UDC 004.89

REINFORCEMENT LEARNING FOR UAV NAVIGATION AND QUAD-ROTOR CONTROL: TRENDS, ACHIEVEMENTS AND RESEARCH CHALLENGES

Honcharenko V.D.
e-mail: vadym.honcharenko@nure.ua
Scientific Supervisor – Cand. Sc. (Tech.), Ass. Prof. Yesilevskyi V.S.
Kharkiv National University of Radio Electronics,
Department of Applied Mathematics
Kharkiv, Ukraine

Навчання з підкріпленням (Reinforcement learning – RL) є перспективним підходом до навігації та керування квадрокоптерами в складних та невизначених умовах. На відміну від класичних методів, RL дозволяє приймати рішення та керувати безпосередньо на основі взаємодії з навколишнім середовищем, зменшуючи залежність від точних моделей та ручного налаштування. Останні дослідження показують його ефективність у навігації за маршрутними точками, енергетично інформованому плануванні місій та перегонах автономних дронів. Водночас залишаються виклики щодо надійності, обчислювальної ефективності та роботи в умовах відсутності GPS.

1. Introduction

Quadrotor navigation can be understood as the overall process that achieves goal of moving vehicle from start point to finish with some constraints (e.g. obstacle, passing waypoint, time optimality). In contrast, quadrotor control is the process of generating command signals that make the vehicle follow the desired motion or trajectory. In practical systems, navigation is therefore a higher-level task that includes perception, localization, path planning, and decision-making, while control is responsible for stabilizing and actuating the vehicle [1].

Classical quadrotor control usually uses an outer control loop that generates high-level commands and an inner loop that tracks them. Common methods include PID, MPC, and sliding-mode control for higher-level motion regulation, together with lower-level PID control for actuator-level execution. However, such approaches require careful modelling, manual tuning, and expert knowledge, and their fixed structures may become less effective when the environment or operating conditions change.

RL is attractive because it offers a more flexible alternative. In RL, an agent observes the state of the environment, takes an action, and receives a reward that reflects success or failure. For UAVs, this means that sensor data can be mapped directly to motion decisions or even actuator commands. Compared with many classical approaches, RL can reduce reliance on detailed prior modelling and can learn policies for tasks where perception, planning, and control are tightly coupled [1, 2].

2. Main RL applications in UAVs and quadrotors

A broad review of the field shows that RL is already used in multiple UAV navigation domains, especially obstacle avoidance, path planning, and autonomous motion in complex environments. Survey [2] notes that model-free RL dominates most navigation studies, while training is often performed in simulation because UAVs have limited onboard computation and crashes are expensive. This makes RL especially relevant for quadrotor research, where safe data collection in the real world is difficult.

Another relevant application area is navigation in GPS-denied or partially observable environments. In such scenarios, quadrotors must rely on onboard sensors such as cameras, inertial measurement units, or range sensors to estimate the environment and make motion decisions. This makes RL attractive because it can learn policies that directly connect perception and action under uncertainty. In this context, RL complements classical navigation pipelines by enabling more adaptive behavior in cluttered and previously unseen environments, where complete maps or accurate localization may not be available [2, 5].

One representative quadrotor application is waypoint navigation. In [1] proposed a TD3-based method for both high-level and low-level quadrotor control and showed that both learned policies can guide the vehicle through assigned waypoints. Their results indicate that the low-level controller is more accurate in nominal conditions, while the high-level controller performs better under disturbances. This is important because it shows that RL can be applied at different layers of the quadrotor control hierarchy.

Another important direction is energy- or endurance-aware mission planning. In [3] authors formulate coverage path planning under varying power constraints and learned a DDQN-based control policy that generalizes across random start positions and varying movement budgets. Their results show that the agent can balance safe landing and coverage objectives, which is especially relevant for UAV missions where available energy directly limits reachable area and mission duration.

At the high-performance end, RL has also demonstrated remarkable capability in autonomous drone racing. In [4] reported that the Swift system competed against three human champions in real-world head-to-head races, won several races, and achieved the fastest recorded race time while relying on onboard sensing and computation. This result shows that RL-based quadrotor autonomy is not limited to slow or simplified tasks but can also operate near the dynamic limits of the platform.

3. Current challenges

Despite latest achievements, challenges remain. Autonomous UAV navigation is a system-level problem, not only a control problem, and must be evaluated using broader criteria such as path length, deviation rate, and exploration efficiency. In addition, current reviews continue to highlight difficulties related to computational power, energy efficiency, robustness, and deployment in realistic

GPS-denied environments. For quadrotors, this means that future RL methods should be assessed not only by tracking success, but also by mission efficiency, safety, and resource consumption [5].

A further challenge concerns sim-to-real transfer and safety assurance. Although simulation is essential for RL training, policies learned in simulation may lose performance in real flights because of modelling errors, sensor noise, actuator limitations, wind disturbances, and mismatch between training and deployment conditions. In addition, RL results strongly depend on reward design, training scenarios, and the balance between exploration and safety constraints. Therefore, for practical quadrotor applications, future work should focus not only on improving task performance, but also on reliable transfer to real platforms, stable behaviour under disturbances, and predictable operation in safety-critical missions [2, 4, 5].

4. Conclusion

Reinforcement learning is now one of the most promising approaches for quadrotor navigation and control. It offers clear advantages for tasks where classical methods require strong modelling assumptions, extensive tuning, or rigid control structures. Existing studies show that RL can support waypoint navigation, energy-constrained mission planning, and high-speed autonomous flight. For a PhD scientific forum, the key conclusion is that RL should be viewed not only as an alternative controller, but as a broader framework for integrating perception, planning, and control in autonomous quadrotor systems.

References:

1. Himanshu K., Kumar H., Pushpangathan J. V. Waypoint Navigation of Quadrotor using Deep Reinforcement Learning. *IFAC-PapersOnLine*. 2022. Vol. 55, no. 22. Pp. 281–286. DOI: 10.1016/j.ifacol.2023.03.047.

2. AlMahamid F., Grolinger K. Autonomous Unmanned Aerial Vehicle Navigation using Reinforcement Learning: A Systematic Review. *Engineering Applications of Artificial Intelligence*. 2022. Vol. 115. P. 105321. DOI: 10.1016/j.engappai.2022.105321.

3. Theile M., Bayerlein H., Nai R., Gesbert D., Caccamo M. UAV Coverage Path Planning under Varying Power Constraints using Deep Reinforcement Learning. *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (Las Vegas, NV, USA, 24 October 2020 – 24 January 2021). Pp. 1444–1449. DOI: 10.1109/IROS45743.2020.9340934.

4. Kaufmann E., Bauersfeld L., Loquercio A., Müller M., Koltun V., Scaramuzza D. Champion-level drone racing using deep reinforcement learning. *Nature*. 2023. Vol. 620, no. 7976. Pp. 982–987. DOI: 10.1038/s41586-023-06419-4.

5. Chang Y., Cheng Y., Manzoor U., Murray J. A review of UAV autonomous navigation in GPS-denied environments. *Robotics and Autonomous Systems*. 2023. Vol. 170. P. 104533. DOI: 10.1016/j.robot.2023.104533.