Lecture 05. Applications of multimodal learning
Socially compliant mobile robot navigation via inverse reinforcement learning

- Henrik Kretzschmar, Markus Spies, Christoph Sprunk, Wolfram Burgard
- IJRR 2016
Problem definition

• Modeling cooperative, socially compliant navigation behavior for mobile robots
Technical challenges

• Decisions can be discrete (high-level or strategic) or continuous (low-level or reactive)
• Human navigation behavior is not deterministic
• Forward problem: computing feature expectation w.r.t. high dimensional space of continuous trajectories can be hard
Approach

Learn from observations of human pedestrians

Assume that observed trajectories are drawn from some probability distribution that depends on trajectory features

Figure 1: Our method is able to learn a model of human cooperative navigation behavior from demonstrations. We learn the model parameters of a mixture distribution over composite trajectories to capture the discrete and continuous aspects of the behavior. The learned model generalizes to new situations and allows us to draw trajectory samples that capture the stochasticity of natural navigation behavior.
Inverse reinforcement learning

- D: demonstrations (observed samples)
- Feature vector $f$: $X \rightarrow \mathbb{R}^n$
- Empirical feature value $f_D$
- Find distribution $p(x)$ that matches $f_D$
- Issue: there can be many that match
Maximum Entropy IRL

[Ziebart et al. 2008]

Principle of maximum entropy: desirable distribution maximizes entropy:

\[
\arg\max_p H(p) = \arg\max_p \int_x -p(x) \log p(x) dx, \\
p_\theta(x) = \frac{1}{Z(\theta)} \exp(-\theta^T f(x)),
\]

Parameter vector defined for features

(9)

Gradient

\[
\frac{\partial}{\partial \theta} h(p, \alpha, \theta) = \mathbb{E}_{p_\theta(x)}[f(x)] - f_D.
\]

(10)

1/31/18
MaxEnt IRL

• Cost function as a weighted sum of feature values

\[ \theta^T f(x) = \sum_i \theta_i f_i(x) \]  

(14)

• Learning navigation behavior: finding feature weights that lead to feature matching

• Gradient-based optimization method
Modeling navigation behavior

• Discrete & continuous navigation behavior
• Mixture distribution over homotopy classes
• Voronoi diagram to find homotopy classes
• Continuous behavior under a homotopy class is consistent (i.e., shares the same feature weights)
Learning navigation behavior

• Discrete behavior to determine homotopy class
  – Features: passing left/right, group behavior, most likely composite trajectory

• Continuous behavior to determine trajectory
  – Features: time, velocity, acceleration, clearance to other agents, collision avoidance w.r.t. static obstacles
Socially compliant robot navigation

• Predict jointly so that the interaction between pedestrians and robot can be taken into account at planning time

• Goal prediction:
  \[ p(\text{goal} \mid \text{trajectory}) = \alpha \ p(\text{trajectory} \mid \text{goal}) \ p(\text{goal}) \]

• Global path planning using A*
Fig. 6. Trajectories observed during one hour of interactions of four persons in our test environment. The depicted area has a size of approximately 11 m × 8 m.
**Turing test**

![Bar chart showing comparison between human and machine perceptions](chart.png)

**Fig. 10.** Turing test to evaluate whether the behaviors induced by our new approach, the approach of Kuderer et al. (2012), and the social forces model by Helbing and Molnar (1995) qualify as human. The results suggest that the behavior induced by our approach most resembles human behavior.
Hallway experiments

[Fig. 14]
Discussion points

- Feature engineering
- Decomposition of the problem into discrete & continuous behavior (organizing distribution space)
- Freezing robot problem
Super blue blood moon tonight!

Jan. 31, 2018