

Predicting Actions to Act Predictably: Cooperative Partial Motion Planning with Maximum Entropy Models

Mark Pfeiffer, Ulrich Schwesinger, Hannes Sommer, Enric Galceran and Roland Siegwart

Abstract—This paper reports on a data-driven motion planning approach for interaction-aware, socially-compliant robot navigation among human agents. Autonomous mobile robots navigating in workspaces shared with human agents require motion planning techniques providing seamless integration and smooth navigation in such. Smooth integration in mixed scenarios calls for two abilities of the robot: predicting actions of others and acting predictably for them. The former requirement requests trainable models of agent behaviors in order to accurately forecast their actions in the future, taking into account their reaction on the robot’s decisions. A human-like navigation style of the robot facilitates other agents—most likely not aware of the underlying planning technique applied—to predict the robot motion vice versa, resulting in smoother joint navigation. The approach presented in this paper is based on a feature-based maximum entropy model and is able to guide a robot in an unstructured, real-world environment. The model is trained to predict joint behavior of heterogeneous groups of agents from onboard data of a mobile platform. We evaluate the benefit of interaction-aware motion planning in a realistic public setting with a total distance traveled of over 4 km. Interestingly the motion models learned from human-human interaction did not hold for robot-human interaction, due to the high attention and interest of pedestrians in testing basic braking functionality of the robot.

I. INTRODUCTION

The problem of robotic navigation in static, known environments is nowadays well-studied and understood. Yet navigating in dynamic environments with agents moving in a varied set of patterns is still an open research topic. In contrast to planning in static environments where optimization of, e.g., the path length or travel time is intuitive, the optimization objective in dynamic and unstructured environments is often not well-defined. If the task is to integrate the robot in a smooth fashion into a workspace shared with other agents, optimization of, e.g., travel time will result in aggressive behavior, perceived as unnatural by other individuals. Eventually we wish to realize a robot navigation style which is close (in some sense) to human behavior, commonly referred to as “social compliance” [1]. Otherwise, a navigation behavior perceived as artificial will always stand out and will cause artificial interaction which diminishes the performance of the robot.

Joint pedestrian navigation behavior is influenced by a multitude of factors, such as their preferred avoidance distance of other agents and static obstacles, their preferred velocity, goal-driven behavior and group formation aspects [2].

This work has received funding from the European Union Seventh Framework Programme FP7/2007-2013, Challenge 2, Cognitive Systems, Interaction, Robotics, under grant agreement No. 610603, EUROPA2.

The authors are with the ETH Zurich, Zurich, Switzerland. {pmark, sculrich, sommerh, enricg, rsiegwart}@ethz.ch.

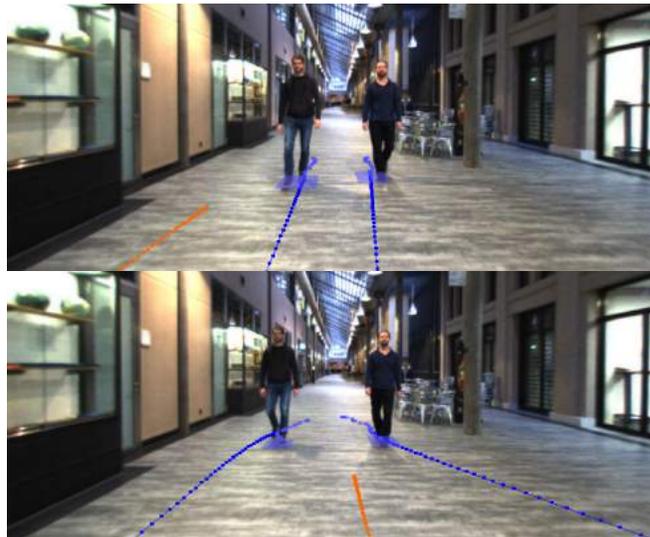


Fig. 1: Predicted trajectories (blue) and resulting plan of robot (orange) with two pedestrians approaching the robot. Top: Neglecting the reaction of the others to the robot’s own future motion through a constant velocity prediction results in over-conservative actions of the robot. Bottom: In our maximum entropy model we build a probability distribution over joint navigation behavior. The cooperative planning approach anticipates the others participation in the collision avoidance and effectively reduces over-conservative behavior.

Perhaps the hardest challenge to achieve social compliance is modeling the reaction of each individual agent on its neighbors. Neglecting these reactions and operating the robot in a two-stage predict-react fashion as depicted in Figure 1 (top) often leads to over-conservative actions of the robot and likewise to behaviors not expected by other agents. Therefore, an effective solution requires a holistic approach where the motion of interacting agents is jointly treated.

The work presented in this paper builds upon the approach developed in [1, 3, 4] in which a maximum entropy distribution over joint navigation behavior is learned from demonstration. The contribution of this paper is the real-world application and evaluation of the approach and modifications to the original formulation to minimize the deployment effort in unknown environments. After the approach showed to outperform state-of-the-art methods in lab environments [1], we determine whether human-like behavior actually improves the navigation performance of the robot among pedestrians unaware of the experiment. We tested the robot’s navigation capabilities with both the interaction-aware model and a two-stage predict-react motion planner in runs in a public building

for several hours. The model was trained from on-board observations of human-human interactions.

In the remainder of the paper we refer to related work in Section II and outline the problem formulation in Section III. The details of our approach are described in Section IV. Results from real-world tests are shown in Section V before a conclusion is drawn in Section VI. We highlight deviations from [1, 3, 4] in the corresponding sections.

II. RELATED WORK

The task of modeling interactions between agents has been approached in numerous fields like robotics, computer graphics, behavioral science and recently the automotive industry, to name a few. Helbing et al. [2] presented the social forces model for pedestrians based on potential fields, where one potential comprised of attractive forces is used to model goal-driven behavior and another potential based on repulsive forces is defined to model interactions. This original model was later refined in [5] to yield more accurate predictions, also for simulations with isolated pedestrians. Furthermore, the model was adapted in [6] to describe the dynamics of interacting vehicles. Similar to the social forces model, Treuille et al. [7] utilize continuum dynamics and potential fields to model naturally moving crowds. The potential field both integrates global navigation capabilities and the avoidance of dynamic obstacles.

A popular approach that accounts for interactions of agents is the Reciprocal Velocity Obstacles (RVO) method presented by van den Berg et al. [8]. It has proven to be a computationally efficient method for interaction-aware motion planning. The RVO method was extended by Alonso-Mora et al. to nonholonomic robots [9]. Velocity Obstacles (VO) [10] and RVO, respectively, compute joint collision-free velocities assuming either constant future velocities, or known shared avoidance policies. A drawback of both the former VO and RVO approach is its inability to handle probabilistic behavior and learn from observed data in order to train the model.

Guy et al. [11] proposed models to measure the similarity between a given set of observed (real-world) data, and a visual simulation technique to generate crowd motions. Lee et al. [12] use an agent decision model governing each agent’s actions based on environment features and the motion of nearby agents in the crowd. While these techniques generate socially compliant motion patterns in simulation, they are not effective for motion predictions in real-world applications. A drawback is that global knowledge of crowd parameters such as motion direction and people density is required [7] which are difficult to obtain online.

Trautman et al. [13] introduced the novel concept of Interacting Gaussian Processes (IGP), where each agent’s trajectory is modeled via a Gaussian Process (GP) and individual GPs are coupled through an interaction potential modeling cooperation between different agent’s trajectories. This work was later extended to include mixture models to account for multimodal behavior in human crowds [14].

A significant amount of work has been conducted on learning policies from demonstrations. Atkeson et al. [15]

presented an approach that learns a task model and a reward function from human demonstrations. The learned information is used to solve an optimal control problem for planning the robot’s actions. Abbeel et al. [16] developed an algorithm based on Markov Decision Processes (MDPs), where the reward function is unknown. In [17], they further develop the algorithm and its performance is demonstrated on a navigation task on a parking lot. Only few demonstrations are required during the learning step. However, these approaches are not able to model interactions between several agents.

Ziebart et al. [18] apply the principle of maximum entropy Inverse Reinforcement Learning (IRL) to specify a structure for the reward function employed by decision-making agents. They demonstrate the ability of the model to learn the reward function from demonstrations of human behavior. In Henry et al. [19], the authors extend maximum entropy IRL to work in partially observable dynamic environments and introduce features that capture aspects of crowd navigation. Kuderer, Kretzschmar et al. [1, 3, 4] leverage this approach to continuous state-spaces and introduce features to capture socially-compliant navigation behavior. They show that the approach outperforms the social forces model [2] and RVO [8] in terms of its predictive qualities of pedestrian motion. The approach is tested both on test datasets as well as on a mobile robotic platform, yet the latter evaluation involves only few pedestrians in a lab environment – most probably aware of the ongoing experiment. For the latter evaluation, the authors also assume known final destination points for each agent and perform a two-dimensional (2D) grid search to create a set of intermediate waypoints for the local interaction-aware planner. This planner then computes the most likely trajectory for each agent (including the robot) given a fixed start and goal position. The strong assumption about known final destination points hinders experimental validation in real-world scenarios. It also imposes a strong commitment towards the high-level solution from the global waypoint planner.

III. PROBLEM FORMULATION

To account for the stochasticity of human motion, we seek a trainable probabilistic model

$$p(\mathbf{x}|\boldsymbol{\theta}) \quad (1)$$

of joint behavior, where \mathbf{x} represents a composite trajectory and $\boldsymbol{\theta}$ constitute parameters of the probability distribution. We define a composite trajectory as the tuple of trajectories of all n (relevant) traffic participants (including the ego agent trajectory \mathbf{x}_R)

$$\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_R, \dots, \mathbf{x}_n).$$

Each agent’s trajectory is parametrized by time. It is defined on the time interval $[a_i, b_i]$ and in the agent’s extrinsic state space

$$\mathbf{x}_i : t \in [a_i, b_i] \mapsto \mathbf{x}_i(t) \in X,$$

where X is the extrinsic state space of a single agent, e.g. \mathbb{R}^2 to represent a 2D position.

During training we seek the model parameters θ^* that best explain a set of observed composite trajectories $\mathcal{X}_{\mathcal{D}}$. Once obtained, the distribution can be used for online motion planning by computing the most likely composite trajectory \mathbf{x}^* and applying \mathbf{x}_R^* to the robot. If we train the model with observations of cooperative behavior, the robot will be able to reproduce such behavior when navigating.

IV. APPROACH

In general, numerous models are able to explain observed data. Yet models may overfit on the training data yielding a large generalization error for newly gathered data. The question arises which model one should choose to avoid the danger of overfitting.

A. Maximum entropy model

A natural choice to avoid overfitting is to select the distribution with the largest entropy. Expressed in casual terms, this distribution has the least commitment to the data and encodes only knowledge of certain statistics of it. The maximum entropy probability distribution – encoding only knowledge about the expected value of a set of measurable functions $\mathbf{f}(\mathbf{x})$ – has an exponential structure given by

$$p(\mathbf{x}|\theta, C(\mathbf{x})) = \frac{1}{z(\theta, C(\mathbf{x}))} \cdot \exp\left(-\theta^T \cdot \mathbf{f}(\mathbf{x})\right), \quad (2)$$

with a normalization constant $z(\theta, C(\mathbf{x}))$, a weight vector θ and a feature vector $\mathbf{f}(\mathbf{x})$ which is defined over a composite trajectory. We represent each trajectory as a time parametrized 2D B-spline. As a novel contribution, we condition the distribution on a “configuration” $C(\mathbf{x})$, describing e.g. the number of agents in the composite trajectory \mathbf{x} . This conditional distribution is the one we need for the planning task later in which we need to predict all agents given their currently observed configuration. The exponent of the distribution is a weighted sum of feature functions and can also be interpreted as weighted sum of cost terms. This interestingly establishes a link to former motion planning techniques such as the dynamic window approach [20], where multiobjective optimization criteria are encoded in the same fashion.

B. The learning problem

In the training stage, we are interested in the feature weights θ^* best explaining a set $\mathcal{X}_{\mathcal{D}}$ of observed composite trajectories. We formulate a loss function as the negative sum of log-likelihoods

$$\ell(\mathcal{X}_{\mathcal{D}}; \theta) := - \sum_{\mathbf{x} \in \mathcal{X}_{\mathcal{D}}} \ln p(\mathbf{x}|\theta, C(\mathbf{x})) \quad (3a)$$

$$= \sum_{\mathbf{x} \in \mathcal{X}_{\mathcal{D}}} \left(\theta^T \cdot \mathbf{f}(\mathbf{x}) + \ln z(\theta, C(\mathbf{x})) \right) \quad (3b)$$

assuming that all $\mathbf{x} \in \mathcal{X}_{\mathcal{D}}$ are independent (this requires that for any two composite agent trajectories, either the time intervals are disjoint, or the trajectories involve different

agents). We then obtain the maximum likelihood parameter estimate θ^* from

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \ell(\mathcal{X}_{\mathcal{D}}; \theta) . \quad (4)$$

To efficiently search for θ^* with gradient-based optimization techniques, we need to evaluate the gradient of the loss function (3a) with respect to the feature weights. For each demonstration $\mathbf{x} \in \mathcal{X}_{\mathcal{D}}$ we get

$$\frac{\partial}{\partial \theta} - \ln p(\mathbf{x}; \theta, C(\mathbf{x})) = \mathbf{f}(\mathbf{x}) + \frac{\partial}{\partial \theta} \ln z(\theta, C(\mathbf{x})) , \quad (5)$$

where the latter term is given by the feature expectations of the distribution

$$\frac{\partial}{\partial \theta} \ln z(\theta, C(\mathbf{x})) = - \mathbb{E}_{p(\cdot|\theta, C(\mathbf{x}))} [\mathbf{f}] . \quad (6)$$

This results in an intuitive equation for the gradient of the loss function (3a)

$$\frac{\partial}{\partial \theta} \ell(\mathcal{X}_{\mathcal{D}}; \theta) = \sum_{\mathbf{x} \in \mathcal{X}_{\mathcal{D}}} \left(\mathbf{f}(\mathbf{x}) - \mathbb{E}_{p(\cdot|\theta, C(\mathbf{x}))} [\mathbf{f}] \right), \quad (7)$$

where feature expectations of the maximum entropy model are matched with the ones from demonstrated composite trajectories. The sample approximation of the expected feature values $\mathbb{E}_{p(\cdot|\theta, C(\mathbf{x}))} [\mathbf{f}]$ is computed with a Hybrid Monte Carlo sampler [21]. Here, we only need to provide a density proportional to the target density, mitigating the problem of computing the normalization constant $z(\theta, C(\mathbf{x}))$. The sample space to compute $\mathbb{E}_{p(\cdot|\theta, C(\mathbf{x}))} [\mathbf{f}]$ is hereby given via the configuration $C(\mathbf{x})$ of the corresponding demonstrated composite trajectory \mathbf{x} . At this point we highlight the difference to [3], equation 5, stating $\mathbb{E}_{p(\mathbf{x})} [\mathbf{f}(\mathbf{x})] = \frac{1}{|\mathcal{X}_{\mathcal{D}}|} \sum_{\mathbf{x}_k \in \mathcal{X}_{\mathcal{D}}} \mathbf{f}(\mathbf{x}_k)$ with no clarification about the sample space for \mathbf{x} of the expectation on the left-hand side.

For each demonstration $\mathbf{x} \in \mathcal{X}_{\mathcal{D}}$ we sample over the composite spline control vectors for $C(\mathbf{x})$. After having computed the Monte Carlo sample approximations for the feature expectations, the gradient given in (7) can be used with a gradient based optimization technique. For this purpose, the Resilient Back-Propagation (RProp) optimizer [22] is used. By only using the gradient direction it provides better resilience to noisy gradient estimates due to the finite sample approximation.

C. The planning problem

The planning framework hinges on the assumption that all agents plan their trajectories by minimizing the negative log-density of the maximum-entropy distribution (2) given their current observed configuration C . The optimization problem for motion planning yields a minimization of the weighted sum of features/cost terms

$$\mathbf{x}^* = \underset{\mathbf{x}, C(\mathbf{x})=C}{\operatorname{argmax}} p(\mathbf{x}|\theta^*, C(\mathbf{x})) = \underset{\mathbf{x}, C(\mathbf{x})=C}{\operatorname{argmin}} \theta^{*T} \cdot \mathbf{f}(\mathbf{x}), \quad (8)$$

since the normalization constant $z(\theta^*, C(\mathbf{x}))$ of the distribution may be neglected as it only depends on θ^* and $C(\mathbf{x}) = C$. The maximum-entropy distribution inherently encodes cooperative behavior due to its definition over a

composite trajectory \mathbf{x} in the time interval given by the planning horizon. Instead of planning paths to a goal position within a fixed distance and optimizing for the travel time as shown in [1], we adapt the formulation to have a fixed planning horizon and a variable end position of the trajectory. This eliminates the need to find a set of fixed waypoints for agents with an unknown destination. This is especially helpful when deploying the robot in unknown environments.

D. Feature design

The selection of meaningful features is essential in order to provide the model the chance to capture relevant statistics of the data. We consider three distinct kinds of features: (i) motion features, (ii) interaction features and (iii) observation features. The first two kinds encode information about independent agent motion and agent-to-agent interaction, respectively. The third kind encodes information gained via observations of the agents' states.

Motion Features in our model act similar to those in [3]. We employ feature functions computing the integrated *velocity* and *acceleration* of the trajectories. We further extend the feature set with a *direction-of-motion* feature capturing the agents' preference to stick with their current velocity. We define this feature being the integrated squared norm of the velocity difference to the initial velocity. It partially replaces the dropped target information for the other agents in our model. *Static obstacle* information is incorporated by a feature accounting for the typical avoidance distance of such. An inverted sigmoid function $1/(1 + \exp(d))$ is applied on the signed 2D distance transform of an occupancy grid representation of the static environment. We employ a bicubic interpolation with Catmull-Rom splines on the discrete distance transform to ensure continuous gradients of the obstacle distance with respect to the positions of the spline. The feature is scaled with the velocity of the agent to provide lower feature values when moving close to an obstacle with lower speed. This effectively reduces the impact of low probability regions arising in narrow corridors.



Fig. 2: Predicting the reaction of other agents to static obstacles in the environment affects the robot's plan (orange) too. Based on the typical avoidance distance learned from demonstrations, the robot leaves the pedestrian more room for navigation by swerving to the left.

Interaction Features model the interaction of pairwise agent trajectories within a composite trajectory \mathbf{x} . As stated in [2], pedestrians tend to dislike other agents entering their

"private sphere". By computing the integral over the inverse of the squared Euclidean distance as in [1, 3, 4], the model prefers pairwise trajectories with larger distance.

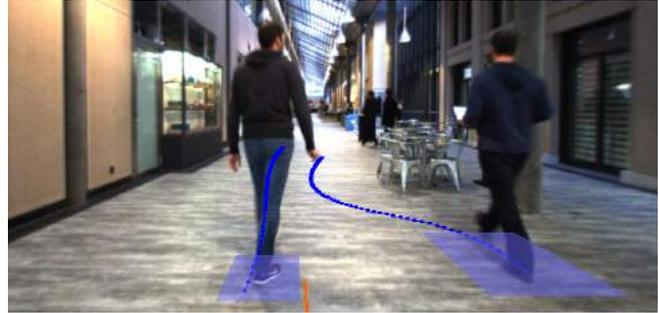


Fig. 3: Interaction features capture typical aspects of pairwise agent motion such as the agents' "private spheres". In conjunction with static obstacle information we can predict complex scenarios such as an agent (left) giving room to another due to a predicted conflict with a static obstacle.

Observation Features comprise a probabilistic measurement model that copes with the noisy observations of the observed agents. We employ a Gaussian noise model for each position and velocity measurement we obtained for an agent. Due to a lack of groundtruth information, we do not adapt the feature weights of observation features in the training stage.

V. EXPERIMENTS

The results presented in this section evaluate the added value of using interaction-aware motion models during online motion planning in real-world environments. We conducted experiments with both unaware (not informed about the experiment) pedestrians and aware (instructed to behave "normally" around the robot) pedestrians. In both cases, the interaction-aware motion planning approach is compared to a common predict-react approach assuming a constant velocity model for the other agents. In the latter case, the robot maximizes the probability (2) only for the own motion and hence does not account for the reaction of the other agents to its actions.



Fig. 4: EUROPA2 platform in its testing environment, surrounded by various pedestrians interacting with the robot.

A. Experimental Setup

The experiments are conducted with the platform of the European EUROPA2 project (Figure 4), targeting the development of a pedestrian assistant robot in urban environments.

Equipped with multiple laser range finders, the differential drive platform extracts and tracks moving objects by clustering dynamic elements in the 2D laser range finders' point clouds. The robot's diameter of 0.8 m is comparable to a human. Our maximum entropy navigation software runs on a standard laptop (Intel Core i7, 2.80 GHz) mounted in the robot body.

We trained the maximum entropy model from demonstrated "optimal"/expert behavior. Figure 5 depicts the two training sets recorded with the laser-based object tracker, teaching the robot especially static and dynamic object avoidance behavior.

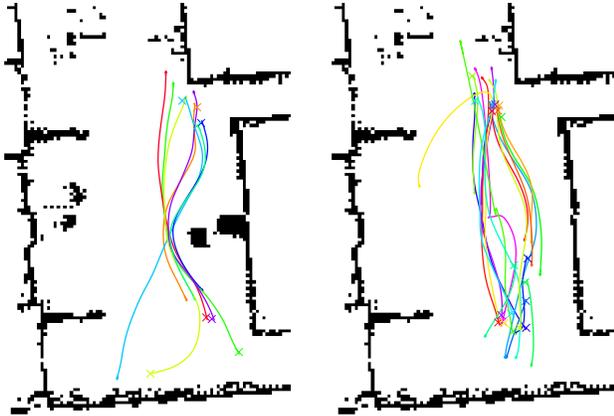


Fig. 5: Training data presented to the robot for learning the feature weights of the maximum entropy model. Left: Evasive maneuvers for static obstacle avoidance, right: agent-to-agent interactions with two pedestrians. The colored tracks depict spline fits to the discrete measurements of the onboard 2D laser-based object tracker. Crosses indicate the measured start positions.

B. Online motion planning

Establishing clear evaluation metrics is difficult for motion planners with fuzzy goals such as behaving in a socially compliant manner. With a sufficiently large planning time horizon, noise-free observations, a perfect model and every agent indeed choosing its trajectory from the most likely composite trajectory, the robot would have to plan its trajectory only once. Plans acquired in subsequent re-planning cycles would not deviate from the initial plan. However, there is no such perfect motion model for pedestrian motion and re-planning is required. Using a more realistic motion model for all participating agents therefore requires less re-planning effort and results in a smoother and more predictable behavior of the robot. We can use the prediction error of the ego motion as a score for how well the model describes joint navigation behavior. A higher prediction error is directly linked to a higher re-planning effort for the robot. For the error statistics shown in the experiments, the error is only evaluated during intervals where at least one other pedestrian is in an area of interest in front of the robot in order to filter out time intervals with no interactions.

The goal of the experiments is to take the presented approach out of the lab environment and analyze its performance and advantages in real-world experiments. In terms of

the quality of the motion plans, we expect the interaction-aware approach to outperform the planner that assumes constant velocity motion of the other agents.

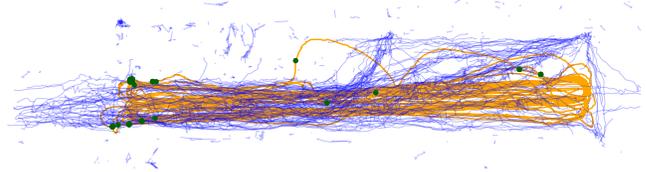


Fig. 6: Path driven by the robot (orange), pedestrian tracks (blue) and emergency stop points (green) for one of our experiment runs. The overall traveled distance by the robot in this run was more than 2 km. Throughout this experiment, 890 pedestrians were detected (multiple counting possible) by the detection system of the robot. The distance between the two landmarks of the alternating trajectories was about 50 m.

Experiments with unaware pedestrians

In order to test the feasibility of the approach in a realistic scenario, passing people were unaware about the ongoing experiments. We conducted two experiments under this premise. One, where the motion planner expects other pedestrians to interact with it based on the learned interaction model. The other assumes a constant velocity model for pedestrians with no adaptations of them in the future. In either experiment the robot collected data of at least 2 km of autonomous driving by alternating between two fixed waypoints. The measured trajectories for the first experiment are shown in Figure 6.

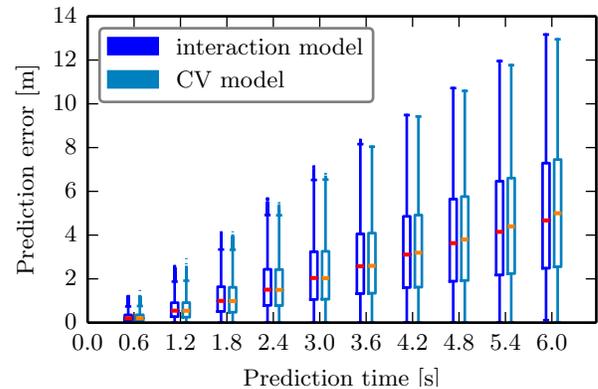


Fig. 7: Experiment with unaware pedestrians. Prediction error statistics for the ego motion at varying time horizons. The participating pedestrians were not informed about the experiments and either avoided the robot at large security distances in order not to disturb it or they tried to aggressively block its path.

Figure 7 depicts the distribution of the Euclidean position error of both experiments at varying time horizons in the future. Surprisingly, no significant difference between the two approaches can be noticed. We observed that pedestrians unaware of the ongoing experiment interact with the robot in a completely different manner than among pedestrians only. Either they try to keep a large security distance to the robot since they don't want to disturb a potential experiment or they try to trick it and test the robot's stopping behavior when they block its path (see Figure 6, green points). Especially

the latter case produces large prediction errors if the robot is stopped abruptly.

Experiments with aware pedestrians

As the previous experiments showed, human-human interaction differs from human-robot interaction. In another experiment we asked the participants to interact with the robot as if they would interact with other people. As before, one run per motion model was conducted with the robot.

Figure 8 shows the evaluation of the prediction error when interacting with aware pedestrians. The motion planning approach using the interaction-aware motion model for pedestrians outperforms the common predict-react approach based on constant velocity predictions in terms of prediction error and therefore re-planning effort. Both the absolute value of the error quantiles and the number and extent of the outliers was reduced significantly. The improvement upon the constant velocity model is based on the fact that aware pedestrians interact similarly with the robot as with other pedestrians. The behavior of aware pedestrians can be better explained with the interaction-aware motion model trained with data from human-human interaction.

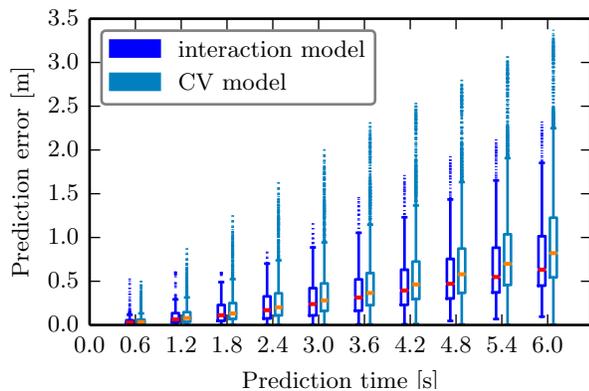


Fig. 8: Experiment with aware pedestrians. Prediction error statistics for the ego motion at varying time horizons. Only if pedestrians were instructed to behave normally when encountering the robot, the interaction-aware model outperforms the simple constant velocity approach.

VI. CONCLUSION

In this paper we analyzed the performance of a data-driven motion planning approach for interaction-aware navigation of a robot among pedestrians. We learn an interaction-aware motion model based on human-human interactions observed by the robot with onboard sensors only. In order to allow for a fast deployment of the robot in a previously unseen and unstructured environment, we present a receding horizon motion planning approach that does not require previous target information for pedestrians. The performance of interaction-aware motion planner is evaluated in thorough real-world experiments with over 4 km of autonomous navigation data. The prediction quality of the motion planner is analyzed and compared to a non-interaction-aware approach. Our results show that the motion models learned from human-human-interaction cannot be transferred directly to robot-human interaction as long as robots constitute a rarity and

attract the people’s attention. Yet if pedestrians are asked to interact naturally with the robot, the interaction-aware approach significantly outperforms the predict-react motion planner.

REFERENCES

- [1] H. Kretzschmar, *et al.*, “Learning to predict trajectories of cooperatively navigating agents,” in *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*, Hong Kong, China, May/June 2014, pp. 4015–4020.
- [2] D. Helbing and P. Molnar, “Social force model for pedestrian dynamics,” *Physical review E*, vol. 51, no. 5, p. 4282, 1995.
- [3] M. Kuderer, *et al.*, “Feature-based prediction of trajectories for socially compliant navigation,” in *Proc. of Robotics: Science and Systems (RSS)*, Sydney, Australia, July 2012, p. 193.
- [4] M. Kuderer, *et al.*, “Teaching mobile robots to cooperatively navigate in populated environments,” in *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, Tokyo, Japan, Nov. 2013, pp. 3138–3143.
- [5] T. I. Lakoba, *et al.*, “Modifications of the helbing-molnar-farkas-vicsek social force model for pedestrian evolution,” *Simulation*, vol. 81, no. 5, pp. 339–352, 2005.
- [6] D. Helbing and B. Tilch, “Generalized force model of traffic dynamics,” *Physical Review E*, vol. 58, no. 1, p. 133, 1998.
- [7] A. Treuille, *et al.*, “Continuum crowds,” in *ACM Trans. on Graphics (TOG)*, vol. 25, no. 3. ACM, 2006, pp. 1160–1168.
- [8] J. D. Van Berg, *et al.*, “Reciprocal velocity obstacles for real-time multi-agent navigation,” in *Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA)*, Pasadena, USA, May 2008, pp. 1928–1935.
- [9] J. Alonso-Mora, *et al.*, “Optimal Reciprocal Collision Avoidance for Multiple Non-Holonomic Robots,” in *Distributed Autonomous Robotic Systems*, ser. Springer Tracts in Advanced Robotics, A. Martinoli, *et al.*, Eds. Springer Berlin Heidelberg, Jan. 2013, no. 83, pp. 203–216.
- [10] P. Fiorini and Z. Shiller, “Motion planning in dynamic environments using velocity obstacles,” *The International Journal of Robotics Research*, vol. 17, no. 7, pp. 760–772, 1998.
- [11] S. J. Guy, *et al.*, “A statistical similarity measure for aggregate crowd dynamics,” *ACM Trans. on Graphics (TOG)*, vol. 31, no. 6, p. 190, 2012.
- [12] K. H. Lee, *et al.*, “Group Behavior from Video: A Data-Driven Approach to Crowd Simulation,” in *Eurographics/SIGGRAPH Symp. on Computer Animation*. The Eurographics Association, 2007, p. 109–118.
- [13] P. Trautman and A. Krause, “Unfreezing The Robot: Navigation in Dense, Interacting Crowds,” in *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*. Taipei, Taiwan: IEEE, Oct. 2010, pp. 797–803.
- [14] P. Trautman, *et al.*, “Robot navigation in dense human crowds: the case for cooperation,” in *Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA)*, Karlsruhe, Germany, May 2013, pp. 2153–2160.
- [15] C. G. Atkeson and S. Schaal, “Robot learning from demonstration,” in *Proc. of the Int. Conf. on Machine Learning (ICML)*, vol. 97, ACM, Corvallis, USA: ACM, June 1997, p. 12–20.
- [16] P. Abbeel and A. Y. Ng, “Apprenticeship Learning via Inverse Reinforcement Learning,” in *Proc. of the Int. Conf. on Machine Learning (ICML)*. Banff, Canada: ACM, July 2004, p. 1–.
- [17] P. Abbeel, *et al.*, “Apprenticeship learning for motion planning with application to parking lot navigation,” in *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, Nice, France, Sept. 2008, pp. 1083–1090.
- [18] B. Ziebart, *et al.*, “Maximum entropy inverse reinforcement learning,” in *Proc. of the National Conference on Artificial Intelligence (AAAI)*, Chicago, USA, July 2008, pp. 1433–1438.
- [19] P. Henry, *et al.*, “Learning to navigate through crowded environments,” in *Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA)*, Anchorage, USA, May 2010, pp. 981–986.
- [20] D. Fox, *et al.*, “The dynamic window approach to collision avoidance,” *IEEE Robotics & Automation Mag.*, vol. 4, no. 1, p. 23–33, 1997.
- [21] S. Duane, *et al.*, “Hybrid monte carlo,” *Physics letters B*, vol. 195, no. 2, pp. 216–222, 1987.
- [22] H. Braun and M. Riedmiller, “RPROP: a fast adaptive learning algorithm,” in *Proc. of the Int. Symp. on Computer and Information Science VII*, 1992.