PROPOSITION DE POST-DOCTORAT

Intitulé : Sensor-based Reinforcement Learning for Multi-Robot Flocking

Référence : **POSTDOC-DTIS-2023-1**
(à rappeler dans toute correspondance)

Début du contrat : janvier 2023	**Date limite de candidature :**

Durée : 12 mois, éventuellement renouvelable une fois - Salaire net : environ 25 k€ annuel

Mots clés

Multi-Robot Systems, Flocking, Reinforcement Learning

Profil et compétences recherchées

PhD in Robotics and Machine Learning with a strong publication track record and a taste for experimental activities. Some prior knowledge in Computer Vision and Control Theory would be appreciated.

Présentation du projet post-doctoral, contexte et objectif

<u>Context and positioning of the problem:</u>

Centralized control architectures combined with external position sensing (mostly GNSS) make it possible to fly hundreds of Micro Air Vehicles (MAVs) at once, but do not reproduce the conditions in which natural swarms move. On the other hand, distributed approaches developed by the control community are getting closer to these conditions, relying on flocking algorithms relying on neighborhoods and local behavior-based rules [1,2,4]. This type of approaches enables the emergence of a global behavior of the swarm, but at the cost of a local need for information on the agents [5]. This information may be obtained from on-board sensors but, in practice, local relative sensing is limited and imperfect, and thus calls for communication capabilities between robots to mitigate its shortcomings [3]. Communication-based approaches also bring their own challenges, depending on the environment or other constraints (e.g. occlusions due to obstacles, delays, packet losses [4,6]) and a combination of communication and sensing capabilities is often to be considered for practical implementation of distributed flocking controllers on multi-robot systems.

Independently of this context, machine learning for perception and control has made significant progress in the last decade, in particular with the introduction of deep neural networks in the common practitioner toolkit. In terms of control, the field of Reinforcement Learning (RL) has also seen strong progress with the rise of neural network based approaches, with recent investigations on possible applications to flocking [7,8,9] and other related formation control tasks [10,11,12]. In terms of sensing, new powerful vision-based capabilities are also available [13] (depth perception, detection, localization, classification, segmentation…), with more and more efficient dedicated hardware such as smart cameras [14].

With these observations in mind, it is tempting to leverage the strengths of modern data-based sensing and control to tackle the challenges of distributed swarm control with local sensing and limited communication. End-to-end approaches in RL could theoretically make it possible to learn a direct mapping between onboard camera inputs and flock control outputs, but these approaches are still deemed immature for such an advanced use-case. The proposed work rather aims at integrating learning-based building blocks in the control and sensing aspects of distributed robotic flocking, while keeping these functions partially decoupled in the overall architecture. Complex scenarios such as cluttered environments with obstacles will be considered.

<u>Proposed work:</u>

In this post doc, the proposed research will consist in studying the possibility, given a set/swarm of robots equipped with smart cameras, of making flocking behaviors emerge by combining:

1) at the intra-swarm level, local relative sensing based on the new capabilities of smart cameras and vision algorithms;

2) at the swarm level: a distributed control architecture based on recent Reinforcement Learning algorithms, e.g. starting from [7].

The work will start with a literature review on the relevant topics and implementation of state-of-the-art selected approaches as a comparison basis. A first milestone consists in learning from fully vision-based sensing and RL a flocking control scheme for multiple robots, with no communications, in order to evaluate the issues caused by their absence. The shaping of the RL reward could be inspired from model-based flocking algorithms such as [2,4]. A second milestone will consist in mitigating the identified shortcomings by gradually re-introducing communication capabilities, and developing the required algorithms to find control schemes in the new setting. A strong effort will be made to put the developed approaches into practice through real-world experiments on swarms of ground mobile robots first and then micro air vehicles, all equipped with appropriate vision sensors and on-board computational capacities (off-the-shelf embedded CPU / GPU) [15].

References:

1. C. W. Reynolds. Flocks, herds, and schools: A distributed behavioral model. *Computer Graphics, 25–34*, 1987.
2. Y. Rochefort, H. Piet-Lahanier, S. Bertrand, D. Beauvois, D. Dumur. Guidance of flocks of vehicles using virtual signposts. *Proceedings of the 18th IFAC World Congress*, 2011.
3. Heng Li, Jun Peng, Weirong Liu, Jing Wang, Jiangang Liu, and Zhiwu Huang. Flocking control for multi-agent systems with communication optimization. *American Control Conference (ACC), pages 2056–2061*, 2013.
4. Bonnefond, A., Simonin, O., Guérin-Lassous, I. “Extension of Flocking Models to Environments with Obstacles and Degraded Communications”. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 9139-9145)*, 2021.
5. Ahmad, A., Walter, V., Petráček, P., Petrlík, M., Báča, T., Žaitlík, D., Saska, M. “Autonomous aerial swarming in GNSS-denied environments with high obstacle density”. *IEEE International Conference on Robotics and Automation (ICRA) (pp. 570-576)*, 2021.
6. B. Fraser, R. Hunjet, and C. Szabo. Simulating the effect of degraded wireless communications on emergent behavior. *Winter Simulation Conference (WSC), pages 4081–4092*, 2017.
7. Wang, Chao, Jian Wang, and Xudong Zhang. "A deep reinforcement learning approach to flocking and navigation of UAVs in large-scale complex environments." *IEEE global conference on signal and information processing (GlobalSIP)*, 2018.
8. Hung, S. M., Givigi, S. N. “A Q-learning approach to flocking with UAVs in a stochastic environment”. *IEEE transactions on cybernetics, 47(1), 186-197*, 2016.
9. Zhu, P., Dai, W., Yao, W., Ma, J., Zeng, Z., Lu, H. “Multi-robot flocking control based on deep reinforcement learning”. *IEEE Access, 8, 150397-150406*, 2020.
10. Sampedro, C., Rodriguez-Ramos, A., Gil, I., Mejias, L., Campoy, P., “Image-based visual servoing controller for multirotor aerial robots using deep reinforcement learning”. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 979-986*, 2018.
11. Schilling, F., Lecoeur, J., Schiano, F., Floreano, D. “Learning vision-based flight in drone swarms by imitation”. *IEEE Robotics and Automation Letters, 4(4), 4523-4530*, 2019.
12. He, S., Xu, R., Zhao, Z., Zou, T. “Vision-based neural formation tracking control of multiple autonomous vehicles with visibility and performance constraints”. *Neurocomputing*, 2021.
13. Carrio, A., Vemprala, S., Ripoll, A., Saripalli, S., Campoy, P. “Drone detection using depth maps”. *IEEE/RSJ international conference on intelligent robots and systems (IROS), 1034-1037, 2018.*
14. Rojas-Perez, L. O., Martinez-Carranza, J. “Towards Autonomous Drone Racing without GPU Using an OAK-D Smart Camera”. *Sensors, 21(22), 7436,* 2021.
15. www.onera.fr/copernic

Collaborations extérieures

Olivier Simonin (INSA Lyon – CITI Lab – Inria CHROMA)

Laboratoire d’accueil à l’ONERA

Département : Traitement de l’Information et Systèmes

Lieu (centre ONERA) : Palaiseau

Contacts : Pierre FOURNIER, Sylvain BERTRAND

Emails : pierre.fournier@onera.fr, sylvain.bertrand@onera.fr