Czech Technical University in Prague

Faculty of Electrical Engineering
Department of Cybernetics
Multi-robot Systems

Learning High-Speed Flight of Unmanned Aerial Vehicle in Cluttered Environments

Bachelor's Thesis

Vít Knobloch

Prague, May 2023

Study programme: Open Informatics
Specialization: Artificial Intelligence and Computer Science

Supervisor: Ing. Robert Pěnička, Ph.D.

Declaration

I declare that the presented work was developed independently and that I have listed all sources of information used within it in accordance with the methodical instructions for observing the ethical principles in the preparation of university theses.

Vt Knobloch
May 24, 2023 in Prague

Acknowledgments

Firstly, I would like to thank my supervisor for his guidance, helpfulness, and kindness with which he led my work throughout the last year. Secondly, I would like to thank my parents and my family, who has supported me in everything my whole life and as well through my university studies. I would also like to thank my classmates who have made all the studying a bit easier and a lot more fun for me. The access to the computational infrastructure of the OP VVV funded project CZ.02.1.01/0.0/0.0/16_019/0000765 \Research Center for Informatics" is also gratefully acknowledged.

Abstract

Unmanned Aerial Vehicles (UAVs) and particularly quadrotors have become increasingly popular in recent years. Their high versatility allows their deployment in many tasks such as infrastructure inspection, farmland monitoring, delivery, and reconnaissance. Consequently, their autonomous capabilities are being researched to decrease the need for human supervision. Many of the tasks where quadrotors are used require high-speed flight such as search and rescue, cinematography, protected airspace guarding, and more. This thesis describes a method using policy trained with deep reinforcement learning to perform perception-aware high-speed flight of a quadrotor in simulated environments cluttered with obstacles. A computationally efficient learning environment is presented and paired with an existing reinforcement learning library to train the policy. The approach is tested in several environments on a number of tracks and the resulting behavior is described. The policy is able to control the quadrotor flight through multiple waypoints in a cluttered environment while using depth sensor information to avoid collisions.

Keywords Unmanned aerial vehicles, Automatic control, Simultaneous planning and control, Deep reinforcement learning, Quadrotor, Learning environment, Simulator, High-speed flight, Perception-aware flight

Abstrakt

Bezpilotn vzdusne prostredky a zejmena kvadrokoptery se v poslednch letech zacaly hojne vyuzvat, jejich obratnost a vsestranost z nich delaj idealn prostredky pro vykon nejruznejsch ukolu jako inspekce, mapovan a prohledavan terenu, monitorovan ruznych prostor, hldan strezeneho vzdusneho prostoru a mnoho dalsch. V navaznosti na to je provaden rozsahly vyzkum v oblasti jejich autonomnch schopnost, aby bylo mozne tyto ukoly automatizovat. Pro mnoho z techto uloh je nezbytna schopnost dronu pohybovat se rychle ve slozitych prostredch. Tato prace popisuje metodu vyuzvajc hluboke posilovane ucen k naucen agenta schopneho ovladat vysokorychlostn let kvadrotoru v simulovanem prostred s prekazkami. Bylo vytvoreno trenovac prostred a napojeno na existujc knihovnu s implementacemi algoritmu hlubokeho posilovaneho ucen. Chovan nauceneho agenta je otestovano v nekolika prostredch a na ruznych trasach. Nauceny agent je schopny ovladat dron pri letu vysokou rychlost ve slozitych prostredch skrze nekolik zadanych clu a vyuzvat data ze simulovaneho hloubkoveho senzoru, aby se vyhnul prekazkam.

Klcova slova Bezpilotn prostredky, Automaticke rzen, Soucasne planovan a rzen, Hluboke posilovane ucen, Kvadrokoptera, Trenovac prostred, Simulator, Vysokorychlostn let, Let s vedomym vnmanm

Abbreviations

FOV	field of view
GPS	global positioning system
LiDAR	light detection and ranging
UAV	unmanned aerial vehicle
PPO	proximal policy optimization
TRPO	trust region policy optimization
KL	Kullback–Leibler divergence
MDP	Markov decision process
POMDP	partially observable Markov decision process
RL	reinforcement learning
DRL	deep reinforcement learning
CNN	convolutional neural network
ReLU	rectified linear unit
FW	fixed-winged
RW	rotary-winged
VTOL	vertical takeoff and landing
SDF	signed distance field
LV	linear velocity and yaw rate
CTBR	collective thrust and bodyrate
SRT	single-rotor thrust
RCI	Research Center for Informatics

Contents

Chapter 1

Introduction

Unmanned aerial vehicles (UAVs) became very popular over the course of the twenty-first century with an ever-increasing amount of research into their hardware and autonomous capabilities. Conventionally UAVs are telemetry monitored and operated by a pilot on the ground, either nearby or far away [2]. Lately, more focus is put on the development of autonomous control and planning algorithms to limit the need for human supervision. First UAVs have been large aircraft with military applications. But with the advancement of motors, GPS receivers, telemetry sensors, and microcontrollers, new smaller UAVs have been created and found applications in many areas including photography and filmmaking, infrastructure inspection, recreation, search-and-rescue, express delivery, and surveillance [2].

There are two main types of UAVs: fixed-winged (FW) and rotary-winged (RW). Fixed-winged design is simpler and more energy efficient and generally capable of carrying larger payloads. However, FW UAVs are incapable of vertical takeoff and landing (VTOL) and have to move forward in order to retain lift, this makes them impractical or unusable for many use cases. Rotary-winged UAVs have rotors that provide thrust, they are more agile, able to take off and land vertically, and hover in one place. Even though RW UAVs are less energy efficient, their advantages have made them the most popular design [2].

The thesis is about the use of reinforcement learning for simultaneous planning and control of perception-aware high-speed flight of an autonomous quadrotor with an onboard depth sensor. We say the agent is perception-aware when it actively uses sensor data about the surrounding environment to gain situational awareness. An introductory video showing some key components of our approach is available in Appendix A and on YouTube[1]. A quadrotor is a rotary-winged UAV with four rotors. Opposing pairs of rotors rotate in the same direction while neighboring rotors rotate in opposite directions. This allows for easy attitude control as yaw, pitch, and roll can be controlled by just changing the rotational speeds of individual rotors. Quadrotors are widely used and probably the most popular and researched design of UAVs. Planning and control for high-speed flight of autonomous quadrotors in unknown environments remains an open problem despite a lot of effort put into the research. Physical models have limited accuracy, and so do sensors onboard the quadrotor. The computational resources available online onboard are also limited and more computation means less battery life and as a result less flight time. Cluttered environments and fast flight pose another difficulty as such navigation requires highly agile maneuvering.

Existing solutions, that are efficient enough to compute onboard in flight, mostly use polynomial trajectory planning [11] or imitation learning [1], [7]. Polynomial trajectories are suboptimal for minimum-time flight because of the inherent smoothness of polynomials. Imitation learning has shown promising results when a sampling-based teacher was used in [7] to train a neural policy network to fly in unknown environments at speeds unreachable by polynomial trajectories. Most of the existing methods output a trajectory to be followed by

---

[1] Introductory video on YouTube: https://youtu.be/7mR8IvWtQWU