Feature-Based Prediction of Trajectories for Socially Compliant Navigation

Markus Kuderer   Henrik Kretzschmar   Christoph Sprunk   Wolfram Burgard

Department of Computer Science, University of Freiburg, Germany

{kudererm, kretzsch, sprunkc, burgard}@informatik.uni-freiburg.de

Abstract—**Mobile robots that operate in a shared environment with humans need the ability to predict the movements of people to better plan their navigation actions. In this paper, we present a novel approach to predict the movements of pedestrians. Our method reasons about entire trajectories that arise from interactions between people in navigation tasks. It applies a maximum entropy learning method based on features that capture relevant aspects of the trajectories to determine the probability distribution that underlies human navigation behavior. Hence, our approach can be used by mobile robots to predict forthcoming interactions with pedestrians and thus react in a socially compliant way. In extensive experiments, we evaluate the capability and accuracy of our approach and demonstrate that our algorithm outperforms the popular social forces method, a state-of-the-art approach. Furthermore, we show how our algorithm can be used for autonomous robot navigation using a real robot.**

Fig. 1.   Encounter of a robot and a pedestrian. A robot that is able to predict the behavior of the pedestrian can navigate in a socially more compliant way.

I. INTRODUCTION

As the scope of application of robots is expanding from industrial labor to domestic services, mobile robots are more and more expected to share their environment with people. As a result, new challenges for their navigation systems arise. The ability to perceive the intention and to predict the behavior of people is of pivotal importance for efficient and socially compliant navigation in human environments.

Common approaches to mobile robot navigation in populated environments compute a path to the target position without taking into account the typical collaborative collision avoidance behavior of humans. Many approaches rather rely on reactive collision avoidance methods [3, 24, 21]. Other methods evaluate trajectories based on the risk of motion conflicts, given predictions about potential future places covered by people [2, 15, 28]. Although such methods effectively avoid collisions with humans, the resulting trajectories are often suboptimal with respect to the expectations of the humans due to awkward and unexpected evasive movements. In this paper, we present an algorithm for learning typical human navigation behavior. Predicting human behavior and the ability to react in a natural way will enable robots to become socially more compatible, which is relevant for mobile robots operating in populated environments including urban scenes and traffic.

Our approach is motivated by recent studies that suggest that the navigation behavior of humans is based on multiple criteria including the time of travel and comfort, which can be expressed in terms of quantities including accelerations and proximity to other humans. These findings suggest that one can accurately model human behavior if the generated trajectories

are similar to those of humans with respect to these criteria. In this paper, we present an approach to learning typical human trajectories emerging from interactive behavior. Therefore, our method considers features that correspond to natural properties of planned trajectories of several people to learn a probability distribution over possible joint trajectories that fits previously observed human behavior. We apply two methods for online trajectory planning during navigation. First, we exploit the fact that the modes of the probability distribution mainly correspond to different topological variants of the trajectories. This provides a natural separation of the search space. Second, we present a hierarchical and efficient optimization scheme to calculate the most likely trajectory.

The contribution of this paper is a novel approach to learn a model of human navigation behavior that is based on the principle of maximum entropy. The approach allows a robot to elicit the underlying probability distribution that captures human navigation behavior from the observations of pedestrians. We believe that, given this model, mobile robots will be enabled to predict the navigation behavior of people in their vicinity and to react appropriately.

II. RELATED WORK

After Goffman [8] had outlined how mutually accepted rules allow humans to navigate safely in shared environments, several authors proposed rule-based approaches that describe human behavior. Recently Müller et al. [19] presented an application of this in the context of navigation where a mobile robot

detects and follows people that walk in the same direction as the robot. In contrast to this, Lerner et al. [16] infer a database of navigation rules from video data.

A more general and highly popular approach to model pedestrian motion behavior is the social forces model by Helbing and Molnar [11], which suggests that human motion can be described by forces that correspond to internal objectives of humans, such as the desire to reach a target and to avoid obstacles. Several authors used parameter learning methods to fit the model to observed crowd behavior [10, 14]. Although the social forces model performs well at simulating crowds, we found that the model poorly predicts the movements of individual pedestrians, particularly during evasive maneuvers.

Several reactive collision avoidance approaches were successfully applied to mobile robot navigation for robots that navigate in crowded or dynamic environments. Such methods include the dynamic window approach by Fox et al. [6] and the velocity obstacles by Fiorini and Shillert [4]. In 2009, the latter approach was extended to reciprocal velocity obstacles (RVO), a local reactive collision avoidance method that is guaranteed to produce collision free paths if all agents apply this algorithm. To allow for more human-like behavior, Guy et al. [9] extend RVO by introducing response and observation time to other agents. Whereas the approaches described above seek to avoid dynamic obstacles such as pedestrians, they do not consider cooperation and human predictive abilities, which sometimes results in unnatural movements.

For this reason, several authors proposed utility-based optimization approaches that consider properties of possible trajectories in a given period of time. For instance, Arechavaleta et al. [1] present a nonholonomic motion model that minimizes the norm of the accelerations along the trajectory, assuming humans navigate as efficiently as possible. Mombaur et al. [18] add an additional cost component, penalizing the time to reach the target. Furthermore, they use inverse optimal control to learn the model parameters that best explain observed data. While these two approaches do not consider interactions between humans, Hoogendoorn and H.L. Bovy [12] propose a model for multi-agent behavior. They argue that cooperative human behavior can be modeled by optimizing a joint cost function, which is similar to the assumptions in the work presented here.

Pellegrini et al. [20] present a motion model to improve the prior estimate of a tracking system. They assume that humans seek to minimize an energy function comprising the predicted minimal distance between two agents. As in the work described here, their approach takes into account the multiple topological structures of the trajectories of the different agents.

Trautman and Krause [25] point out that joint collision avoidance is crucial for mobile robot navigation to prevent the robot from "freezing" and getting stuck in densely populated environments. Their work assumes humans to be utility-optimizing agents that prefer trajectories with low cost. In contrast to many other approaches, their method is the only work that takes into account mutual interaction to plan joint trajectories. In this paper we are also interested in the most likely joint trajectories of robots and humans. To tackle the

optimization problem, the authors propose sampling from a joint density function, whereas we exploit the topological structure of the problem. Furthermore, as opposed to Trautman and Krause [25], we learn our model from observations of pedestrians to achieve more natural and human-like trajectories.

The work presented in this paper is also related to the elastic bands approach, which was introduced by Quinlan and Khatib [22] and also considers continuous trajectories to the goal position. Elastic bands aim to compute collision free, smooth paths by gradually deforming an initially coarse path to the goal. Fraichard and Delsart [7] extend elastic bands to deform the entire trajectory in the configuration-time space, similar to our optimization approach, although they assume constant velocity models for all obstacles.

The learning approach described in this paper is inspired by maximum entropy inverse reinforcement learning, introduced by Ziebart et al. [27], who also used the approach for learning pedestrian behavior on a grid world [28]. In contrast to their approach, we do not assume an underlying Markov decision process and we consider continuous trajectories, which allows for introducing features that capture physical aspects of the trajectories as well as topological properties.

III. LEARNING HUMAN-LIKE NAVIGATION

The objective of this work is to accurately learn human navigation behavior from observations. Some authors [1, 18] suggest to minimize the Euclidean distance between the generated and the observed trajectories. According to our experience in practical experiments, this approach often leads to overfitting since the algorithms try to match the exact shape of the demonstrated trajectories and often do not generalize seamlessly to different situations. Lerner et al. [16] try to overcome this problem by using a huge amount of training data. In contrast to that, we use features that capture important properties of human motion behavior to generate behavior that is similar with respect to these features. In particular, the features may capture principles underlying human navigation behavior found by psychological studies, such as the attempt to move efficiently. Therefore, our approach is based on a continuous representation of the trajectories that allows us to define appropriate features. Furthermore, our model explicitly accounts for the decisions of the agents regarding which sides they choose when passing each other to avoid collisions by reasoning about topological variants of the trajectories.

A. Preliminaries

The trajectory x^a of agent a is defined as a continuous function

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

that maps each point in time t to a configuration $x \in \mathcal{X}$. We furthermore consider the joint trajectory $\mathbf{x}$ that defines the behavior of all the agents as the Cartesian product of the trajectories of all N agents:

$$\mathbf{x}(t) = x^1(t) \times x^2(t) \times \ldots \times x^N(t) \in \mathcal{X}^N. \qquad (2)$$

We refer to a topological variant ψ of the joint trajectories as a subset of $\mathcal{X}^N$ in which all pairs of agents choose the same sides when passing each other to avoid collisions. With this definition, we can partition the space $\mathcal{X}^N$ of joint trajectories into its topological variants $\psi \in \Psi$ such that

$$\bigcup_{\psi \in \Psi} \psi = \mathcal{X}^N, \; \forall \psi_i, \psi_j \in \Psi : \psi_i \cap \psi_j = \emptyset. \qquad (3)$$

A feature f_i is a function that maps joint trajectories to feature values

$$f_i : \mathcal{X}^N \rightarrow \mathbb{R}, \qquad (4)$$

and we refer to a feature vector that comprises several features as $\mathbf{f}$.

B. Principle of Maximum Entropy and Feature Matching

We assume that the behaviors of the agents can be described by a probability distribution of the joint trajectories, and, in particular, that it is a function of the feature values. Therefore, our goal is to find the probability distribution p that induces behavior such that the expected feature values of the agents match the empirical features $\mathbf{f}_\mathcal{D}$ of the demonstrated human joint trajectories $\mathcal{D}$. Hence, we want

$$\mathbb{E}_{p(\mathbf{x})}[\mathbf{f}(\mathbf{x})] = \mathbf{f}_\mathcal{D} = \frac{1}{|\mathcal{D}|} \sum_{\mathbf{x}_k \in \mathcal{D}} \mathbf{f}(\mathbf{x}_k). \qquad (5)$$

Following Ziebart et al. [27], we are interested in the distribution that matches the feature expectations, as given in Eq. (5), without implying any further assumptions. In this section, we outline how their approach can be applied to continuous spaces. The principle of maximum entropy states that the desired distribution maximizes the differential entropy

$$\underset{p}{\operatorname{argmax}} \, H(p) = \underset{p}{\operatorname{argmax}} \int_{\mathbf{x}} -p(\mathbf{x}) \log p(\mathbf{x}) d\mathbf{x} \qquad (6)$$

subject to the constraints

$$\int_{\mathbf{x}} p(\mathbf{x}) \, d\mathbf{x} = 1, \qquad (7)$$

$$\forall i \; f_{i\mathcal{D}} = \mathbb{E}_{p(\mathbf{x})}[f_i(\mathbf{x})] = \int_{\mathbf{x}} p(\mathbf{x}) f_i(\mathbf{x}) d\mathbf{x}. \qquad (8)$$

Introducing Lagrangian multipliers α and θ_i for these constraints yields the maximization problem

$$p^\star, \alpha^\star, \boldsymbol{\theta}^\star = \underset{p,\alpha,\boldsymbol{\theta}}{\operatorname{argmax}} \int_{\mathbf{x}} -p(\mathbf{x}) \log p(\mathbf{x}) \, d\mathbf{x}$$
$$- \alpha \Big(\int_{\mathbf{x}} p(\mathbf{x}) \, d\mathbf{x} - 1 \Big) - \sum_i \theta_i \Big(\int_{\mathbf{x}} p(\mathbf{x}) f_i(\mathbf{x}) \, d\mathbf{x} - f_{i\mathcal{D}} \Big). \qquad (9)$$

Applying the Euler-Lagrange equation from calculus of variations (see [5, Sec. 1.4]) to Eq. (9) implies that the probability distribution $p^\star(\mathbf{x})$ has the structure

$$p_{\boldsymbol{\theta}}(\mathbf{x}) = \frac{1}{Z(\boldsymbol{\theta})} e^{-\boldsymbol{\theta}^T \mathbf{f}(\mathbf{x})}, \qquad (10)$$

where $Z(\boldsymbol{\theta})$ is a normalization factor to satisfy Eq. (7).

Interestingly, the term $\boldsymbol{\theta}^T \mathbf{f}(\mathbf{x}) = \sum_i \theta_i f_i(\mathbf{x})$ can be interpreted as a cost function that depends on a weighted sum of feature values. Hence, the model assumes that the agents are exponentially more likely to choose a trajectory with lower cost. In the following, we refer to $\boldsymbol{\theta}$ as the feature weights.

The log-likelihood of the observed behavior $\mathcal{D}$ is given by

$$L_{p_{\boldsymbol{\theta}}}(\mathcal{D}) = \log \frac{1}{Z(\boldsymbol{\theta})} e^{-\boldsymbol{\theta}^T \mathbf{f}_\mathcal{D}}, \qquad (11)$$

and its derivative with respect to $\boldsymbol{\theta}$ is given by

$$\frac{\partial}{\partial \boldsymbol{\theta}} L_{p_{\boldsymbol{\theta}}}(\mathcal{D}) = \int_{\mathbf{x}} p(\mathbf{x}) \mathbf{f}(\mathbf{x}) \, d\mathbf{x} - \mathbf{f}_\mathcal{D}. \qquad (12)$$

Eq. (12) reveals that the derivative of the log-likelihood with respect to $\boldsymbol{\theta}$ vanishes when the expected feature values match the empirical feature values. In other words, we obtain the constraints that we introduced in Eq. (5). Consequently, the problem of finding the maximum entropy distribution of joint trajectories subject to feature matching is equivalent to maximizing the likelihood of the training data when assuming an exponential family distribution [13].

Maximizing the log-likelihood of the demonstrations to find the distribution $p^\star(\mathbf{x})$ translates to computing the weights θ_i. The resulting optimization problem cannot be solved analytically. However, using the gradient given in Eq. (12), we can apply gradient based optimization techniques.

C. Learning the Feature Weights of Human Navigation

Given the continuous representation of joint trajectories (see Sec. IV-A), it is not feasible to compute the expected feature values

$$\mathbb{E}_{p_{\boldsymbol{\theta}}(\mathbf{x})}[\mathbf{f}(\mathbf{x})] = \int_{\mathbf{x}} p_{\boldsymbol{\theta}}(\mathbf{x}) \mathbf{f}(\mathbf{x}) \, d\mathbf{x} \qquad (13)$$

for a given $\boldsymbol{\theta}$ because integrating over the high-dimensional space of all trajectories is neither feasible analytically, nor by means of sampling techniques, such as Markov chain Monte Carlo methods. Hence, we need to resort to an efficient approximation of Eq. (13).

Our experiments suggest that, given the features we propose, the modes of the probability distribution coincide with the different topological variants. We therefore propose to approximate the feature expectations using the partition given in Eq. (3). For each topological variant $\psi \in \Psi$, we compute the most likely joint trajectory $\mathbf{x}_\psi^\star$ using Eq. (10). We estimate the feature expectations as

$$\mathbb{E}_{p_{\boldsymbol{\theta}}(\mathbf{x})}[\mathbf{f}(\mathbf{x})] \approx \hat{\mathbf{f}}_{\boldsymbol{\theta}} = \frac{1}{\sum_\psi p_{\boldsymbol{\theta}}(\mathbf{x}_\psi^\star)} \sum_{\psi \in \Psi} p_{\boldsymbol{\theta}}(\mathbf{x}_\psi^\star) f(\mathbf{x}_\psi^\star), \quad (14)$$

where

$$\mathbf{x}_\psi^\star = \underset{\mathbf{x}}{\operatorname{argmax}} \, p_{\boldsymbol{\theta}}(\mathbf{x}), \; \text{s.t. } \mathbf{x} \in \psi. \qquad (15)$$

This means that we approximate the distribution of joint trajectories using a weighted sum of Dirac delta functions at the modes of the distribution that correspond to the topological variants. This approximation can be interpreted as that the agents are assumed to reason about a topological variant

Fig. 2. Recording real-world training data: The figures show snapshots of three human subjects avoiding each other in order to proceed to their next target location. Using a motion capture system, we recorded one hour of human interactions by tracking markers that were attached to hats worn by the subjects. To prevent the subjects from focussing on their walking behavior, which may result in awkward trajectories, we made them read and memorize newspaper articles at various locations saying that we were conducting psychological research on the effects on attention when frequently changing locations.

in terms of its most likely joint trajectory. This is a rather coarse approximation of the feature expectations. However, in our practical experiments, we never observed that this approximation had a negative influence on the overall result.

IV. EFFICIENT REPRESENTATION AND LEARNING OF TRAJECTORIES IN CONTINUOUS ENVIRONMENTS

A. Spline-Based Representation

In this paper, we utilize a continuous representation of joint trajectories that represents the positions of the agents over time. More precisely, we utilize cubic splines in $\mathbb{R}^2$ to represent the x and y positions of the agents, as splines have several properties that are desirable for representations of human trajectories and at the same time have a finite dimensionality. Each spline segment

$$s_j^a : [t_j, t_{j+1}] \to \mathbb{R}^2 \qquad (16)$$

defines the position of the agent a in the time interval $[t_j, t_{j+1}]$. Thus, the position of the agent at time t is

$$x^a(t) = s_j^a(t), \text{ where } t \in [t_j, t_{j+1}]. \qquad (17)$$

The control points $x_j^a := x^a(t_j)$ and $v_j^a := \dot{x}^a(t_j)$ fully specify the trajectory, since cubic splines have four degrees of freedom. This allows for computing the positions $x^a(t)$, the velocities $\dot{x}^a(t)$, and the accelerations $\ddot{x}^a(t)$ in closed form, which is convenient for efficiently computing the values of the features we will describe in the following section.

B. Features

According to recent studies, pedestrians seem to navigate in crowded environments considering various criteria, such as time of travel, accelerations, and proximity to other pedestrians [12]. In our case, we express these criteria in terms of features that map joint trajectories to real numbers. To account for cooperative human behavior, we sum up the features of the individual agents according to

$$f_i(\mathbf{x}) = \sum_a f_i^a(\mathbf{x}). \qquad (18)$$

In our current implementation, we use the following features to capture human navigation behavior:

1) Time: The incentive of a pedestrian typically is to reach a certain target position as fast as possible [18]. We therefore introduce a feature

$$f_{\text{travel time}}^a = t_{\text{travel time}}^a \qquad (19)$$

that is given by the time of travel to the target position. Note that to compute this feature, we need to estimate the target positions of all agents considered.

2) Acceleration: Pedestrians typically aim to walk efficiently, avoiding unnecessary accelerations [12, 18]. Integrating the squared acceleration over the trajectory yields the feature

$$f_{\text{acceleration}}^a = \int_t \|\ddot{x}^a(t)\|^2 \ dt. \qquad (20)$$

3) Velocity: Pedestrians have a desired walking velocity that is uncomfortable to exceed [11]. We therefore use the feature

$$f_{\text{velocity}}^a = \int_t \|\dot{x}^a(t)\|^2 \ dt. \qquad (21)$$

4) Collision Avoidance: A further typical objective of human navigation is to evade other pedestrians. We assume that evasive maneuvers depend on the distance between the agents and their velocities:

$$f_{\text{distance}}^a = \sum_{b \neq a} \int_t \frac{v_t^a}{\|x^a(t) - x^b(t)\|^2} \ dt, \qquad (22)$$

where the pairwise components are summed up for all agents $b \neq a$. A similar feature could account for static obstacles. Note that these features give rise to the modes of the probability distribution corresponding to the topological variants.

5) Topological Variants: Our model also allows for using high-level features that explicitly account for the decisions of the agents which sides they choose when passing each other to avoid collisions. For instance, a feature may count the frequency of passing left versus passing right to account for cultural differences. A feature that captures groups of people may prevent the robot from moving between associated people, if this was observed in human behavior.

C. Computing the Maximum Likelihood Trajectories

As mentioned above, we approximate the distribution of joint trajectories using a weighted sum of Dirac delta functions at the modes of the distribution that correspond to the topological
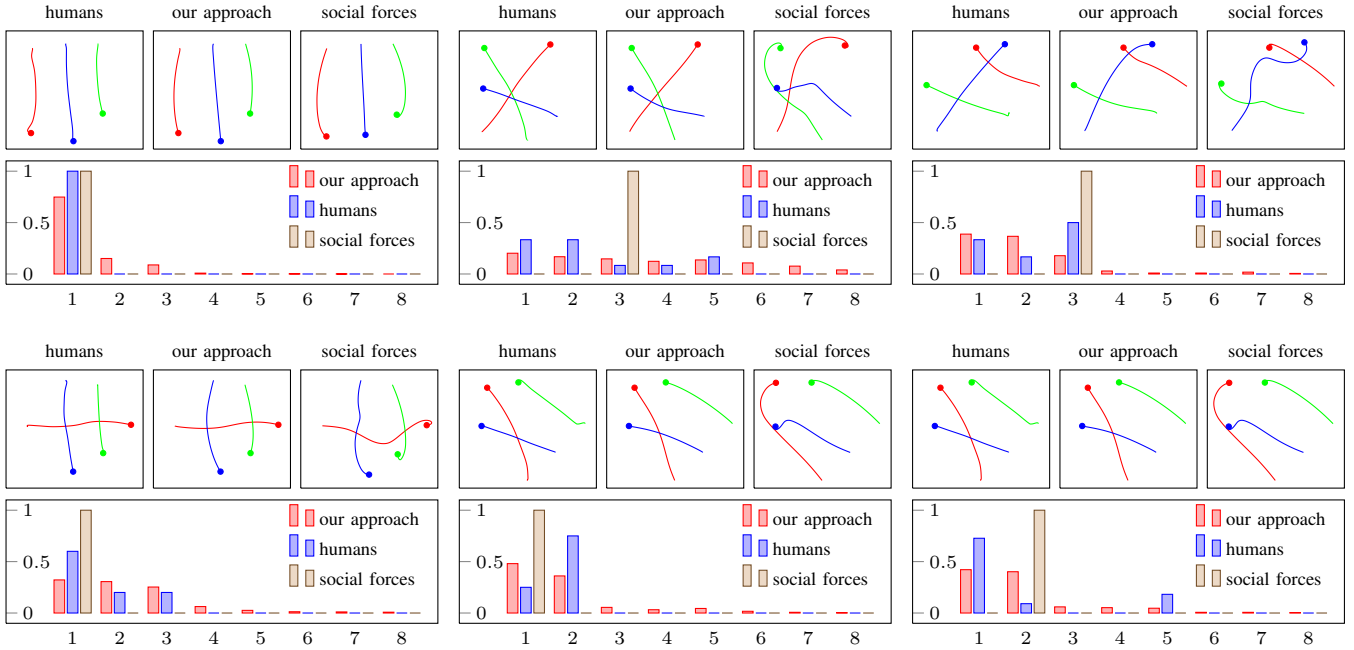
Fig. 3. Comparison of recorded human behavior and the corresponding predictions of our approach and the social forces method in six scenarios. For each scenario, the plots show representative trajectories, where the agents' target positions are indicated by dots. The bar plots visualize the probability of the eight possible topological variants ordered by the probability estimated by our approach.

variants. We generate one representative of each $\psi \in \Psi$ by listing all combinations of sides the agents pass each other in a pairwise fashion, i.e., left or right. To compute the maximum likelihood joint trajectory for each of the topological variants, we apply gradient-based optimization techniques. We therefore computed the partial derivatives of $p_{\boldsymbol{\theta}}(\mathbf{x})$ with respect to x_j^a and v_j^a introduced in Sec. IV-A and evaluated several gradient-based optimization techniques, namely gradient descent, the Levenberg-Marquardt algorithm [17], and RPROP [23]. In empirical experiments, we found that RPROP performed best. To further improve convergence speed, we apply a hierarchical optimization scheme that first runs on a coarse discretization of the trajectories based on just a few spline control points, which quickly computes an approximate shape of the maximum likelihood joint trajectory. Our algorithm then subsequently refines the estimate using increasingly finer-grained discretizations in time. During optimization, we must take care to not violate the constraint $\mathbf{x} \in \Psi$ given in Eq. (15). In other words, we must not change $\mathbf{x}$ such that it belongs to a different topological variant. We therefore check in each step of the optimization whether the integral over the derivative of the angle of the vectors between all pairs of agents changes the sign. Hence, we have

$$\kappa(a, b) = \int_t \frac{d}{dt} \alpha_a^b(t) \ dt, \qquad (23)$$

where α_a^b refers to the angle between the vector $x^b - x^a$ and the vector $(1, 0)^T$. We reject optimization steps that violate the constraint $\mathbf{x} \in \psi$, as indicated by κ, and continue the optimization with a decreased step size.

V. NAVIGATION IN POPULATED ENVIRONMENTS

A socially compliant robot that navigates in a populated environment has to reason about future paths the humans are likely to follow. The humans, however, will react to the actions of the robot themselves, thus the robot has to adopt its behavior, which in turn affects the humans. To break up this infinite loop, our approach reasons about the joint trajectories that are likely to be followed by all the agents, including the robot itself. To do so, the robot incorporates the current poses of the humans and itself into the prediction process and implicitly "predicts" its own trajectory, which it then follows. Our approach enables the robot to efficiently evaluate multiple topological variants of the trajectories online. Hence, it allows the robot to choose the most appropriate one given the current situation.

However, humans are likely to react differently to robots than to other humans, depending on the characteristics of the robot. To solve this problem, we could record large data sets of people interacting with a an already installed robot. Alternatively, the robot could be tele-operated or joysticked by a human with the goal to teach it an appropriate navigation behavior.

VI. EXPERIMENTAL EVALUATION

Our experimental evaluation demonstrates that our approach accurately predicts human navigation behavior, generalizing to different scenarios. We conduct an experiment with a real robot that uses our approach for socially-compliant robot navigation. We furthermore present a comparison that suggests that our approach outperforms the popular social forces method.

A. Acquisition of Real-World Training Data

To fit our motion model to human navigation behavior, we recorded trajectories of human subjects using a motion capture system (see Fig. 2). To prevent the subjects from focussing on their walking behavior, which may result in awkward trajectories, we made them read and memorize newspaper articles at various locations saying that we were conducting psychological research on the effects on attention when frequently changing locations.

We set up eight differently numbered locations in an area of $7\,\mathrm{m} \times 7\,\mathrm{m}$, where each location was equipped with a different newspaper article. We repeatedly asked the three subjects to read the articles and simultaneously change their locations such that their paths crossed each other, which forced the subjects to avoid each other. Our setup gave rise to eight different scenarios that repeatedly occurred after each cycle. In total, we recorded one hour of human interaction, resulting in 96 runs, i.e., 12 runs per scenario. Fig. 3 shows the trajectories recorded during the first occurrence of the scenarios. We then applied the learning algorithm described in Sec. III to estimate the feature weights that best capture the navigation behavior of the human subjects.

B. Cross Validation on Real-World Data and Comparison to the Social Forces Method

We performed a 10-fold cross validation on the recorded data set described in the previous section to evaluate how well our approach generalizes to different situations and how it compares to the social forces method [10].

To allow for a fair comparison, we trained and evaluated both models given the same data sets in each fold of the cross validation. We applied stochastic gradient descent to compute the parameters of the social forces model that minimize the norm of the distance between the resulting feature values and the feature values of the behavior that was to be imitated.

We analyzed the norm of the discrepancy between the behavior predicted by the two methods and the recorded human behavior averaged over all ten folds and obtained $\|\mathbf{f}_{\mathcal{D}} - \mathbf{f}_{\boldsymbol{\theta}}\|_2 = 1.7$, and $\|\mathbf{f}_{\mathcal{D}} - \mathbf{f}_{\text{social forces}}\|_2 = 7.2$. This indicates that our approach has a substantially better capability to capture and reproduce human behavior given the features we suggest.

Fig. 3 provides insight into the results, which suggest that the trajectories predicted by our approach are substantially more human-like than the predictions of the social forces method. For instance, the jerky evasive movements predicted by the social forces method do not replicate well the recorded behavior of the humans. These movements can be attributed to the fact that the social forces method is reactive. In contrast to that, our approach also reasons about future interactions of the agents and is thus able to accurately predict the human behavior.

Our model explicitly accounts for the topological variants the humans are likely to choose. The bar plots in Fig. 3 suggest that our approach is able to predict a probability distribution over the outcomes of these decisions. In contrast to that, the social forces method only yields a single estimate.
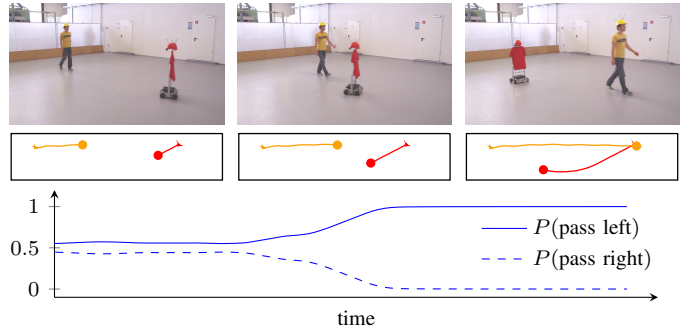

Fig. 4. An example encounter of a robot (red) and a human (yellow) where the robot can pursue its initial intent to pass on the left side.
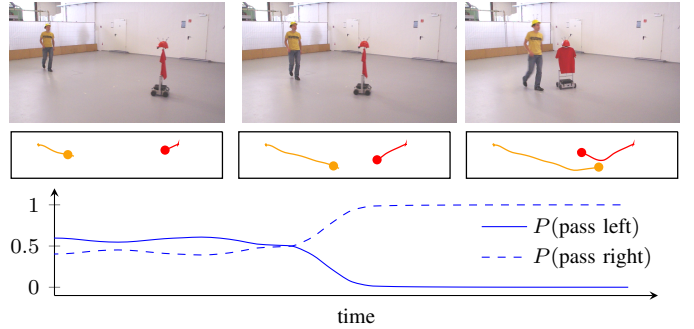

Fig. 5. The robot changes its initial intent to pass on the left side because the human's actions differ from the robot's initial predictions.

C. Real-World Robot Experiments

We implemented our approach on a mobile robot and carried out experiments in which a human and a robot pass each other while moving to their target positions. We tracked the poses of the robot and the humans using a visual motion capture system. Fig. 4 and Fig. 5 visualize the recorded trajectories of the robot and the humans. The bottom plot of each figure visualizes how the estimated probability distribution of the topological variants evolves during the encounters.

In the first image of Fig. 4, the belief of the robot reflects the symmetry of the situation. Both topological variants, i.e., passing left or passing right, seem more or less equally likely to the robot in terms of its probability distribution estimate over joint trajectories. Nevertheless, a slight tendency towards the left side makes the robot commit to this topological variant. As a consequence, the robot expects the human to evade to the opposite side. The human, however, refuses to evade to any side but instead walks on a straight line towards his goal. The robot adapts its expectations to the new situation but still decides to pass on the left side. After the situation has been resolved, the robot directly proceeds to its target position.

Fig. 5 illustrates a similar initial situation, although the human now behaves differently from the previous example as the human insists on passing on the right side regardless of the robot's visible intent to pass left. The robot thus re-evaluates the probability distribution over the topological variants and decides to pass on the other side.
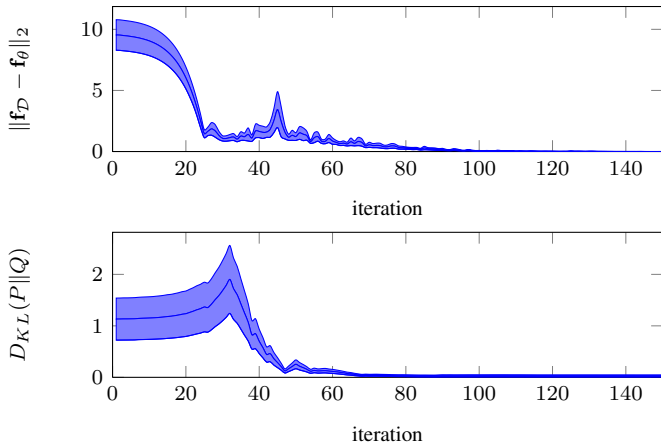
Fig. 7. Computing the maximum likelihood trajectories given the feature weights $\boldsymbol{\theta}$. The plot shows the evolution of the mean and the standard deviation of the distance between the current feature values and the resulting feature values after convergence. Our approach carries out each iteration to optimize 60 spline control points in approximately 1 ms on a standard desktop computer.

Fig. 6. Learning the feature weights averaged over 10 randomly generated data sets. Top: The evolution of the norm of the difference between the feature values of the behavior that is to be imitated and the expected feature values of our approach. Bottom: The evolution of the Kullback-Leibler divergence between the true probability distribution P of the topological variants and the probability distribution Q estimated by our approach. The plots illustrate that both error measures converge to zero.

D. Simulation Experiments

In addition to the experiments based on real-world data, we present a series of simulation experiments that aims to evaluate the performance of our learning algorithm based on synthetic data that adhere to our assumptions, in particular, the approximation expressed in Eq. (14). We therefore randomly generated a set of 100 scenarios each with three agents by uniformly sampling the start and goal positions of the agents in an area of $10\,\mathrm{m} \times 10\,\mathrm{m}$. We randomly set the feature weights $\boldsymbol{\theta}$ that subsequently gave rise to the behavior that is to be imitated. In the following, we demonstrate that our method was able to accurately recover the generating model.

Fig. 6 evaluates our approach during learning when the exact feature expectations of the behavior to be imitated are known. We express the performance in terms of the norm of the difference between the feature values of the observed behavior and the expected feature values of our approach. The figure furthermore visualizes the Kullback-Leibler divergence between the estimated and the true probability distribution of the topological variants. Both errors converge to zero, which suggests that our approach is able to exactly reproduce the behavior in terms of the features and the decisions of which sides the agents choose to pass each other.

We conducted another simulation experiment to evaluate the performance of our approach for situations in which only few observations of human behavior are available. We therefore sampled a set $\mathcal{D}$ of n joint trajectories from the synthetic data set described above, where $n \in \{10, 25, 50, 75, 100\}$. To generate the training samples, we first uniformly sampled one of the scenarios and then sampled a topological variant from the corresponding probability distribution. We then applied our approach to recover the feature weights $\boldsymbol{\theta}$ based on the sampled observations $\mathcal{D}$ and their empirical feature values $\mathbf{f}_{\mathcal{D}}$. We performed a 10-fold cross validation, which, averaged over
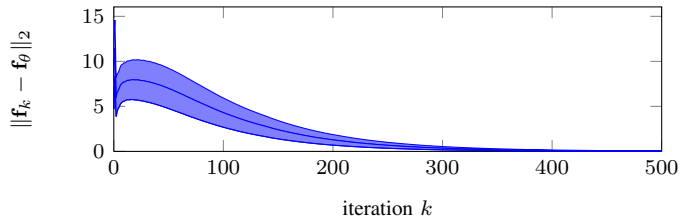
all folds, resulted in the following values: $\|\mathbf{f}_{\mathcal{D}_{10}} - \mathbf{f}_{\boldsymbol{\theta}}\|_2 \approx 0.64$, $\|\mathbf{f}_{\mathcal{D}_{25}} - \mathbf{f}_{\boldsymbol{\theta}}\|_2 \approx 0.44$, $\|\mathbf{f}_{\mathcal{D}_{50}} - \mathbf{f}_{\boldsymbol{\theta}}\|_2 \approx 0.33$, $\|\mathbf{f}_{\mathcal{D}_{75}} - \mathbf{f}_{\boldsymbol{\theta}}\|_2 \approx 0.25$, and $\|\mathbf{f}_{\mathcal{D}_{100}} - \mathbf{f}_{\boldsymbol{\theta}}\|_2 \approx 0.21$.

E. Computing the Maximum Likelihood Joint Trajectories Corresponding to the Topological Variants

We require our gradient-based optimization approach described in Sec. IV-C to converge to a unique maximum within each topological variant. To evaluate the convergence of our approach, we computed the maximum likelihood joint trajectories of the real-world data set given a set of randomly sampled initial guesses. For each topological variant, our algorithm found a unique solution irrespective of the initial guess with a maximal deviation of less than $2\,\mathrm{cm}$. This suggests that our algorithm indeed finds the modes of the probability distribution corresponding to the topological variants.

F. Runtime

To use our approach for autonomous robot navigation, it is necessary to compute the maximum likelihood trajectories in an efficient manner. We evaluated the convergence behavior of our approach and provide results averaged over 100 random scenarios. Fig. 7 visualizes the distance between the current feature values and the feature values of the maximum likelihood trajectory after convergence. On a standard desktop computer, our implementation optimizes several topological variants in parallel and carries out one iteration to optimize 60 spline control points in approximately $1\,\mathrm{ms}$. This suggests that our approach is able to predict human behavior online in situations similar to our experiments. In more densely populated environments, however, it is not feasible to compute all topological variants. We are currently investigating means for efficiently pruning the graph of plausible topological variants.

VII. Conclusion

In this paper, we presented a novel approach to predicting socially compliant trajectories in populated environments. Our approach is based on features that capture relevant aspects of the trajectories. It applies the principle of maximum entropy to elicit the probability distribution that governs the navigation behavior of people from real-world observations. We have implemented and tested our method with extensive data sets and real robots. The experimental results demonstrate the efficacy

of our method and suggest that it leads to natural, socially compliant trajectories. In particular, our algorithm seems to better capture characteristics of human trajectories than the social forces method, a state-of-the-art approach.

REFERENCES

[1] G. Arechavaleta, J.-P. Laumond, H. Hicheur, and A. Berthoz. An optimality principle governing human walking. *IEEE Transactions on Robotics*, 24(1):5–14, 2008.

[2] M. Bennewitz, W. Burgard, G. Czielniak, and S. Thrun. Learning motion patterns of people for compliant robot motion. *Int. Journal of Robotics Research*, 24(1):31–48, 2005.

[3] W. Burgard, A.B. Cremers, D. Fox, D. Hähnel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun. Experiences with an interactive museum tour-guide robot. *Artificial Intelligence*, 114(1-2):3–55, 2000.

[4] P. Fiorini and Z. Shillert. Motion planning in dynamic environments using velocity obstacles. *Int. Journal of Robotics Research*, 17:760–772, 1998.

[5] C. Fox. *An Introduction to the Calculus of Variations*. Courier Dover Publications, 1987.

[6] D. Fox, W. Burgard, and S. Thrun. The dynamic window approach to collision avoidance. *IEEE Robotics Automation Magazine*, 4(1):23–33, 1997.

[7] T. Fraichard and V. Delsart. Navigating dynamic environments with trajectory deformation. *Journal of Computing and Information Technology*, 17(1):27–36, 2009.

[8] E. Goffman. *Relations in Public: Microstudies of the Public Order*. New York: Basic Books, 1971.

[9] S.J. Guy, M.C. Lin, and D. Manocha. Modeling collision avoidance behavior for virtual humans. In *Proc. of the 9th Int. Conference on Autonomous Agents and Multi-Agent Systems*, pages 575–582, 2010.

[10] D. Helbing and A. Johansson. Pedestrian, crowd and evacuation dynamics. In Robert A. Meyers, editor, *Encyclopedia of Complexity and Systems Science*, pages 6476–6495. Springer New York, 2009.

[11] D. Helbing and P. Molnar. Social force model for pedestrian dynamics. *Physical Review E*, 51:4282–4286, 1995.

[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] E. T. Jaynes. Where do we stand on maximum entropy. *Maximum Entropy Formalism*, pages 15–118, 1978.

[14] A. Johansson, D. Helbing, and P. Shukla. Specification of the social force pedestrian model by evolutionary adjustment to video tracking data. *Advances in Complex Systems (ACS)*, 10:271–288, 2007.

[15] E. Kruse and F.M. Wahl. Camera-based monitoring system for mobile robot guidance. In *Proc. of the IEEE/RSJ Int. Conference on Intelligent Robots and Systems (IROS)*, volume 2, pages 1248–1253, 1998.

[16] A. Lerner, Y. Chrysanthou, and D. Lischinski. Crowds by example. *Computer Graphics Forum (Proc. of Eurographics)*, 26(3):655–664, 2007.

[17] D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441, 1963.

[18] K. Mombaur, A. Truong, and J.-P. Laumond. From human to humanoid locomotionan inverse optimal control approach. *Autonomous Robots*, 28:369–383, 2010.

[19] J. Müller, C. Stachniss, K.O. Arras, and W. Burgard. Socially inspired motion planning for mobile robots in populated environments. In *Int. Conference on Cognitive Systems (COGSYS)*, pages 85–90, 2008.

[20] S. Pellegrini, A. Ess, M. Tanaskovic, and L. Van Gool. Wrong turn - no dead end: A stochastic pedestrian motion model. In *Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 15–22, 2010.

[21] R. Philippsen and R. Siegwart. Smooth and efficient obstacle avoidance for a tour guide robot. In *Proc. of the IEEE Int. Conference on Robotics and Automation (ICRA)*, pages 446–451, 2003.

[22] S. Quinlan and O. Khatib. Elastic bands: Connecting path planning and control. In *Proc. of the IEEE Int. Conference on Robotics and Automation (ICRA)*, pages 802–807, 1993.

[23] M. Riedmiller and H. Braun. A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In *Proc. of the IEEE Int. Conference on Neural Networks (ICNN)*, pages 586–591, 1993.

[24] S. Thrun, M. Beetz, M. Bennewitz, W. Burgard, A.B. Cremers, F. Dellaert, D. Fox, D. Haehnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz. Probabilistic algorithms and the interactive museum tour-guide robot Minerva. *Int. Journal of Robotics Research*, 19(11):972–999, 2000.

[25] P. Trautman and A. Krause. Unfreezing the robot: Navigation in dense, interacting crowds. In *Proc. of the IEEE/RSJ Int. Conference on Intelligent Robots and Systems (IROS)*, pages 797–803, 2010.

[26] J. van den Berg, S.J. Guy, M.C. Lin, and D. Manocha. Reciprocal n-body collision avoidance. In *Int. Symposium on Robotics Research (IJRR)*, pages 3–19, 2009.

[27] B.D. Ziebart, A. Maas, J.A. Bagnell, and A.K. Dey. Maximum entropy inverse reinforcement learning. In *Proc. of the National Conference on Artificial Intelligence (AAAI)*, pages 1433–1438, 2008.

[28] 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 *Proc. of the IEEE/RSJ Int. Conference on Intelligent Robots and Systems (IROS)*, pages 3931–3936, 2009.