UNIVERSITEIT VAN AMSTERDAM

Active Perception for Person Tracking

Y. Satsangi

Summary

Active perception is the ability of an agent to take actions to reduce its uncertainty when it is uncertain about the world it is acting in. This thesis tackles the challenge of active perception for tracking people in multi-camera networks. Multi-camera systems are routinely used for security, surveillance and person tracking. A key challenge in the design of such networks is the efficient allocation of scarce resources such as the *bandwidth* required to communicate the collected data to a central server, the *CPU cycles* required to process that data, the *energy* costs of the entire system or the *manpower* required to manually monitor all the collected data. Maintaining surveillance is an example of an *active perception* task where an agent must select k out of the n available cameras to allocate the scarce resources to minimize its uncertainty about the state of the world. To this end, in this thesis we propose and give multiple results and methods for resource allocation in multi-camera networks, that in principle, reduce the uncertainty about the about the position of each person in the scene, in turn, enabling an agent to take actions to track people with a resource constrained multi-camera network.

Specifically, we propose:

- Equivalence of ρPOMDP and POMDP-IR, two frameworks that allow to expressing the reward in a partially observable Markov decision process (POMDP) as a function of the *belief* of the agent. This equivalence shows that entropy (uncertainty) of a probability distribution over a hidden variable can be approximated by giving the agent a choice to make predictions about the hidden state and then rewarding the agent for making right predictions.

- Greedy PBVI: a new POMDP planning methods that uses *greedy* maximization to scale in the large combinatorial action space of an active perception POMDP that consists of $\binom{n}{k}$ subset of cameras.

- Probably approximately correct (PAC) greedy maximization that is an approximate but computationally cheaper version of greedy maximization that requires access to only *upper* and *lower* confidence bounds on a submodular function to maximize it.

- PartiMax: a particle filter based algorithm that greatly reduces the computational cost of tracking people in ultra high resolution images by applying a trained person detector only to the k most relevant part of an image.

For each of the above mentioned methods, we prove results that guarantee firm error bounds and establish the conditions for application of those error bounds. To test the em-

pirical performance of our methods we conduct experiments on multiple real-life datasets that show that greedy PBVI and PAC greedy maximization achieve similar performance as the existing methods but at a fraction of the computational cost. PartiMax when applied on a real-life tracking problem retains 80% of the original tracking performance while processing only 10% of the original image while running in real-time.

Finally, we propose deep anticipatory network (DAN) that enable an model-free approach for active perception, thus, enabling an agent to learn the optimal policy from its experience given some ground truth data. We show that a DAN agent can potentially be trained in a supervised fashion to take actions to reduce its uncertainty.

In summary, this thesis pushes the state-of-the-art in active perception for person tracking, by proposing multiple methods and results applicable in general and especially for tracking people in resource constrained multi-camera networks.

Yash Satsangi
University of Amsterdam
November 2018