Trajectory Planning of Biplane Quadcopter using Reinforcement Learning

Ayush Gupta (170191)[1] and Vivek Agrawal (170808)[2]

I. PROBLEM STATEMENT

We propose to find a methodology for a biplane-quadcopter to learn a forward transition maneuver using reinforcement learning methods iteratively. Tailsitter biplanes can work in two different modes: hover and high-speed forward flight. The transition between these two stages is challenging due to the complexity of the aerodynamic model's accurate modeling. We take inspiration from iterative learning methods and combine it with the superior non-linear function approximation provided by neural networks to model the autonomous flight properly. Simulation for the physics modeling of the biplane would be created for training the actor-critic RL models.

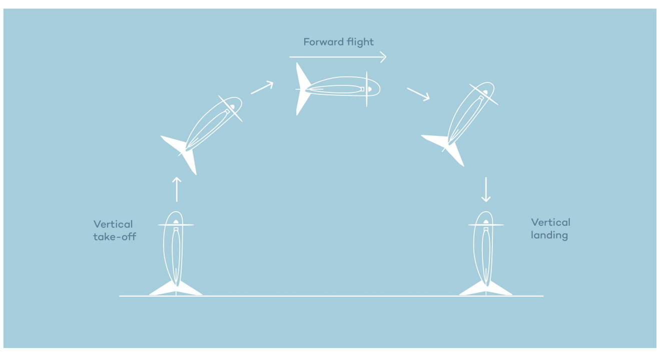

Fig. 1. Transition of biplane-tailsitter from hover to cruise mode and back to hover

II. LITERATURE SURVEY

The detailed literature survey comprising of all papers looked into is attached here.

A. *Reinforcement Learning for Aerial Vehicles*

There is much literature on aerial vehicles' path-planning, where we wish to find an optimal path between the source and the destination. [1] enlists the various techniques for path-planning algorithms. Finding the optimal path for a swarm of drones is an advanced problem for which reinforcement learning has been effectively utilised. [11, 17, 3, 4] talk about the application of deep reinforcement learning in the path-planning of autonomous base stations.

Attitude control of quadcopters using reinforcement learning is also a well-researched problem. [7, 12, 13] talk about using RL for attitude control while performing tight maneuvers such as flips at high-velocities. [8] develops a complete RL system integrated with Gazebo for teaching control to high-speed aerial drones. [9] develops a universal control strategy based on non-linear function approximation that can allow us to control hybrid UAVs. They evaluate their results on different non-conventional aerial vehicles such as double-wing, asymmetric tail sitter, and quad-planes. To overcome the sim-to-reality gap, they use a combination of the orientation cost and a novel integral block structure.

[5] generates trajectories for autonomous helicopter using an expectation-maximization inspired algorithm called policy learning by weighted exploration with the returns. They test the algorithm against trajectory planning tasks such as catching a moving target, obstacle avoidance, and coordinated formation flight. [14] develops a grey-wolf based optimizer algorithm for path planning across multiple 3D obstacles in a tight space using reinforcement learning.

B. *Transition control for biplane quadcopters*

[15] develops an iterative-learning based feed-forward control for the transition of a biplane quadrotor tail sitter. They use a geometric attitude controller, valid for all flight models, to track the pitch angle. The maneuver is controlled by regulating the pitch angle and propeller thrust according to feed-forward control laws parameterized by polynomials. [2] systematically develops an optimal control strategy for quadcopter biplanes, where the transition maneuver is challenging to achieve reliably with traditional controllers due to its exquisite dynamics.

Methods for modeling the trajectory control of a transitioning quadrotor biplane tailsitter are developed in [16, 6]. [10] develops a method for automating the flight of a fixed-wing UAV through highly constrained environments.

III. SOLUTION APPROACH

A reinforcement learning solution would mainly depend on modeling the environment, the agent, and the reward function. Using non-linear methods allows us to relax the complex modeling of the biplane dynamics, and will be intrinsically learned by the entire algorithm. Thus, we need to model primary gravitational and drag forces and proper implementation for forces and moments. The reward function would be the key to generating a smooth transition from state H (hover) to state F (forward flight) and will need to be experimented with.

Our goal is to reach a particular point, starting from a random point, velocity, and orientation. The plane reaches the goal position while taking care of all kinematic and

[1]Dept. Of Mechanical Engineering ayushgup@iitk.ac.in
[2]Dept. Of Electrical Engineering vivekagr@iitk.ac.in

Milestone	Time Required	Date
Simulator Design	1 week	30th September
Basic RL setup with basic implementation	2 weeks	14th October
Optimisation of implemented algorithms	2 weeks	31st October
Finalised results and reports	2 weeks	14th November
Buffer time and Mid-Semester exam	1-1.5 weeks	-

TABLE I
PROPOSED TIMELINE

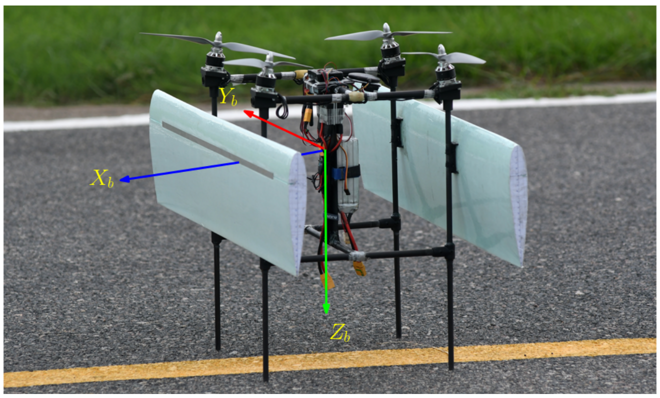

Fig. 2. Biplane quadrotor with body-frame axes

dynamic constraints. For this first, we need a deep learning algorithm that can process continuous states and can give continuous outputs. We can generate random targets for training purposes and let our Deep learning agent learn to generate trajectory from the current state vector to the target position using the Deep Deterministic Policy Gradient (DDPG) algorithm in the simulator.

IV. SIMULATION & CODE

A. Simulator

For easy integration with the reinforcement learning pipeline, we will be writing the code in Python. We will build upon the necessary code for the aerodynamic model available on MATLAB and convert it to Python. This would be used for simulating the environment's effect on the aircraft, with gravity and drag included in it. Since we are focusing on the control process's outer-loop mechanics, we will use a simple control scheme to move from one via-point to another. Our environment (simulation setup) needs to maintain decent fidelity at low computational costs to speed up each experiment's training times. It would be a large bounding box where the agent (or biplane) can properly navigate between via-points.

B. Reinforcement Learning

Reinforcement learning would be done on PyTorch or Tensorflow and trained using publicly available OpenAI baselines for state-of-the-art implementations of these algorithms. Tensorflow Board would be used for proper visualization of the training procedures. We would be using our own Graphical Processing Units for training these models, which can cause a delay due to lower memory and FLOPs supported by them.

ACKNOWLEDGMENT

We thank Prof. Mangal Kothari for giving us the opportunity to work on this exciting idea.

References

[1] Shubhani Aggarwal and Neeraj Kumar. "Path planning techniques for unmanned aerial vehicles: A review, solutions, and challenges". In: *Computer Communications* 149 (2020), pp. 270–299. ISSN: 1873703X. DOI: 10.1016/j.comcom.2019.10.014. URL: https://doi.org/10.1016/j.comcom.2019.10.014.

[2] Mangal Kothari Animesh Kumar Shastry. In: (2019).

[3] Harald Bayerlein, Paul De Kerret, and David Gesbert. "Trajectory Optimization for Autonomous Flying Base Station via Reinforcement Learning". In: *IEEE Workshop on Signal Processing Advances in Wireless Communications, SPAWC* 2018-June (2018). DOI: 10.1109/SPAWC.2018.8445768.

[4] Yun Chen et al. "A Deep Reinforcement Learning Approach to Efficient Drone Mobility Support". In: (2020), pp. 1–9. arXiv: 2005.05229. URL: http://arxiv.org/abs/2005.05229.

[5] Xiao Guo. "Trajectory generation using reinforcement learning for autonomous helicopter with adaptive dynamic movement primitive". In: *Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering* 231.6 (2017), pp. 495–509. ISSN: 20413041. DOI: 10.1177/0959651816684427.

[6] Vikram Hrishikeshavan, Derrick Yeo, and Inderjit Chopra. "Onboard flow sensing in a quad rotor biplane micro air vehicle for transition between hover and steady-level flight". In: *American Helicopter Society International - 6th AHS International Specialists' Meeting on Unmanned Rotorcraft System 2015: Platform Design, Autonomy, Operator Workload Reduction and Network Centric Operations* August (2015), pp. 152–165.

[7] J. Hwangbo et al. "Control of a Quadrotor with Reinforcement Learning". In: *IEEE Robotics and Automation Letters* 2.4 (2017), pp. 2096–2103. ISSN: 23773766. DOI: 10.1109/LRA.2017.2720851. arXiv: 1707.05110.

[8] William Koch et al. "Reinforcement Learning for UAV Attitude Control". In: *ACM Transactions on Cyber-Physical Systems* 3.2 (2019), pp. 1–21. ISSN: 2378962X. DOI: 10.1145/3301273. arXiv: 1804.04154.

[9] *Learning To Fly*. http://cfg.mit.edu/content/learning-fly-computational-controller-design-hybrid-uavs-reinforcement-learning.

[10] Joshua M. Levin, Meyer Nahon, and Aditya A. Paranjape. "Real-time motion planning with a fixed-wing UAV using an agile maneuver space". In: *Autonomous Robots* 43.8 (2019), pp. 2111–2130. ISSN: 15737527. DOI: 10.1007/s10514-019-09863-2. URL: https://doi.org/10.1007/s10514-019-09863-2.

[11] Kai Li et al. "Deep Reinforcement Learning for Real-Time Trajectory Planning in UAV Networks". In: *2020 International Wireless Communications and Mobile Computing, IWCMC 2020* (2020), pp. 958–963. DOI: 10.1109/IWCMC48107.2020.9148316.

[12] Sergei Lupashin and Raffaello D'Andrea. "Adaptive fast open-loop maneuvers for quadrocopters". In: *Autonomous Robots* 33.1-2 (2012), pp. 89–102. ISSN: 09295593.

[13] Huy Xuan Pham et al. "Reinforcement Learning for Autonomous UAV Navigation Using Function Approximation". In: *2018 IEEE International Symposium on Safety, Security, and Rescue Robotics, SSRR 2018* (2018), pp. 1–6. DOI: 10.1109/SSRR.2018.8468611.

[14] Chengzhi Qu et al. "A novel reinforcement learning based grey wolf optimizer algorithm for unmanned aerial vehicles (UAVs) path planning". In: *Applied Soft Computing Journal* 89 (2020), p. 106099. ISSN: 15684946. DOI: 10.1016/j.asoc.2020.106099. URL: https://doi.org/10.1016/j.asoc.2020.106099.

[15] Nidhish Raj et al. "Iterative Learning based feedforward control for Transition of a Biplane-Quadrotor Tailsitter UAS". In: *IEEE International Conference on Robotics and Automation (ICRA)* (2020), pp. 321–327.

[16] Jean Paul Reddinger et al. "Modeling and trajectory control of a transitioning quadrotor biplane tailsitter". In: *The Vertical Flight Society - Forum 75: The Future of Vertical Flight - Proceedings of the 75th Annual Forum and Technology Display* (2019).

[17] Tianze Zhang et al. "Hybrid path planning of a quadrotor UAV based on Q-Learning algorithm". In: *Chinese Control Conference, CCC* 2018-July (2018), pp. 5415–5419. ISSN: 21612927. DOI: 10.23919/ChiCC.2018.8482604.