Inference and Prediction Insights for Safe Vehicle-Pedestrian Interaction

Katie Driggs-Campbell
Assistant Professor
Dept. of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
How can we ensure safety in autonomous systems that operate with people in the real-world?
Robust Decision-Making and Control

We wish to safely and efficiently control a vehicle, despite uncertainty and disturbances:

\[ x[k + 1] = f(x[k], u[k]) + v(x[k], u[k], d[k]) \]
Robust Decision-Making and Control

We wish to safely and efficiently control a vehicle, despite uncertainty and disturbances:

\[ x[k + 1] = f(x[k], u[k]) + v(x[k], u[k], d[k]) \]
Robust Decision-Making and Control

We wish to safely and efficiently control a vehicle, despite uncertainty and disturbances:

\[ x[k + 1] = f(x[k], u[k]) + v(x[k], u[k], d[k]) \]

Sensor Noise & Environment Uncertainty

Model Mismatch

Long-Term Interactions

Are pedestrian-vehicle interactions any different?

Simulations and datasets are not readily available.

Are pedestrian-vehicle interactions any different?

Simulations and datasets are not readily available.

Experiments are a bit of a production.
Are pedestrian-vehicle interactions any different?

Simulations and datasets are not readily available.

Experiments are a bit of a production.

In validation, pedestrians are the most challenging.


Today’s Destinations

Effective Pedestrian Prediction Methods for Safe Vehicle-Pedestrian Interactions

Inferring Occluded Pedestrians
Pedestrian Prediction Methods

Physics-based Methods
\[ \dot{x}_t = f(x_t, u_t, t) + w_t \]
- \( x_t \): human state
- \( u_t \): control input
- \( w_t \): process noise

Pattern-based Methods
\[ P[x_{t+1}|x_{1:t}] = P[x_{t+1}|x_{t-(n-1):t}] \]
\[ x_{t+1} = f(x_{t-(n-1):t}) \]
Pedestrian Prediction Methods

<table>
<thead>
<tr>
<th>Method (rank)</th>
<th>Final Disp.</th>
<th>Mean Disp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vanilla LSTM (1)</td>
<td>1.191</td>
<td>0.355</td>
</tr>
<tr>
<td>Physical Comp (6)</td>
<td>1.229</td>
<td>0.366</td>
</tr>
<tr>
<td>Social Forces (10)</td>
<td>1.266</td>
<td>0.371</td>
</tr>
</tbody>
</table>


Trajectory Forecasting Benchmark: trajnet.stanford.edu
Intention-Aware Predictions

Problem 1: How to infer human’s intent?
Problem 2: How to predict future trajectory?

Intent-Guided Prediction

- When we observe motion, we usually care very little about the surface behaviors

→ *Intentions* determine how we understand, recall, react, and predict

- When observing continuous motion, humans often agree where *boundaries separating distinct actions* lie, corresponding to intent

- In human motion prediction, incorporating intent (or goals) significantly improves predictions

K. Driggs-Campbell, et al., Identifying Modes of Intent from Driver Behaviors in Dynamic Environments, ITSC 2015.
Intention-Aware Predictions

Filter Comparisons

Prediction Samples

Intent Prob. Distribution
Mutable Intention Filter
Nominal pedestrian predictions are modeled using the Generalized Potential Field Approach:

- A set of sources represent obstacles in the map and generate a repulsive force
- A set of sinks represent goal locations and pose an attractive force on the pedestrian

The sum of the forces act as a control input to a linear state space model of the pedestrian:

\[ x_t = F x_{t-1} + G u_{t-1} + w_t \]
\[ y_t = H x_{t-1} + v_t \]

where \( x_t \) is the pedestrian state, \( y_t \) is the measurement, \( w_t \) and \( v_t \) are Gaussian white noise tuned to match real-world data.

Intention-Aware Predictions

- Ground Truth
- Social force (iLM)
- LSTM
- Warping Method
Trajectory Warping with Residuals

Residual Method: Warping LSTM

\[ \hat{x}_{1:t+T_g} = [x_{1:t} \; \tilde{x}_{t+1:t+T_g}] \]
Intention-Aware Predictions
Prediction Results

Effectively warps the nominal prediction to capture human tendencies
Simultaneously estimating intent with prediction provides improved performance and bonus utility

<table>
<thead>
<tr>
<th>Method</th>
<th>AOE</th>
<th>FOE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF</td>
<td>3.124</td>
<td>3.909</td>
</tr>
<tr>
<td>LSTM</td>
<td>2.132</td>
<td>3.005</td>
</tr>
<tr>
<td>SLSTM</td>
<td>1.524</td>
<td>2.510</td>
</tr>
<tr>
<td>ALSTM</td>
<td>0.986</td>
<td>1.311</td>
</tr>
<tr>
<td>SGAN</td>
<td>1.042</td>
<td>2.088</td>
</tr>
<tr>
<td>MIF-WLSTM₁</td>
<td>0.665</td>
<td>1.236</td>
</tr>
<tr>
<td>MIF-WLSTM₄</td>
<td><strong>0.636</strong></td>
<td><strong>1.179</strong></td>
</tr>
</tbody>
</table>
Online Monitoring for Vehicle-Pedestrian Interaction

Online Monitoring for Vehicle-Pedestrian Interaction


Online Reachability with DryVR

• Compute the reachable set $\text{Reach}_A(\Theta, p, T)$, given system $A$ with state space $X$, mode $p \in P$, set of initial states $\Theta$, and look-ahead time $T$

• DryVR uses a sensitivity function $\beta$ to bounds distance between trajectories

• $\Theta$ is partitioned into regions from which a numerical simulation $\xi(x_i, t)$ is computed for $T$ time

• An over-approximation of the reachable set $\text{Reach}_A(\Theta, p, T_{\text{Look}})$ is obtained by bloating $\xi(x_i, t)$ with $\beta$ and then taking a union

\[
\dot{x} = v \cos \theta, \quad \dot{y} = v \sin \theta, \quad \dot{\phi} = u, \quad \dot{v} = a, \quad \dot{\theta} = \frac{v}{L} \tan \phi
\]
where $(x, y)$ is the position, $v$ is the speed, $\theta$ is the steering angle, $\phi$ is the heading angle, and $L$ is the length of the vehicle

<table>
<thead>
<tr>
<th>$T_{\text{Look}}$ (s)</th>
<th>3.0</th>
<th>3.5</th>
<th>4.0</th>
<th>4.5</th>
<th>5.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compute Time (s)</td>
<td>0.096</td>
<td>0.103</td>
<td>0.129</td>
<td>0.136</td>
<td>0.163</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$T_{\text{Look}}$ (s)</th>
<th>3.0</th>
<th>3.5</th>
<th>4.0</th>
<th>4.5</th>
<th>5.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Conf. (%)</td>
<td>98.905</td>
<td>98.338</td>
<td>98.617</td>
<td>97.825</td>
<td>96.851</td>
</tr>
<tr>
<td>Med Conf. (%)</td>
<td>98.161</td>
<td>96.629</td>
<td>97.138</td>
<td>95.917</td>
<td>94.874</td>
</tr>
<tr>
<td>Low Conf. (%)</td>
<td>95.825</td>
<td>94.078</td>
<td>94.672</td>
<td>93.416</td>
<td>92.078</td>
</tr>
</tbody>
</table>

DryVR: Data-driven verification and compositional reasoning for automotive systems, C. Fan, B. Qi, S. Mitra & M. Viswanathan, CAV 2017
Online Monitoring for Vehicle-Pedestrian Interaction

1. Perception for non-experts remains a challenge
2. Reachability (and often decision-making and control) assumes single (most likely) trajectory prediction
3. Scaling to multiple agents adds complexity

Online Monitoring for Vehicle-Pedestrian Interaction

Challenges
1. Perception for non-experts remains a challenge
2. Reachability (and often decision-making and control) assumes single (most likely) trajectory prediction
3. Scaling to multiple agents adds complexity

Today’s Destinations

Effective Pedestrian Prediction Methods for Safe Vehicle-Pedestrian Interactions

Inferring Occluded Pedestrians
Can we make improved inferences about the state of the environment by observing other agents’ behaviors?
Consider the occluded region as a map we must estimate.

Occupancy Grid Approach:

\[ p_t(m|x_{1:t}, z_{1:t}) = \prod_{i=1}^{n} p_t(m_i|x_t, z_t) \]

Use people as sensors to improve mapping:

\[ p_t(m|x_{1:t}, a_{1:t}) = p_t(a_{t-\tau}m_t|x_t, z_t) \]

---

People as Sensors

Imputing Maps from Human Actions


---

Image Similarity Metric

- Our Method
- Vanilla Map
Real-World Tests and Additional Complexities
Today’s Destinations

Effective Pedestrian Prediction Methods for Safe Vehicle-Pedestrian Interactions

Inferring Occluded Pedestrians