Coupling Deep Learning and Advanced Control in UAV navigation

Supervisor : Philippe Martinet, Co-supervisor : Guillaume Allibert, Patrick Rives

Inria Team CHORALE* in Sophia Antipolis

Research team:

Inria team CHORALE studies autonomous robotics systems (autonomous vehicles, mobile robots, UAV, and the association of heterogeneous robots), from the point of view perception and control, interacting in dynamic and complex environments often unknown a priori. The research objective is to propose and study new paradigms and concepts allowing autonomous systems i) to acquire and share a compact representation of the environment taking explicitly into account the constraints of the task to be done, ii) to act and interact in a collaborative way in order to realize safely and efficiently the task in shared environments with human respecting social behaviour rules. The CHORALE team follows the LAGADIC Sophia team which has ended in december 2017. The team members skills mainly concern perception, control, autonomous navigation and architecture for the control of the interactions with the environment. With a long expertise in model based approaches, CHORALE has been interested for two years in the contributions made by Deep Learning in the perception and navigation aspects of mobile robots [1].

Context :

Deep learning is at the heart of recent advances in the perception and control of mobile robots. The integration of these techniques into an autonomous robotic system with a perception system such as a 2D or 3D camera is an important issue. In so-called *end2end* approaches, deep learning is used to estimate the control of the system directly from the perception signal (image (s), Lidar, ...) [2]. Conventional control strategies are no longer used and the learning-based controller implicitly determines which information (s) is used to determine the command to be executed. If the results obtained with this kind of approach seem to be relevant and efficient, it is difficult to finely analyze the behavior of the perception/control coupling performed in the task (very high dimensional parameter space, difficulty of defining metrics). On the other hand, model-based approaches have, for their part, been studied for a long time and offer reliable and well-controlled performance analysis tools. However, they rely on a simplification of models (linearization, geometric sensors, Lambertian photometric model, etc.) which drastically limits their field of application. From a methodological point of view, one of the objectives of this thesis will be to design new approaches integrating both the benefits of deep learning and model-based methods, on the one hand, to broaden their field of application and to define new performance analysis tools that can integrate aspects of deep learning.

PhD subject :

The thesis work will be based on a particular application framework, namely, the navigation of an AUV in a forest environment not mapped beforehand. Very few works can be currently identified for this type of application [5] [6]. As a first step, a work on perception will have to be carried out to extract relevant information from a vision sensor useful for the navigation of the UAV. Various information will be evaluated (optical flow, disparity, semantics, pose, etc.) and will define the perception space used for control of the UAV. Model-based image processing methods will be compared to recent methods based on deep learning. In a second step, a prediction model for safe navigation of the UAV will have to be developed. This model will make it possible to determine free spaces of navigation allowing the evolution of the UAV according to the control vector. Several models will be considered in this thesis according to the information obtained during the perception phase. Multi-scale grids such as OctoMap will be studied in priority because they are particularly interesting in terms of calculation time and simplicity of representation. Like some recent works, we propose that the determination of this model is also realized by a deep learning. Finally, the control part will focus on combining the previously obtained model with an advanced control strategy of the predictive control type. Because of its efficiency and advantages, this control strategy [3] is regularly used to control robotic systems such as quadrotors. It has also already been extended to the case of visual servoing [4] where its interest has clearly been put forward. Work will then be carried out to define cost functions making it possible to exploit the data of the prediction model, of different nature and dimension. The navigation approach proposed in this thesis will then be compared with the *end2end* approach. A metric (distance traveled, percentage of success, number of crashes, etc.) as well as a protocol (distance to be covered, number of different routes, etc.) must be defined in order to evaluate the performance of the proposed approach.

*https://project.inria.fr/chorale/

Work plan :

The work plan of this thesis will take place in the following chronological way:

- Bibliographic study on perception and control parts based on Deep Learning;
- Development of perception algorithms based on Deep Learning or Model-Based, and performance analysis;
- Determination of prediction models for the safe navigation of the UAV;
- Study of visual predictive control strategy with a model obtained by learning;
- Realization of a demonstrator and implementation of the proposed approach and approach *end2end*;
- Comparison of results and performance evaluation.

Prerequisites :

The candidate, holder of a Master in Robotics or Automatic, will have to have solid skills in development (MATLAB, C / C ++, Python, LINUX, ROS, Git, OpenCV, TensorFlow). He will also have a taste for research, for fundamental work and multidisciplinary opening.

Contact :

Interested candidates must send a detailed CV, their Master's results (1 and 2) and one or more letters of recommendation to Philippe. Martinet@inria.fr, Allibert@unice.fr and Patrick.Rives@inria.fr.

Bibliography :

[1] E. Fernandez-Moral, R. Martins, D. Wolf, P. Rives, *A new metric for evaluating semantic segmentation: leveraging global and contour accuracy*, in IEEE- Intelligent Vehicle Symposium, June 2018.

[2] S. Levine, C. Finn, T. Darrell, P. Abbeel, *End-to-end Training of Deep Visuomotor Policies*, in Journal Machine Learning Researsh,1:17, 1334-1373, January 2016.

[3] D. Q. Mayne. *Model predictive control: Recent developments and future promise.* Automatica, 50(12):2967 – 2986, 2014.

[4] G. Allibert, E. Courtial, F. Chaumette, *Predictive Control for Constrained Image-Based Visual Servoing*, IEEE Trans. on Robotics, 26(5):933-939, October 2010.

[5] S. Ross, N. Melik-Barkhudarov, K. S. Shankar, A. Wendel, D. Dey, J. A. Bagnell, and M. Hebert. *Learning monocular reactive uav control in cluttered natural environments*, in IEEE International Conference on Robotics and Automation, pages 1765–1772, May 2013.

[6] M. Kisantal. *Deep reinforcement learning for goal-directed visual navigation*, 2018.