Intent-Aware Long-Term Prediction of Pedestrian Motion

Vasiliy Karasev Alper Ayvaci Bernd Heisele Stefano Soatto

***Abstract*— We present a method to predict long-term motion of pedestrians, modeling their behavior as jump-Markov processes with their goal a hidden variable. Assuming approximately rational behavior, and incorporating environmental constraints and biases, including time-varying ones imposed by traffic lights, we model intent as a policy in a Markov decision process framework. We infer pedestrian state using a Rao-Blackwellized filter, and intent by planning according to a stochastic policy, reflecting individual preferences in aiming at the same goal.**

I. Introduction

Safe operation of autonomous systems co-mingling with humans could benefit from some level of understanding of human behavior by the robot. One of the simplest requirements of safe operation is collision avoidance: The robot must avoid regions of space likely to be concurrently occupied by a human. This requires predicting human motion beyond a short time-horizon. While short-term prediction is sufficiently informed by recent past history, longer time-scale prediction depends on intents or goals, which are not manifest to the viewer. We posit that knowledge of a person's intents or goals, in conjunction with contextual knowledge can help prediction beyond a few seconds into the future. Since in reality we do not know a person's goal, we can infer it along with a prediction of human behavior or marginalize it as a nuisance variable. In this paper we explore the problem of long-term prediction of pedestrian behavior in the autonomous (or assisted) driving scenario.

A. Related work and contributions

Common approaches to prediction of time series assume they are samples from a process generated by a linear model (*e.g.* ARMA) driven by a white, zero-mean, Gaussian input. Traditional tools from filtering and system identification can then be used to infer the state and model parameters [1]. When the order of the model is restricted, it yields predictions that are accurate at short time-scales, but insufficient beyond a few steps, as the processes it represents are stationary. It is also difficult to embed context, such as partial knowledge of the environment, into these models while preserving their structure. Extension to switching models (*e.g.* PWARX) have been used in constrained scenarios [2], [3], [4] but again their predictive ability has only been demonstrated on short horizons. A nonparametric approach based on Gaussian processes proposed in [5] is accurate on longer time-scales, but requires significant computational efforts. A number of models of "social" or "socially compliant" behavior have been proposed [6], [7], [8], [9], [10], [11]. These models permit prediction of actors' joint behavior, yet have been verified only on relatively short-term horizons. In this paper we focus on each pedestrian individually, neglecting their interactions with other traffic participants.

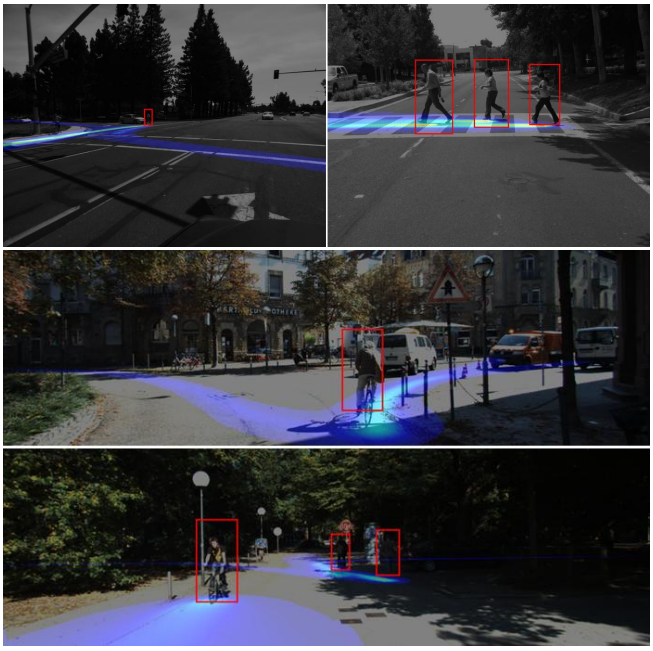

Fig. 1: *Sample long-term predictions of traffic participants' motion generated by our model. Warmer colors indicate more probable paths. Notice that the predictions are multi-modal and obey constraints of the environment.*

A common approach to improve accuracy of long-term prediction has been to postulate "goals" or "destinations", and to assume that agents navigate toward them by approximately following stochastic shortest paths [12], [13], [14], [15], [16], [17]. This framework allows one to represent rich dynamics, and to naturally incorporate environment constraints – that the agent cannot navigate through "obstacles" [12], [14], [17], and that regions may be more or less preferred depending on their semantic category [16]. During training, this framework requires one to identify the shortest path policy used by the agents – *i.e.* to solve an inverse reinforcement learning problem [18], [19], [20]. At test time, one uses the agent's recent measured motion to estimate the goal, whose

V. Karasev is with the Department of Electrical Engineering, University of California Los Angeles. E-mail: karasev00@ucla.edu. A. Ayvaci and B. Heisele are with the Honda Research Institute. E-mail: aayvaci@hra.com, bheisele@hra.com. S. Soatto is with the Department of Computer Science, University of California Los Angeles. E-mail: soatto@cs.ucla.edu. This work was supported by Honda Research Institute and grants ARO-W911NF-15-1-0564, ONR-N00014-15-1-226, AFRL-FA8650-11-1-7156.

knowledge then completely reveals future trajectories. We show that these models of agents' behavior can be interpreted as switching nonlinear dynamical systems, with the latent goal variable governing the switches, and the policy describing the nonlinear motion dynamics.

In this paper, we provide an interpretation of models in [13], [14], [16] as jump-Markov processes. The large body of work on this topic provides guidance on how estimation/prediction can be approached, while systematically accounting for the uncertainty in dynamics or measurements. We apply the framework to the problem of long-term pedestrian motion prediction. Discrete-space models are applied to a problem in continuous space by adding speed as an additional latent variable. Further, we show how the model can be naturally extended to handle dynamic environments (traffic lights) and additional measurements (pedestrian orientation).

Sec. II describes the framework: Sec. II-A is dedicated to estimation and prediction, Sec. II-B focuses on the underlying optimal control problems, Sec. II-C explains the extension to dynamic environments. Sample results of our framework are shown in Fig. 1, and we further verify the performance of our method in Sec. III.

II. Formalization

The state of the pedestrian is described by her position, orientation, and speed in the global coordinate frame $(x, \theta, s) \in SE(2) \times \mathbb{R}_+$, and the (unknown) goal $\mathbf{g} \subset \mathbb{R}^2$ (possibly just a point) from a finite set $\mathbf{G}$ which is known a-priori.

We define *intent* to be a function π mapping goals to *actions*: given $\mathbf{g}$ and a current state $\mathbf{x} = (x, \theta, s)$, an intent π yields future physical state trajectories from the current time t to a future time when the goal is achieved. Because different individuals having the same goal will act differently, the map π is stochastic. In the language of Markov decision processes (MDPs), an intent π is called a *policy*, a sample from it, integrated over time, is called a *plan*, which at each instant of time corresponds to an *action*. For simplicity, we assume that the stochastic component of the intent is represented by a process w_π, given which the policy π has a known functional form. So, we write $\pi(\mathbf{x}, \mathbf{g}, w_\pi)$ to indicate that, given a sample w_π and a goal $\mathbf{g}$, the policy uniquely determines an action.

We assume that the goal $\mathbf{g}$ is slowly time-varying; specifically that at each time instant it switches to another, uniformly chosen goal with a small probability. This can be summarized by a distribution $p(\mathbf{g}_{t+1}|\mathbf{g}_t)$, or by the relation $\mathbf{g}_{t+1} = \mathbf{g}_t \oplus w_{g,t}$, where $\mathbf{g}$ is interpreted as an integer[1], w_g is an integer-valued random process, and the operation is addition modulo $|\mathbf{G}|$. This allows us to write the model as a discrete-time jump-Markov process:

$$\begin{cases} \mathbf{x}_{t+1} = f(\mathbf{x}_t, \pi(\mathbf{x}_t, \mathbf{g}_t, w_{\pi,t})) \\ \mathbf{g}_{t+1} = \mathbf{g}_t \oplus w_{g,t} \end{cases} \tag{1}$$

[1]Throughout the paper we will exploit finiteness of $\mathbf{G}$ to interpret $\mathbf{g}$ either as a region in $\mathbb{R}^2$, or as an *index* in $\{1, \dots, |\mathbf{G}|\}$. The use should be clear from context.

where the low-level pedestrian dynamics are abstracted into f, the state transition map. This is a model of a stochastic hybrid system with continuous states $\mathbf{x}$, discrete state $\mathbf{g}$, and feedback π. As a further simplification, we assume that π is only affected by the positional component x of $\mathbf{x}$. This is equivalent to saying that the goal only affects the *direction* of heading, whereas the speed at which the pedestrian walks is unknown, but otherwise unaffected by the goal. This is clearly a simplification (one may walk at different speeds depending on where s/he is heading), but sufficient for our purpose. We therefore summarize the model as:

$$\begin{cases} x_{t+1} &= x_t + s_t \begin{bmatrix} \cos(\theta_t) \\ \sin(\theta_t) \end{bmatrix} \\ \theta_{t+1} &= \pi(x_t, \mathbf{g}_t, w_{\pi,t}) \\ s_{t+1} &= s_t + w_{s,t} \\ \mathbf{g}_{t+1} &= \mathbf{g}_t \oplus w_{g,t} \end{cases} \tag{2}$$

where $(w_{\pi,t}, w_{s,t}, w_{g,t}) \sim P_w$ are the state transition processes, with distribution P_w. The pose of our vehicle in the global coordinate frame is described by $(x_{\text{ego},t}, \theta_{\text{ego},t})$ and is assumed to be measured. The observation $\mathbf{y}_t$ consists of the pedestrian position measured relative to the vehicle:

$$\mathbf{y}_t = \mathsf{R}(\theta_{\text{ego},t})^T (x_t - x_{\text{ego},t}) + n_t \tag{3}$$

where R maps orientation to a rotation matrix and $n_t \sim P_Q$ is the measurement noise.

If we knew the goal state $\mathbf{g}_t$, assuming rational behavior of the agent, we could infer her intent by computing the (future) state trajectory that minimizes expected time-to-goal or, equivalently in our case, path cost. Unfortunately, we do not know the goal state, which we must instead infer using the past state trajectory, which is only known through the history of measurements up to time t, i.e. $\mathbf{y}_1, \dots, \mathbf{y}_t$ or $\mathbf{y}^t$. The goal state could be inferred along with the physical state by estimating (filtering) the posterior $p(\mathbf{x}_t, \mathbf{g}_t|\mathbf{y}^t)$. A filter maintains an estimate of the filtering density by computing the prediction $p(\mathbf{x}_{t+1}, \mathbf{g}_{t+1}|\mathbf{y}^t)$ using Chapman-Kolmogorov's equation and knowledge of $P_w, \mathbf{G}$, and updating the prediction once measurements $\mathbf{y}_{t+1}$ become available: $p(\mathbf{x}_{t+1}, \mathbf{g}_{t+1}|\mathbf{y}^{t+1})$ using Bayes' rule and knowledge of P_Q.

The filtering density can be approximated by a generic particle filter, which maintains a sample-based representation of the posterior. However, in our case we exploit the structure of the model and apply a lower-complexity Rao-Blackwellized filter, as described in Sec. II-A. Once we have the filtering density, we can infer intent by solving the same stochastic optimal control problem that, presumably, a rational agent solves. To do that, we must know the reward function, which we obtain by leveraging the contextual information in the form of a semantic map, as described in Sec. II-B. In Sec. II-C we extend the model in (2) to account for pedestrians changing their behavior with respect to changes in the environment. In an assisted driving scenario, an important example of this involves pedestrian behavior being influenced by traffic rules, including time-varying ones enforced by traffic lights. In Sec. II-D we show that the inference is improved via additional

Algorithm 1 Rao-Blackwell particle filter

for $i = 1 \dots |\mathbf{G}|$ **do**
$p(\mathbf{x}_t|\mathbf{y}^{t-1}, \mathbf{g}_{t-1} = i) \leftarrow \mathbf{predict}\{p(\mathbf{x}_{t-1}|\mathbf{y}^{t-1}, \mathbf{g}_{t-1} = i)\}$
end for
Compute marginal likelihood:
$p(\mathbf{y}_t|\mathbf{y}^{t-1}, \mathbf{g}_t = i) = \int p(\mathbf{y}_t|\mathbf{x}_t, \mathbf{g}_t = i)p(\mathbf{x}_t|\mathbf{y}^{t-1}, \mathbf{g}_t = i)d\mathbf{x}$
Predict $\mathbf{g}$:
$p(\mathbf{g}_t|\mathbf{y}^{t-1}) = \sum_{\mathbf{g}_{t-1}} p(\mathbf{g}_t|\mathbf{g}_{t-1})p(\mathbf{g}_{t-1}|\mathbf{y}^{t-1})$
Update posterior over $\mathbf{g}$:
$p(\mathbf{g}_t|\mathbf{y}^t) \propto p(\mathbf{y}_t|\mathbf{y}^{t-1}, \mathbf{g}_t)p(\mathbf{g}_t|\mathbf{y}^{t-1})$
for $i = 1 \dots |\mathbf{G}|$ **do**
$p(\mathbf{x}_t|\mathbf{y}^t, \mathbf{g}_t = i) \leftarrow \mathbf{update}\{\mathbf{y}_t, p(\mathbf{x}_t|\mathbf{y}^{t-1}, \mathbf{g}_t = i)\}$
end for

Algorithm 2 Sampling state trajectory $\hat{\mathbf{x}}_t^{t+\tau}, \hat{\mathbf{g}}_t^{t+\tau}$

Generate $\hat{\mathbf{x}}_t, \hat{\mathbf{g}}_t \sim p(\mathbf{x}_t|\mathbf{y}^t, \mathbf{g}_t)p(\mathbf{g}_t|\mathbf{y}^t)$
for $k = 1 \dots \tau$ **do**
$\hat{\mathbf{g}}_{t+k} \sim p(\mathbf{g}_{t+k}|\hat{\mathbf{g}}_{t+k-1})$
$\hat{\mathbf{x}}_{t+k} \sim p(\mathbf{x}_{t+k}|\hat{\mathbf{x}}_{t+k-1}, \hat{\mathbf{g}}_{t+k-1})$
end for

measurements of the pedestrian's orientation. Performance of our proposed model is compared to multiple baselines in Sec. III.

A. Filtering and prediction

To infer the state posterior we use a Rao-Blackwellized particle filter (RBPF) [21]. In our case, the posterior is represented by a discrete distribution over $\mathbf{G}$, $p(\mathbf{g}_t|\mathbf{y}^t)$, and by a set of $|\mathbf{G}|$ filters (Kalman or particle filters) that approximate each distribution $p(\mathbf{x}_t|\mathbf{g}_t = i, \mathbf{y}^t)$. In Alg. 1 we list the RBPF filtering steps. The "predict" and "update" steps are the standard steps performed by either Kalman or particle filters.

Once we have the posterior $p(\mathbf{g}_t, \mathbf{x}_t|\mathbf{y}^t)$, we can perform prediction. We are interested in the state τ steps into the future, i.e. $p(\mathbf{x}_{t+\tau}, \mathbf{g}_{t+\tau}|\mathbf{y}^t)$ (or just the position, *i.e.* $p(x_{t+\tau}|\mathbf{y}^t)$). This is a multi-modal distribution and the integration needed to compute it is intractable. However, state trajectories $\hat{\mathbf{x}}_t^{t+\tau}, \hat{\mathbf{g}}_t^{t+\tau}$ can readily be sampled as described in Alg. 2, and can be used to approximate $p(x_{t+\tau}|\mathbf{y}^t)$. As a byproduct it allows us to compute an "occupancy map", which gives a probability that a region $r \subset \mathbb{R}^2$ is occupied between t and $t + \tau$:

$$P_{occ}(r) \doteq \mathbb{P}\big(\cup_{k=1}^{\tau} \{x_{t+k} \in r\}\big) \tag{4}$$

This quantity enables visualizing future trajectories by discounting time.

B. Goals and environment context

In this section we describe how to model the *intent* of an agent as a solution to a planning problem. As noted above, the "rational" navigation is abstracted into π, which decomposes into independent functions $\{\pi_{\mathbf{g}}\}_{\mathbf{g}\in\mathbf{G}}$. Each is defined by $\pi_{\mathbf{g}} : \mathbb{R}^2 \to \mathbb{S}$, with $\mathbb{S}$ being the action space. This function specifies the optimal "move direction" for reaching the goal $\mathbf{g}$ quickly, while satisfying environment constraints and pedestrian's contextual preferences. Specifically, it is a solution to the optimal control problem, which can be formulated as a Markov decision process (MDP) [22], as follows:

$$\pi_{\mathbf{g}}^* = \arg\max_{\pi} \sum_{t=0}^{\infty} \gamma^t R_{\mathbf{g}}(x_t, \pi(x_t)) \tag{5}$$

$$\text{subject to } x_{t+1} = x_t + \pi(x_t) \tag{6}$$

where the objective is the sum of γ-discounted rewards $R_{\mathbf{g}}$ for being at location x and applying action $\pi(x)$. In our case, the transition dynamics given by (6) are deterministic; in the general MDPs they can be stochastic, as we will also explore in Sec. II-C. This sum in the objective is also denoted by $V^{\pi_{\mathbf{g}}}(x_0)$ – the *value* of starting at x_0 and applying policy $\pi_{\mathbf{g}}$ thereafter. Using the value function, it becomes possible to rewrite the optimization problem recursively as

$$\begin{cases} V^{\pi_{\mathbf{g}}}(x) & = \max_u Q^{\pi_{\mathbf{g}}}(x, u) \\ Q^{\pi_{\mathbf{g}}}(x, u) & = R_{\mathbf{g}}(x, u) + \gamma V^{\pi}(x + \pi_{\mathbf{g}}(x)). \end{cases} \tag{7}$$

Within this formalism, the optimal policy attains the maximum of Q at every x, that is, $\pi_{\mathbf{g}}^*(x) = \arg\max_u Q^{\pi_{\mathbf{g}}^*}(x, u)$. To fully describe the MDP, we need to specify $R_{\mathbf{g}}$, and to do that we leverage availability of the *semantic map*. Each point in $\mathbb{R}^2$ is classified in one of several categories (namely, "building", "sidewalk", "crosswalk", "road", "grass"), and the reward $R_{\mathbf{g}}$ is taken to be a function of the environment category. Specifically, $R_{\mathbf{g}}(x, u) = \theta^T \phi(x)$ where θ is a vector of preferences and $\phi(x)$ is a vector specifying probabilities of different semantic categories at x. The reward is low in regions corresponding to buildings (since those are "obstacles"), high on sidewalks and crosswalks, and is lower on roads. Moreover, for $x \in \mathbf{g}$, $R_{\mathbf{g}}(x, u) = 0$ and elsewhere $R_{\mathbf{g}}(x, u) < 0$, which implies that the optimal policy $\pi_{\mathbf{g}}^*$, the solution to (5) describes a shortest path to $\mathbf{g}$. We show a region of the world with overlaid semantic categories and associated reward preferences in Fig. 2.

Commonly, one solves an MDP to obtain a deterministic optimal policy. However, for prediction, it is advantageous to use a stochastic, possibly suboptimal policy. This is because whenever shortest paths to a certain $\mathbf{g}$ are not unique, nonzero probability should be assigned to each. Moreover, an optimal policy assumes that pedestrians are completely rational and assigns zero probability to any deviations from the optimal trajectory. To allow small deviations from optimality, we use a stochastic Boltzmann policy, also used in [13], [14], [16]:

$$u \sim \pi_{\mathbf{g}}(x) \text{ w.p. } \propto \exp\big(\alpha(Q_g(x, u) - V_g(x))\big) \tag{8}$$

As $\alpha \to \infty$, the policy assigns nonzero probability only to the optimal actions: $u \sim \pi_{\mathbf{g}}(x)$ w.p. $\propto \mathbf{1}\{u \in \arg\max_{\tilde{u}} Q_{\mathbf{g}}(x, \tilde{u})\}$. As $\alpha \to 0$, the policy becomes uniformly random: $u \sim \pi_{\mathbf{g}}(x)$ w.p. $\frac{1}{|\mathcal{U}|}$. Fig. 2 shows several stochastic shortest paths for different goals; note that the shortest path is not necessarily unique.

Fig. 2: *Semantic map, goals, and shortest paths toward them in two regions. The semantic categories are: ■≒"building", ■≒"grass", ■≒"road", ■≒"sidewalk",■≒"crosswalk". Pedestrian preferences are taken to be: ■ ≥ ■ ≥ ■ ≥ ■ ≥ ■. The goals are marked with red "⋆". Stochastic shortest paths toward sample goals are shown in the same two areas, overlaid on the satellite image. Pedestrian is marked with "∘" and the destination – with "⋆".*

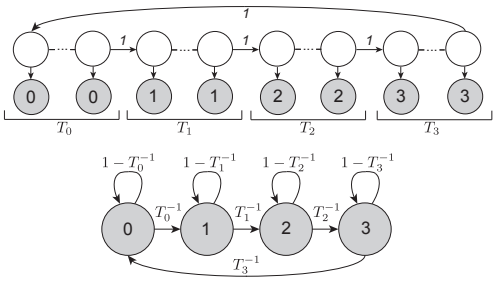

Fig. 3: Top*: original representation of traffic light as a hidden Markov chain.* Bottom*: our compressed representation. Shaded nodes are observed, light nodes are hidden.*

C. Environment dynamics

Accurately predicting pedestrian behavior is particularly important near intersections. The framework outlined above inadequately describes this behavior, as it does not account for traffic signals, the simplest instance of environment dynamics. In this section, we describe a simple extension to the framework that more accurately models pedestrians waiting for the light signal near intersections. To achieve this, we include the state of the traffic signal into the policy π.

A traffic signal can itself be represented by a dynamical system. The output is one of four values $\{0,1,2,3\}$. The four states correspond to the most typical signal configuration with red/green in either direction (2 states) and red/yellow in either direction (2 additional states). Extensions to multi-way signals are possible but not addressed here. If $\mathcal{T}_0$,$\mathcal{T}_1$,$\mathcal{T}_2$,$\mathcal{T}_3$ are the durations that each output is observed, and $\sum_{i=0}^{3}\mathcal{T}_i = \mathcal{T}$ is the period, then the state dynamics is a timer: $c_{t+1} = \text{mod}(c_t+1,\mathcal{T})$. This model is deterministic and neglects feedback, due for instance to loop sensors detecting proximity of vehicles, including the own-vehicle. Nevertheless, even for this simplified model, directly adding c_t into π would yield a dramatic increase in state-space dimension, since $\mathcal{T}$ is typically very large, and since the pedestrian must operate on the belief about c_t, as it is not directly observed. We instead propose a simplified, "compressed", and fully-observable model where $c_t \in \{0,1,2,3\}$ and dynamics are stochastic:

$$p(c_{t+1}=j|c_t=i) = \begin{cases} 1-\frac{1}{T_i} & \text{if } j=i \\ \frac{1}{T_i} & \text{if } j=\text{mod}(i+1,4) \end{cases} \quad (9)$$

The diagram in Fig. 3 illustrates the compression: effectively, the multinomial state transition distribution is replaced by a geometric one [23]. To include c_t into the model of pedestrian behavior, we augment $\mathbf{x}_t$ with c_t (so that $\mathbf{x}_t = (x_t, s_t, \theta_t, c_t)$) and modify $\{\pi_{\mathbf{g}}\}_{\mathbf{g}\in\mathbf{G}}$ to depend on c_t in addition to x_t. Furthermore, we modify the reward $R_{\mathbf{g}}$ to be $R_{\mathbf{g}}(x,c,u) = \theta_c^T\phi(x)$; in other words we modify the vector of preferences to be dependent on the traffic light state. Specifically, the state of the signal affects the reward at the crosswalks: if the signal is "red", the cost of being in the region is high, and if the signal is "green", the cost is low. The net effect of these changes is that shortest paths on the modified MDPs involve the pedestrian waiting at intersections for the appropriate light signal, and avoiding moving in traffic. We assume that traffic light state is directly observed, i.e. $\mathbf{y}_t = (y_{1,t}, y_{2,t})$ with $y_{1,t} = \mathsf{R}(\theta_{\text{ego},t})^T(x_t - x_{\text{ego},t}) + n_t$ (as before), and $y_{2,t} = c_t$. In Sec. III-D we show an example of how this simple extension improves prediction accuracy near intersections.

D. Pedestrian orientation

The model described above uses pedestrian *motion* to estimate the latent goal, and when the pedestrian is initially detected, the distribution over $\mathbf{g}$ is uniform. However, since a pedestrian does not typically walk backwards, we can leverage its orientation to infer which goals are more likely even at the first detection time. We do this by incorporating an additional measurement of the pedestrian orientation. At each time we observe $\mathbf{y}_t = (y_{1,t}, y_{2,t})$ with $y_{1,t} = \mathsf{R}(\theta_{\text{ego},t})^T(x_t - x_{\text{ego},t}) + n_{1,t}$ (as before) and $y_{2,t} = \phi(\theta_t - \theta_{\text{ego},t} + n_{2,t})$, where $\phi(x) = \mathbf{floor}\left(\frac{N}{2\pi}x\right)$ models the output of a pose detector quantized to N orientation values. The uncertainty of the pose detector response is accounted for by $n_{2,t}$. The additional measurement makes it possible to quickly reduce uncertainty over goals, yielding a more accurate prediction earlier. We show an example of this in Sec. III-E. This improvement is especially important in situations when the pedestrian is tracked only for a small number of frames and an early prediction is required.

III. Experiments

We compare against the following baselines:

$$\text{(RW)}\ x_{t+1} = x_t + n_t,\ n_t \sim \mathcal{N}(0,\sigma^2) \quad (10)$$

$$\text{(LM)}\ \begin{cases} x_{t+1} & = x_t + v_t \\ v_t & = v_t + n_t,\ n_t \sim \mathcal{N}(0,\sigma^2) \end{cases} \quad (11)$$

$$\text{(MM)}\ x_{t+1} = x_t + u_t,\ u_t \sim p(u|x_t) \quad (12)$$

$$\text{(DM)}\ x_{t+1} = x_t + \pi(x_t, \mathbf{g}, n_t) \quad (13)$$

"RW" is a simple random-walk. "LM" is a linear "constant-velocity" model commonly used for tracking, and used

in comparisons in [2]. "MM" ("Markov model") learns a location-dependent probability distribution over actions, but does not reason about "goals", used in [16]. "DM" ("discrete model") is described in [13], [14], [16], and is similar to ours, except that it assumes a fixed goal and constant-length steps – it does not adapt to the observations of the pedestrian speed. In our experiments, step lengths for MM and DM were taken to be the average pedestrian speed in the training set.

In sections III-B-III-C, we evaluate our "basic" model that does not use the orientation measurement. For this reason, θ can be omitted from the state in (2), and the model can be simplified to:

$$\begin{cases} x_{t+1} &= x_t + s_t \begin{bmatrix} \cos(\theta_t) \\ \sin(\theta_t) \end{bmatrix}, \theta_t \doteq \pi(x_t, \mathbf{g}_t, w_{\pi,t}) \\ s_{t+1} &= s_t + w_{s,t} \\ \mathbf{g}_{t+1} &= \mathbf{g}_t \oplus w_{g,t} \end{cases} \tag{14}$$

To perform inference with a RBPF, in this case, we approximated $\{p(\mathbf{x}_t|\mathbf{y}^t, \mathbf{g}_t = i)\}_{i=1,\ldots,|\mathbf{G}|}$ using a set of Kalman filters. This is possible if one assumes that $\nabla_x \pi(x_t, \mathbf{g}_t, w_{\pi,t}) = 0$, and that the constraint $s_t \in \mathbb{R}_+$ can be dealt with by projecting the variable onto the feasible set after each update.

The orientation-aware model is evaluated in Sec. III-E. There, the system dynamics are nonlinear, and to approximate $\{p(\mathbf{x}_t|\mathbf{y}^t, \mathbf{g}_t = i)\}_{i=1,\ldots,|\mathbf{G}|}$ we used a set of particle filters.

Error metrics: The difficulty of prediction depends both on prediction horizon and on the period of observation (since some time is required to infer the pedestrian location and the goal). A natural measure that evaluates the model's ability is the expected ℓ_2 error between the model's predictions and the actual pedestrian locations (denoted by $\{x_t^{obs}\}_{t=1}^T$), for a given prediction horizon τ and observed sequence length t:

$$\mathcal{E}(\tau, t) \doteq \mathbb{E}_{p(x_{t+\tau}|\mathbf{y}^t)}\left[\|x_{t+\tau} - x_{t+\tau}^{obs}\|\right]. \tag{15}$$

This quantity depends on two variables and makes it inconvenient to perform comparisons. A summary error for a given prediction horizon can be computed as the average error over all observation periods, as $\mathcal{E}(\tau) \doteq \frac{1}{T-\tau} \sum_{t=1}^{T-\tau} \mathcal{E}(\tau, t)$.

A. Dataset and implementation details

To verify the effectiveness of our model, we captured a dataset on a vehicle that is equipped with two cameras, LiDAR, and DGPS, each acquiring data at 15Hz. The streams are synchronized with best-effort which produces satisfactory data batches for the long-term prediction task at low vehicle speed. Data from LiDAR and DGPS makes it possible to accurately localize the pedestrians in world coordinates.

Sample data sequences acquired by the vehicle are shown in Fig. 4. For our experiments, we manually annotated 17 video sequences, ranging from 30 to 900 frames, and containing 67 pedestrian trajectories, used in the "fully-observable" experiment. In the "partially-observable" experiment, pedestrians in those videos were automatically detected and tracked; we used a total of 50 of the resulting trajectories (some were lost due to poor detections). In addition to this dataset, we evaluated our approach on a subset of videos from the KITTI benchmark [24].

Fig. 4: Left: *Sample frames from the front camera, with LiDAR returns (colored by depth) and the detected pedestrians.* Right: *Bird's-eye-view of the scene. The vehicle is in north-west corner and pedestrians, shown by circles, are in the center. LiDAR returns are shown in black. Note that the satellite image is outdated and does not contain the pedestrian crossing seen in the images.*

For learning model parameters, we used 80/20% train/test splits and five-fold cross-validation. Pedestrian detection was performed using a cascade classifier [25] using HoG [26] features. The location of the detections with respect to the vehicle is determined by clustering the LiDAR measurements that are projected into the detection window. Once the relative position of the target object is known, its world-coordinates can be found by addition of the vehicle location given by the DGPS. In the experiment reported in Sec. III-D, the traffic light behavior was manually annotated.

Goals, MDPs, and rewards: Destination points were placed near the map boundaries and near building entrances (due to lack of a large enough dataset to learn them); each map region contained less than 15 goals. We used standard value iteration to solve MDPs and produce policies π. Note that MDPs are typically formulated and solved in discrete spaces, but throughout the paper we treated the MDP state-space (and action-space) as being continuous (i.e. $\mathbb{R}^2$ and $\mathbb{S}$). Since the state space is fairly low-dimensional, we can solve the discretized MDP and perform interpolation thereafter: at test-time π is bilinearly interpolated from the state's nearest neighbors. We discretize $\mathbb{S}$ to 16-25 directions. To specify the reward in MDPs, formally, one should learn the preference vector θ in $R_{\mathbf{g}}$ using a large enough dataset. We found that the order relations between the cost of each category inferred via IRL [19] matches the obvious ones appointed manually.

Error computation: At very short observation periods, the state estimate (pedestrian location, velocity, goal) and the predictions are unreliable, and do not accurately reflect the model performance. For this reason, in computing the prediction errors for all models, we ignored all errors with observation period $t < 10$. The error during this "initialization stage" is shown in the error surfaces $\mathcal{E}(t, \tau)$. In all experiments we used 5000 samples to approximate the predictive $p(\mathbf{x}_{t+\tau}|y^t)$ and to calculate prediction errors, for $\tau = 1, \ldots, 350$ (about 23 seconds).

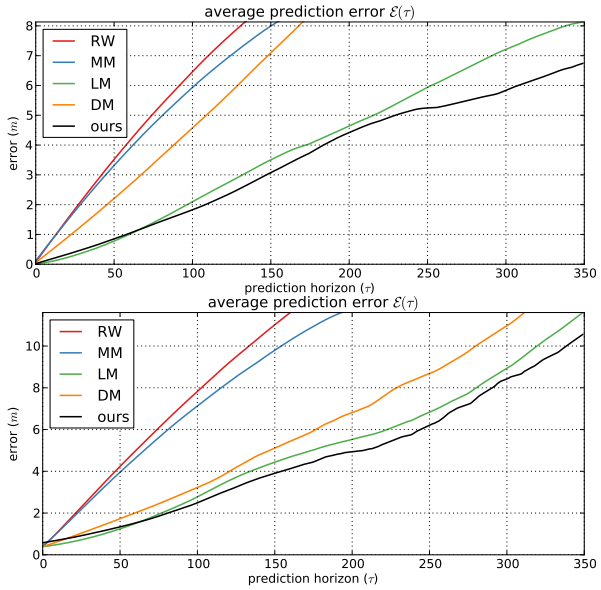

Fig. 5: *Summary prediction error $\mathcal{E}(\tau)$ for all methods, on the entire dataset.* Top: *prediction with ground truth trajectories (perfectly observed case: $n_t = 0$).* Bottom: *prediction with noisy trajectories (imperfectly observed case: $n_t \sim P_Q$)*

B. Prediction in the perfectly observed case

We first evaluated our framework on manually annotated "ground-truth" trajectories. In this case, the measurement noise $n_t = 0$, and observations fully disclose pedestrian locations. Quantitative performance results are summarized in Fig. 5, where we show prediction errors for up to 350 steps. Each time step corresponds to 67 ms, i.e. the camera frame rate is 15 Hz. As is well known, MM degenerates to random walk and is often unable to generate meaningful predictions, similarly to RW. LM accurately predicts short-term behavior but cannot predict inevitable direction changes. These baselines do not utilize environment semantics, which strongly affect behavior of traffic participants. DM generates meaningful predictions, but ones that are not temporally accurate, since the model does not estimate pedestrian speed. In particular, the predictions are erroneous when the pedestrian is stationary (e.g. waiting for the signal at the intersection) and the model predicts motion. Our model utilizes environment semantics, estimates the pedestrian speed, and performs best.

C. Prediction in the imperfectly observed case

We also tested the framework in the imperfectly observed setting (with $n_t \sim P_Q$), where observations are obtained from noisy detector responses. These results are qualitatively similar and are shown in Fig. 5 (bottom). Qualitative examples of predictions generated by our model and by the baseline are shown in Fig. 6. In that figure DM is not shown, as it generates an occupancy map nearly identical to ours. Unfortunately, its error is high due to predicting motion at the incorrect speed.

D. Intent near intersections

We compared the basic framework with the traffic-aware extension described in Sec. II-C. The extended framework is

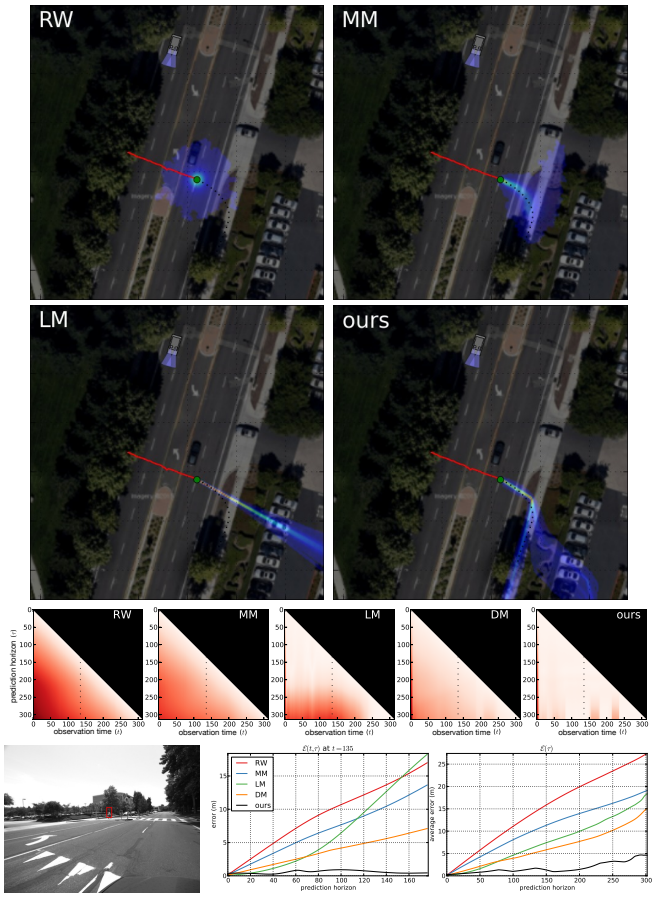

Fig. 6: Top row: *occupancy maps predicted by the four methods ("DM" not shown, as it generates occupancy map identical to "ours").* Middle row: *quantitative prediction errors. Light colors correspond to small errors, and dark red – to large errors. Time is on the abscissa, and prediction horizon is on the ordinate. The cross-section of the error surface at $t = 135$ (indicated by a vertical dashed black line) is shown in the bottom panel.* Bottom row: *frame from the video sequence, and average prediction error $\mathcal{E}(\tau)$. Note that in this example, the satellite image used for visualization is outdated and does not contain the crosswalk seen in the sample frame. The crosswalk is present in our semantic map (see Fig. 2).*

expected to reduce the probability that a pedestrian violates traffic rules (e.g. the probability that she walks on the crosswalk on "red"). An appropriate measure that illustrates the difference between the two models is the probability that the pedestrian is on the road at $t + \tau$, given the observations $\mathbf{y}^t$, i.e. $P_{road}(\tau, t) = \mathbb{E}_{p(x_{t+\tau}|\mathbf{y}^t)}[\mathbf{1}\{x_{t+\tau} \in \text{road}\}]$. In Fig. 7 we show that the extended model correctly predicts that pedestrian does not walk out on the road (doing so would violate traffic rules), and improves the overall prediction error.

E. Pedestrian orientation

To leverage orientation, we trained a linear classifier with HoG features that discretized the orientation θ to four values {"left","front","right","back"}, and learned a response likelihood $p(y_2|\theta)$. The improvement due to the additional measurement occurs primarily during the first few detection

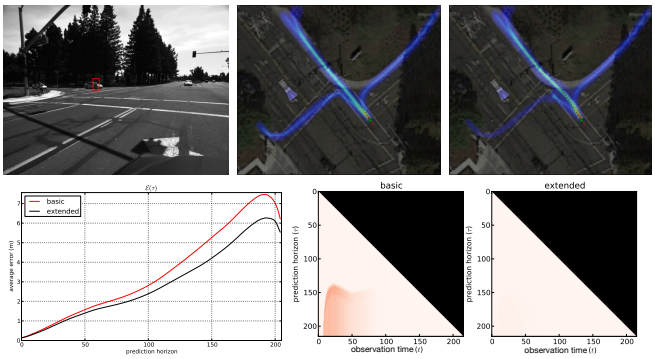

Fig. 7: Top row: *Sample frame, and the trajectories predicted by the "basic" (left) and "extended" (right) models. Notice that the "extended model" assigns a much lower probability to pedestrian making a left turn – which would violate traffic rules.* Bottom row: *prediction error, and the probability of the pedestrian being on the road* $P_{road}(\tau, t)$. *Note that the model does not predict that the probability of left turn is zero: a nonzero probability is assigned because* c_t *may switch to "green", in which case a left turn is allowed.*

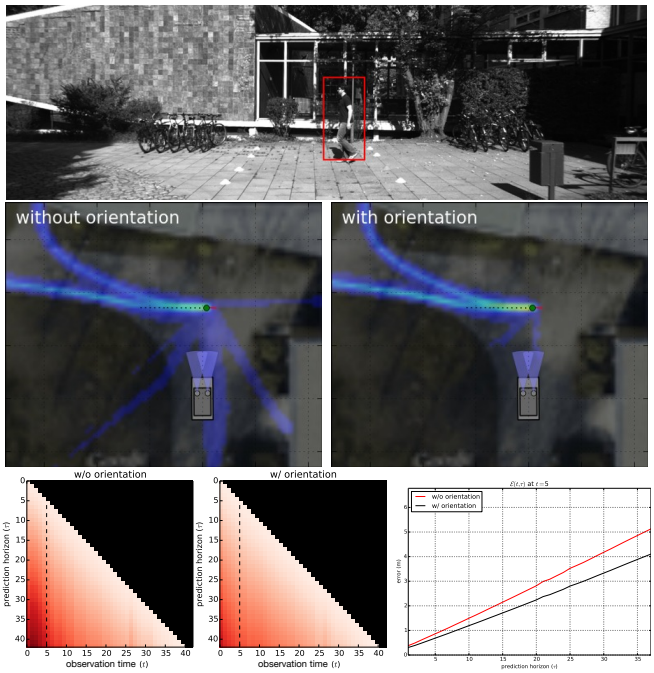

Fig. 8: Top row: *Sample frame with a tracked pedestrian.* Middle row: *Predictions generated by the orientation-agnostic and orientation-aware models, at* $t = 5$ *(i.e. 500 ms after detection).* Bottom row: *Prediction errors* $\mathcal{E}(\tau, t)$*; the plot on the right is the cross-section at* $t = 5$ *(shown with dashed line).*

steps. In Fig. 8 we show an example on a video sequence from KITTI dataset. The orientation-aware model estimates goals faster than the "basic" model, and is able to reduce the prediction error as a result.

IV. Discussion

Our method for long-term prediction of pedestrians has several limitations. We discard feedback and interactions, due to the influence of agents on each other and by the ego-vehicle, including feedback by the environment, *e.g.* adaptive traffic lights, and changes in the environment, *e.g.* parked cars. These could be incorporated at a significant increase in computational cost that would prevent real-time operation. We also assume rational agent behavior, necessary to even define a predictive strategy, which means that we cannot predict *rare* events such as a pedestrian suddenly jumping onto the road. This limitation is shared with the majority of existing methods.

References

[1] A. H. Jazwinski, *Stochastic processes and filtering theory*, ser. Mathematics in science and engineering. Academic Press, 1970.

[2] N. Schneider and D. Gavrila, "Pedestrian path prediction with recursive bayesian filters: A comparative study," in *Pattern Recognition*. Springer, 2013.

[3] J. F. P. Kooij, N. Schneider, F. Flohr, and D. M. Gavrila, "Context-based pedestrian path prediction," in *ECCV*, 2014.

[4] A. Bissacco and S. Soatto, "Hybrid dynamical models of human motion for the recognition of human gaits," *IJCV*, May 2009.

[5] D. Ellis, E. Sommerlade, and I. Reid, "Modelling pedestrian trajectory patterns with Gaussian processes," in *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, 2009.

[6] S. Pellegrini, A. Ess, K. Schindler, and L. van Gool, "You'll never walk alone: Modeling social behavior for multi-target tracking," in *International Conference on Computer Vision*, 2009.

[7] P. Trautman and A. Krause, "Unfreezing the robot: Navigation in dense, interacting crowds," in *IROS*, 2010.

[8] K. Yamaguchi, A. Berg, L. Ortiz, and T. Berg, "Who are you with and where are you going?" in *CVPR*, 2011.

[9] M. Kuderer, H. Kretzschmar, C. Sprunk, and W. Burgard, "Feature-based prediction of trajectories for socially compliant navigation," in *RSS*, 2012.

[10] W. Choi, C. Pantofaru, and S. Savarese, "A general framework for tracking multiple people from a moving camera," *TPAMI*, 2013.

[11] G. Ferrer and A. Sanfeliu, "Bayesian human motion intentionality prediction in urban environments," *Pattern Recognition Letters*, vol. 44, no. 0, 2014.

[12] A. F. Foka and P. E. Trahanias, "Predictive autonomous robot navigation," in *IROS*, 2002.

[13] C. L. Baker, R. Saxe, and J. B. Tenenbaum, "Action understanding as inverse planning," *Cognition*, 2009.

[14] 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 *IROS*, 2009.

[15] T. Bandyopadhyay, Z. J. Chong, D. Hsu, M. Ang, D. Rus, and E. Frazzoli, "Intention-aware pedestrian avoidance," in *International Symposium on Experimental Robotics (ISER)*, 2012.

[16] K. M. Kitani, B. D. Ziebart, J. A. Bagnell, and M. Hebert, "Activity forecasting," in *ECCV*, 2012.

[17] D. Xie, S. Todorovic, and S. Zhu, "Inferring "dark matter" and "dark energy" from videos," in *ICCV*, 2013.

[18] A. Y. Ng and S. J. Russell, "Algorithms for inverse reinforcement learning," in *ICML*, 2000.

[19] B. D. Ziebart, A. Maas, J. A. Bagnell, and A. K. Dey, "Maximum entropy inverse reinforcement learning," in *UAI*, 2008.

[20] D. Ramachandran and E. Amir, "Bayesian inverse reinforcement learning," in *IJCAI*, 2007.

[21] A. Doucet, N. De Freitas, K. Murphy, and S. Russell, "Rao-blackwellised particle filtering for dynamic bayesian networks," in *UAI*, 2000.

[22] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY, USA: John Wiley & Sons, Inc., 1994.

[23] T. V. Duong, H. H. Bui, D. Q. Phung, and S. Venkatesh, "Activity recognition and abnormality detection with the switching hidden semi-markov model," in *CVPR*, 2005.

[24] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *IJRR*, 2013.

[25] Q. Zhu, M. Yeh, K. Cheng, and S. Avidan, "Fast human detection using a cascade of histograms of oriented gradients," in *CVPR*, 2006.

[26] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *CVPR*, 2005.