PhD Proposal
AUV Fault Detection and Control with Deep Reinforcement Learning

Benoit Clement, Gilles Le Chenadec, Eva Artusi, Karl Sammut, Paulo Santos

starting 2022

1 Introduction

Operating in a constantly perturbed marine environment, autonomous underwater vehicles (AUVs) must compensate for wave and current induced forces acting on their body. In this context, a common practice is to employ a PID type control law with fixed parameters that are typically obtained using model-based optimization theory. Although PID controllers can be made to work reasonably well under static known environment conditions, the performance of the fixed parameter controller may diminish under adverse conditions when exposed to large dynamic variations in the process. The performance of the PID regulators can be improved considerably by accommodating these variations using online PID regulator tuning techniques, such as adaptive control theory.

Adaptive control methods are widely used in the context of dynamic processes and provide, what seems to be, an ideal framework for automatic tuning of regulators. New tuning techniques are emerging from the development of data-driven theory. The well-known model-free adaptive method Extremum Seeking has been proven to compensate for uncharacterised and unmodelled process variations. Machine Learning based model-free adaptive methods have been suggested to tune PID regulators. Deep Reinforcement Learning (DRL) techniques in particular, have shown great performance as optimization methods for model-free adaptive scheme. They exploit the strong abilities of artificial neural networks (ANNs) to perform nonlinear mapping between sensor feedback and control inputs directly in closed-loop systems. Nevertheless, the stability analysis remains difficult to derive in the Artificial Neural Networks framework. There is no theoretical guarantee of performance and ANNs tends to be over parameterized which results in poor performance under real conditions.

A current joint Ph.D project [4, 20] conducted between ENSTA Bretagne and Flinders University in collaboration with Naval Group, is exploring new applications of Reinforcement Learning for Underwater vehicle control.

2 Fault detection and tolerance

Fault detection and control is an important issue for maintaining the reliability of an AUV. In the case of most AUVs, the vehicle scrubs its mission and resurfaces whenever any fault is detected, even non-critical failures such as a faulty depth sensor or inertial navigation sensor, etc. The objective of this project is to scope out the state-of-the-art and develop new algorithms for fault detection and control in AUVs that would enable a vehicle to continue its mission undettered even in the event of minor/major faults with the sensors and control surfaces. More concisely, the end goal for this project is to propose a learning-based approach that is able to discover new control policies to overcome thruster/steering-plane failures as they happen. The proposed approach can be:

- a model-based direct policy search that learns from a high fidelity simulated model of the AUV.
- a model-free approach based on a deep learning model.

The associative function of the ANN is used to recognize and detect the failures by observing the various changing parameters of the dynamic vehicle. The inferencing ability of an intelligent autonomous system suggests ways to control the failure and indicates the subsequent status of the vehicle. The entire system could be used as a low-level diagnostic tool in an overall control system for AUVs.

The proposed approach should be platform independant and thus capable of being applied to different types of underwater vehicles, because the theoretical aspect will be independent of the choice of the vehicle. As long as a dynamic model of the vehicle and its related hydrodynamic parameters are available [24], this proposed approach will be applicable.

3 Sim to Real

The use of high-fidelity simulation environments in robotics for learning suitable control behaviour or response actions using deep reinforcement learning techniques has in principle proven to be effective. In simulation, one can fully control the simulated environment, develop complex scenarios for learning and testing, all without compromising the physical integrity of the robot. Another advantage where robots are involved, is that conducting the learning within the simuator and then transferring the learned behaviour to the physical robot reduces the amount of time needed for the physical robot to learn a task. This could be highly advantageous for AUVs where field trials in open water can be both expensive to conduct and possibly dangerous. Nevertheless, while the simulation environment attempts to mimic the real environment as closely as possible, it never completely achieves this, hence it it still necessary for the transfer learning process to conduct some fine-tuning of the learned behaviour once it is tranferred to the real robot. For tasks that involve complex (nonlinear) dynamics, the fine-tuning itself may take a substantial amount of time. In order to reduce the amount of fine-tuning we propose to make the controllers robust through learning within the simulation environment. Controllers can be made robust through learning by exploiting the ability to change simulation parameters (both appearance and dynamics) for successive training episodes. An additional benefit of this approach is that it alleviates the precise determination of the physics parameters for the simulator, which is a non-trivial task.

4 Working plan

As a 3-year project, the PhD will focus on:

1. Reviewing the state of the art for fault detection and control using Reinforcement Learning techniques within a simulation environment;
2. Sim-to-Real theoretical issue and demonstration for an underwater vehicle;
3. investigation in Deep Reinforment Learning in order to answer the fault detection and control issue.
4. Experiments will be conducted in order to demonstrate how the proposed approaches can make the algorithms more efficient.

5 Research Training Environment

The PhD study will be jointly conducted between:

- Flinders University in Adelaide, SA;
- UMR CNRS Lab-STICC at ENSTA Bretagne in Brest;
- IRL CNRS CROSSING in Adelaide, SA;
- Naval Group Research in Ollioules.

The candidate will share his/her time between France (ENSTA Bretagne in Brest and Naval Group in Ollioules) and Australia (Flinders University and CROSSING in Adelaide, SA).

- Prof. Benoît CLEMENT, PhD, HDR
  ENSTA Bretagne - Lab-STICC UMR CNRS 6285
  `http://www.ensta-bretagne.fr/clement/`
- Prof. Karl SAMMUT, PhD
  Centre for Maritime Engineering, Flinders University
  `https://www.flinders.edu.au/people/karl.sammut`
- Eva ARTUSI, PhD, Naval Group

Gilles LE CHENADEC and Paulo SANTOS are also involved in the project.

6 Application

The PhD will start in October 2022 and the application documents are:

- a resume
- a cover letter
- a list of referees (2 or 3)

Please send your application to `benoit.clement@ensta-bretagne.fr` **before March 30th, 2022.**
Applications will be processed on a continuous basis.

References

[1] Steven L. Brunton and J. Nathan Kutz. *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press, 2019.

[2] Ignacio Carlucho, Mariano De Paula, Sen Wang, B. menna, Y. Petillot, and G. Acosta. Auv position tracking control using end-to-end deep reinforcement learning. *OCEANS 2018 MTS/IEEE Charleston*, pages 1–8, 2018.

[3] T. Chaffre, Gilles Le Chenadec, K. Sammut, E. Chauveau, and B. Clement. Direct adaptive pole-placement controller using deep reinforcement learning: Application to auv control. In *IFAC Conference on Control Applications in Marine Systems, Robotics and Vehicles*, Germany, 2021.

[4] Thomas Chaffre, J. Moras, A. Chan-Hon-Tong, J. Marzat, K. Sammut, G. L. Chenadec, and B. Clement. Learning-based vs model-free adaptive control of a MAV under wind gust. *ArXiv*, abs/2101.12501, 2021.

[5] Thor I. Fossen. *Handbook of Marine Craft Hydrodynamics and Motion Control*. Wiley, 2011.

[6] Wang Fu-li, Li Ming-zhong, and Yang Yinghuai. A neural network-based adaptive pole placement controller for nonlinear systems. *International Journal of Systems Science*, 28:415–421, 1997.

[7] T. Haarnoja, Aurick Zhou, P. Abbeel, and S. Levine. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *ICML*, 2018.

[8] T. Haarnoja, Aurick Zhou, Kristian Hartikainen, G. Tucker, Sehoon Ha, J. Tan, V. Kumar, H. Zhu, A. Gupta, P. Abbeel, and S. Levine. Soft actor-critic algorithms and applications. *ArXiv*, abs/1812.05905, 2018.

[9] Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *Proceedings of the 35th International Conference on Machine Learning*, pages 1861–1870. PMLR, 2018.

[10] Ronald A Howard. *Dynamic programming and markov processes*. John Wiley, 1960.

[11] A. Jayachitra and R. Vinodha. Genetic algorithm based pid controller tuning approach for continuous stirred tank reactor. *Adv. Artif. Intell.*, 2014, 2014.

[12] Diederik P. Kingma and J. Ba. Adam: A method for stochastic optimization. 2015.

[13] Kristoffer Borgen Knudsen, Mikkel Cornelius Nielsen, and I. Schjølberg. Deep learning for station keeping of auvs. In *MTS/IEEE OCEANS*, Seattle, 2019.

[14] T. Lillicrap, J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *CoRR*, abs/1509.02971, 2016.

[15] Timothy P. Lillicrap, Jonathan J. Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning, 2019.

[16] L. Lin. Reinforcement learning for robots using neural networks. 1992.

[17] Musa Manhães, Sebastian Scherer, Martin Voss, Luiz Douat, and Thomas Rauschenbach. UUV simulator: A gazebo-based package for underwater intervention and multi-robot simulation. In *MTS/IEEE OCEANS*, Monterey, 2016.

[18] Aidan O'Dwyer. *Handbook of PI and PID controller tuning rules*. Imperial College Press, 2006.

[19] M. Quigley. ROS: an open-source robot operating system. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2009.

[20] Y. Sola, T. Chaffre, K. Sammut, Gilles Le Chenadec, and B. Clement. Robust guidance and control of autonomous underwater vehicles with deep reinforcement learning. In *IEEE Oceans Conference*, Singapore, 2020.

[21] R. Sutton and A. Barto. Reinforcement learning: An introduction. *IEEE Transactions on Neural Networks*, 16:285–286, 2005.

[22] J. Tobin, Rachel H Fong, Alex Ray, J. Schneider, W. Zaremba, and P. Abbeel. Domain randomization for transferring deep neural networks from simulation to the real world. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 23–30, 2017.

[23] C. Wang, L. Wei, Zhaohui Wang, Min Song, and N. Mahmoudian. Reinforcement learning-based adaptive trajectory planning for auvs in under-ice environments. In *MTS/IEEE OCEANS*, Charleston, 2018.

[24] R. Yang, B. Clement, A. Mansour, M. Li, and N. Wu. Modeling of a complex-shaped underwater vehicle for robust control scheme. *Journal of Intelligent and Robotic Systems*, 2015.

[25] S. Zhang and R. Sutton. A deeper look at experience replay. *ArXiv*, abs/1712.01275, 2017.