Learning for aerial robots - Learning to coordinate

B(E)3M33MRS — Aerial Multi-Robot Systems

Ing. Robert Pěnička, Ph.D.

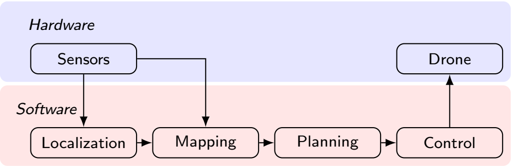
Multi-Robot Systems group, Faculty of Electrical Engineering
Czech Technical University in Prague

FACULTY OF ELECTRICAL ENGINEERING CTU IN PRAGUE | MRS MULTI-ROBOT SYSTEMS GROUP

Autonomous Aerial System Navigation Pipeline

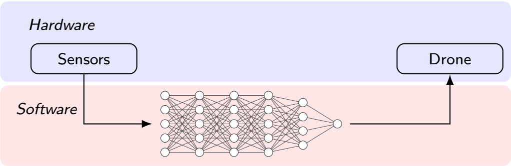
Remember lecture **4** — **Single aerial robot mapping and planning**.

Classical navigation pipeline



- **Separation of individual autonomy sub-tasks.**
- **Modularity** with respect to changes in individual parts.
- Modular to changes in hardware.

End-to-End navigation



- Replacement of some parts or entire navigation autonomy by **learned policies**.
- Usually **specialized** to certain scenarios and hardware.
- Learning need training environment (simulation) or **data**.

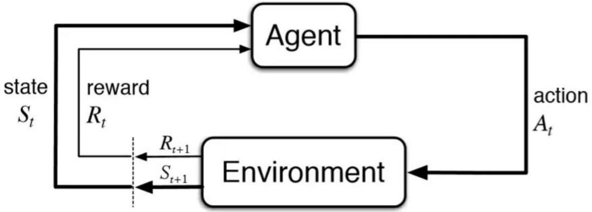
This time let us focus on the learning-based approaches.

Learning types

- **Supervised learning** — the learning approach is presented with example input and outputs and **tries to approximate such teacher**.
  Examples: human **labeled data** in computer vision tasks or **imitation** of model-based UAV navigation.

- **Unsupervised learning** — **learns patterns** in data without labels.
  Examples: **clustering** or learning distribution function (or Markov models for robot navigation).

- **Reinforcement Learning** — the robot/agent/program interacts and observes simulated environment and **learns to maximize some reward** to fulfill a certain goal.
  Examples: outplay opponent in game of GO, outrace other pilots in drone racing.

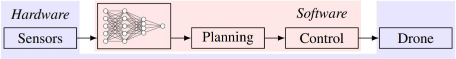


Reinforcement learning paradigm

Learning in aerial robotics

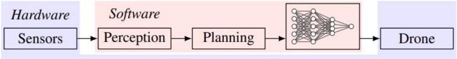
- **In many areas of the pipeline** used for the aerial robots.
- Localization and Mapping can be summarized as **Perception**.

Learned Perception



- Learned perception is **very common** in even standard navigation pipeline.
- Basically application of **Computer Vision** tasks in robotics.
- Examples: detection of **landmarks for localization**, detection of **humans to follow in cinematography**, **semantic segmentation** for scene understanding.

Learned Control

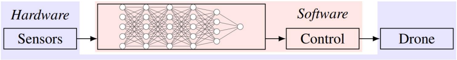


- Replaces the low-level control algorithms like PID or MPC by a **learned policy that follows given trajectory**.
- Can be faster than optimization-based solutions such as MPC.
- **Can optimize more complex objectives**, e.g., specified as reward in Reinforcement Learning.

[1] D. Hanover, A. Loquercio, L. Bauersfeld, A. Romero, R. Penicka, Y. Song, *et al.*, *Autonomous drone racing: A survey*, 2023. [Online]. Available: https://arxiv.org/abs/2301.01755

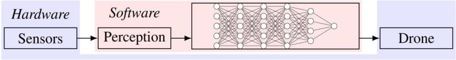
Learning in aerial robotics

Learned Planning & Perception



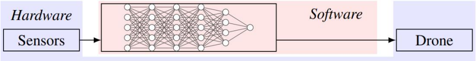
- Learns to **map sensory measurements directly to a plan of future motion**.
- Can simplify perception task without having globally-consistent metric state.
- Can be **robust against noise in perception**.

Learned Planning & Control



- Learns to directly **output control commands from estimated/given robot state** and environment knowledge.
- Does not require intermediate representation of trajectory which is then followed by a **control with usually different objective**.

End-to-End Flight



- **Replaces entire navigation pipeline with learned policy.**

Benefits and Disadvantages of Learning-based Approaches

Benefits

- Can easily identify patterns in training data.
- Can optimize/approximate even very complicated objective/functions.
- Can handle large-dimensional input/output data.
- Can be applied even without deeper understanding of data/phenomena.

Disadvantages

- Huge data sets required to learn.
- Need for annotation or realistic simulation.
- Almost impossible result interpretation.
- Certification only through testing.
- Usually requires GPUs which are not suitable to be on board UAV.

Test ChatGPT to see how far the learning got recently `https://chat.openai.com/chat`**. Does it pass your personal Turring test?**

Moravec's paradox

- **Moravec's paradox** is the observation that learning sensorimotor and perception skills is more challanging than reasoning.
- Formulated by Hans Moravec, Rodney Brooks, Marvin Minsky and others in 1980s.
- Original formulation: 'it is comparatively easy to make computers exhibit adult level performance on intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility' [2].

[2] H. Moravec, *Mind children: The future of robot and human intelligence*. Harvard University Press, 1988

Learned UAV acrobatic maneuvers

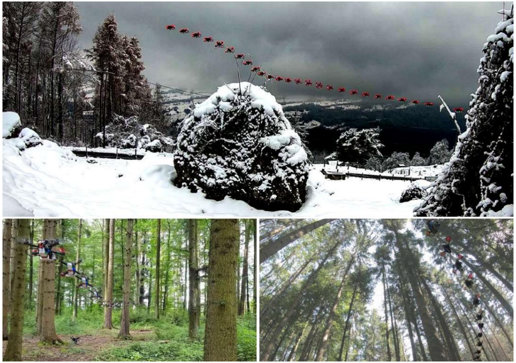
- **Learned Acrobatic manuevers** using only onboard sensors - Camera and IMU.
- **End-to-End approach** learned fully in simulation.
- Trained using **imitation learning** where optimal controller with privileged information serves as the teacher.

[3] E. Kaufmann, A. Loquercio, R. Ranftl, M. Müller, V. Koltun, and D. Scaramuzza, "Deep drone acrobatics," in *Proceedings of Robotics: Science and Systems*, Corvalis, Oregon, USA, 2020. DOI: 10.15607/RSS.2020.XVI.040

Video: https://youtu.be/2N_wKXQ6MXA

Learned high-speed collision avoidance

- **Learned Planning and Perception** where few distinct trajectories are created from sensory measurements.
- Map noisy sensory measurements directly to trajectory planning in receding horizon fashion.
- **Outperforms model-based methods in high-speed flight in cluttered environment.**
- The control is solved by using MPC on polynomial trajectory.
- Trained using **imitation learning** with sampling-based expert.

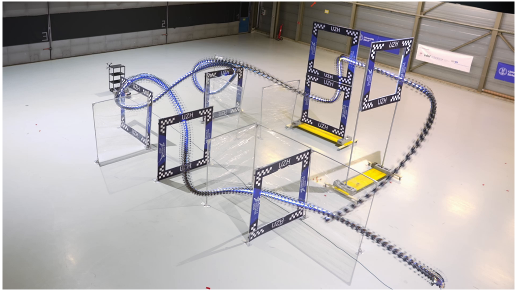


Video: https://youtu.be/m89bNn6RFoQ

[4] A. Loquercio, E. Kaufmann, R. Ranftl, M. Müller, V. Koltun, and D. Scaramuzza, "Learning high-speed flight in the wild," in *Science Robotics*, 2021

Reinforcement learning for minimum-time collision-free flight

- **Learned Planning and Control** for minimum-time flight in cluttered environment for drone racing.
- Uses **Reinforcement Learning** to learn policies in simulation that avoid collisions while minimizing the time of flight.
- Only a global guiding path is given and the policy outputs directly the body rate command and collective thrust of the UAV.

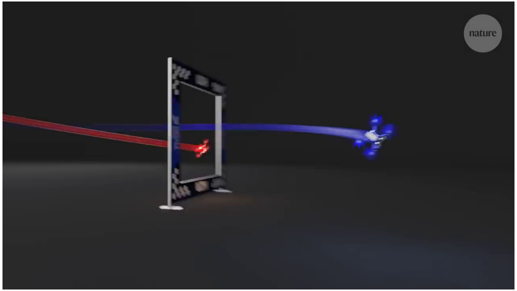


Video: https://youtu.be/wR1niZvI3pI

[5] R. Penicka, Y. Song, E. Kaufmann, and D. Scaramuzza, "Learning minimum-time flight in cluttered environments," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 7209–7216, 2022

Champion-level drone racing using deep reinforcement learning

- **Learned Planning and Control** using Deep Reinforcement Learning for drone racing.
- Uses known track layout to estimate drone's state using VIO and detection of the gates.
- **First AI-powered machine to beat human competitive sport.**
- Great example that the AI can not only dominate Chess, Go or computer games.

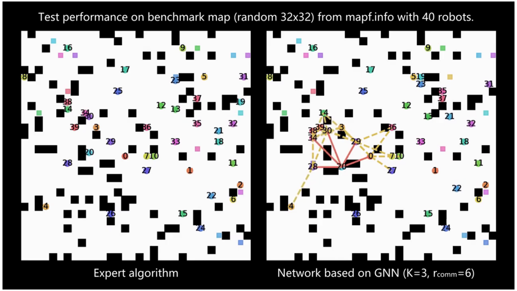


Video: `https://youtu.be/pq53uCDZelQ`

[6] E. Kaufmann, L. Bauersfeld, A. Loquercio, M. Müller, V. Koltun, and D. Scaramuzza, "Champion-level drone racing using deep reinforcement learning," *Nature*, vol. 620, no. 7976, pp. 982–987, 2023

Learned decentralized multi-robot path planning

- **Learned Planning & Control** for multi-robot path planning.
- Convolutional Neural Network (CNN) for creating features from local observations and Graph Neural Network (GNN) that communicates the features among robots.
- Trained to **imitate** an expert algorithm - Conflict-Based Search algorithm.
- Can solve the NP-hard multi-robot path planning problem with only local data.
- Can reach comparable performance to centralized solvers and scales better.
- Classical **Multilayer perceptron** (MLP) for creating the action — probability density of motion primitives.

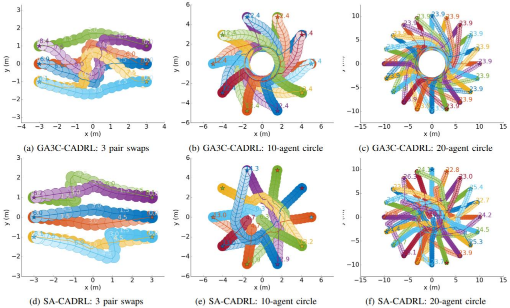


Video: `https://youtu.be/AGDk2RozpMQ`

[7] Q. Li, F. Gama, A. Ribeiro, and A. Prorok, "Graph neural networks for decentralized multi-robot path planning," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 11 785–11 792

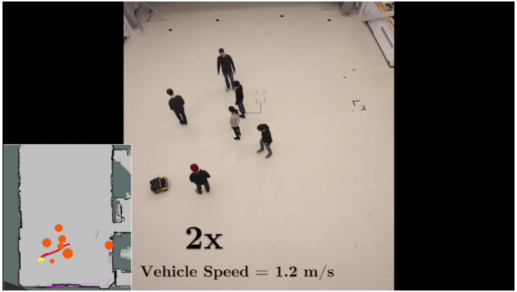
Reinforcement learning for multi-robot collision avoidance

- Trained in simulation with **deep reinforcement learning** to avoid dynamic obstacles such as other robots or humans.
- **Long short-term memory (LSTM)** allows to observe variable number of agents.
- Tested in real environment, with successful **sim-to-real transfer**.



(a) GA3C-CADRL: 3 pair swaps (b) GA3C-CADRL: 10-agent circle (c) GA3C-CADRL: 20-agent circle

(d) SA-CADRL: 3 pair swaps (e) SA-CADRL: 10-agent circle (f) SA-CADRL: 20-agent circle



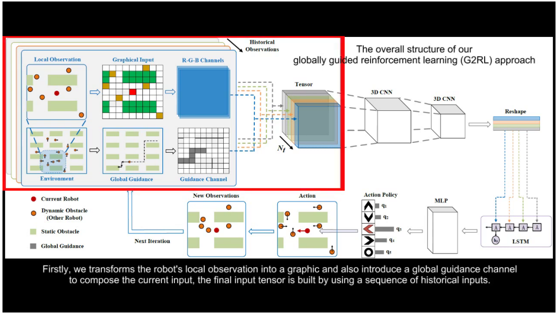
Video: `https://youtu.be/XHoXkWLhwYQ`

[8] M. Everett, Y. F. Chen, and J. P. How, "Motion planning among dynamic, decision-making agents with deep reinforcement learning," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Madrid, Spain, Sep. 2018

Learned decentralized multi-robot path planning

- **Learned Planning & Control** for multi-robot path planning.
- **Global-guided path** during training, i.e. A* guiding path.
- Uses **Reinforcement Learning** to stay close to the guiding path
- Learns to avoid other robots **without communication**.

[9] B. Wang, Z. Liu, Q. Li, and A. Prorok, *Mobile robot path planning in dynamic environments through globally guided reinforcement learning*, May 2020



Video: `https://youtu.be/JW39VQZwxKw`

Quadrotor Swarms control with Deep RL

- Drone swarm learned using RL to do flocking and collision avoidance.
- **Learned Planning and Control** for distributed swarming.
- Use of state encoders to create a representation of quadrotor's own state and state of neighboring quadrotors.
- Directly controls motor thrusts.

[10]S. Batra, Z. Huang, A. Petrenko, T. Kumar, A. Molchanov, and G. S. Sukhatme, "Decentralized control of quadrotor swarms with end-to-end deep reinforcement learning," in *CoRL*, 2021. [Online]. Available: https://arxiv.org/abs/2109.07735

Video: https://youtu.be/H2XI-Y_GCt0

Exam form

- **Exam will contain three questions!**
- **One question** will require deep knowledge of the topic that you are suppose to write down and explain in detail. These topics are listed in the next slides.
- **Two questions** will be more about having overview of the field and can be from any topics described during lectures.
- **There will be preparation period, where you can write down notes about the first know-in-detail question, before the oral exam starts.**

Exam topics

Topics for detailed question:

- **Multirotor UAV system dynamics and control:** Describe the dynamics system of a multirotor UAV. What notable states do we define in the system description, and how are they connected? Attempt to draw the diagram.
- **State estimation:** How are the various states in the UAV dynamics chain typically estimated? What sensors are used to measure the states?
- **Kalman filters:** LKF, EKF and UKF. Describe their properties and differences. How would you use a Kalman filter to estimate a hidden state?
- **Odometry:** What is odometry? What are the possible ways of obtaining odometry on a drone? What are the limitations and benefits of odometry?
- **Localization:** What are possible ways to localize a single drone in 3D? What is SLAM? How can we solve the SLAM problem?
- **MRS:** Make a sketch and explain decentralized and centralized systems. Discuss their pros and cons.
- **Formations:** Show how to compute desired states of followers derived from a leader position. Consider holonomic and none-holonomic (e.g. car-like) models.
- **Swarms:** Explain Boids model and show how to compute future motion direction for a focal agent within a swarm in an environment with obstacles.
- **Swarms:** Explain a local neighborhood in swarms. Sketch and explain two models of the local neighborhood.

Exam topics

- **Cooperative localization:** Show how to estimate distance of an object using a monocular camera model. Discuss possible errors and show how to compute them.
- **Failure recovery:** Draw and explain examples of robust, partially vulnerable, and fully vulnerable multi-rotor platforms.
- **Differential flatness:** Make a sketch of quadrotor and world reference frames. Explain how we can get quadrotor thrust and attitude from derivatives of its position.
- **Mapping:** How is a map stored in occupancy maps. Write how is the map updated with new sensory measurement. How does it relates to sensor model.
- **Multi-robot planning:** Make a sketch of velocity obstacle and explain how to create Optimal Reciprocal Collision Avoidance (ORCA) lines. How do you find optimal collision avoidance speed with ORCA.
- **Exploration:** How does the frontier-based exploration works. Explain Wavefront frontier detection algorithm.
- **Task assignment:** Write down and explain individual steps of the Hungarian method.
- **Learning for aerial robots:** What types of learning you can use on aerial robots? What data are typically used for training and what parts of navigation pipelines can be replaced with learned policies?

Now lets do the lab tour

References I

[1] D. Hanover, A. Loquercio, L. Bauersfeld, A. Romero, R. Penicka, Y. Song, G. Cioffi, E. Kaufmann, and D. Scaramuzza, *Autonomous drone racing: A survey*, 2023. [Online]. Available: https://arxiv.org/abs/2301.01755.

[2] H. Moravec, *Mind children: The future of robot and human intelligence*. Harvard University Press, 1988.

[3] E. Kaufmann, A. Loquercio, R. Ranftl, M. Müller, V. Koltun, and D. Scaramuzza, "Deep drone acrobatics," in *Proceedings of Robotics: Science and Systems*, Corvalis, Oregon, USA, 2020. DOI: 10.15607/RSS.2020.XVI.040.

[4] A. Loquercio, E. Kaufmann, R. Ranftl, M. Müller, V. Koltun, and D. Scaramuzza, "Learning high-speed flight in the wild," in *Science Robotics*, 2021.

[5] R. Penicka, Y. Song, E. Kaufmann, and D. Scaramuzza, "Learning minimum-time flight in cluttered environments," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 7209–7216, 2022.

[6] E. Kaufmann, L. Bauersfeld, A. Loquercio, M. Müller, V. Koltun, and D. Scaramuzza, "Champion-level drone racing using deep reinforcement learning," *Nature*, vol. 620, no. 7976, pp. 982–987, 2023.

[7] Q. Li, F. Gama, A. Ribeiro, and A. Prorok, "Graph neural networks for decentralized multi-robot path planning," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 11 785–11 792.

[8] M. Everett, Y. F. Chen, and J. P. How, "Motion planning among dynamic, decision-making agents with deep reinforcement learning," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Madrid, Spain, Sep. 2018.

[9] B. Wang, Z. Liu, Q. Li, and A. Prorok, *Mobile robot path planning in dynamic environments through globally guided reinforcement learning*, May 2020.

[10] S. Batra, Z. Huang, A. Petrenko, T. Kumar, A. Molchanov, and G. S. Sukhatme, "Decentralized control of quadrotor swarms with end-to-end deep reinforcement learning," in *CoRL*, 2021. [Online]. Available: https://arxiv.org/abs/2109.07735.