Agile Drone Path Planning Based on Reinforcement Learning Algorithms

Summary

Candidate:
Afshin Zeinaddini Meymand
Supervisors:
prof. Giorgio Guglieri
Dr. Francesco Marino

July 2023

Over the last decade, the use of autonomous drone systems has seen a significant increase in various industries such as surveying, search and rescue, and last-mile delivery. These systems rely on various algorithms for trajectory planning, which are designed to navigate in different environments. However, most of the algorithms developed for trajectory planning are dedicated to static environments, where all objects other than the autonomous vehicle remain fixed during the system's operation.

One of the major challenges that arise when working with autonomous drone systems is the dynamic nature of the environment. When the environment is not completely static, and other objects such as the goal object are moving, it requires the implementation of different algorithms for various tasks such as object detection, state estimation, and trajectory planning. These algorithms must be able to accurately detect and track the moving objects, estimate their state, and plan a trajectory that avoids collisions while still reaching the goal object. This is a complex task that requires advanced techniques and algorithms.

There are several solutions available for state estimation of moving objects in dynamic environments, one of which is the use of Visual-Inertial Odometry (VIO) cameras. VIO cameras are specialized cameras that are specifically designed for state estimation tasks by providing precise and accurate 3D tracking data. They work by using multiple cameras to capture images of markers placed on the object of interest. These markers are small, highly reflective, and typically placed on the object in a known pattern. The VIO system uses advanced algorithms and image processing techniques to track the markers in the camera images, even in challenging conditions such as low light or fast motion. The system compares the images captured by each camera and calculates the distance between the markers by using the principle of triangulation, thus providing a precise and accurate 3D position and orientation of the object. Additionally, VIO cameras also use an Inertial Measurement Unit (IMU) which is a device that measures linear and angular accelerations, and magnetic fields. The IMU sensor works in conjunction with the cameras to provide additional information about the object's movement. The IMU sensor measures the angular velocity and the linear acceleration of the object, and then the data is fused with the visual data obtained from the cameras to improve the estimate of the object's state. Overall, the use of VIO cameras in autonomous drone systems provides a robust and efficient solution for state estimation of moving objects in dynamic environments, as it is able to track the object's position and orientation in real-time, even under challenging conditions.

The use of Visual Odometry sensors for state estimation in model predictive control for trajectory planning of autonomous drones in dynamic environments can be challenging, as there are two main problems that need to be addressed. These problems are:

- Handling the continuous action spaces: The action space for controlling the drone's trajectory is continuous and high-dimensional, making it difficult to find the optimal policy using traditional techniques.

- Dealing with uncertainty and non-stationary environments: The environment in which the drone operates is dynamic and uncertain, making it challenging to predict the future state of the system and plan a trajectory that avoids collisions by VIO systems.

One solution to these problems is to use Kalman Filters for state estimation, by assuming that the position of the gate centers can be extracted by computer vision algorithms without knowing anything about the dynamic model of the gates. Kalman Filters are a powerful tool for state estimation in dynamic systems, as they can handle non-linear systems and estimate the state of the system even in the presence of uncertainty and noise. It is widely used in various fields such as control systems, navigation, and robotics. Once the position of the gate centers is estimated using Kalman Filters, the trajectory of the drone for passing through that center can be derived. This trajectory can then be fed to different Advanced Actor-Critic Reinforcement Learning Algorithms such as Deep Deterministic Policy Gradient(DDPG), Soft-Actor Critic (SAC), and Proximal Policy Optimization (PPO) to derive the best policy. These algorithms are well-suited for trajectory planning tasks in dynamic environments as they can handle continuous action spaces, uncertainty, and non-stationary environments. RL algorithms are used to learn the optimal policy for an agent to make decisions based on the system's state. RL algorithms use a trial-and-error approach to learn the optimal policy by exploring different actions and receiving feedback in the form of rewards. When Kalman filtering and RL algorithms are combined, the Kalman filter can provide accurate and precise state estimates, while the RL algorithm can learn the optimal policy for the agent to make decisions based on these estimates. The combination of these two techniques can lead to improved performance in tasks such as trajectory planning, control, and decision-making. One of the key challenges in RL is defining the reward function, which can be different for each task and environment. The mixed Kalman Filter and RL algorithms address this challenge by defining the reward function based on the distance of the current position of the drone with respect to the predicted position of the center of the gate.

The built model of the KFRL algorithm is a continuous state-action environment, which means that the state of the system and the actions taken by the agent are continuous variables. In this model, the agent can take any action in a continuous range of values, rather than only a discrete set of actions. This allows for more flexibility and precision in controlling the drone's trajectory.