Proceedings of AIAA SCITECH 2022 Forum, 3-7 January 2022, San Diego, CA and Virtual Event
Paper number AIAA 2022-0879, DOI:10.2514/6.2022-0879

Transition Flight Control System Design for Fixed-Wing VTOL UAV: A Reinforcement Learning Approach

Burak Yuksek* and Gokhan Inalhan†
Cranfield University, United Kingdom, MK43 0AL

Tilt-rotor vertical takeoff and landing aerial vehicles have been gaining popularity in urban air mobility applications because of their ability in performing both hover and forward flight regimes. This hybrid concept leads energy efficiency which is quite important to obtain a profitable and sustainable operation. However, inherent dynamical nonlinearities of the aerial platform requires adaptation capability of the control systems. In addition, transition flight phase should be planned carefully not only for a profitable operation but also for a safe transition between flight regimes in the urban airspace. In this paper, transition flight phase of a tilt-rotor vertical-takeoff-and-landing unmanned aerial vehicle (UAV) is studied. Low-level flight control systems are designed based on adaptive dynamic inversion methodology to compensate aerodynamic effects during the transition phase. Reinforcement learning method is utilized to provide safety and energy efficiency during the transition flight phase. An actor-critic agent is utilized and trained by using deep deterministic policy gradient algorithm to augment the collective channel of the UAV. This augmentation on the collective input is used to adjust flight path angle of the UAV which results in adjusting the angle of attack when pitch angle is zero. By using this relationship, it is proposed to generate aerodynamic lift force and perform transition flight with minimum altitude change and energy usage. Simulation results show that the agent reduces the collective signal level as the aerodynamic lift force is created in the descent flight phase. This affects overall system efficiency, reduces operational costs and makes the enterprise more profitable.

I. Introduction

As the operational concepts and infrastructure developments in the urban area have been evolving in the last decade, integration of the unmanned aerial vehicles (UAVs) into the daily life has been accelerated. These developments have been initiators of an emerging multi-disciplinary research topic called urban air mobility (UAM) which defines the aerial operations such as passenger and cargo transportation that are performed in the urban airspace.

From the operational safety point of view, trustworthiness of the UAM applications is a crucial requirement as the aerial vehicles operate in or near the cities where the population density is high. This means that the aerial vehicles should have a certain degree of reliability and they should be able to operate safely even in the emergency situations as a result of faults and failures. Flight control systems should be designed to meet several resilience requirements to handle with the unexpected situations such as effectiveness loss on actuators or sensors when the vehicle is airborne. Fault tolerant and adaptive control algorithms and estimation methods provide powerful and proven techniques to obtain flight safety of these aerial vehicles.

In addition to the safety considerations, UAM applications should be sustainable and profitable for an operator company which means that operational and infrastructure costs should be at a manageable level to provide a continuous service. This requires optimized mission plans and efficient mobile platforms. When we look closer to this requirement, it can be divided into subtopics such as path planning, air traffic management and optimal usage of the infrastructure facilities (i.e. charging stations, etc.) and energy sources (i.e. batteries, hybrid power plants, etc.). Optimal usage of the onboard energy sources is directly related to the optimality of the generated control signal by the flight control system which should be designed by considering both energy consumption and flying quality requirements of the aerial vehicle.

Fixed-wing vertical-takeoff-and-landing (VTOL) design is one of the most effective aerial vehicle concepts for the UAM applications. They combine efficiency of the fixed-wing platforms and hover capability of the rotary-wing platforms in one airframe [1, 2]. An example design for fixed-wing VTOL UAV is given in Figure (1). Here, fixed-wing

*Postdoctoral Research Fellow, School of Aerospace, Transport and Manufacturing, burak.yuksek@cranfield.ac.uk, AIAA Professional Member.
†BAE Systems Chair, Professor of Autonomous Systems and Artificial Intelligence, School of Aerospace, Transport and Manufacturing, inalhan@cranfield.ac.uk, AIAA Associate Fellow.

Published by AIAA. This is the Author Accepted Manuscript issued with: Creative Commons Attribution Non-Commercial License (CC:BY:NC 4.0).
The final published version (version of record) is available online at DOI:10.2514/6.2022-0879. Please refer to any applicable publisher terms of use.

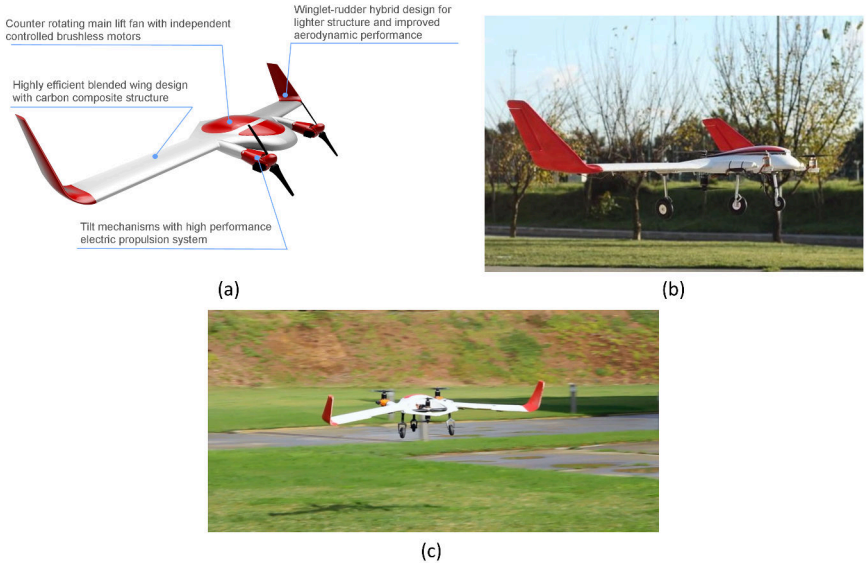


Fig. 1 General overview of the fixed-wing VTOL UAV. a) Conceptual design, b) Flight test of the 1/2 scale prototype, c) Flight test of the 1/3 scale prototype.

concept is beneficial for long-range flights with minimum energy consumption. On the other hand, rotary-wing concept has vertical-takeoff and landing capability that provides ability to land limited-size landing areas in the urban environment. Although there are several advantages of the fixed-wing VTOL UAVs, they always require more attention to model their complex dynamics and to design a suitable flight control system that should be able to handle inherent nonlinearities of the aerial vehicle. Tilting propeller group, which is mounted on the leading edge of the wings, is the most dominant nonlinearity source on this aerial vehicle concept. It generates high-speed, propeller-induced airflow which affects on the wing and/or fuselage section of the airframe. As a result of this accelerated airflow, transition flight maneuver becomes a complex problem and requires an optimal approach to minimize the energy usage while providing the flight safety. Hence, before trying to solve the transition flight problem, it is important to generate a detailed mathematical model of the aerial platform which represents these nonlinearities and complex dynamics of the aerial platform. In this context, detailed mathematical modeling process of a fixed-wing VTOL UAV was performed by the authors and it is given in [3, 4].

Although tilt-rotor aircraft concept has been commonly used in UAM applications such as air taxis because of its effectiveness, it can be quite complicated to perform transition flight phase due to safety and efficiency considerations. The first issue that should be considered is adaptation capability of the low-level flight control systems. As aircraft accelerates and dynamic pressure increases, aerodynamic forces and moments become effective on the airframe. Flight control system should be able to compensate these dynamical changes to provide safe flight in the transition phase. The second issue about the operational safety is maneuver planning in the urban airspace which requires consideration of the minimum distance from urban infrastructures and noise regulations. If it is required to perform transition flight in a dense environment, these should be considered while planning the transition maneuver. The third issue is about system efficiency which is directly connected with profitability of the enterprise. If onboard energy sources could be utilized efficiently during the transition flight phase, operational costs would be reduced significantly which is quite critical for continuity of the enterprise. Beside of the safety and efficiency restrictions during the transition flight, passenger comfort should also be considered by avoiding aggressive maneuvers for hover-to-transition and transition-to-cruise flight regimes. Several transition strategies have been used in practical applications for UAVs to perform these switches. In these strategies, UAVs are accelerated horizontally by pitching-down and tilt-rotors are set to cruise flight configuration immediately after reaching a certain level of airspeed. This sudden change in angle of the tilt-rotors will also create additional pitching moment on the aircraft body. This strategy could be a simple and effective solution for many tilt-rotor UAVs but it may not provide the most comfortable flight experience for the passengers in the context of urban air mobility operations.

In literature, there are many studies on transition flight that have been conducted for the last three decades. The common aim of these studies is designing a flight control system for the aerial platform to perform an effective hover-

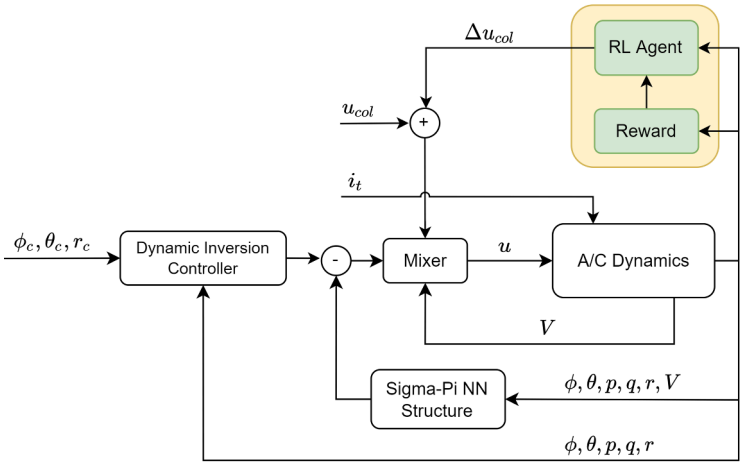


Fig. 2 General structure of the RL-aided transition flight control system.

to-cruise (forward transition) and cruise-to-hover (back transition) maneuvers. In [5], dynamic inversion (DI)-based flight control systems are designed for the transition phase of the two different VTOL aircraft. Nonlinear models are generated and proposed closed-loop systems are analyzed. In [6], it is proposed to develop an efficient transition method for a small variable-incidence-wing UAV to perform the transition with minimum altitude variation.Transition between the hover and forward flight is defined as fixed-time, two-point boundary value problem and it is solved by utilizing sequential quadratic programming technique. In [7], autonomous transition is proposed for a mid-scale fixed-wing UAV by utilizing a hybrid control architecture. For hover and cruise flight phases, linear optimal controllers are designed. A nonlinear controller is developed for the transition flight phase to provide the stability of the system. In addition, to provide the robustness of the system against the exogenous disturbances, reference maneuvers are generated for the transition flight phase. In [8], as a continuation study of the [7], a recovery controller is also developed and integrated with the previous study to provide the flight safety in the transition phase. In [9], authors focused on the transition phase control system for a quadrotor fixed-wing UAV. A Lyapunov-based nonlinear control system is designed. It is shown that the proposed controller provides stability in the sense of Lyapunov theory for the whole envelope from hover to forward flight. In [10], a continuous and easy-to-implement control algorithm is proposed to stabilize the transition flight of a tail sitter UAV.

In recent years, developments in the computer technologies have been accelerated and resulted in productions with high power density. It means that researchers and engineers have been able to solve complex problems on powerful computers with small physical dimensions. This has made it possible to use these powerful and small-size computers on the aerospace vehicles and produce fast solutions for the problems. Transition flight is an interesting and complex problem to be solved and it has also been investigated by the researchers from the artificial intelligence (AI) domain. In [11], it is proposed a mode-free, model-agnostic neural network controller design for hybrid UAVs. Training of the NN is performed by utilizing the reinforcement learning method. In this application, the developed controller structure could be used on different types of hybrid UAV concepts such as tail-sitter, X-wing and quad-wing designs. Developed controllers are implemented on the Pixhawk autopilot hardware which uses the Ardupilot flight software. Several flight tests are performed on quad-wing, tail-sitter and X-wing UAVs to evaluate the system performance. Although it is one of the first studies in which the transition flight problem is solved by utilizing AI-based tools, it is not applied on a tilt-rotor VTOL UAV which has highly nonlinear dynamics because of the propeller-induced airflow. In addition, no adequate analysis of the proposed method is performed and no comparison study is given with a baseline transition flight control algorithm.

In this study, we focused on developing an efficient and safe transition strategy for hover-to-cruise flight regime of a tilt-rotor aerial vehicle. As we mentioned before, two main issues arise about the transition flight problem of the tilt-rotor concepts. The first one is adaptation of the flight control systems in the presence of aerodynamic effects and the second one is designing an energy efficient transition flight strategy. For these issues, a) adaptive dynamic inversion (ADI)-based attitude-control-attitude-hold (ACAH) and rate-control-attitude-hold (RCAH) flight control systems are implemented to provide adaptation ability for the low-level closed-loop structure in the presence of aerodynamic effects

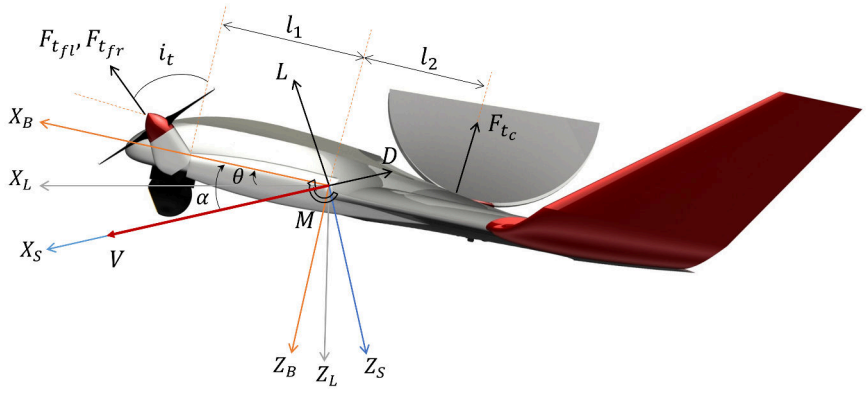


Fig. 3 Aerodynamic and propulsion forces/moments on the UAV.

on the airframe. b) an RL-aided transition flight strategy for tilt-rotor aerial vehicles is developed in which collective input is utilized to control the altitude of the vehicle and to generate aerodynamic lift force by adjusting flight path angle while the pitch attitude is kept zero. In this strategy, tilt-angle is controlled manually and follows a ramp input during the transition phase. The RL-agent is responsible to generate aiding collective input to increase and decrease the total collective signal. If sufficient lift force is created, the agent decreases the total collective signal level to decrease power consumption. General structure of the RL-based transition flight control system is given in Figure (2).

The main contribution of this paper is developing a reinforcement learning (RL)-aided transition flight control system for a fixed-wing, tilt-rotor VTOL platform. To the authors' knowledge, this is one of the first study in which RL methodology is utilized to solve the optimal transition problem in tilt-rotor aerial vehicle concepts. The proposed algorithm is applied on a 6 degrees-of-freedom (DoF) nonlinear mathematical model of a fixed-wing tilt-rotor UAV for conceptual studies but it could also be extended for full-scale aerial platforms used in UAM applications.

The remaining of the paper is structured as follows; in Section II, preliminaries are given about mathematical model, ADI control system and reinforcement learning method. In Section III, general structure of the RL-aided transition flight controller is introduced, reward function, action and observation vectors are explained. In Section IV, simulation results are presented and discussed. In Section V, concluding remarks are given and possible future works are explained.

II. Preliminaries

Transition phase of tilt-rotor airframes have highly nonlinear dynamics due to the propeller induced airflow effects on the UAV airframe. As the front propellers are being tilted down, accelerated airflow passes through the airframe and generates additional aerodynamic forces and moments on the UAV body. After a certain tilt angle, propeller induced airflow becomes dominant on the UAV dynamics and this directly affects the flight characteristics and transition performance of the UAV. To avoid excitation of the propeller-induced airflow effects on the UAV, the Assumption-1 is defined as;

Assumption A1 *During the transition regime, operation interval of the tilt angle is set as* $i_t \in [0, 30°]$. *It follows a ramp signal profile as given in Eq. (1).*

$$i_t(t) = a\,t \tag{1}$$

where $a = 3\,deg/sec$ defines the slope of the ramp signal and t is time in seconds. According to this relationship, tilt angle achieves its maximum value for the transition phase (i.e. 30°) in 10 seconds.

A. Longitudinal Dynamics

In the transition flight phase of the UAV, dynamical uncertainties are quite dominant on the longitudinal motion of the vehicle because of large variations in total airspeed, angle-of-attack and propeller-induced airflow direction. As a result, deviations in aerodynamic forces and moments become larger as the tilt angle and total airspeed increase. Propulsion force generated by the tilt-rotor system is also affected by the airspeed and directly creates additional

challenge in control effectiveness of the propulsion system. Because of these situations, this study primarily focuses on the longitudinal dynamics of the UAV, as given in Eq. (2-7), to solve the transition flight problem.

$$\dot{U} = -QW - g\sin(\Theta) + \frac{F_{A_x} + F_{T_x}}{m} \tag{2}$$

$$\dot{W} = QU + g\cos(\Phi)\cos(\Theta) + \frac{F_{A_z} + F_{T_z}}{m} \tag{3}$$

$$\dot{Q} = \frac{m_{A_y} + m_{T_y}}{I_{yy}} \tag{4}$$

$$\dot{\Theta} = Q \tag{5}$$

$$\dot{p}_n = U\cos(\Theta) + W\sin(\Theta) \tag{6}$$

$$\dot{p}_d = -U\sin(\Theta) + W\cos(\Theta) \tag{7}$$

where g is gravity acceleration, m is mass of the UAV, I_{yy} is moment of inertia on Y_b axis, $U, W \in \mathbb{R}$ are linear velocity vectors on X_b and Z_b axes, $Q \in \mathbb{R}$ is angular velocity on Y_b axis, $\Theta \in \mathbb{R}$ is pitch attitude, $p_n, p_d \in \mathbb{R}$ are north and east position in local frame. $F_{T_x}, F_{T_z}, m_{T_y} \in \mathbb{R}$ are thrust forces and moments on the body axis and they are defined in Eq. (8-10).

$$F_{T_x} = (F_{t_{f_l}} + F_{t_{f_r}})\sin(i_t) \tag{8}$$

$$F_{T_z} = (F_{t_{f_l}} + F_{t_{f_r}})\cos(i_t) + F_{t_c} \tag{9}$$

$$m_{T_y} = (F_{t_{f_l}} + F_{t_{f_r}})\cos(i_t)l_1 - F_{t_c}l_2 \tag{10}$$

where l_1 is distance between tilt-rotor assembly hinge and center of gravity (cg), l_2 is distance between coaxial fan group and cg, $F_{t_{f_l}}, F_{t_{f_r}} \in \mathbb{R}$ are thrust forces generated by left and right tilt rotors, respectively. $F_{t_c} \in \mathbb{R}$ is thrust force generated by coaxial fan group, $i_t \in [0°, 90°]$ is tilt angle of the front propeller group. In the equations of motion, $F_{A_x}, F_{A_z}, m_{A_y} \in \mathbb{R}$ are aerodynamic forces and moments on the longitudinal plane and defined as given in Eq. (11-13).

$$F_{A_x} = -D\cos\alpha + L\sin\alpha \tag{11}$$

$$F_{A_z} = -D\sin\alpha - L\cos\alpha \tag{12}$$

$$m_{A_y} = M \tag{13}$$

where $\alpha \in \mathbb{R}$ is angle of attack. $D, L, M \in \mathbb{R}$ are aerodynamic drag force, lift force and pitching moment, respectively, and they are defined in Eq. (14-16).

$$D = \left(\frac{1}{2}\rho V^2\right) S\, C_D \tag{14}$$

$$L = \left(\frac{1}{2}\rho V^2\right) S\, C_L \tag{15}$$

$$M = \left(\frac{1}{2}\rho V^2\right) S\, \bar{c}\, C_M \tag{16}$$

where ρ is air density, $V \in \mathbb{R}$ is total airspeed, S is reference area, $\bar{c}$ is mean aerodynamic chord. C_D, C_L and C_M are aerodynamic drag force, lift force and pitching moment, respectively, and they are defined as given in Eq. (17-19).

$$C_D = C'_D(\alpha) + C_{D_q}\frac{c}{2V}q + C_{D_{\delta_e}}\delta_e \tag{17}$$

$$C_L = C'_L(\alpha) + C_{L_q}\frac{c}{2V}q + C_{L_{\delta_e}}\delta_e \tag{18}$$

$$C_M = C'_M(\alpha) + C_{m_q}\frac{c}{2V}q + C_{m_{\delta_e}}\delta_e \tag{19}$$

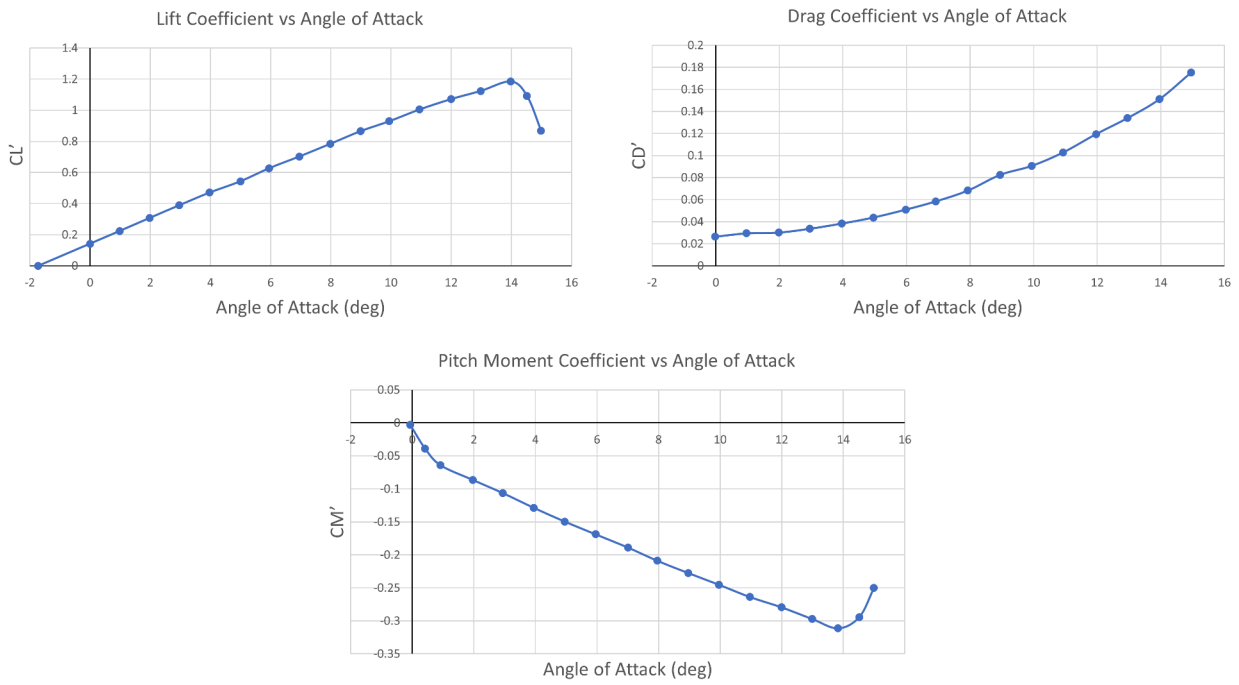


Fig. 4 $C'_D(\alpha), C'_L(\alpha), C'_M(\alpha)$ **aerodynamic coefficients of the UAV.**

Here, $C'_D(\alpha), C'_L(\alpha)$ and $C'_M(\alpha)$ are calculated by utilizing computational fluid dynamics (CFD) analysis and they are given in Figure (4). Rest of the aerodynamic stability and control coefficients are obtained by utilizing XFLR5 and Advanced Aircraft Analysis (AAA) software packages. For more information about mathematical modeling of the tilt-rotor UAV, readers may refer to [3].

B. Adaptive Dynamic Inversion for Low-level Control Loop

Tilt-rotor aerial platforms have quite complex dynamics when compared to conventional fixed-wing and rotary-wing systems from many perspectives. First, aerodynamic and propulsion system characteristics are quite different in hover and cruise flight regimes which requires a proper mixing logic of the control effectors. Second, in hover flight, aerodynamic forces and moments created by the fuselage could be neglected because of low dynamic pressure. However, aerodynamic effects become dominant in the transition and cruise flight regimes as dynamic pressure increases because of the increased total airspeed. These dynamical variations affect reference signal tracking performance and should be considered in design phase of the low-level attitude controllers.

Neural-network (NN) augmented dynamic model inversion control structure, i.e. adaptive dynamic inversion (ADI), is utilized as the low-level attitude controller for the UAV [12, 13]. Here, dynamic inversion controller design process is applied for the hover flight phase at the sea level conditions of the tilt-rotor UAV. If the UAV operates in any other flight conditions, i.e. transition and cruise flight phases, this will result in inversion error which will be compensated by the neural network on the feedback path. While attitude-command/attitude-hold (ACAH) augmentation is utilized for roll and pitch attitude control systems, yaw axis control system is developed based on rate-control/attitude-hold (RCAH) augmentation. General structure of the ADI-based flight control system is given in Figure (5).

1. Command Filter Design

Command filters are used in roll, pitch and yaw channels to produce desired angular rates and angular accelerations for given reference inputs. Second-order filters are used for roll and pitch channels (Eq. (20-21)) as they are designed in ACAH structure.

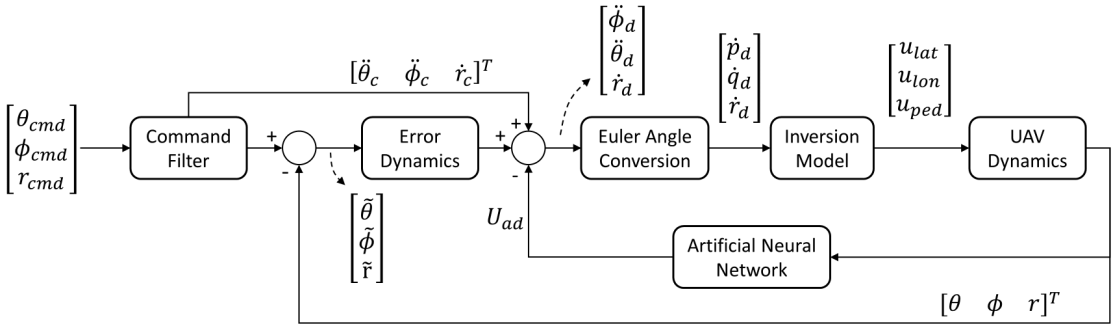


Fig. 5 ADI-based attitude controller structure.

$$\ddot{\phi}_c + 2\xi_\phi \omega_\phi \dot{\phi}_c + \omega_\phi^2(\phi_c - \phi_{cmd}) = 0 \tag{20}$$

$$\ddot{\theta}_c + 2\xi_\theta \omega_\theta \dot{\theta}_c + \omega_\theta^2(\theta_c - \theta_{cmd}) = 0 \tag{21}$$

where θ_{cmd}, ϕ_{cmd} are applied command signals and θ_c, ϕ_c are filtered command signals. Here, damping ratios and natural frequencies in the command filter are selected as $\xi_\phi = \xi_\theta = 0.7$ and $\omega_\phi = \omega_\theta = 3\,rad/s$ for roll and pitch axis, respectively.

Command filter for the yaw channel is designed as first-order filter (Eq. (22)) as the structure of the yaw channel is rate-command-attitude hold (RCAH).

$$\tau_r \dot{r}_c + (r_c - r_{cmd}) = 0 \tag{22}$$

where $\tau_r = 0.25$ is time constant, r_{cmd} is commanded reference signal and r_c is filtered reference signal. Here, it is important to select command filter parameters, such as damping ratio, natural frequency and time constant, to meet handling qualities specifications. Reader may refer to [14] for more information about handling and flying quality requirements of rotorcraft platforms.

2. DI-based Flight Control System Design

DI-based attitude control system is designed based on linearized model for the hover flight conditions as given in Eq. (23).

$$\begin{bmatrix}\dot{p}\\ \dot{q}\\ \dot{r}\end{bmatrix} = [A]\begin{bmatrix}p\\ q\\ r\end{bmatrix} + [B]\begin{bmatrix}u_{lat}\\ u_{lon}\\ u_{ped}\end{bmatrix} \tag{23}$$

Required control inputs can be obtained as given in Eq.(24) by using the definition in Eq. (23).

$$\begin{bmatrix}u_{lat}\\ u_{lon}\\ u_{ped}\end{bmatrix} = [B]^{-1}\left(\mathbf{U} - [A]\begin{bmatrix}p\\ q\\ r\end{bmatrix}\right) \tag{24}$$

where desired angular acceleration vector $\mathbf{U}$ is obtained by using the ADI algorithm and given in Eq. (25).

$$\mathbf{U} = \begin{bmatrix}\dot{p}_d & \dot{q}_d & \dot{r}_d\end{bmatrix}^T \tag{25}$$

Reference tracking error signals for roll, pitch and yaw axis are defined as the difference between the commanded and measured attitude signals and they are given in Eq. (26-28).

$$\tilde{\phi} = \phi_c - \phi \tag{26}$$

$$\tilde{\theta} = \theta_c - \theta \tag{27}$$

$$\tilde{r} = r_c - r \tag{28}$$

To compensate the tracking error signals, proportional-derivative (PD) controllers are utilized for roll and pitch channels which have ACAH structure. For the yaw channel, proportional-integral (PI) controller is used as it has RCAH structure. By using these definitions, desired angular acceleration signals can be calculated as given in Eq. (29).

$$\begin{bmatrix} \ddot{\phi}_d \\ \ddot{\theta}_d \\ \dot{r}_d \end{bmatrix} = \begin{bmatrix} \ddot{\phi}_c \\ \ddot{\theta}_c \\ \dot{r}_c \end{bmatrix} + \begin{bmatrix} K_{P_\phi}\tilde{\phi} + K_{D_\phi}\dot{\tilde{\phi}} \\ K_{P_\theta}\tilde{\theta} + K_{D_\theta}\dot{\tilde{\theta}} \\ K_{P_r}\tilde{r} + \int K_{I_r}\tilde{r}dt \end{bmatrix} - \begin{bmatrix} U_{ad_\phi} \\ U_{ad_\theta} \\ 0 \end{bmatrix} \tag{29}$$

here U_{ad_ϕ} and U_{ad_θ} are adaptive control signals generated by NN structure to compensate the aerodynamic uncertainties in roll and pitch dynamics.

In Eq. (29), desired angular acceleration vector is calculated in Euler angle representation form. Hence, it is required to transform into the body angular acceleration form as given in Eq. (30-31).

$$\dot{p}_D = \ddot{\phi}_D - (\ddot{\theta}_D \sin\phi \sin\theta + \dot{r}_D \sin\theta + \dot{\psi}\dot{\theta}\cos\phi + \dot{\theta}\dot{\phi}\cos\phi\sin\theta + \dot{\psi}\dot{\phi}\sin\theta\sin\phi\cos\theta)/(\cos\phi\cos\theta) \tag{30}$$

$$\dot{q}_D = (\ddot{\theta}_D + \dot{r}_D \sin\phi + \dot{\psi}\dot{\phi}\cos\theta)/\cos\phi \tag{31}$$

3. Adaptation Mechanism

As mentioned before, DI-based flight control system is designed for hover flight conditions. In the transition phase, free and propeller-induced airflows act on the UAV body and aerodynamic forces and moments become effective on the airframe as dynamic pressure increases. This situation leads inversion error and reduces signal tracking performance of the closed-loop system. To compensate the aerodynamic effects during the transition phase, DI-based flight control system is augmented by an adaptive signal which is generated by a single-layer sigma-pi NN structure. Parameter update law of the NN is developed based on Lyapunov stability analysis in [15].

Augmentation signals for roll and pitch channels $U_{ad_\phi}, U_{ad_\theta}$ are defined as given in Eq. (32). For simplification, only pitch axis augmentation signal is derived here. Adaptation signal for the roll axis can be obtained by using the same structure.

$$U_{ad_\theta} = \mathbf{W}_\theta^T \beta \tag{32}$$

Here, $\mathbf{W}_\theta^T \in \mathbb{R}^{1\times 36}$ is NN weight vector for pitch channel and $\beta \in \mathbb{R}^{36}$ is basis vector (Eq. (33)) which is created by using input vectors of the NN.

$$\beta = Kron\left[Kron(C_1, C_2), C_3\right]^T \tag{33}$$

where *Kron* is used for Kronecker product and C_1, C_2, C_3 are input vectors defined in Eq. (34-36).

$$C_1 = \left[0.1, V, V^2\right] \tag{34}$$

$$C_2 = \left[0.1, u, w, q, \theta, \ddot{\theta}_d\right] \tag{35}$$

$$C_3 = \left[0.1, \theta\right] \tag{36}$$

The first group of the inputs C_1 consists of airspeed measurements to cover the effects of airspeed variation. The second group C_2 consists of state measurements and pseudo-control signals to cover nonlinear effects in states and control signal adjustments. The third group C_3 consists of Euler angles to cover the nonlinear effects of large attitude angles [13].

Update law for the NN weights is obtained by using Lyapunov stability analysis and given in Eq. (37) [15].

$$\dot{\mathbf{W}}_\theta = -\gamma_\theta\, \mathbf{e}_\theta^T\, \mathbf{P}_\theta\, \mathbf{b}\, \beta - \mu_\theta \left\|\mathbf{e}_\theta^T\, \mathbf{P}_\theta^T\, \mathbf{b}\right\| \mathbf{W}_\theta \tag{37}$$

where $\gamma_\theta > 0$ is learning rates and μ_θ is e-modification term. $\mathbf{e}_\theta^T$ is error vector as given in Eq. (38).

$$\mathbf{e}_\theta = [\tilde{\theta}, \dot{\tilde{\theta}}]^T \tag{38}$$

Here, the vector b is defined as $[0, 1]^T$. P_θ is obtained by utilizing the Lyapunov stability analysis and given in Eq. (39).

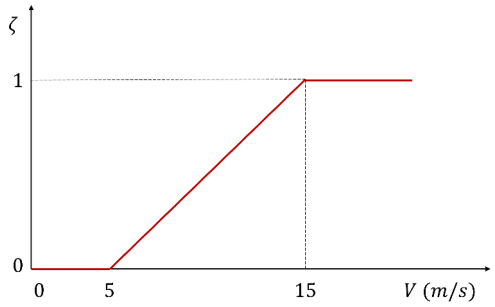


Fig. 6 Mixer parameter as a function of total airspeed.

$$P_\theta = \begin{bmatrix} \frac{K_{D_\theta}}{K_{P_\theta}} + \frac{1}{2K_{P_\theta}} & \frac{1}{2K_{D_\theta}} \\ \frac{1}{2K_{P_\theta}} & \frac{1+K_{P_\theta}}{2K_{P_\theta}K_{D_\theta}} \end{bmatrix} \tag{39}$$

4. Control Mixer Design

At the beginning of the transition phase, the propulsion system (i.e. coaxial fan group and tilt rotors) are used to control the attitude of the UAV as the dynamic pressure is quite low to create control force and moment by utilizing the aerodynamic control surfaces. However, using only the propulsion system group for control purposes during the transition results in increased power consumption which is not desired from the efficiency point of view. Hence, a mixing algorithm is designed to distribute the generated control signal to propulsion group and aerodynamic control surfaces as a function of airspeed.

In the hover flight, calculated control signal should be distributed to the propellers to provide attitude and rate reference tracking. This is done by the mixer matrix M' (in Eq. (40)) which is constructed as a function of a) thrust and drag coefficients of the propellers and tilt angle of the front propulsion group (i_t).

$$M' = \begin{bmatrix} k_{t_f} l_1 \cos i_t & -k_{t_f} l_1 \cos i_t & 0 & 0 \\ k_{t_f} l_2 \cos i_t & k_{t_f} l_2 \cos i_t & -k_{t_c} l_3 & -k_{t_c} l_3 \\ k_{d_f} \cos i_t & -k_{d_f} \cos i_t & -k_{d_c} & k_{d_c} \\ k_{t_f} \cos i_t & k_{t_f} \cos i_t & k_{t_c} & k_{t_c} \end{bmatrix} \tag{40}$$

where k_{t_f}, k_{t_c} are thrust coefficient of the tilt-rotor propellers and coaxial fan group, k_{d_f}, k_{d_c} are drag moment coefficient of the tilt-rotor propellers and coaxial group and i_t is tilt angle of the front propulsion group. Here, all of the propellers on the UAV are identical. So, thrust force and drag moment coefficients of the tilt-rotor group and coaxial-fan group could be assumed equal to each other. Hence, notation can be simplified as given in Eq. (41-42).

$$k_{t_f} = k_{t_c} = k_t \tag{41}$$

$$k_{d_f} = k_{d_c} = k_d \tag{42}$$

By using linear definitions of propeller-generated force and drag moment, i.e. $F_{t_n} = k_t \omega_n^2$ and $m_{d_n} = k_d \omega_n^2$, square of the rotation speed of each BLDC motor can be calculated as given in Eq. (43).

$$\omega_m^2 = M'^{-1} \begin{bmatrix} u_{lat}(1-\zeta) & u_{lon}(1-\zeta) & u_{ped}(1-\zeta) & u_{col} \end{bmatrix} \tag{43}$$

where $m \in [1, 2, 3, 4]$ indicates BLDC motor number and ζ is control mixer parameter which is utilized to allocate control signal between BLDC motors and aerodynamic control surface as a function of total airspeed (Figure (6)). This function assigns full authority on the aerodynamic control surfaces after $15\,m/s$ total airspeed is achieved which is approximately 1.2 times of the stall speed of the UAV. By using this definitions, control signal allocation on the aerodynamic control surfaces can be performed as given in Eq. (44-46).

$$\delta_a = u_{lat}\,\zeta\ /(q_{bar}\,S\,b\,C_{l_{\delta a}}) \tag{44}$$

$$\delta_e = u_{lon}\,\zeta\ /(q_{bar}\,S\,\bar{c}\,C_{m_{\delta e}}) \tag{45}$$

$$\delta_r = u_{ped}\,\zeta\ /(q_{bar}\,S\,b\,C_{n_{\delta r}}) \tag{46}$$

where q_{bar} is dynamic pressure and b is wing span. $C_{l_{\delta a}}, C_{m_{\delta e}}, C_{n_{\delta r}}$ are primary control derivatives on roll, pitch and yaw axes, respectively.

C. Deep Deterministic Policy Gradient Algorithm

Deep deterministic policy gradient (DDPG) is an off-policy reinforcement learning method and developed based on actor-critic structure. It is a widely used reinforcement learning algorithm in control and guidance applications as it is able to handle continuous action space [16].

Structure of the agent consists of two neural networks, i.e. actor network and critic network. Here, actor network generates the action signal and the critic network evaluates the action signal. Then, it produces temporal difference (TD) error to update the actor network parameters.

Policy is defined as the agent behavior and maps the states to probability distribution of actions, $\pi : S \rightarrow P(A)$. Markov decision process is modeled with a state space S, action space $A \in \mathbb{R}^N$, initial state distribution $p(s_1)$, transition dynamics from timestep t to $t+1$ as $p(s_{t+1}|s_t, a_t)$ and reward function for timestep t as $r(s_t, a_t)$.

If the target policy is deterministic, it can be described as a function $\mu : S \rightarrow A$. By doing so, inner expectation can be avoided and Bellman equation can be rewritten as given in Eq. (47).

$$Q^{\mu}(s_t, a_t) = \mathbb{E}_{r_t, s_{t+1}\sim\Psi}\left[r(s_t, a_t) + \gamma[Q^{\mu}(s_{t+1}, \mu(s_{t+1}))]\right] \tag{47}$$

where $\gamma \in [0, 1]$ is discount factor. Eq. (47) means that the expectation depends on the environment and Q^{μ} can be learnt off-policy. Critic network is used as a function approximator parameterized by θ^Q and optimized by minimizing the loss function given in Eq. (48)

$$L(\theta^Q) = \mathbb{E}_{s_t\sim\rho^{\psi}, a_t\sim\psi, r_t\sim\Psi}\left[(Q(s_t, a_t|\theta^Q) - y_t)^2\right] \tag{48}$$

y_t is defined as given in Eq. (49).

$$y_t = r(s_t, a_t) + \gamma Q(s_{t+1}, \mu(s_{t+1})|\theta^Q) \tag{49}$$

The policy of the actor network is updated by using the aid from the critic network. In this process, gradient update of the policy is given in Eq. (50).

$$\nabla_{\theta^{\mu}} J \approx \mathbb{E}_{s_t\sim\rho^{\psi}}\left[\nabla_a Q(s, a|\theta^Q)|_{s=s_t, a=\mu(s_t)}\ \nabla_{\theta^{\mu}}\ \mu(s|\theta^{\mu})|_{s=s_t}\right] \tag{50}$$

where θ^{μ} is actor network parameters. By using the gradients $Q(s, a|\theta^Q)|_{s=s_i, a=\mu(s_i)}$ and $\nabla_{\theta^{\mu}}\mu(s|\theta^{\mu})|_{s=s_i}$, actor network can be updated by using the approximated policy update gradient given in Eq. (51).

$$\nabla_{\theta^{\mu}} J \approx \frac{1}{N}\sum_i\left[\nabla_a Q(s, a|\theta^Q)|_{s=s_i, a=\mu(s_i)}\ \nabla_{\theta^{\mu}}\ \mu(s|\theta^{\mu})|_{s=s_i}\right] \tag{51}$$

After these steps, target networks in actor and critic structure are updated by utilizing Eq. (52) and (52).

$$\theta^{Q'} \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'} \tag{52}$$

$$\theta^{\mu'} \leftarrow \tau\theta^{\mu} + (1-\tau)\theta^{\mu'} \tag{53}$$

where τ is a small constant.

Exploration is one of the major challenges in reinforcement learning applications. To treat the exploration problem independently from the learning process, an exploration policy μ' is obtained by adding noise to the actor policy as given in Eq. (54). For more information about the DDPG algorithm, readers may refer to [16].

$$\mu'(s_t) = \mu(s_t|\theta_t^{\mu}) + \mathcal{N} \tag{54}$$

III. Problem Formulation - RL Approach in Transition Flight

During the transition phase, especially in the urban airspace, it is critical to maintain minimum altitude change while accelerating towards the cruise flight conditions. Acceleration can be obtained either by tilting forward the front propeller group or pitching down the nose of the UAV. As the airspeed increases, dynamic pressure increases and lift force is generated by the airframe.

In this study, we focused on a transition strategy in which pitch attitude is fixed as zero, $\theta = 0^\circ$, during the transition maneuver. So, excess force to accelerate the vehicle is produced by tilting the front propeller group. Tilt angle is controlled manually and limited to 30° to avoid propeller-induced airstream effects on the UAV airframe. Coaxial fan group is utilized to produce thrust force on the Z_b axis to maintain constant altitude until sufficient lift force is created by the airframe.

To minimize the altitude deviation and provide high energy efficiency during the transition maneuver, it is critical to generate aerodynamic lift force at a sufficient level. Angle of attack and total airspeed are key parameters in this phase. If the UAV can produce sufficient lift force, collective input u_{col} could be reduced. In other words, thrust level of the whole propulsion system is decreased to a certain level if sufficient aerodynamic lift force can be generated and this results in energy saving. To generate sufficient lift force in zero pitch angle conditions, positive angle of attack could be generated by increasing velocity on the Z_e axis. This means that flight path angle γ could be controlled to generate lift force in $\theta = 0^\circ$ flight conditions. According to this relation, positive angle of attack could be generated in descend flight conditions, i.e. $\gamma < 0^\circ$ (Eq. (55)).

$$\gamma = \theta - \alpha \tag{55}$$

General structure of the RL-aided transition flight control system is shown in Figure (2). In this structure, $\phi_c, \theta_c, r_c, i_t$ are applied manually by the pilot, u_{col} is fixed for hover flight condition, i.e. $u_{col} = mg$. Actor signal is defined as normalized deviation in the collective input $\Delta u_{col_n} = \Delta u_{col}/mg$ and it is used as a supporting element for the whole transition process. It has limited authority on the collective signal and lower and upper boundaries of the actor signal is defined as $\Delta u_{col_n} \in [-0.2, 0.2]$ to avoid instability during the transition flight. In the training process, RL agent receives observation vector and reward value to adjust NN parameters in the actor-critic structure. After the training process, it just uses the observation vector to create optimal policy for the transition, i.e. Δu_{col_n}.

To obtain an optimal transition strategy, it is quite critical to define a proper reward function which reflects requirements for the transition phase. For the transition strategy that is defined in this study, there are two main requirements for the transition phase; a) maintaining constant altitude and b) minimizing collective usage to generate excess thrust on Z_b axis. These requirements are directly stated in the reward function as given in Eq. (56).

$$R = R_{step} - w_{e_h} R_{e_h} - w_{u_{col}} R_{u_{col}} \tag{56}$$

where R_{step} is a constant reward given for each time step, $w_{e_h}, w_{u_{col}}, R_{e_h}, R_{u_{col}}$ are weights and rewards for altitude error and collective input, respectively. R_{e_h} defines how the UAV follows given altitude reference and $R_{u_{col}}$ is used to minimize the control action as given in Eq. (57) and (58).

$$R_{e_h} = clip\left(\frac{|h_{c_n} - h_n|}{\xi_h}, \mu_h^l, \mu_h^u\right) \tag{57}$$

$$R_{u_{col}} = clip\left(\frac{\int u_{col_n}}{\xi_u}, \mu_c^l, \mu_c^u\right) \tag{58}$$

where $h_{c_n} = h_c/100$ is normalized altitude command, $h_n = h/100$ is normalized altitude measurement, $u_{col_n} = u_{col}/mg$ is normalized collective command applied to the mixer. ξ_h, ξ_c are constants to scale down normalized altitude error and collective signal for clip function. $\mu_h^l, \mu_h^u, \mu_c^l, \mu_c^u$ are lower and upper bounds of clip functions for altitude and collective input, respectively.

Observation vector of the RL agent is another important component for a successful learning process. It should be defined in a problem-specific way according to the requirements in the problem. For the transition flight problem, it is defined as given in Eq. (59).

$$O = \left[h_{e_n}, \int h_{e_n}, V_n, \alpha_n, i_{t_n}, \zeta\right] \tag{59}$$

Table 1 Hyperparameters for learning process.

Hyperparameters	Value
Sample Time	0.02
Target Smooth Factor	1e-3
Experience Buffer Length	1e6
Discount Factor	0.99
Mini Batch Size	256
Actor Learning Rate	0.008
Critic Learning Rate	0.005
Standard Deviation	0.14
Standard Deviation Decay Rate	1e-5

Table 2 Network layer properties.

Layer	Critic Network	Actor Network
Input layer	7	6
Hidden layer 1	400	400
Hidden layer 2	300	300
Hidden layer 3	300	-
Output layer	1	1

where $h_{e_n} = h_e/100$ is normalized altitude command following error, $V_n = V/15$ is normalized airspeed, $\alpha_n = \alpha/(\pi/4)$ is normalized angle-of-attack measurement, $i_{t_n} = i_t/(\pi/6)$ is normalized tilt angle of the front propeller group and $\zeta \in [0, 1]$ is control allocation parameter.

Selecting hyperparameters and neural network structure is another important issue in deep reinforcement learning applications. The agent may not be able to learn with inappropriate NN-structure and learning parameters even if the problem is well-defined by using suitable observation vector and reward function. Hence, to provide reproducibility of the study, hyperparameters and NN structure of the RL agent are given in Table 1 and Table 2, respectively. Training process of the RL agent is performed for 1000-runs and learning curve is given in Figure (7). Here, Q_0 is estimation of the long-term discounted reward and provided by the critic network.

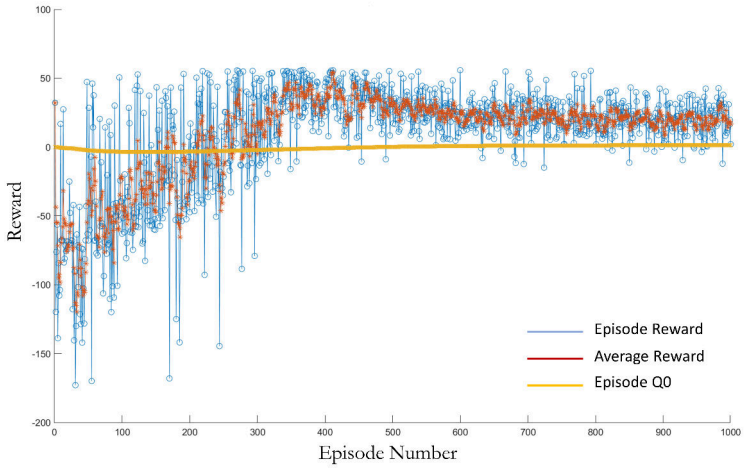


Fig. 7 Learning curve.

IV. Simulation Results

After completing the learning process, simulation studies are conducted and closed-loop system performance is evaluated. Here, it could be helpful to remind the expectations from the proposed system; a) completing the transition flight phase with minimum change in altitude, and b) minimizing the collective input (u_{col}) usage to minimize the energy consumption. Minimizing the altitude deviation is important to provide flight safety especially in the urban environment while minimizing the energy consumption which is crucial for sustainability and profitability of the aerial operations.

While evaluating the system behaviour and performance on the simulation plots, we utilized two time regions for ease of analysis; a) yellow region for the time interval $t \in [0, 8)$ sec and b) green region for the time interval $t \in [8, 30]$ sec. Transition flight phase begins by tilting the front propellers forward at $t = 0\,sec$. Tilt angle (i_t) follows a ramp

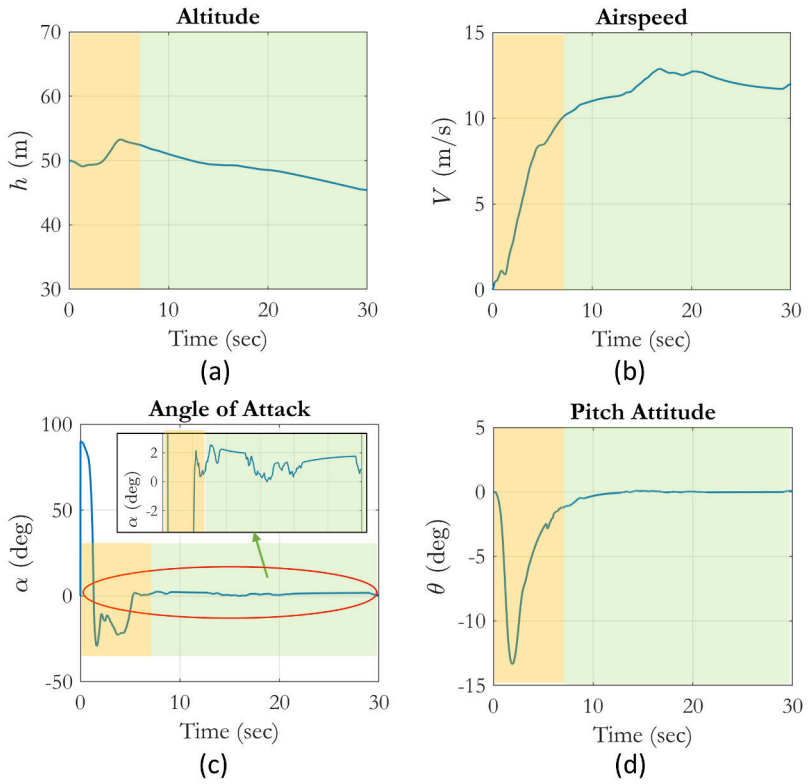


Fig. 8 Simulation results; time histories for a) altitude, b) airspeed, c) angle of attack, d) pitch attitude.

command as given in Figure (9).a to obtain horizontal acceleration in the local frame. Attitude of the UAV is controlled by the propulsion system between $t = [0, 5]$ sec time interval as the dynamic pressure is not sufficient and elevator control surface cannot produce considerable amount of aerodynamic moment on the body. In this initial time segment, RL agent increases the collective signal magnitude as shown in Figure (9).c to increase the altitude of the UAV to $54m$ as given in Figure (8).a. By doing so, potential energy of the system is also increased which will be used in the rest of the transition maneuver. In this phase, airspeed is relatively low and angle of attack is not appropriate to generate aerodynamic lift force.

At the beginning of the time interval $t \in [8, 30]$ seconds, i.e. green region on Figure (8) and (9), RL agent decreases the collective input level and this results in altitude decrement. Indirectly, this creates positive angle of attack in the interval of $\alpha \in [0°, 2°]$ as given in Figure (8).c as the UAV maintains $\theta = 0°$ pitch attitude condition. While the tilt angle continues to increase in this region, airspeed also increases and settles in the $V \in [11, 14]m/s$ interval. Aerodynamic control surfaces (i.e. elevator) is also used in the green region beside of the propulsion system as given in Figure (9).b to control the pitch attitude of the UAV. This is obtained by the control allocation algorithm which is mentioned in Section (II.B.4).

As the dynamic pressure increases with the airspeed, aerodynamic forces and moments become effective on the airframe. Here, aerodynamic lift force is utilized to complete the transition flight phase safely, with minimum energy consumption and minimum altitude deviation. However, as we mentioned in Section (II.B), these effects should be compensated in the low-level flight control system to obtain attitude reference signal tracking with minimum tracking error. Sigma-Pi NN structure of the ADI flight control system provides adaptation in this region and compensates dynamical changes and uncertainties which are inserted into the closed-loop system by the aerodynamic effects. During the transition, NN weights are updated continuously as given in Figure (10) to adapt for the changing flight conditions. It is also important to note that the ADI flight control system provides a smooth adaptation between two flight conditions without high frequency oscillations on the control signal and system outputs. From the safety point of view, it is a desired capability for the aerial vehicles that are used in the urban environment in which it is quite critical to provide safe aerial operations.

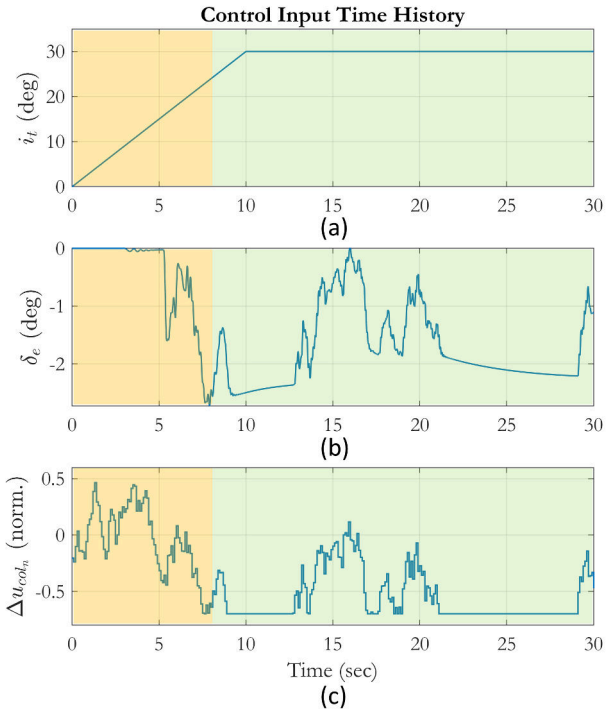


Fig. 9 Control signal time histories for; a) tilt angle, b) elevator deflection, c) normalized collective input created by the RL agent.

Discussions

Generated aerodynamic lift force is an important issue that should be evaluated in the transition phase. Sufficient level of lift force should be generated by the airframe to complete a safe transition to cruise flight. Time history of the generated lift force is given in Figure (11). Here, it is shown that the level of aerodynamic lift force is at about 40 N between time intervals $t_1 \in [9, 13]$ sec and $t_1 \in [21, 29]$ sec which is lower than the actual weight of the UAV (44.12 N). There are several causes which result in this relatively low lift force; during the transition a) the UAV is not able to achieve 15 m/s airspeed which is 1.2 times of the stall speed, b) generated maximum angle of attack is 2° in these flight conditions.

As given in Figure (8).a, initial and final altitudes of the UAV is 50 m and 46 m, respectively. There is 4 m altitude loss when compared to the initial conditions. In addition, if we consider the whole transition phase, total altitude change is almost 8 m as the UAV ascends to 54 m and then descends to 46 m altitude. Although altitude deviation is decreased by the developed algorithm, it could still be high for the UAM applications. It is evaluated that the main reason of this altitude deviation is insufficient level of aerodynamic lift force. Here, it is required to perform further improvements on the proposed transition logic and reinforcement learning process and they are addressed in the future works.

V. Conclusion

As the tilt-rotor aerial platforms have been developed and utilized in the urban airspace, flight safety and efficiency should also be considered to provide sustainability of the operations. Hence, adaptation of the flight control systems and optimization of the flight paths take important roles for aerial operations in the urban airspace.

In this study, RL method is utilized to train an actor-critic agent which is used as an aiding system to obtain optimal transition for the tilt-rotor VTOL aerial vehicle. RL agent has been integrated on the collective channel which controls the normal force on the vehicle body generated by the propulsion system. The agent has limited authority and it is able to increase and decrease the collective level of the aerial vehicle in a certain interval. By integrating the RL agent into the transition flight control logic, we propose to maximize the usage of aerodynamic lift force and minimize the usage of the collective signal to increase the operational efficiency. Observation vector and reward function are defined as a part of the design process. DDPG algorithm is utilized to update the neural network parameters in the actor-critic

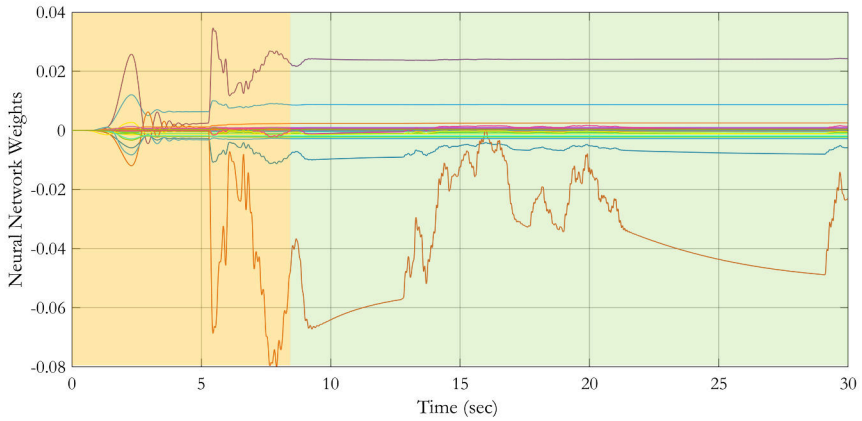


Fig. 10 Neural network weights in ADI flight control system.

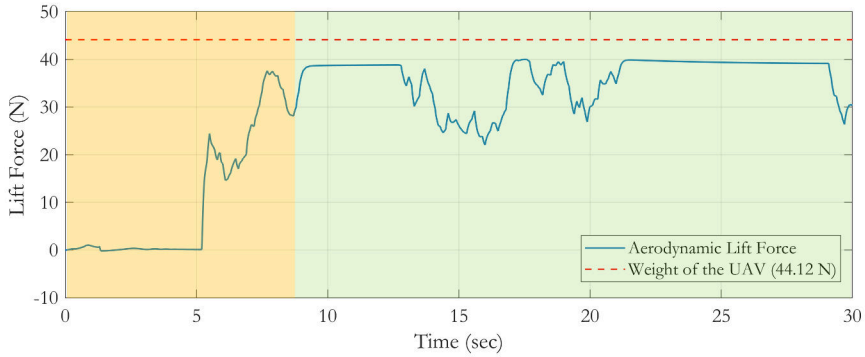


Fig. 11 Time history of the aerodynamic lift force.

structure. Low-level ACAH and RCAH flight control systems are designed based on ADI algorithm to improve signal tracking ability of the closed-loop system by enabling adaptation capability in the presence of aerodynamic effects on the airframe.

Simulation results show that, after the learning process, the proposed algorithm is able to manage the collective signal usage during the transition flight. In the initial phase of the flight, front propellers start to tilt forward to increase the airspeed. RL agent increases the level of collective input to increase altitude and potential energy. Then, it reduces the collective input level to provide descending flight in $\theta = 0^\circ$ conditions which creates positive angle of attack on the airframe. Combination of the increased angle of attack and airspeed results in aerodynamic lift force which is required for a safe and efficient transition flight phase to minimize altitude change and energy usage.

In future studies, further improvements will be performed on the proposed algorithm to minimize the altitude variation and maximize the generated aerodynamic lift force in the transition phase. Also, RL-based methodology will be compared with other studies in the literature. A comprehensive analysis will be performed to evaluate the system performance in the presence of aerodynamic uncertainties. In addition, effects of environmental conditions, such as wind effects, on the transition flight performance will also be evaluated. Finally, different types of RL-based transition flight control methodologies will be investigated. Future studies will especially focus on developing an RL agent with full authority not only on the collective channel, but also on pitch attitude and tilt angle command channels.

Acknowledgments

This work is supported, in parts, by the Engineering and Physical Sciences Research Council [Grant number: EP/V026763/1]

References

[1] Aktas, Y. O., Ozdemir, U., Dereli, Y., Tarhan, A. F., Cetin, A., Vuruskan, A., Yuksek, B., Cengiz, H., Basdemir, S., Ucar, M., et al., "Rapid prototyping of a fixed-wing VTOL UAV for design testing," *Journal of Intelligent & Robotic Systems*, Vol. 84, No. 1, 2016, pp. 639–664.

[2] Ozdemir, U., Aktas, Y. O., Vuruskan, A., Dereli, Y., Tarhan, A. F., Demirbag, K., Erdem, A., Kalaycioglu, G. D., Ozkol, I., and Inalhan, G., "Design of a commercial hybrid VTOL UAV system," *Journal of Intelligent & Robotic Systems*, Vol. 74, No. 1, 2014, pp. 371–393.

[3] Yuksek, B., Vuruskan, A., Ozdemir, U., Yukselen, M., and Inalhan, G., "Transition flight modeling of a fixed-wing VTOL UAV," *Journal of Intelligent & Robotic Systems*, Vol. 84, No. 1, 2016, pp. 83–105.

[4] Vuruskan, A., Yuksek, B., Ozdemir, U., Yukselen, A., and Inalhan, G., "Dynamic modeling of a fixed-wing VTOL UAV," *2014 international conference on unmanned aircraft systems (ICUAS)*, IEEE, 2014, pp. 483–491.

[5] Xili, Y., Yong, F., and Jihong, Z., "Transition flight control of two vertical/short takeoff and landing aircraft," *Journal of guidance, control, and dynamics*, Vol. 31, No. 2, 2008, pp. 371–385.

[6] Maqsood, A., and Go, T. H., "Optimization of hover-to-cruise transition maneuver using variable-incidence wing," *Journal of Aircraft*, Vol. 47, No. 3, 2010, pp. 1060–1064.

[7] Casau, P., Cabecinhas, D., and Silvestre, C., "Autonomous transition flight for a vertical take-off and landing aircraft," *2011 50th IEEE Conference on Decision and Control and European Control Conference*, IEEE, 2011, pp. 3974–3979.

[8] Casau, P., Cabecinhas, D., and Silvestre, C., "Hybrid control strategy for the autonomous transition flight of a fixed-wing aircraft," *IEEE Transactions on control systems technology*, Vol. 21, No. 6, 2012, pp. 2194–2211.

[9] Flores, G., and Lozano, R., "Transition flight control of the quad-tilting rotor convertible MAV," *2013 International Conference on Unmanned Aircraft Systems (ICUAS)*, IEEE, 2013, pp. 789–794.

[10] Flores, A., de Oca, A. M., and Flores, G., "A simple controller for the transition maneuver of a tail-sitter drone," *2018 IEEE Conference on Decision and Control (CDC)*, IEEE, 2018, pp. 4277–4281.

[11] Xu, J., Du, T., Foshey, M., Li, B., Zhu, B., Schulz, A., and Matusik, W., "Learning to fly: computational controller design for hybrid uavs with reinforcement learning," *ACM Transactions on Graphics (TOG)*, Vol. 38, No. 4, 2019, pp. 1–12.

[12] Rysdyk, R. T., and Calise, A. J., "Adaptive model inversion flight control for tilt-rotor aircraft," *Journal of guidance, control, and dynamics*, Vol. 22, No. 3, 1999, pp. 402–407.

[13] Sahani, N. A., and Horn, J. F., "Adaptive model inversion control of a helicopter with structural load limiting," *Journal of guidance, control, and dynamics*, Vol. 29, No. 2, 2006, pp. 411–420.

[14] ADS-33E-PRF, "Aeronautical Design Standard Performance Specification: Handling Requirements for Military Rotorcraft," Tech. rep., United States Army Aviation And Missile Command Aviation Engineering Directorate, Redstone, Arsenal, AL, 2000.

[15] Leitner, J., Calise, A., and Prasad, J., "Analysis of adaptive neural networks for helicopter flight control," *Journal of Guidance, Control, and Dynamics*, Vol. 20, No. 5, 1997, pp. 972–979.

[16] Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., and Wierstra, D., "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.



https://dspace.org/

Transition flight control system design for fixed-wing VTOL UAV: a reinforcement learning approach

Yuksek, Burak; Inalhan, Gokhan

2021-12-29
AIAA
https//doi.org/10.2514/6.2022-0879
Yuksek B, Inalhan G. (2021) Transition flight control system design for fixed-wing VTOL UAV: a reinforcement learning approach. In: AIAA SciTech 2022 Forum, 3-7 January 2022, San Diego, CA, USA and Virtual Event

Downloaded from a DSpace repository. https://dspace.org/