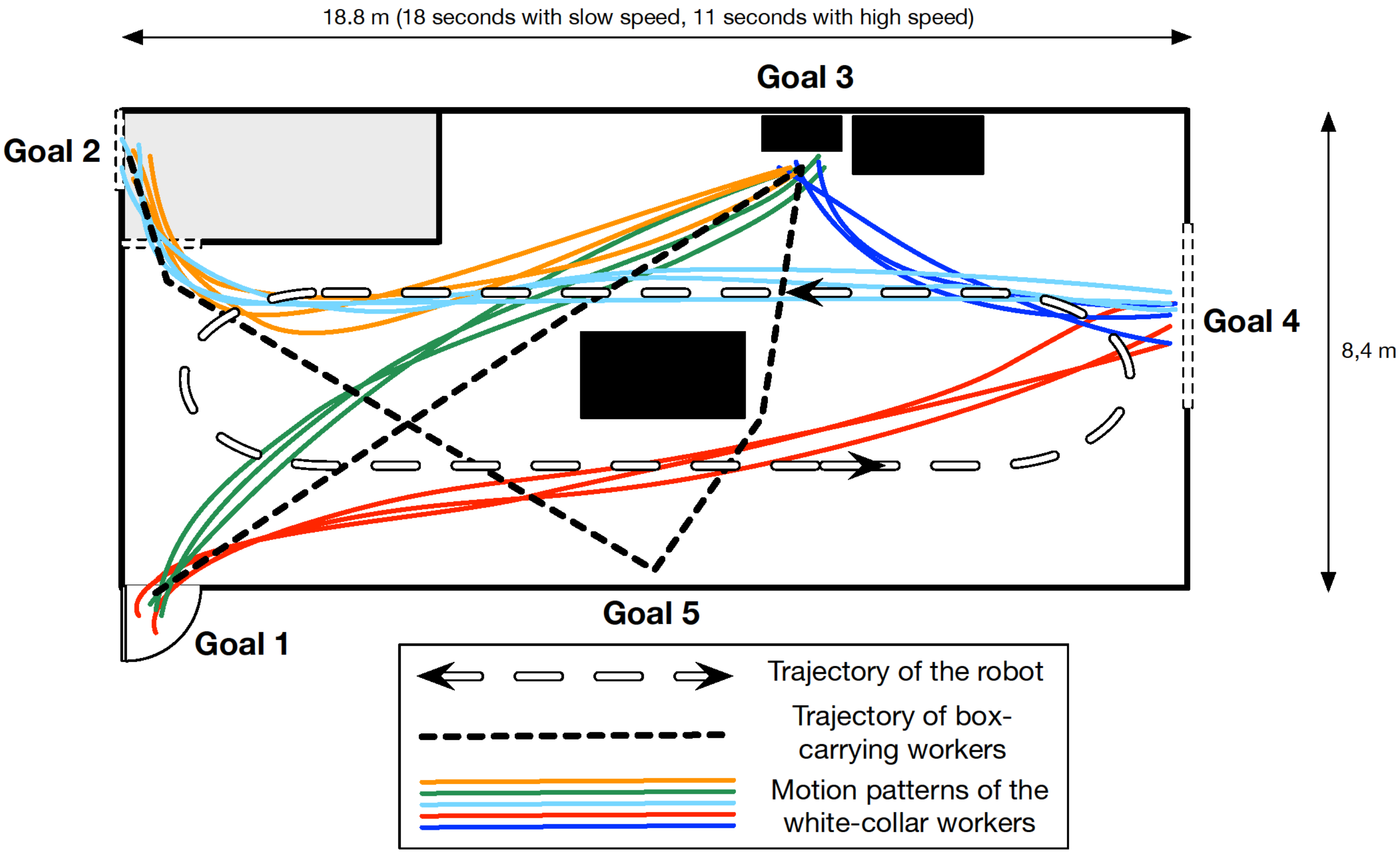
BOSCH

A new dataset of indoor human trajectories
Overview

- Contributors: A. Rudenko, T. Kucner, A. Lilienthal et al.
- **Motivation and key features of the dataset:**
  - **Controlled indoor experiment** in a large open-space environment
  - Experiment and instruction designed to ensure **natural walking patterns**
  - People moving **alone and in groups** with **various velocities** between several goal positions
  - **Long** trajectories, avoiding both **static and dynamic obstacles**
  - **Robot navigating** along the humans
  - **Precise ground truth position estimation** with the motion capture system

A new dataset of indoor human trajectories
Experiment design and recording

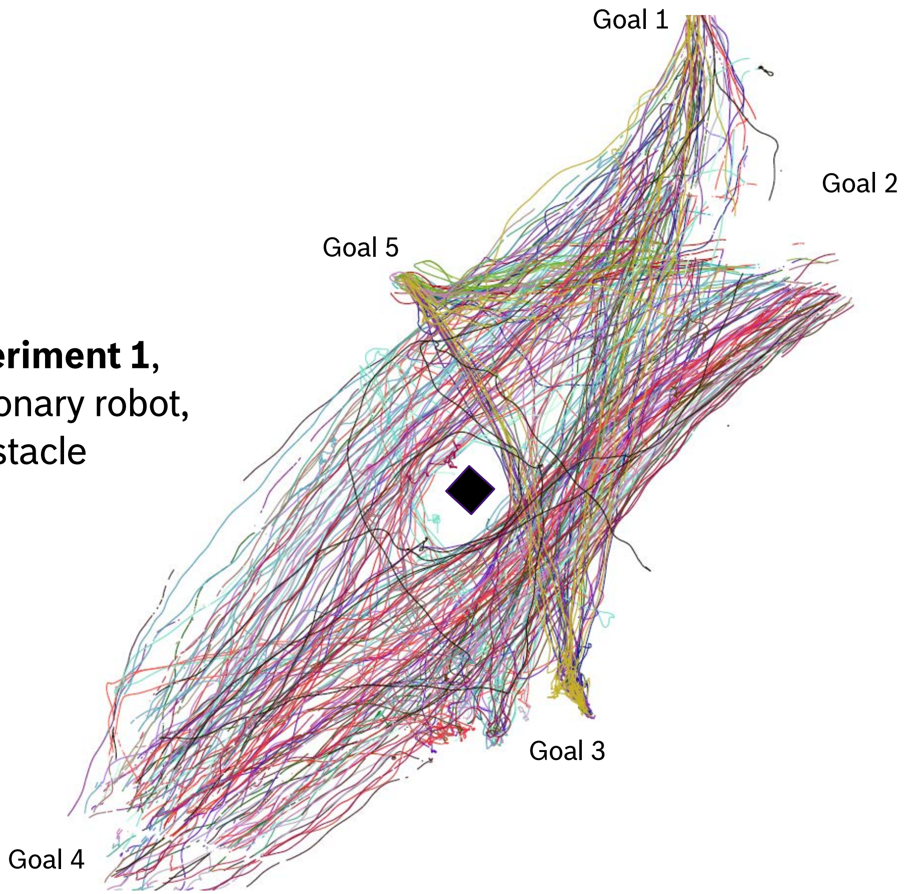
- Details:
  - **9 participants**: 2 carrying boxes, 6 walking between goals, 1 walking off-patterns
  - 4 participants were wearing **eye-tracking glasses**
  - Recording:
    - Qualysis motion capture MATLAB file
    - ROS stream of detection events
    - **RGB** and **Velodyne recordings** from stationary sensors
  - Experiment 1: stationary robot
  - Experiment 2: moving robot
  - Experiment 3: stationary robot, three obstacles
  - Total 39-52 minutes of motion recorded

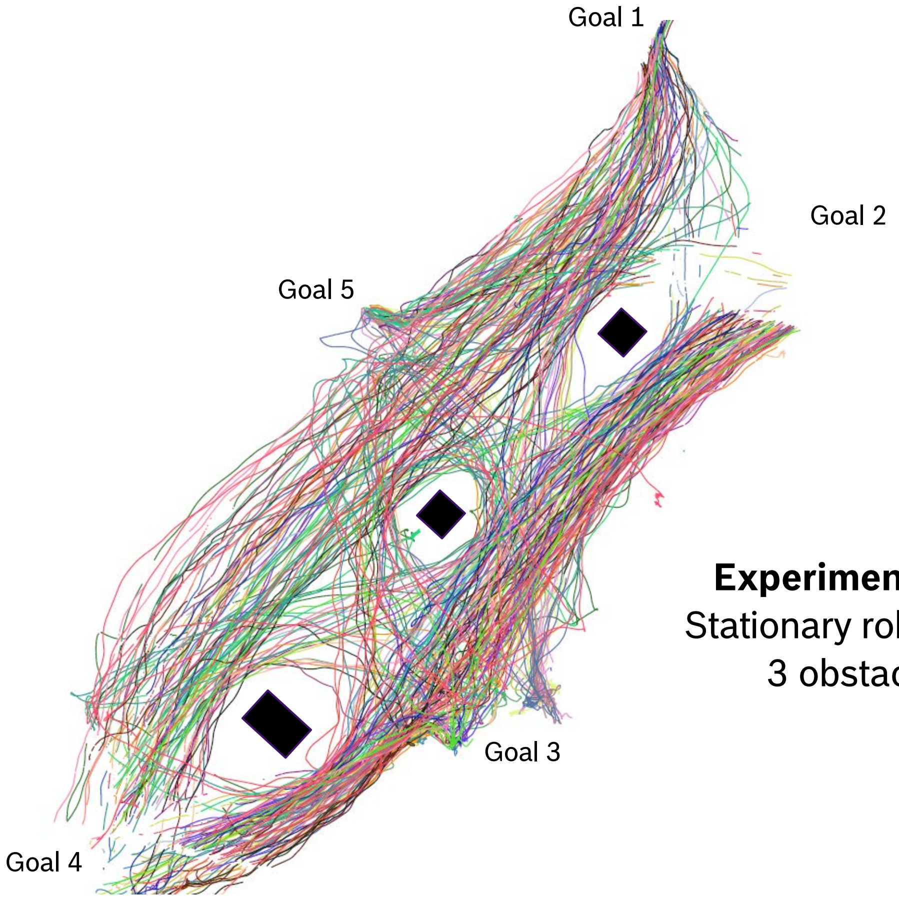


A new dataset of indoor human trajectories
Experiment setup

A new dataset of indoor human trajectories
Collected data

Experiment 1,
Stationary robot,
1 obstacle





Experiment 3,
Stationary robot,
3 obstacles

AASS www.mrolab.eu MR&O MOBILE ROBOTICS & OLFACTION
ÖREBRO UNIVERSITY
BOSCH

THANK YOU

AASS www.mrolab.eu
MR&O
MOBILE ROBOTICS & OLFACTION
ÖREBRO UNIVERSITY
BOSCH