

# Teaching Mobile Robots to Cooperatively Navigate in Populated Environments

Markus Kuderer

Henrik Kretschmar

Wolfram Burgard

**Abstract**—Mobile service robots are envisioned to operate in environments that are populated by humans and therefore ought to navigate in a socially compliant way. Since the desired behavior of the robots highly depends on the application, we need flexible means for teaching a robot a certain navigation policy. We present an approach that allows a mobile robot to learn how to navigate in the presence of humans while it is being tele-operated in its designated environment. Our method applies feature-based maximum entropy learning to derive a navigation policy from the interactions with the humans. The resulting policy maintains a probability distribution over the trajectories of all the agents that allows the robot to cooperatively avoid collisions with humans. In particular, our method reasons about multiple homotopy classes of the agents' trajectories, i.e., on which sides the agents pass each other. We implemented our approach on a real mobile robot and demonstrate that it is able to successfully navigate in an office environment in the presence of humans relying only on on-board sensors.

## I. INTRODUCTION

Mobile robots are expected to populate our human environments for numerous applications in the near future. These service robots should navigate in a socially compliant way that does not disturb nearby humans. The desired navigation behavior of a robot highly depends on the application at hand. For example, a cleaning robot should be unobtrusive and not unnecessarily hinder people, whereas a transportation robot that supplies an emergency room in a hospital must not delay its task by being overly cautious. Furthermore, since humans are likely to react differently to a robot than to a human, it is not always desirable to simply make the robot replicate human navigation strategies. Therefore, we need flexible means for non-experts to teach a robot how to navigate in its particular application.

Autonomous mobile robot navigation in an environment that is populated by humans is a challenging problem. Humans tend to cooperatively avoid collisions by mutually evading each other. To enable socially compliant human-robot interaction, mobile robots need to be able to engage in such cooperative navigation. However, since the behavior expected from robots may differ from typical human behavior, robots cannot learn how they are supposed to interact with humans by observing interacting humans. Consequently, teaching the desired behavior to a robot is difficult without actually having the robot experience interactions with humans in its designated environment.

All authors are with the Department of Computer Science, University of Freiburg, Germany.

This work has partly been supported by the EC under FP7-ICT-248873-RADHAR, by the German Research Foundation (DFG) under contract number SFB/TR-8, and by the Hans L. Merkle-Stiftung.



Fig. 1. Left: We teach a mobile robot an appropriate navigation behavior when interacting with humans via tele-operation. Right: Our approach learns a policy that allows the robot to autonomously navigate in a socially compliant way predicting the interaction behavior of nearby pedestrians.

This paper proposes a method that allows a mobile robot to learn an appropriate navigation policy from interactions with people that the robot experiences while being tele-operated in its designated environment. The robot thereby observes its own behavior, the behavior of nearby pedestrians, and, in particular, the interactions with these pedestrians. Our approach learns preferences of the robot and preferences of the humans in terms of accelerations, velocities, and timing. Furthermore, our method learns the expected human-robot interaction behavior, such as an adequate clearance between the agents. The robot can then use the learned policy to predict interactions with humans and to behave cooperatively, as illustrated in Fig. 1.

This paper builds on our previous work on feature-based prediction of human trajectories [14], which applies inverse reinforcement learning (IRL) to capture the navigation behavior of pedestrians in terms of a probability distribution over their trajectories. In this paper, we apply our behavior learning framework to mobile robot navigation in realistic environments with obstacles that are populated by humans. Our method enables teaching a mobile robot a navigation policy that explicitly takes into account human-robot interaction. Our model reasons about multiple homotopy classes of the agents' trajectories, i.e., on which sides the agents pass each other. Unfortunately, in general, the number of these homotopy classes grows exponentially with the number of agents, which hinders mobile robot navigation in realistic settings. To overcome this limitation, we propose to restrict the number of homotopy classes to consider by ignoring interactions that are highly unlikely in the current situation. Our experimental evaluation demonstrates a mobile robot that uses our approach to successfully navigate in an office environment in the presence of humans using only on-board sensors.

## II. RELATED WORK

There is a wide range of literature on learning policies from demonstrations [1, 3, 4, 21]. In contrast to direct approaches to imitation learning that infer a mapping from state features to actions [4], inverse reinforcement learning (IRL) recovers a cost function that explains observed behavior. In particular, Abbeel and Ng [1] suggest to match features that capture relevant aspects of the behavior. However, feature matching does not yield a unique cost function. To resolve this ambiguity, Maximum Entropy IRL [21] computes the policy with the highest entropy subject to feature matching.

The approach presented by Ziebart et al. [22] and our previous work [14] apply Maximum Entropy IRL to infer models of pedestrian navigation behavior. Kitani et al. [13] use Maximum Entropy IRL to infer human preferences with respect to vision-based physical scene features, such as sidewalks, and to predict the trajectories of pedestrians including their destinations. In addition to these probabilistic learning approaches, many researchers have proposed various models to capture the navigation behavior of humans [6–8]. Steering models describe human navigation behavior as a dynamical system, in which a set of rules determines the agent’s immediate action given its current state in the environment [9, 10, 20]. Optimization models cast pedestrians as utility-optimizing agents that minimize a cost function comprising relevant properties of human navigation [2, 12, 15].

The abovementioned approaches aim at modeling the navigation behavior of humans and can therefore be used to foster efficient and socially compliant mobile robot navigation in populated environments. The objective of this paper is to teach a robot a specific navigation behavior that is suitable for its task and that takes into account the behavior of the pedestrians in the vicinity of the robot. In particular, our approach implicitly learns how pedestrians typically react when interacting with robots.

Trautman and Krause [18] demonstrate that mobile robot navigation fails in densely populated environments unless the robot takes into account the interaction between the robot and the humans. Van den Berg *et al.* [19] present an approach to reciprocal collision avoidance that allows a set of mobile robots to navigate without collisions. Our approach also models cooperative navigation behavior.

Representing homotopy classes of trajectories is an important problem in motion planning [5, 11, 16]. Bhattacharya et al. [5] propose to integrate an obstacle marker function, which is given rise to by a representative point per obstacle. They show that two trajectories are homotopic if and only if they yield the same integral. Our approach reasons about homotopy classes of trajectories with respect to multiple agents.

## III. LEARNING NAVIGATION POLICIES FOR MOBILE ROBOTS IN POPULATED ENVIRONMENTS

The objective of this work is to teach a robot how to interact with humans by tele-operation in the designated environment. The robot observes its behavior and the reaction of nearby pedestrians and learns a policy of the desired interaction.

To derive the policy from the demonstrations, we extend our previous work [14] of learning cooperative navigation behavior of pedestrians, which we will shortly recap in the next section.

### A. Feature-Based Prediction of Trajectories

In this section, we summarize our previous work [14] on feature-based prediction of human trajectories. Given observations of interacting pedestrians, we learn a policy that induces a distribution over the trajectories of all the agents. This distribution depends on the current situation, i.e., the start and goal positions, and on a weight vector  $\theta$ . This weight vector determines the importance of features  $f$  that characterize relevant aspects of the agents’ behavior. During the learning process, we compute the value for  $\theta$  such that the expected feature values of the distribution match their corresponding empirical values of the observations.

More precisely, we define the trajectory  $x^a$  of an agent  $a$  as a continuous function

$$t \mapsto x^a(t) \in \mathcal{X} \quad (1)$$

that maps each point in time  $t$  to a configuration. We use cubic splines to represent these trajectories. The trajectories of all  $N$  agents traveling from their start positions to their target positions are captured by the composite trajectory

$$\mathbf{x} = (x^0, \dots, x^N) \in \mathcal{X}^N. \quad (2)$$

To characterize relevant properties of the behavior, we use features

$$f_i : \mathcal{X}^N \rightarrow \mathbb{R} \quad (3)$$

that map composite trajectories  $\mathbf{x}$  to feature values  $f_i(\mathbf{x})$ . We refer to the vector of all features as  $\mathbf{f}(\mathbf{x})$ . Following the idea of maximum entropy inverse reinforcement learning [21], we assume that the behavior of the agents can be described by a probability distribution over composite trajectories

$$p_{\theta}(\mathbf{x}) = \frac{1}{Z(\theta)} e^{-\theta^T \mathbf{f}(\mathbf{x})} \quad (4)$$

that depends on these feature values. In this model, learning the interaction behavior of the agents translates to finding values of  $\theta$  such that the expectations of the features  $\mathbb{E}_{p_{\theta}(\mathbf{x})}[\mathbf{f}(\mathbf{x})]$  match their empirical values  $\tilde{\mathbf{f}}$  from the observations. The gradient of  $\theta$  that leads to feature matching corresponds to the difference between the empirical and the expected feature values  $\tilde{\mathbf{f}} - \mathbb{E}_{p_{\theta}(\mathbf{x})}[\mathbf{f}(\mathbf{x})]$ , which allows us to apply gradient based methods to learn the desired interaction behavior.

Evaluating the gradient requires repeatedly computing the expected feature values, which is infeasible in practice. However, our experiments suggest that the modes of the probability distribution  $p_{\theta}(\mathbf{x})$  coincide with the homotopy classes of the interactions, which are subsets of  $\mathcal{X}^N$  in which all pairs of agents choose the same side when passing each other. Our method approximates the distribution using a weighted sum over Dirac delta functions at the most likely interaction of each homotopy class. Unfortunately, the number of homotopy classes increases exponentially with the number

**Algorithm 1** Computing the set of relevant homotopy classes

---

```

1: for all agents  $a$  do
2:    $x^a \leftarrow \text{globalPath}(x_{\text{start}}^a, x_{\text{goal}}^a)$ 
3:  $\mathbf{x} \leftarrow (x^1, \dots, x^N)$ 
4:  $\Psi' \leftarrow \{\mathbf{x}\}$ 
5: while unresolvedCollisions( $\Psi'$ ) do
6:   for all  $\mathbf{x} \in \Psi'$  do
7:     for all  $a, b$  with  $a \neq b$  do
8:       if  $a$  and  $b$  interact in  $\mathbf{x}$  and  $\mathbf{x}_{\pm 2\pi ab} \notin \Psi'$  then
9:          $\Psi' \leftarrow \Psi' \cup \mathbf{x}_{\pm 2\pi ab}$ 

```

---

of agents, which hinders mobile robot navigation in realistic applications. In this paper, we propose an approach that disregards unlikely homotopy classes and, therefore, is able to cope with many more agents.

### B. Reasoning About Relevant Homotopy Classes

The policies learned by our approach induce a probability distribution over interactions  $\mathbf{x} \in \mathcal{X}^N$  that comprise the trajectories of all the agents from their start to their target positions. As described in the previous section, we approximate this distribution using the most likely interaction of each homotopy class. The number of these homotopy classes increases exponentially with the number of agents. However, typically, most of the homotopy classes are highly unlikely. In this section, we propose a method to consider only relevant homotopy classes.

We refer to a homotopy class  $\psi \in \Psi$  as a set of interactions in which all pairs of agents pass each other on the same sides as they travel to their target positions. We denote the homotopy class of interaction  $\mathbf{x}$  as  $\psi(\mathbf{x})$ . To capture on which side two agents  $a$  and  $b$  pass each other, we integrate the derivative of the angle  $\alpha_a^b(t)$  of the vector  $x^b(t) - x^a(t)$  over time, leading to

$$\kappa_a^b(\mathbf{x}) = \int_t \dot{\alpha}_a^b(t) dt. \quad (5)$$

All the interactions that are homotopy equivalent yield the same  $\kappa_a^b$  for all pairs of agents. Thus, computing  $\kappa_a^b$  for all pairs of agents  $a$  and  $b$  yields a fingerprint

$$F(\psi) := F(\mathbf{x}) = (\kappa_a^b(\mathbf{x}))_{a \neq b} \text{ for } \mathbf{x} \in \psi \quad (6)$$

that allows us to describe and recognize the homotopy classes. Furthermore, we can enumerate the homotopy classes since we have

$$\kappa_a^b(\psi_1) - \kappa_a^b(\psi_2) \equiv 0 \pmod{2\pi}. \quad (7)$$

For example, two agents  $a$  and  $b$  that swap positions by passing each other on the right side yield  $\kappa_a^b = \pi$ , whereas passing on the left yields  $\kappa_a^b = -\pi$ .

In the following, we describe how our approach maintains a set  $\Psi' \subseteq \Psi$  of relevant homotopy classes for given start and target positions of all agents. The key idea is to identify potential evasive maneuvers of the agents and to consider both outcomes, i.e., the agents passing each other on the left

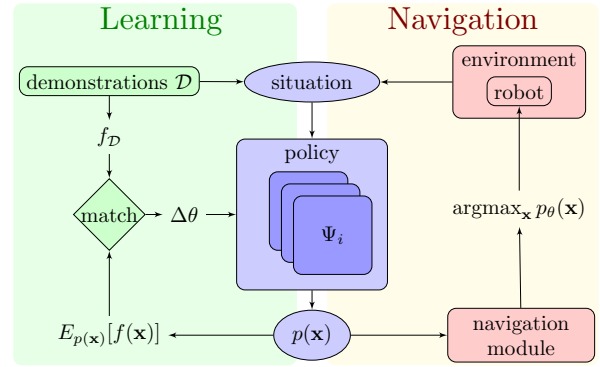


Fig. 2. Learning the policy (left) and using the policy for navigation (right). The learning approach computes the weight vector  $\theta$  such that the expected feature values match the empirical feature values of the observations. When navigating, the robot uses the policy to determine the most likely interaction in each time step and behaves accordingly.

or on the right side. In this way, interactions between agents that are highly unlikely are ignored.

Algorithm 1 outlines how our method computes  $\Psi'$ . For brevity of notation, we use  $\mathbf{x}$  and  $\psi(\mathbf{x})$  interchangeably. First, our algorithm computes global paths for all agents from their start positions to their target positions, which leads to the initial interaction  $\mathbf{x}_0$ . Subsequently, our algorithm identifies a potential evasive maneuver when two agents  $a$  and  $b$  come close to each other at some point in an interaction  $\mathbf{x} \in \Psi'$ . For such a potential evasive maneuver, we want to reason about both possible outcomes, i.e., passing left or passing right. We can efficiently compute the fingerprint of the homotopy class where the agents pass each other on the other side by adding  $\pm 2\pi$  to  $\kappa_a^b(\mathbf{x})$ . As a consequence, we can efficiently check whether our algorithm already takes into account the corresponding homotopy class  $\tilde{\psi}$ . If this is not the case, we compute an interaction of  $\tilde{\psi}$  in which the agents  $a$  and  $b$  pass on the other side as compared to  $\psi$ . Our algorithm repeatedly looks for potential evasive maneuvers in  $\Psi'$  as described above until there are no unresolved collisions.

### C. Learning Human-Robot Interaction in Realistic Environments

Our previous method [14] learns the interaction behavior of pedestrians by averaging over all the agents. Instead, in this paper, we learn individual behavior patterns for the robot and the humans. To do so, we use features that individually capture the preferences of the robot and the preferences of the pedestrians. Furthermore, we use features that capture the interaction behavior between the agents. To derive the policy from the demonstrations, we apply the learning method described in Sec. III-A, as illustrated in Fig. 2. We approximate the expected feature values using a weighted sum of Dirac delta functions that correspond to the homotopy classes. We apply the technique presented in the previous section to only consider relevant homotopy classes.

In our previous work, we proposed features accounting for velocities, accelerations, the time to reach the target, and distances between the agents, which allow us to reason about

pedestrian interactions in free space. In this paper, to navigate a robot in realistic environments with static obstacles, we additionally propose to use a feature

$$f_{\text{obstacle}}(\mathbf{x}) = \sum_a \frac{1}{f_{\text{time}}^a} \int_t \frac{1}{\|x^a(t) - o_{\text{min}}^a(t)\|^2} dt \quad (8)$$

that represents the agents' intent to keep clear of obstacles, such as walls, where  $o_{\text{min}}^a$  denotes the position of the obstacle that is closest to agent  $a$  at time  $t$ . The integral in Eq. (8) is normalized by the travel time of agent  $a$  to achieve invariance with respect to the agent's velocities.

#### IV. ROBOT NAVIGATION IN POPULATED ENVIRONMENTS

This section describes how the learned policy allows a mobile robot to navigate in populated environments, as illustrated in Fig. 2. The robot utilizes the learned policy to maintain a probability distribution over interactions of all the agents from their current positions to their target positions. Given this distribution, the robot computes the most likely interaction  $\arg\max_{\mathbf{x}} p_{\theta}(\mathbf{x})$  of all the agents and behaves accordingly.

##### A. Global Path Planning

The time to compute the probability distribution over interactions for a certain situation scales linearly to the overall travel time of the agents. To allow for real time planning, we utilize the learned policy to represent the immediate, joint behavior of all the agents and represent more distant behavior by globally planned paths independently for each agent. At time  $t_0$ , we evaluate the learned policy in the time interval  $[t_0, t_0 + t_p]$  and represent the behavior in the time interval  $(t_0 + t_p, t_{\text{end}}]$  by a globally planned trajectory for each agent. More specifically, for a pedestrian  $a$  that is detected by the robot the first time, we estimate its target position and generate an initial trajectory  $x_{\text{global}}^a$  using A\* search. The cost function for global path planning comprises the time to reach the target and the distance to obstacles, which is a subset of the features used for the learned policy, as described in Sec. III-C. Based on this global path, we set intermediate target positions given by  $x_{\text{global}}^a(t_0 + t_p)$  and utilize the learned policy to compute a probability distribution of the composite trajectory of all agents from their current position to their intermediate target positions.

In each planning cycle, the robot updates the intermediate targets along this global path and executes the trajectory that corresponds to the most likely interaction. Our current implementation allows the robot to replan at a frequency of 5hz, which means that the robot adapts its plan in real time to changes in the environment.

##### B. Perception

In order to perceive the environment, we localize the robot in a static map using laser-based Monte Carlo Localization [17]. To estimate the position and velocity of nearby pedestrians, we extract objects in the laser image that show typical characteristics of pedestrians. We then assign these observations to existing tracks of pedestrians, or add



Fig. 3. Left: The trajectory that the robot has planned to travel to its target position. The trajectory from the current position of the robot to its intermediate goal is computed by our policy. The trajectory from the intermediate goal to the target position of the robot is computed by an A\*-planner. Right: Trajectories driven by the robot using the policy learned by our approach. The robot successfully reached its goal position in 10 successive runs.

a new track if the observation cannot be assigned to any of the existing tracks. To prevent false positive detections, we disregard all laser observations in occupied space according to the static map. Our implementation allows the robot to observe its environment during tele-operation as well as during autonomous navigation only relying on on-board sensors.

#### V. EXPERIMENTAL EVALUATION

The goal of this section is to demonstrate that our approach allows a mobile robot to learn a navigation policy from demonstrations and to use it to autonomously navigate in an office environment shared with humans. In the experiments, we used a Pioneer-3 robot as well as a robotic wheelchair that were equipped with laser range finders. The robots used their laser ranger finder to track pedestrians and to localize themselves in the environment with a particle filter.

##### A. Cooperative Navigation in an Office Environment

We tele-operated the robot in our office environment to teach the robot a particular behavior, as illustrated in Fig. 1. In the process, the robot repeatedly encountered people and engaged in joint collision avoidance. In total, the robot recorded 12 minutes of interactions. During the first half of the demonstrations, we controlled the robot at a speed of  $0.25 \frac{m}{s}$ , in the remaining time at a speed of  $0.5 \frac{m}{s}$ . For each of these two demonstration sets the robot learned a policy using the method proposed in this paper.

In a first set of experiments, we evaluated the ability of our approach to replicate the demonstrated navigation behavior. The robot applied the learned policies to repeatedly navigate through our office environment. Both policies allowed the robot to reach its goal position in 10 successive runs without colliding with an obstacle. Fig. 3 (right) depicts the trajectories driven by the robot. The policies resulted in average velocities of  $0.25 \frac{m}{s}$ , and  $0.46 \frac{m}{s}$ , respectively. These results suggest that our approach accurately replicates the demonstrated behavior.

In a second set of experiments, we evaluated the ability of our approach to cooperatively navigate in an office environment in the presence of humans. An autonomous robotic wheelchair controlled by the method proposed in this paper passes two pedestrians in a hallway. Fig. 4 depicts the

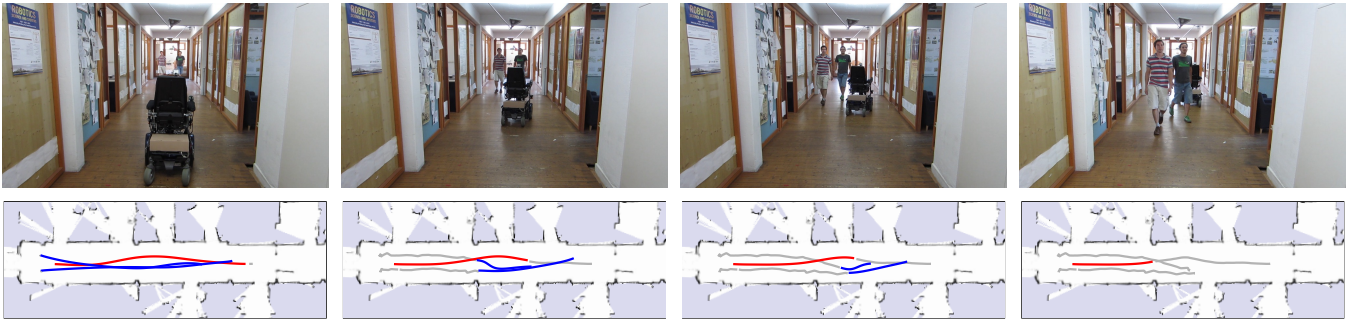


Fig. 4. Autonomous mobile robot navigation in a realistic scenario in which a robotic wheelchair passes two pedestrians in a hallway using the approach presented in this paper. The bottom images depict the driven trajectories (gray) and the interaction of the robot (red) with the pedestrians (blue) that is considered to be most likely by the robot at different time steps. At first, the pedestrians block the hallway such that a traditional path planning algorithm would be unable to find a path to the target position. In contrast, our method expects the pedestrians to cooperatively engage in joint collision avoidance with the wheelchair, and, hence, is able to find a path to the target position. Left: The robot assigns highest probability to evading on the right before the pedestrians begin to evade to either side. Middle left and middle right: The robotic wheelchair and the pedestrians evade each other in a natural manner. Right: After the situation has been resolved, the wheelchair proceeds to its target position.

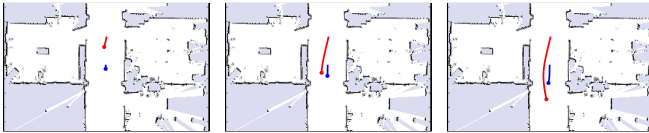


Fig. 5. Example interaction of the robot (red) with a pedestrian (blue) who walks at a low velocity in front of the robot. As a result, the robot decides to overtake the pedestrian.

composite trajectory that the robotic wheelchair predicted to be most likely at four different timesteps during the encounter. First, the pedestrians walk side by side, blocking the corridor. Note that a traditional path planner would not be able to find a path to the target position in such a situation. In contrast, our method expects the humans to cooperatively engage in joint collision avoidance. During the encounter, the robot repeatedly computes the most likely cooperative interaction with the pedestrians, which allows the wheelchair to engage in natural joint collision avoidance.

In a third set of experiments, a robot has learned to move faster than the pedestrians. We evaluated the behavior of the robot when it observes a pedestrian that walks in front of it in the same direction. Fig. 5 shows the resulting interaction, where the robot decides to overtake the pedestrian.

In the experiments described above, we assumed known target positions of the pedestrians. Currently, we are investigating methods to infer the target position of pedestrians based on the observed trajectories, which will broaden the applicability of our approach.

### B. Comparison in Simulation

We furthermore conducted a set of experiments in simulation to compare the performance of our method with the performance of a global path planning algorithm that is designed for dynamic environments. We therefore implemented an A\* planner in configuration-time space that predicts the motion of the pedestrians with a constant velocity model. We set the maximum velocity to  $0.5 \frac{m}{s}$ , which is similar to the velocity learned by the policy used for our approach. To

TABLE I  
COMPARISON OF OUR APPROACH TO AN A\* PLANNER.

algorithm	min dist	avg velocity	% blocked
our approach	0.34	0.56	0
A*, $r_{\min} = 0.2$	0.16	0.49	1
A*, $r_{\min} = 0.8$	0.38	0.27	60

acquire realistic test data, we used the abovementioned laser-based people tracker to record natural evasive movements of two pedestrians that evaded a third person in a hallway. To allow for a fair comparison, we fixed the trajectories of the two pedestrians and let the method proposed in this paper and the A\* planner control the third person, respectively.

Tab. I summarizes the results of the two methods averaged over 10 different scenarios. The results suggest that our method learned a policy that safely evaded the other pedestrians at a minimum clearance of 0.34 m and reached its target at an average velocity of  $0.5 \frac{m}{s}$ . In contrast, it turned out that it is difficult to tune the A\*-planner to obtain satisfactory results. Setting the minimal planned distance between the robot and dynamic obstacles to a low value such as 0.2 m, the planner did not sufficiently evade the pedestrians. To achieve an acceptable clearance, the minimal planned distance to dynamic obstacles needed to be set to 0.8 m since the pedestrians do not always comply with the constant velocity assumption. However, a value this large prevents the A\*-planner from finding a path to the target in many situations. In our experiments, the A\*-planner failed and needed to stop the robot in 60 % of the time steps, resulting in a rather low average velocity of  $0.27 \frac{m}{s}$ . This problem is referred to as the “freezing robot problem” [18]. Our experiment demonstrates the shortcomings of A\*-like path planners when navigating in the presence of humans. The approach presented in this paper is able to predict cooperative behavior of nearby pedestrians and is therefore able to navigate a robot efficiently and in a socially compliant way.

## VI. CONCLUSION

In this paper, we presented an approach that allows a mobile robot that is tele-operated in its designated environment to learn from observation how it is supposed to navigate in the presence of humans. The robot learns a model of its own navigation behavior, of the navigation behavior of the pedestrians, and, most importantly, of the cooperative interactions with the pedestrians. Our method relies on the principle of maximum entropy inverse reinforcement learning to compute a policy of the desired interaction behavior. The policy maintains a probability distribution over the trajectories of all the agents, which allows the robot to engage in cooperative collision avoidance with the pedestrians. The number of homotopy classes of the interactions between the agents, however, scales exponentially with the number of agents. We propose a technique to efficiently disregard unlikely homotopy classes, which increases the number of agents the robot is able to cooperate with. We demonstrated the advantage of our cooperative approach over global path planning algorithms. Furthermore, we carried out extensive experiments with a real robot that successfully navigated in an office environment shared with humans using only on-board sensors.

## REFERENCES

- [1] P. Abbeel and A.Y. Ng. Apprenticeship learning via inverse reinforcement learning. In *ICML*, 2004.
- [2] G. Arechavaleta, J.-P. Laumond, H. Hicheur, and A. Berthoz. An optimality principle governing human walking. *IEEE Transactions on Robotics (T-RO)*, 24(1): 5–14, 2008.
- [3] Brenna D Argall, Sonia Chernova, Manuela Veloso, and Brett Browning. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57 (5):469–483, 2009.
- [4] Christopher G Atkeson and Stefan Schaal. Robot learning from demonstration. In *Proceedings of the Fourteenth International Conference on Machine Learning (ICML)*, pages 12–20, 1997.
- [5] Subhrajit Bhattacharya, Vijay Kumar, and Maxim Likhachev. Search-based path planning with homotopy class constraints. In *In Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence*, Atlanta, Georgia.
- [6] S. Bitgood and S. Dukes. Not another step! economy of movement and pedestrian choice point behavior in shopping malls. *Environment and Behavior*, 38(3):394–405, 2006.
- [7] H.I. Christensen and E. Pacchierotti. Embodied social interaction for robots. *AISB-05*, pages 40–45, 2005.
- [8] E.T. Hall. *The Hidden Dimension*. Doubleday, New York, 1966.
- [9] D. Helbing and A. Johansson. Pedestrian, crowd and evacuation dynamics. In *Encyclopedia of Complexity and Systems Science*, pages 6476–6495. Springer New York, 2009.
- [10] D. Helbing and P. Molnar. Social force model for pedestrian dynamics. *Physical Review E (PRE)*, 51: 4282–4286, 1995.
- [11] John Hershberger and Jack Snoeyink. Computing minimum length paths of a given homotopy class. *Computational geometry*, 4(2):63–97, 1994.
- [12] S. Hoogendoorn and P.H.L. Bovy. Simulation of pedestrian flows by optimal control and differential games. *Optimal Control Applications and Methods*, 24(3):153–172, 2003.
- [13] Kris M. Kitani, Brian D. Ziebart, Drew Bagnell, and Martial Hebert. Activity forecasting. In *European Conference on Computer Vision (ECCV 2012)*.
- [14] M. Kuderer, H. Kretschmar, C. Sprunk, and W. Burgard. Feature-based prediction of trajectories for socially compliant navigation. In *Proceedings of Robotics: Science and Systems (RSS)*, 2012.
- [15] Q.-C. Pham, H. Hicheur, G. Arechavaleta, J.-P. Laumond, and A. Berthoz. The formation of trajectories during goal-oriented locomotion in humans. ii. a maximum smoothness model. *European Journal of Neuroscience*, 26:2391–2403, 2007.
- [16] E Schmitzberger, JL Bouchet, M Dufaut, D Wolf, and R Husson. Capture of homotopy classes with probabilistic road map. In *Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on*, volume 3, pages 2317–2322, 2002.
- [17] S. Thrun, D. Fox, W. Burgard, and F. Dellaert. Robust monte carlo localization for mobile robots. *Artificial Intelligence*, 128(1-2):99–141, 2000.
- [18] P. Trautman and A. Krause. Unfreezing the robot: Navigation in dense, interacting crowds. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 797–803, 2010.
- [19] J. van den Berg, S.J. Guy, M.C. Lin, and D. Manocha. Reciprocal n-body collision avoidance. In *Proceedings of the International Symposium of Robotics Research (ISRR)*, pages 3–19, 2009.
- [20] W. Warren. The dynamics of perception and action. *Psychological Review*, 113:358–389, 2006.
- [21] B.D. Ziebart, A. Maas, J.A. Bagnell, and A.K. Dey. Maximum entropy inverse reinforcement learning. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, pages 1433–1438, 2008.
- [22] B.D. Ziebart, N. Ratliff, G. Gallagher, C. Mertz, K. Peterson, J.A. Bagnell, M. Hebert, A.K. Dey, and S. Srinivasa. Planning-based prediction for pedestrians. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2009.