Reinforcement Learning Approaches for Autonomous Navigation and Control in Robotic and Aerial Systems

T.Nikil Prakash, gajanan
Vishwanath Ghuge, Anup Ingle
St. Joseph's College, Vishwakarma Institute
of technology

Reinforcement Learning Approaches for Autonomous Navigation and Control in Robotic and Aerial Systems

[1]T.Nikil Prakash, Assistant Professor, Information Technology, St. Joseph's College (Autonomous)Tiruchirappalli-02,MailID: nikilprakash1992@gmail.com/nikilprakasht_it2@mail.sjctni.edu

[2]gajanan Vishwanath Ghuge, assistant professor /DESH , Vishwakarma Institute of technology pune-37, Mail ID: gajanan.ghuge@vit.edu

[3]Anup Ingle, Assistant Professor, Department of Electronics and Telecommunication Engineering, Vishwakarma Institute of Technology, Pune, India, Mail ID: anup.ingle@vit.edu

Abstract

Autonomous navigation and control in robotic and aerial systems have witnessed substantial advancements through reinforcement learning, enabling intelligent decision-making in dynamic and uncertain environments. This chapter presents a comprehensive investigation of model-free, model-based, hierarchical, and multi-agent reinforcement learning frameworks for robotic and aerial applications. Key theoretical foundations, including Markov decision processes, policy optimization, value functions, and exploration-exploitation trade-offs, are analyzed to provide a structured understanding of learning dynamics. Practical implementations encompassing Q-learning, deep deterministic policy gradient, actor-critic architectures, and policy gradient methods are discussed in the context of trajectory optimization, adaptive control, and task decomposition. Simulation environments, including Gazebo, AirSim, and PyBullet, are examined for training and evaluation, with a focus on domain randomization and sim-to-real transfer to address challenges in real-world deployment. Sensor integration strategies leveraging LiDAR, cameras, and IMUs, coupled with hardware-in-the-loop experiments, are highlighted to ensure robustness, precision, and reliability. Challenges in scalability, communication, and coordination in multi-agent systems are addressed, providing insights into cooperative UAV navigation and swarm robotics. The chapter concludes by outlining future directions for enhancing sample efficiency, safety, and adaptability in reinforcement learning-driven autonomous systems.

Keywords: Reinforcement Learning, Autonomous Navigation, Robotic Systems, Aerial Systems, Multi-Agent Coordination, Sim-to-Real Transfer

Introduction

Autonomous navigation and control in robotic and aerial systems have become increasingly critical in modern applications ranging from industrial automation and warehouse logistics to unmanned aerial surveillance and disaster response [1]. The growing complexity of operational environments demands adaptive decision-making strategies capable of handling dynamic,

uncertain, and partially observable conditions [2]. Reinforcement learning (RL) has emerged as a powerful paradigm for addressing these challenges by enabling agents to learn optimal policies through iterative interactions with the environment [3]. Unlike traditional control methods that rely on precise models or pre-defined heuristics, RL provides flexibility to develop policies that adapt to non-linear dynamics, sensor noise, and unmodeled disturbances [4]. The integration of RL into robotic and aerial platforms facilitates high-dimensional control, real-time decision-making, and autonomous task execution, laying the foundation for systems capable of performing complex maneuvers without human intervention. Research in this domain emphasizes both theoretical understanding and practical implementation, ensuring that learned policies can be generalized across multiple scenarios while maintaining reliability and safety in real-world deployments [5].

Theoretical foundations form the backbone of reinforcement learning approaches, providing formal tools to model decision-making under uncertainty [6]. Markov decision processes (MDPs) define the structure for sequential decision problems, capturing state transitions, actions, and associated rewards in a probabilistic framework [7]. Policy optimization and value function estimation form core components of RL, guiding agents to maximize cumulative expected rewards while balancing exploration and exploitation [8]. Exploration strategies allow agents to discover effective behaviors, whereas exploitation ensures that learned policies capitalize on successful actions [9]. Bellman equations and temporal-difference learning provide mathematical formulations for evaluating policies and propagating value estimates across states, enabling efficient learning in both discrete and continuous domains. These foundational concepts underpin advanced RL methodologies, including model-free and model-based approaches, which differ in how they incorporate knowledge of system dynamics and plan for long-term objectives. A deep understanding of these principles is essential for designing reinforcement learning architectures capable of robust performance in high-dimensional robotic and aerial systems [10].