Karlsruhe Institute of Technology

Autonome Lernende Roboter (ALR)

Prof. Gerhard Neumann

Project Type

- ☑ Master Thesis
- ☑ Bachelor Thesis
- ☑ Research Project

Supervisors

Philipp Becker
philipp.becker@kit.edu
Vaisakh Shaj

Difficulty

Algorithmic

Math

Application

Deep Reinforcement Learning under Partial Observability

Description

Since the real world is usually not fully observable, Reinforcement Learning (RL) approaches capable of dealing with partial observability are crucial for many realistic scenarios, in particular for autonomous systems. For example, a robot can only perceive the environment through its sensors, from which it has to infer its belief about the state of the world, e.g., location and properties of different objects. A common approach for deep RL under partial observability is working with a belief state, i.e., using state-estimation to infer the current state from partial observations and then applying standard RL using this inferred state. For deep RL approaches, this usually means adding recurrency, usually in the form of an LSTM [2], to the network describing the agents behaviour.

We recently proposed a novel recurrent architecture, Recurrent Kalman Networks (RKN), [1] and showed that it can yield more accurate state estimates. Additionally, the RKN does not only provides a state estimate but also a principled estimate of the uncertainty. This additional uncertainty information is very valuable under partial observability as it provides information about whether the agent can safely act or it should continue gathering information to become more aware of its surroundings.

Figure 1: Nearly all tasks with image-inputs are partial observable as no velocity information can be obtained from a single image, the observations are limited to the cameras field of view and relevant aspects of the environment can be occluded.

Tasks

The tasks in this project will involve:

- Literature Research. Literature for deep RL in partial observable environments is to be reviewed and the most promising approaches to continue are to be identified.
- Employing the RKN. One (or multiple) of those approaches are to be augmented by the RKN and evaluated on partially observable environments in simulation.
- (for Master Thesis) Principled usage of Uncertainty. The RKN provides explicit uncertainty estimates. The RL approach is extended such that those estimates are used in a principled manner.

References

[1] Philipp Becker, Harit Pandya, Gregor Gebhardt, Cheng Zhao, C. James Taylor, and Gerhard Neumann. Recurrent kalman networks: Factorized inference in high-dimensional deep feature spaces. In *International Conference on Machine Learning*, pages 544–552, 2019.

[2] Matthew Hausknecht and Peter Stone. Deep recurrent q-learning for partially observable mdps. In *2015 AAAI Fall Symposium Series*, 2015.