Reinforcement Learning for Robotic Exploration

Context and Objectives

State-of-the-art algorithms are nowadays able to provide solutions to most elementary robotic problems like exploration, mapless navigation or Simultaneous Localization And Mapping (SLAM), under reasonable assumptions [Cadena2016]. Robotic pipelines are usually an assembly of several modules, each one dealing with an elementary function (e.g. control, planning, localization, mapping) dedicated to one technical aspect of the task. Each of these modules usually requires expert knowledge to be integrated, calibrated, and tuned and some of the elementary functions can raise issues in hard cases (e.g., computer vision in weakly textured environment or varying illumination conditions). Besides, this architecture makes complex behaviors difficult to model and synthesize, in particular, those that need to interact with the environment.

Combining several elementary functions into a single gray box module is a challenge but is an extremely interesting alternative. Indeed, an unified structure could achieve a better performance trade-off than a system splitted into a nearly optimal control module processing a coarse computer vision mapping output. Moreover, it allows to reduce calibration needs or expertise dependency. For these reasons, there is a large academic effort to try to combine several robotic functions into learning-based modules, in particular using a deep reinforcement strategy as in [Zamora2016]. Following this line of research, we would like to investigate the interest of Reinforcement Learning strategies to enhance exploration missions of autonomous robots in the real world, possibly in a multi-robot setup. In this context, preliminary work has been conducted at ONERA to evaluated the relevance of such learning approach on a simple task in a sim-to-real scenario. In this experiment, the robot should reach a target location (by controlling its actuators), given its absolute localization, the target location, and the depth measurements provided by a sensor. The Soft Actor-Critic algorithm was used, trained in simulation and deployed successfully on the real robot. DLR recently worked on a technique to train RL agents directly on real robots [Raffin2020]. This technique was applied to the control of an elastic robotic neck, using as input current 6D pose, desired 6D pose, tendons forces and outputting desired force. DLR also gained expertise in RL by developing an open source library which implements reference RL algorithms within the PyTorch framework. http://github.com/DLR-RM/stable-baselines3}{github.com/DLR-RM/stable-baselines3.

Scientific Challenges

The objective would be to study further and validate experimentally RL algorithms for autonomous robot exploration, along the following guidelines.

- Combine Reinforcement Learning and graph-based (PRM,RRT and friends) Motion Planning algorithms, possibly in a multi-robot setup [Faust2018, Matheron2020, Camci2020].
- Compare End-to-End Reinforcement Learning strategies with pipeline techniques combining model-based or learning-based separate algorithms (i.e. for depth computation, localization, mapping, planning, control) for perception-guidance loops embedded on ground or aerial

robots. The perfect result would be a network structure that takes as inputs directly the sensor measurements and the mission objectives (e.g., explore a given unknown area and build a model of the environment, or locate targets in this environment), and provides robot control variables in output.

- Using multiple RL agents to perform open exploration of unknown environments leads to a number of possibilities and challenges. For example, how agents should communicate to aim at fast and maximal coverage of the environment? or how should one agent reuse the experience of another agent that may have a different body (correspondence problem [Nehaniv2002]) or that may have acquired its experience following different behaviors (off-policy reinforcement learning [Foerster2017])?

References

[Cadena 2016] Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I. D., and Leonard, J. J. (2016). Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. IEEE Transactions on Robotics, 32:1309–1332.

[Zamora 2016] Zamora, I., Lopez, N. G., Vilches, V. M., and Cordero, A. H. (2016). Extending the OpenAI Gym for robotics: a toolkit for reinforcement learning using ROS and Gazebo. arXiv preprint arXiv:1608.05742.

[Chaffre 2020] Chaffre, T., Moras, J., Chan-Hon-Tong A., Marzat J. (2020) Sim-to-Real Transfer with Incremental Environment Complexity for Reinforcement Learning of Depth-Based Robot Navigation, ICINCO.
Video: https://w3.onera.fr/copernic/sites/w3.onera.fr.copernic/files/videos/video_sim2real_rl_depthbasedrobotnavigation.mp4

[Faust 2018] Faust, A., Oslund, K., Ramirez, O., Francis, A., Tapia, L., Fiser, M., & Davidson, J. (2018, May). PRM-RL: Long-range robotic navigation tasks by combining reinforcement learning and sampling-based planning. ICRA 2018 (pp. 5113-5120).

[Matheron 2020] Matheron, G., Perrin, N., & Sigaud, O. (2020). PBCS: Efficient Exploration and Exploitation Using a Synergy between Reinforcement Learning and Motion Planning. arXiv preprint arXiv:2004.11667.

[Camci 2020] Camci, E. (2020). Reinforcement learning based motion planning of quadrotors using motion primitives. Master Thesis, NTU Singapore.

[Nehaniv 2002] Nehaniv, Chrystopher L., Kerstin Dautenhahn, and Kerstin Dautenhahn, eds. Imitation in animals and artifacts. MIT press, 2002.

[Foerster 2017] Foerster, Jakob, et al. "Stabilising experience replay for deep multi-agent reinforcement learning." Proceedings of the 34th International Conference on Machine Learning-Volume 70. JMLR. org, 2017.