2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

Learning to Predict Trajectories of Cooperatively Navigating Agents

Henrik Kretzschmar Markus Kuderer Wolfram Burgard

***Abstract*— The problem of modeling the navigation behavior of multiple interacting agents arises in different areas including robotics, computer graphics, and behavioral science. In this paper, we present an approach to learn the composite navigation behavior of interacting agents from demonstrations. The decision process that ultimately leads to the observed continuous trajectories of the agents often also comprises discrete decisions, which partition the space of composite trajectories into homotopy classes. Therefore, our method uses a mixture probability distribution that consists of a discrete distribution over the homotopy classes and continuous distributions over the composite trajectories for each homotopy class. Our approach learns the model parameters of this distribution that match, in expectation, the observed behavior in terms of user-defined features. To compute the feature expectations over the high-dimensional continuous distributions, we use Hamiltonian Markov chain Monte Carlo sampling. We exploit that the distributions are highly structured due to physical constraints and guide the sampling process to regions of high probability. We apply our approach to learning the behavior of pedestrians and demonstrate that it outperforms state-of-the-art methods.**

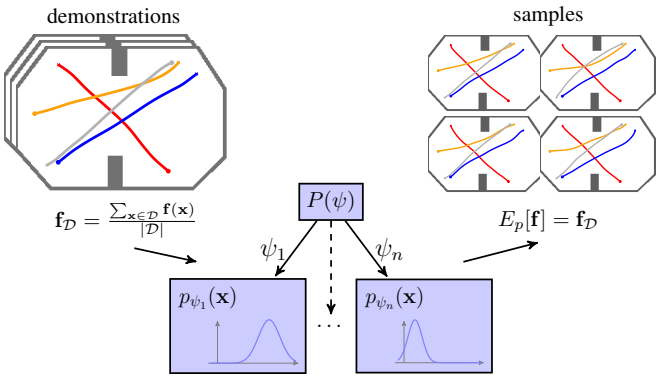


Fig. 1. Our method learns a model of the behavior of cooperatively navigating agents from demonstrations. We learn the model parameters of a mixture distribution over composite trajectories to capture the discrete and continuous aspects of the agents' behavior. The learned model generalizes to new situations and allows us to draw samples that capture the stochasticity of natural navigation behavior.

I. Introduction

Accurate models of the navigation behavior of interacting agents are a prerequisite for a variety of applications in diverse fields including robotics, computer graphics, and behavioral science. For instance, mobile robots that have models of pedestrian navigation behavior can predict the trajectories of nearby pedestrians and thus can navigate among pedestrians in a safe, efficient, and socially compliant way [25, 21], or, alternatively, they can mimic the behavior of pedestrians [13]. In computer graphics, Guy et al. [6] proposed such models for generating realistic simulations of pedestrians and traffic.

We propose an approach to learn a probabilistic model of the navigation behavior of cooperatively navigating agents from observations of their trajectories. The decision process that ultimately leads to the observed trajectories in continuous state spaces often comprises discrete decisions such as whether to pass other agents on the left or on the right side. Furthermore, the navigation behavior of agents such as humans and animals is not a deterministic process. They rather exhibit stochastic properties, i. e., their trajectories vary from run to run when they repeatedly navigate in the same situation. Our approach therefore learns a model that captures both the discrete decision process of the agents as well as the distributions over their continuous trajectories to explain the observed behavior. It reasons about composite trajectories of the agents, i.e., the joint behavior of all the agents, by means of a mixture distribution. As illustrated in Fig. 1, this distribution comprises a discrete distribution that captures the discrete choices in the navigation process of the agents. The distribution furthermore consists of continuous distributions over trajectories that capture the variance of the observed trajectories. We propose a feature-based maximum entropy learning approach to infer the distribution that matches the observed behavior of the agents in expectation.

A key challenge of such a learning approach is the so-called forward problem, i.e., computing for a given model the expected feature values with respect to the distribution over the high-dimensional space of continuous trajectories. We propose to use Markov chain Monte Carlo (MCMC) sampling and exploit that the distributions over observed trajectories of interacting agents are highly structured. The use of a spline-based representation of the trajectories makes it possible to efficiently compute the gradient of the probability density, which allows our method to guide the sampling process to regions of high probability using the Hybrid Monte Carlo (HMC) [5] algorithm. Therefore, our method is able to capture the stochasticity of observed trajectories, which is in contrast to existing approaches that learn deterministic models that do not replicate well the stochastic behavior of natural agents such as humans and animals.

We applied our method to learning a model of the navigation behavior of pedestrians. Our experimental evaluation suggests that our approach outperforms three state-of-the-art methods. Furthermore, a Turing test demonstrates that our method generates trajectories that appear more human-like than those created by previous methods.

All authors are with the Department of Computer Science, University of Freiburg, Germany. This work has partly been supported by the German Research Foundation (DFG) under contract number SFB/TR-8, and by the EC under contract number FP7-610603.

II. Related Work

Atkeson and Schaal [3] developed one of the pioneering approaches to infer a mapping from state features to actions to directly mimic observed behavior. More recently, Ng and Russel applied inverse reinforcement learning (IRL) to recover a cost function that explains observed behavior [16]. Later, Ratliff et al. [19], Ziebart et al. [25], Kitani et al. [12] used these techniques to address a variety of problems including route planning for outdoor mobile robots and learning pedestrian navigation behavior. In particular, Abbeel and Ng [1] suggest to match features that capture relevant aspects of the behavior that is to be imitated. However, feature matching does not lead to a unique cost function. To resolve this ambiguity, Maximum Entropy IRL [24] relies on the principle of maximum entropy [10] and, hence, aims at finding the policy with the highest entropy subject to feature matching. These learning methods work well in discrete state-action spaces of low dimensionality. However, they do not scale well to continuous trajectories, especially when taking into account higher-order dynamics such as velocities and accelerations. Inspired by Abbeel and Ng [1] and Ziebart et al. [24], our approach aims to find maximum entropy distributions that match the observed feature values. However, in contrast to the abovementioned techniques, our approach reasons about continuous trajectories including their higher-order dynamics.

To learn a model from observed trajectories, several authors assume utility-optimizing agents that try to minimize a parametrized cost function that captures relevant properties of their continuous trajectories. For instance, whereas Hoogendoorn and Bovy [9] penalize accelerations, drift from planned trajectories, and closeness to other agents, Arechavaleta et al. [2] minimize the derivative of the curvature to model pedestrian navigation behavior. In contrast to these approaches, we do not assume utility optimizing agents. Instead, we model the navigation behavior using a probability distribution over the trajectories since the observed trajectories are rarely optimal but rather exhibit stochastic variations.

Our previous work [13, 14] models cooperative behavior by jointly predicting the trajectories of all interacting agents. The approach to reciprocal collision avoidance presented by van den Berg et al. [22] ensures that a set of agents navigates without collisions. Helbing and Molnar [8] learn the parameters of a potential field to model interaction behavior from observations. Trautman and Krause [21] model the coupled behavior of agents using interactive Gaussian processes. Lerner et al. [15] present a data-driven approach to modeling the behavior of crowds, where agents behave according to examples that are stored in a database of previously recorded crowd movements.

In general, estimating the feature expectations with respect to a high-dimensional probability distribution over continuous trajectories is a challenging problem. Our previous approach [13] approximates the feature expectations using Dirac delta functions at the modes of the distribution. This approximation, however, leads to suboptimal models when learning from imperfect human navigation behavior since its stochasticity is not sufficiently captured. In practice, this method under-estimates the feature values and thus favors samples from highly unlikely homotopy classes. Vernaza and Bagnell [23] constrain the features to have a certain low-dimensional structure. Kalakrishnan et al. [11] assume the demonstrations to be locally optimal and sample continuous trajectories by adding Gaussian noise to the model parameters. In this paper, we estimate the feature expectations using Hybrid Monte Carlo sampling [5], which allows for arbitrary differentiable features of the continuous trajectories. Ratliff et al. [18] use Hybrid Monte Carlo ideas for globally robust trajectory optimization. In summary, our method is capable of accurately modeling natural and potentially non-optimal real-world behavior.

III. Learning to Predict Composite Trajectories of Interacting Agents from Demonstrations

The objective of our approach is to learn a model of the navigation behavior of interacting agents from samples of their trajectories. To this end, we consider the product space of the trajectories of all the agents $a_i \in A$, i. e., the space of composite trajectories $\mathbf{x} = (x^{a_1}(t), \ldots, x^{a_n}(t))$, where $x^{a_i}(t)$ is the trajectory of agent a_i that continuously maps time t to the continuous configuration of agent $a_i \in A$ at time t. We use cubic splines in $\mathbb{R}^2$ to represent the x and y positions of the agents. We refer to the space of all composite trajectories as $\mathcal{X}$. We assume that the behavior of the agents can be described in terms of a joint probability distribution $p(\mathbf{x})$ that depends on a feature vector $\mathbf{f}$ comprising features that capture relevant properties of the composite trajectories.

A. Maximum Entropy and Feature Matching

To learn observed behavior, we aim to model the distribution that underlies the empirical sample trajectories. Following Abbeel and Ng [1], we aim to find a model that induces distributions that match, in expectation, the feature values $\mathbf{f}_\mathcal{D}$ of the empirical trajectories $\mathcal{D}$, yielding

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

In general, however, there is not a unique distribution that matches the features. Ziebart et al. [24] resolve this ambiguity by applying the principle of maximum entropy [10], which states that the distribution with the highest entropy represents the given information best since it does not favor any particular outcome besides the observed constraints. The resulting distribution and the gradient of its parameters are well known in information theory and have been used to infer trajectories in discrete [25, 12] and continuous [13, 23] spaces. In short, our goal is to find the distribution $p^\star$ that maximizes the differential entropy

$$H(p) = \int_\mathbf{x} -p(\mathbf{x}) \log p(\mathbf{x}) d\mathbf{x} \tag{2}$$

subject to feature matching. Applying constrained optimization maximizes the Lagrangian

$$H(p) - \sum_i \theta_i (\mathbb{E}_p[f_i] - f_{i,\mathcal{D}}) - \alpha(\int_x p(\mathbf{x})\, d\mathbf{x} - 1) \tag{3}$$

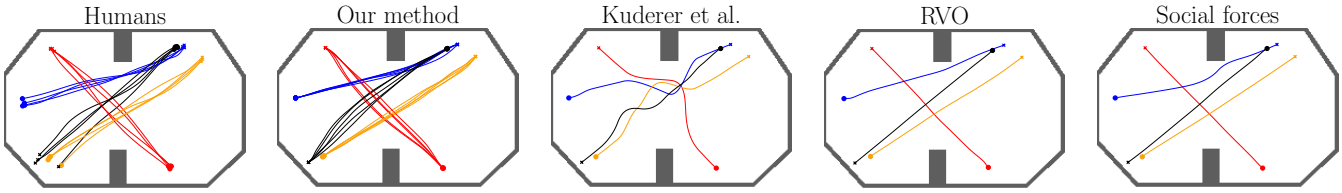


Fig. 2. Trajectories of four pedestrians predicted by four different methods in the same situation. Humans: Observed trajectories recorded in the test environment shown in Fig. 3. Our method: Samples drawn from the policy learned by our method replicate the stochasticity of the observed trajectories. Kuderer et al. [13]: The Dirac approximation favors samples from highly unlikely homotopy classes. RVO and social forces: Deterministic predictions.

with respect to the distribution p and the Lagrangian multipliers $[\theta_1, \ldots, \theta_n] = \boldsymbol{\theta}$ and α, where the last term ensures that the probability integrates to one. Taking derivatives with respect to p reveals that the desired probability distribution belongs to the exponential family $p_{\boldsymbol{\theta}}(x) \propto e^{-\boldsymbol{\theta}^T \mathbf{f}(\mathbf{x})}$.

One can interpret the term $\boldsymbol{\theta}^T \mathbf{f}(\mathbf{x})$ as a cost function that comprises the feature weights $\boldsymbol{\theta}$, where trajectories with higher cost are less likely. Computing the parameters $\boldsymbol{\theta}^\star$ that maximize the Lagrangian is analytically not feasible. However, we can use the gradient of the dual function with respect to $\boldsymbol{\theta}$, which is given by $\mathbf{f}_{\mathcal{D}} - \mathbb{E}_p[\mathbf{f}(\mathbf{x})]$, to apply gradient-based optimization. The resulting distribution $p^\star$ equals the exponential-family distribution that maximizes the likelihood of the empirical data $\mathcal{D}$. The derivation of the discrete maximum entropy distribution follows the same idea, where integrals are substituted by sums.

B. Modeling the Navigation Behavior of Interacting Agents Using Mixture Distributions

The decision process of interacting agents that ultimately leads to their trajectories in continuous spaces often comprises continuous and discrete decisions. The continuous decisions affect physical properties of the composite trajectories such as velocities, accelerations, and clearances when passing obstacles and other agents. The discrete decisions made by the agents, such as choices to pass obstacles and other agents on the left or on the right side, determine the homotopy class of the resulting composite trajectories.

Our approach aims to learn a model of the behavior that captures both the continuous and the discrete decisions of the agents. To this end, our method models the behavior using a mixture distribution that comprises a discrete probability distribution $P(\psi)$ over the homotopy classes $\psi \in \Psi$ of the composite trajectories, and continuous distributions $p_\psi(\mathbf{x})$ over the composite trajectories $\mathbf{x}$ of the corresponding homotopy classes ψ. Note that the number of homotopy classes increases exponentially with the number of agents. However, techniques that only consider relevant homotopy classes [14] scale to many agents. We assume that the discrete mixture proportions $P(\psi)$ and the component densities $p_\psi(\mathbf{x})$ depend on features $f^{\text{hom}} : \Psi \mapsto \mathbb{R}$ and $f^{\text{phys}} : \mathcal{X} \mapsto \mathbb{R}$, respectively, that capture relevant properties of the navigation behavior. According to the previous section, to learn the navigation behavior, we aim to find a model that induces distributions over the continuous composite trajectories as well as over homotopy classes that match the feature expectations of the observed composite trajectories $\mathcal{D}$, which leads to

$$\mathbb{E}_{p_\psi}[\mathbf{f}^{\text{phys}}] = \mathbf{f}^{\text{phys}}_{\mathcal{D}} = \sum_{\mathbf{x} \in \mathcal{D}} \frac{\mathbf{f}^{\text{phys}}(\mathbf{x})}{|\mathcal{D}|}, \tag{4}$$

$$\mathbb{E}_P[\mathbf{f}^{\text{hom}}] = \mathbf{f}^{\text{hom}}_{\mathcal{D}} = \sum_{\mathbf{x} \in \mathcal{D}} \frac{\mathbf{f}^{\text{hom}}(\psi_{\mathbf{x}})}{|\mathcal{D}|}, \tag{5}$$

where $\psi_{\mathbf{x}}$ is the homotopy class of $\mathbf{x}$. Applying the principle of maximum entropy subject to feature matching leads to $p_\psi(\mathbf{x}) \propto e^{-\boldsymbol{\theta}^{\text{phys}} \mathbf{f}^{\text{phys}}(\mathbf{x})}$ and $P(\psi) \propto e^{-\boldsymbol{\theta}^{\text{hom}} \mathbf{f}^{\text{hom}}(\psi)}$.

To capture the homotopy class of a composite trajectory $\mathbf{x}$, we integrate the derivative of the angle $\alpha^b_a(t)$ between the vector $x^b(t) - x^a(t)$ and the vector $(1, 0)^T$ over time for all agents a and b, which leads to $\kappa^b_a = \int_t \dot{\alpha}^b_a(t)\, dt$. This function effectively accounts for the rotation of the vectors $x^b(t) - x^a(t)$, which is an invariant of all the composite trajectories that belong to the same class.

The following features capture physical properties of the navigation behavior of interacting agents in terms of an intent to reach the target positions energy efficiently, taking into account velocities and clearances when avoiding obstacles and other agents:

$$f^{\text{phys}}_{\text{time}}(\mathbf{x}) = \sum_{a \in A} \int_t 1 \; dt, \tag{6}$$

$$f^{\text{phys}}_{\text{acceleration}}(\mathbf{x}) = \sum_{a \in A} \int_t \|\ddot{x}^a(t)\|^2 \; dt, \tag{7}$$

$$f^{\text{phys}}_{\text{velocity}}(\mathbf{x}) = \sum_{a \in A} \int_t \|\dot{x}^a(t)\|^2 \; dt, \tag{8}$$

$$f^{\text{phys}}_{\text{obstacle}}(\mathbf{x}) = \sum_{a \in A} \int_t \frac{1}{\|x^a(t) - o^a_{\text{closest}}(t)\|^2} \; dt, \tag{9}$$

$$f^{\text{phys}}_{\text{distance}}(\mathbf{x}) = \sum_{a \in A} \sum_{b \neq a} \int_t \frac{1}{\|x^a(t) - x^b(t)\|^2} \; dt, \tag{10}$$

where o^a_{closest} is the position of the closest obstacle to agent a at time t. Our approach aims to find the continuous distributions $p_{\boldsymbol{\theta}}(\mathbf{x})$ over the composite trajectories that, in expectation, match these features. The feature

$$f^{\text{hom}}_{\text{angle}}(\psi) = \sum_{a \in A} \sum_{b \neq a} \kappa^b_a. \tag{11}$$

captures decisions to avoid other agents on the left or on the right. Similarly, in case we are able to recognize group memberships of agents, the following feature indicates if an agent moves through such a group. An agent that passes two

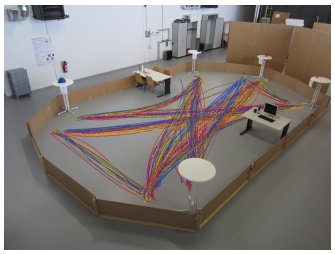


Fig. 3. Trajectories observed during one hour of interactions of four persons in our test environment.

members of a group on different sides moves through the corresponding group. Therefore, we have

$$f^{\text{hom}}_{\text{group}}(\psi) = \sum_{a \in A} |\{G \in \mathcal{G} \mid \exists b, c \in G : b, c \neq a \wedge \kappa^b_a \kappa^c_a < 0\}|,$$

where $\mathcal{G}$ is the set of groups of agents, which allows our approach to learn to which extent the agents avoid to move through groups. Furthermore, we allow the features $\mathbf{f}^{\text{hom}}$ to depend on the distribution over composite trajectories of the corresponding homotopy class. For example, the feature

$$f^{\text{hom}}_{\text{ml_cost}}(\psi) = \min_{\mathbf{x} \in \psi} \boldsymbol{\theta}^T \mathbf{f}(\mathbf{x}) \tag{12}$$

captures the cost of the most likely composite trajectory $\mathbf{x}$ of homotopy class ψ, which allows the model to reason about the homotopy class the agents choose in terms of the cost of the composite trajectory that is most likely according to the learned distribution $p_\psi(\mathbf{x})$. Based on the results of our previous experiments [13], we assume that we can compute this most likely composite trajectory using gradient based optimization techniques. We furthermore found the optimization algorithm RPROP [20] to perform best. After having learned the physical behavior in terms of the distributions $p_\psi(\mathbf{x})$, we can evaluate the features f^{hom} to learn the discrete aspects of the behavior.

C. Computing Feature Expectations with Respect to Distributions over Continuous Trajectories of High Dimensionality

In general, inference about distributions over continuous trajectories requires a finite-dimensional representation of the trajectories. However, even computing the feature expectations with respect to a finite-dimensional representation is not analytically tractable in general. Monte Carlo sampling methods yet provide means to approximate the expectations using a set of sample trajectories drawn from the distribution. In particular, Markov chain Monte Carlo (MCMC) methods allow to obtain samples from high-dimensional distributions. These methods aim to explore the state space by constructing a Markov chain whose equilibrium distribution is the target distribution. Most notably, the widely-used Metropolis-Hastings algorithm [7] generates a Markov chain in the state space using a proposal distribution and a criterion to accept or reject the proposed steps. This proposal distribution and the resulting acceptance rate, however, have a dramatic effect on the mixing time of the algorithm, e. g., the number of steps after which the distribution of the samples can be considered to be close to the target distribution. In general, it is difficult to design a proposal distribution that leads to satisfactory mixing. As a result, efficient sampling from complex high-dimensional distributions is often not tractable in practice.

We exploit the structure of the distributions over composite trajectories to enable efficient sampling. First, the navigation behavior of physical agents shapes the trajectories according to certain properties such as smoothness and goal-directed navigation, which are captured by the features. As a result, the distributions over the composite trajectories of the same homotopy class are highly peaked. Exploiting the gradient of the probability densities allows us to guide the sampling process towards these regions of high probability. To this end, assuming that the physical features are differentiable with respect to the parameterization of the trajectories allows us to compute the gradient of the density $p_\psi(\mathbf{x})$. In particular, we use the Hybrid Monte Carlo algorithm [5], which takes into account the gradient of the density to sample from the distributions $p_\psi(\mathbf{x})$. The algorithm considers an extended target density $p_\psi(\mathbf{x}, \mathbf{u})$ to simulate a fictitious physical system, where $\mathbf{u} \in \mathbb{R}^n$ are auxiliary momentum variables. The method constructs a Markov chain by alternating Hamiltonian dynamical updates and updates of the auxiliary variables, utilizing the gradient of the density $p_\psi(\mathbf{x})$ with respect to $\mathbf{x}$. After performing a number of these "frog leaps", Hybrid Monte Carlo relies on the Metropolis-Hastings algorithm to accept the candidate samples $\mathbf{x}^\star$ and $\mathbf{u}^\star$ with probability

$$\min\left(1, \frac{\tilde{p}_\psi(\mathbf{x}^\star) e^{-\frac{1}{2}\mathbf{u}^{\star T}\mathbf{u}^\star}}{\tilde{p}_\psi(\mathbf{x}^{(\tau)}) e^{-\frac{1}{2}\mathbf{u}^{\tau T}\mathbf{u}^\tau}}\right), \tag{13}$$

where the normalizer Z_p in the distribution $p_\psi(\mathbf{x}) = \tilde{p}_\psi(\mathbf{x})/Z_p$ vanishes.

The resulting generative model allows sampling composite trajectories from the mixture distribution by means of ancestral sampling [4], as illustrated in Fig. 1. In practice, we initialize a Markov chain for each homotopy class at its most likely composite trajectory, which we can compute efficiently using gradient-based optimization. Being able to efficiently sample from the continuous distributions $p_\psi(\mathbf{x})$ allows us to compute the feature expectations, which is necessary for learning.

D. Inferring the Target Positions of the Agents

When using our model to predict the trajectories of the agents in new situations, the target positions of the agents might be unknown. Applying Bayes' theorem allows our model to reason about their target positions [24]. After having observed the agents traveling from the composite positions A to B along the composite trajectory $\mathbf{x}_{A\to B}$, the probability that the agents proceed to composite target C is given by

$$\begin{aligned} P_{\boldsymbol{\theta}}(C \mid \mathbf{x}_{A\to B}) &\propto p_\theta(\mathbf{x}_{A\to B} \mid C) P_{\boldsymbol{\theta}}(C) \\ &\propto \frac{e^{-\boldsymbol{\theta}^T \mathbf{f}(\mathbf{x}_{A\to B})} + \int_{\mathbf{x}\in\mathcal{X}_{B\to C}} e^{-\boldsymbol{\theta}^T \mathbf{f}(\mathbf{x})} d\mathbf{x}}{\int_{\mathbf{x}\in\mathcal{X}_{A\to C}} e^{-\boldsymbol{\theta}^T \mathbf{f}(\mathbf{x})} d\mathbf{x}} P_{\boldsymbol{\theta}}(C), \end{aligned} \tag{14}$$

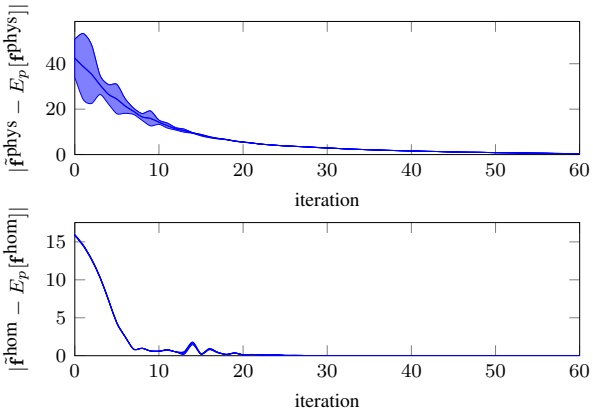
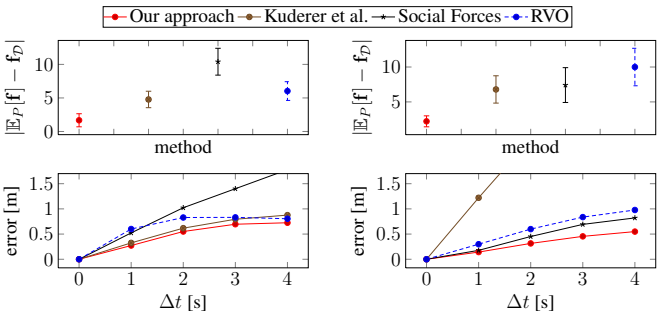


Fig. 4. Evolution of the norm and variance over the five folds of the discrepancy between the feature expectations of the model and the empirical feature values while learning pedestrian navigation behavior. Top: Learning the physical properties of the trajectories. Bottom: Learning the discrete decisions that determine the homotopy classes of the composite trajectories.



(a) Dataset by Pellegrini et al. (b) Motion capture dataset

Fig. 5. Cross validation. The results suggest that our approach better captures pedestrian navigation behavior in terms of the features (top) and the prediction error in meters (bottom) compared to the approach of Kuderer et al. [13], the social forces method [8], and reciprocal velocity obstacles [22]. Left: Results on the dataset provided by Pellegrini et al. [17]. Right: Results on the dataset recorded using our motion capture system.

where $\mathcal{X}_{A\to C}$ and $\mathcal{X}_{B\to C}$ refer to the set of all composite trajectories that lead the agents from A to C, and from B to C, respectively.

IV. Experimental Evaluation

We applied our approach to the problem of learning a model of pedestrian behavior. We considered two datasets of trajectories of interacting pedestrians. The first dataset, depicted in Fig. 3, comprises one hour of interactions of four persons that we recorded using a motion capture system, leading to 81 individual composite trajectories. To distract the persons from the navigation task, we made them read and memorize newspaper articles at different locations that were consecutively numbered. At a signal, they simultaneously walked to the subsequent positions, which repeatedly gave rise to situations where the participants had to evade each other. The second dataset [17] comprises 12 minutes of trajectories of pedestrians interacting in a hotel entrance, leading to 71 composite trajectories with three to five pedestrians each.

A. Cross Validation

We conducted a five-fold cross validation on the aforementioned datasets to evaluate how well the models learned by our approach generalize to new situations. We compared our approach to the approach of Kuderer et al. [13], the social forces algorithm by Helbing and Molnar [8], and the reciprocal velocity obstacles (RVO) introduced by van den Berg et al. [22]. Fig. 2 shows example trajectories of the different methods. We evaluated the evolution of the discrepancy between the feature expectations and the empirical feature values on the training sets while learning the model parameters of our approach. Fig. 4 shows that our method is able to replicate the observed behavior in terms of the features. To allow for a fair comparison, we used the same set of features for all the methods. To optimize the parameters of the social forces method and RVO, we minimized the norm of the discrepancy between the feature values as induced by the methods and the empirical feature values using stochastic gradient descend. We additionally evaluated the parameters provided by Helbing and Molnar [8] and Guy et al. [6], which turned out to not perform better than the learned parameters. For all methods, we assumed that the target positions of the agents were the positions last observed in the datasets. The results of the cross validation, depicted in Fig. 5, suggest that our method is able to capture human behavior more accurately than the other methods in terms of features and in terms of the prediction error in Euclidean distances.

B. Turing Test

When using models of human navigation behavior to simulate agents in the context of computer graphics, it is often important that the generated trajectories appear human-like. We carried out a Turing test to evaluate how human-like the behavior generated by our approach compares to other methods. We asked ten human subjects to distinguish recorded human behavior from behavior generated by one of the algorithms. We evaluated how well the subjects performed on a set of runs that was randomly drawn from recorded human demonstrations. We showed them animations of trajectories that were either recorded from the human demonstrations or from the prediction of one of the algorithms. In particular, we presented 40 runs to each of the human subjects, where the trajectories were equally drawn from the human demonstrations, from the predictions computed by our approach, by the approach of Kuderer et al. [13], and Helbing and Molnar [8]. Fig. 6 summarizes the results. The human subjects correctly identified 79 % of all the human demonstrations, but they mistook 68 % of the predictions of our approach, 40 % of the predictions of the approach of Kuderer et al. [13], and 35 % of the predictions of the social forces algorithm for human behavior. In summary, the results of this Turing test indicate that the behavior induced by our approach is significantly more human-like than the behavior induced by the other two methods according to a one-sided paired sample t-test.

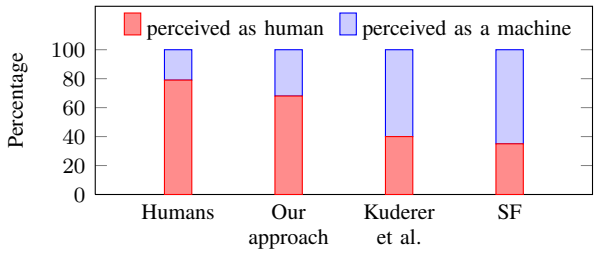


Fig. 6. Turing test to evaluate whether the behaviors induced by our new approach, our previous work [13], and the social forces model by Helbing and Molnar [8] qualify as human. The results suggest that the behavior induced by our approach most resembles human behavior.

C. Robot Navigation

The learned models are applicable to mobile robot navigation in populated environments. To interact with pedestrians, the robot constantly maintains a distribution over the composite trajectories of the pedestrians and itself in the current situation and acts according to the predicted cooperative behavior. This allows the robot to mimic the behavior of the pedestrians and to behave according to their expectations. Our current implementation computes the most likely composite trajectory comprising 3 agents in less than 100 milliseconds, which enables mobile robot navigation in real time.

V. Conclusion

We presented a novel approach to learning a model of the navigation behavior of cooperatively navigating agents such as pedestrians from demonstrations. Based on observations of their continuous trajectories, our approach infers a model of the underlying decision process. To cope with the discrete and continuous aspects of this process, our model uses a joint mixture distribution that captures the discrete decisions regarding the homotopy classes of the composite trajectories as well as continuous properties of the trajectories, such as higher-order dynamics. Our method computes the feature expectations with respect to the continuous high-dimensional probability distributions using Hamiltonian Markov chain Monte Carlo sampling. A Turing test suggests that the pedestrian trajectories induced by our approach are perceived as highly human-like. Furthermore, a cross validation demonstrates that our method generalizes to new situations and outperforms three state-of-the-art techniques.

References

[1] P. Abbeel and A. Ng. Apprenticeship learning via inverse reinforcement learning. In *Proc. of the International Conference on Machine Learning (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] Ch. Atkeson and S. Schaal. Robot learning from demonstration. In *Proc. of the Fourteenth International Conference on Machine Learning (ICML)*, 1997.

[4] C.M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., 2006.

[5] S. Duane, A.D. Kennedy, B.J. Pendleton, and D. Roweth. Hybrid monte carlo. *Physics Letters B*, 195(2):216–222, 1987.

[6] S.J. Guy, J. van den Berg, W. Liu, R. Lau, M.C. Lin, and D. Manocha. A statistical similarity measure for aggregate crowd dynamics. *ACM Transactions on Graphics (TOG)*, 31 (6):190, 2012.

[7] W.K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

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

[9] 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.

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

[11] M. Kalakrishnan, P. Pastor, L. Righetti, and S. Schaal. Learning objective functions for manipulation. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2013.

[12] K.M. Kitani, B.D. Ziebart, D. Bagnell, and M. Hebert. Activity forecasting. In *Proc. of the European Conference on Computer Vision (ECCV)*, 2012.

[13] M. Kuderer, H. Kretzschmar, C. Sprunk, and W. Burgard. Feature-based prediction of trajectories for socially compliant navigation. In *Proc. of Robotics: Science and Systems (RSS)*, 2012.

[14] M. Kuderer, H. Kretzschmar, and W. Burgard. Teaching mobile robots to cooperatively navigate in populated environments. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2013.

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

[16] A. Ng and S. Russell. Algorithms for inverse reinforcement learning. In *Proc. of the International Conference on Machine Learning (ICML)*, 2000.

[17] S. Pellegrini, A. Ess, K. Schindler, and L.J. Van Gool. You'll never walk alone: Modeling social behavior for multi-target tracking. In *Proc. of the IEEE International Conference on Computer Vision (ICCV)*, 2009.

[18] Nathan Ratliff, Matt Zucker, J. Andrew Bagnell, and Siddhartha Srinivasa. Chomp: Gradient optimization techniques for efficient motion planning. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2009.

[19] N.D. Ratliff, J.A. Bagnell, and M.A. Zinkevich. Maximum margin planning. In *Proc. of the International Conference on Machine Learning (ICML)*, 2006.

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

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

[22] J. van den Berg, S.J. Guy, M.C. Lin, and D. Manocha. Reciprocal n-body collision avoidance. In *Proc. of the International Symp. of Robotics Research (ISRR)*, 2009.

[23] P. Vernaza and D. Bagnell. Efficient high dimensional maximum entropy modeling via symmetric partition functions. In *Advances in Neural Information Processing Systems (NIPS)*, volume 25, pages 584–592. 2012.

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

[25] 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 International Conference on Intelligent Robots and Systems (IROS)*, 2009.