



Reinforcement learning-based controller design for a proposed octorotor with tilt-arm angles

D. Sharifi¹, M. Irani Rahaghi¹, K. Torabi², H. Shahbazi²

¹ Department of Mechanical Engineering, University of Kashan, Kashan, Iran

² Department of Mechanical Engineering, Faculty of Engineering, University of Isfahan, Isfahan, Iran

ABSTRACT: The maneuverability of a quadrotor or octorotor UAV is limited in the standard configuration because the force vectors of the propellers are parallel and only have four active degrees of freedom. Therefore, they lack the controllability of six independent degrees of freedom. This study designs a novel configuration for an octorotor capable of hovering with roll or pitch angles in a specific position, contrary to UAVs with a standard configuration that can only hover in a horizontal position. In other words, in this octorotor, orientation tracking is also added to the octorotor's targets in addition to position tracking. The proposed model can be controlled by altering the velocity of the eight rotors and the tilt angle of the four arms. Such alterations in velocity and tilt angle are such that they can provide the aerial vehicle with the most optimum maneuverability. After deriving the proposed dynamic octorotor model, a controller is proposed using neural networks and reinforcement learning, capable of controlling the proposed octorotor with six independent degrees of freedom. Finally, trajectory tracking, octorotor position, and controller robustness to possible motor malfunctions are examined, and numerical simulation results are provided.

Review History:

Received: Nov. 02, 2022

Revised: Oct. 18, 2023

Accepted: Jan. 08, 2024

Available Online: Jan. 21, 2024

Keywords:

octorotor with tilt-arm angle

reinforcement learning

trajectory tracking

independent degrees of freedom

motor malfunctions

1- Introduction

most of the studies conducted on multirotors can be divided into two parts: (1) proposing a novel configuration for UAV's rotors and (2) designing a controller consistent with its expected performance. Below is a brief history of these studies.

Tadokoro et al. (2017) analyzed the maneuverability of UAVs concerning the tilt angles and configuration of a fully actuated hexa-rotor's propellers [1]. Zhu et al. (2020) proposed a new octocopter configuration based on preliminary aerodynamic considerations to rectify the drawbacks of conventional octocopters with monolayer and coaxial configurations. This study analyzed five different octorotor UAV models with three different configurations, including a conventional configuration, a coaxial configuration, and a new configuration, to achieve a precise aerodynamic analysis [2]. Quintana et al. (2018) presented a methodology for designing a hexa-rotor with disturbance rejection capability using tilted propellers. This method proposes using the robustness index as a measure of the ability to reject external disturbances and the energy index as a measure of the energy consumed by the hovering hexa-rotor[3]. Kase and Oya (2020) designed and implemented an adaptive trajectory tracking control scheme for hexacopters with wind disturbances [4]. Pi et al. (2020) proposed a quadrotor control algorithm using NNs with

the model-free RL method. This study examines quadrotor training in two hovering and tracking modes of the proposed trajectory. Then, training is also implemented on a real-world quadrotor[5].

This study seeks to construct a novel rotor configuration for an octorotor and design a robust RL-based controller to create the capability of controlling the position and angle of the octorotor UAV separately.

2- Proposed octorotor configuration

This study considers the octorotor configuration (Fig. 1) in a way that its four rotors are positioned like a conventional octorotor, and the other four rotors are placed on four arms with tilt angles to achieve 6-DoF independently. All four angular arms are positioned so that one side is placed on the octorotor's CoM and has a tilt angle with the plane of the four fixed arms. Based on this configuration, the rotation axis of the four rotors positioned in the same plane is parallel, and the axis of the other four rotors is in different directions. The angles of the two facing arms are assumed to be the same to keep the CoM constant while tilting the angles. It is expected that by changing the rotor velocity and tilting the arm angles, it will be possible to create independent 6-DoF motion.

*Corresponding author's email: k.torabi@eng.ui.ac.ir



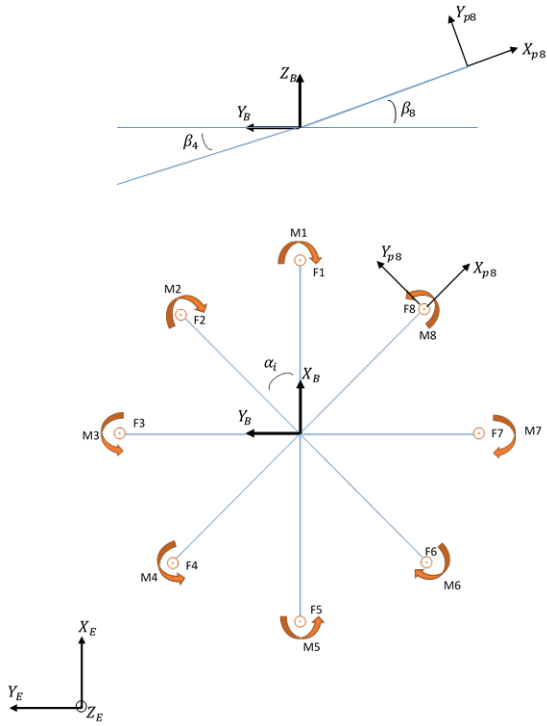


Fig. 1. Rotor configuration, thrust, and torque generated by octorotor rotors and reference frame display

3- The proposed controller configuration

This study aims to design an NN controller using the RL method to independently control the octorotor UAV in 6-DoF. Therefore, the NN training phase uses the following general equations as six independent dynamic equations to consider the training environment and six networks are trained separately to monitor them.

$$m\ddot{X} = \begin{bmatrix} F_x & F_y & F_z \end{bmatrix} \quad (1)$$

$$I\dot{\omega} + \omega \times I\omega = \begin{bmatrix} \tau_x & \tau_y & \tau_z \end{bmatrix} \quad (2)$$

In these equations, F_x , F_y and F_z denote the required force in the x, y, and z directions to achieve the desired position, and τ_x , τ_y and τ_z indicate the required torques around the x, y, and z axes to achieve the desired position.

One actor network and two critic networks are required for each degree of freedom to design and train the NNs using the TD3 algorithm. In network training, it is assumed that if an octorotor UAV is mounted on any position with any velocity, it is directed towards zero, i.e., positioned at the origin of coordinates, and remains to hover at that location. During training, the position within a radius of 2 meters and the position within a radius of 60 degrees are randomly

selected. It is noteworthy that this learning will also create a robust controller in the octorotor UAV because external factors cause it to deviate from the target trajectory and the trained network returns to the target trajectory from any position with any velocity. The reward function in network training is also considered to achieve the objective with the least control effort and move towards the target. For instance, for training along the x direction, the reward function is defined as follows:

$$r = -\rho(x - x_d)^2 - (1 - \rho)u^2 \quad (3)$$

where $x_d = 0$ represents the target position in the x direction, u denotes the control effort in the x direction, and ρ indicates the influence coefficient.

It is also worth noting that there is no need to retrain the network for a target position other than the origin of the coordinates. The octorotor UAV can be directed towards the desired targets only by making a variable change in the problem dynamics.

4- Results and Discussion

As mentioned earlier, in network training, it is assumed that if an octorotor UAV is mounted on any position with any velocity, it is directed towards zero, i.e., positioned at the origin of coordinates, and remains to hover at that location. Each network training episode is performed in four seconds, which includes 400 time steps. This study performed the full network training (for robustness in hovering mode) within less than 200-thousand-time steps for each degree of freedom. In contrast, other studies used RL for quadrotor control. For example, Hwangbo et al. (2017) [1] needed about 2150 million time steps and Pi et al. (2020) [5] needed about 10 million time steps to achieve proper robustness in hovering mode. The proposed model and independent degrees of freedom have made it possible to reduce network dimensions compared to previous studies and significantly minimize training time, which is one of the advantages of the present study.

The graphs drawn in Fig. 2 represent the proposed octorotor capability in the independent control of 6-DoF in the trajectories and orientations defined in Table 3 within ten seconds, demonstrating the capabilities of the configuration and controller proposed for the octorotor. The maximum deviation in Fig. 2 (a) corresponds to the deviation of the octorotor UAV in the x-direction from the trajectory determined within eight seconds, which is due to a sudden change in the velocity. The drawn graph indicates the controller's capability to direct the UAV toward the predetermined target trajectory quickly.

5- Conclusions

This study proposed a novel configuration for an octorotor. The octorotor's four propellers were positioned like a conventional octorotor, and the other four propellers were mounted on four arms with a tilt angle. One of the capabilities

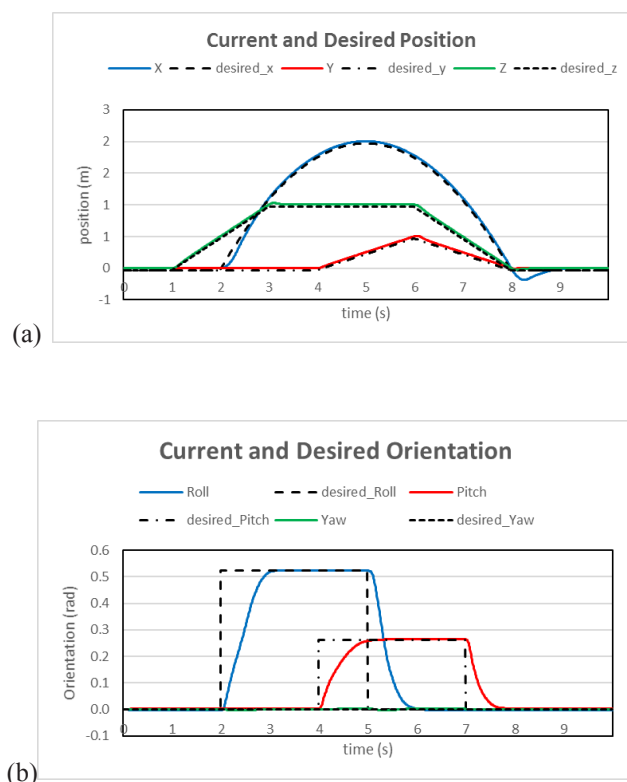


Fig. 2. (a) Trajectory tracking and (b) orientation graphs of the octorotor in independent 6-DoF

of this configuration is to add orientation tracking to UAV targets and position tracking. A controller based on NNs and RL was also designed for this UAV, which is robust to one or more motor malfunctions and the octorotor controllability

in independent six independent degrees of freedom. Due to the independent degrees of freedom, the network dimensions considered for this controller were smaller than those of similar studies. The training time was significantly decreased without reducing the capabilities and precision of the designed controller.

References

- [1] Y. Tadokoro, T. Ibuki, M. Sampei, Maneuverability analysis of a fully-actuated hexrotor UAV considering tilt angles and arrangement of rotors, IFAC-PapersOnLine, 50(1) (2017) 8981-8986.
- [2] H. Zhu, H. Nie, L. Zhang, X. Wei, M. Zhang, Design and assessment of octocopter drones with improved aerodynamic efficiency and performance, Aerospace Science and Technology, 106 (2020) 106206.
- [3] V.M. Arellano-Quintana, E.A. Portilla-Flores, E.A. Merchán-Cruz, Multi-objective design optimization of a hexa-rotor with disturbance rejection capability using an evolutionary algorithm, IEEE Access, 6 (2018) 69064-69074.
- [4] S. Kase, M. Oya, Adaptive tracking controller for hexacopters with a wind disturbance, Artificial Life and Robotics, 25(2) (2020) 322-327.
- [5] C.-H. Pi, K.-C. Hu, S. Cheng, I.-C. Wu, Low-level autonomous control and tracking of quadrotor using reinforcement learning, Control Engineering Practice, 95 (2020) 104222.
- [6] J. Hwangbo, I. Sa, R. Siegwart, M. Hutter, Control of a quadrotor with reinforcement learning, IEEE Robotics and Automation Letters, 2(4) (2017) 2096-2103.

HOW TO CITE THIS ARTICLE

D. Sharifi, M. Irani Rahaghi, K. Torabi, H. Shahbazi, Reinforcement learning-based controller design for a proposed octorotor with tilt-arm angles, Amirkabir J. Mech Eng., 55(10) (2024) 245-248.

DOI: [10.22060/mej.2024.21904.7538](https://doi.org/10.22060/mej.2024.21904.7538)





طراحی کنترل کننده به روش یادگیری تقویتی برای یک اوکتاروتور پیشنهادی با زاویه بازوهای متغیر

داود شریفی^۱، محسن ایرانی رهقی^۱، کیوان ترابی^{۲*}، حامد شهبازی^۲

۱- دانشکده مهندسی مکانیک، دانشگاه کاشان، کاشان، ایران
۲- دانشکده مهندسی مکانیک، دانشگاه اصفهان، اصفهان، ایران.

تاریخچه داوری:

دریافت: ۱۴۰۱/۰۸/۱۱
بازنگری: ۱۴۰۲/۰۷/۲۶
پذیرش: ۱۴۰۲/۱۰/۱۸
ارائه آنلاین: ۱۴۰۲/۱۱/۰۱

کلمات کلیدی:

اوکتاروتور با زاویه بازوهای متغیر
یادگیری تقویتی
تعقیب مسیر
درجات آزادی مستقل
اختلال در عملکرد موتورها

خلاصه: تحرک یک پرنده‌ی کوادروتور یا اوکتاروتور در پیکربندی استاندارد به دلیل موازی بودن بردارهای نیروی روتورها محدود است و آنها فقط دارای چهار درجه آزادی فعال هستند. بنابراین قابلیت کنترل شش درجه‌ی آزادی به صورت مستقل برای آنها وجود ندارد. در این پژوهش یک ساختار و پیکره‌بندی جدید برای یک اوکتاروتور طراحی می‌گردد تا برخلاف پرنده‌های با ساختار استاندارد که تنها در حالت افقی می‌توانند در یک محل مشخص معلق بمانند، قابلیت معلق ماندن با زوایای رول یا پیچ در یک موقعیت مشخص را داشته باشند، به معنی دیگر در این پرنده علاوه بر دنبال کردن موقعیت، دنبال کردن جهت‌گیری نیز به اهداف پرنده اضافه شده است. در مدل پیشنهادی می‌توان از تغییر در سرعت هشت روتور و تغییر در زاویه‌ی چهار بازو برای کنترل بهره برد، در واقع این سرعت‌ها و زوایا به نحوی تعیین می‌گردد تا قابلیت مانورپذیری پرنده به نحو مطلوب تامین شود. پس از استخراج مدل دینامیکی اوکتاروتور پیشنهادی، یک کنترل کننده با استفاده از شبکه‌های عصبی و روش یادگیری تقویتی پیشنهاد می‌گردد که قابلیت کنترل اوکتاروتور پیشنهادی با شش درجه‌ی آزادی مستقل را خواهد داشت. در نهایت قابلیت تعقیب مسیر و وضعیت اوکتاروتور بررسی شده و نتایج شبیه‌سازی‌های عددی ارائه می‌گردد.

۱- مقدمه

حوزه شده است. در ادامه به صورت مختصر تاریخچه‌ای از این مطالعات بیان می‌گردد.

بارباراسی با ایجاد یک ساختار هندسی متغیر، دینامیک یک کوادکوپتر را مورد بررسی قرار داده و مزایای استفاده از تغییرات هندسه را مورد بررسی قرار می‌دهد. ایشان تاثیر زاویه‌ی بازوها (بازوها در یک صفحه قرار دارند و حول محور بازو دوران می‌کنند) بر دینامیک پرنده را مورد بررسی قرار داده است. نتایج شبیه‌سازی بیانگر کاهش قابل توجه دامنه‌ی ارتعاشات این کوادکوپتر در مقایسه با یک کوادکوپتر با زاویه‌ی بازوی ثابت می‌باشد [۱].

برسیانینی و همکاران به طراحی یک پرنده با هشت پروانه پرداختند به نحوی که پروانه‌ها در زوایای مختلف جایگذاری شده‌اند. قرارگیری پروانه‌ها به نحوی است که قابلیت ایجاد نیرو و گشتاور در تمامی جهات و در نتیجه حرکت پرنده در ۶ درجه آزادی را داشته باشد. سیستم کنترلی در نظر گرفته شده برای این پرنده به نحوی است که قابلیت کنترل موقعیت و وضعیت به صورت مجزا قابل اجرا باشد [۲]. تادوکورو و همکاران با توجه به آرایش و زاویه‌ی پروانه‌های یک هگزاروتور، به تحلیل قدرت مانور این پرنده‌ی بدون

در سال‌های اخیر انواع مختلف ربات‌های هوایی از زوایای متنوع مورد بررسی قرار گرفته و پیشرفت‌های آنها مورد توجه پژوهشگران واقع شده است. کاربرد فراوان پرنده‌های بدون سرنشین در موقعیت‌ها و محیط‌های خطرناک را می‌توان از دلایل مورد توجه قرار گرفتن این پرنده‌ها دانست چراکه انسان‌ها اغلب توانایی نزدیک شدن به محیط‌های خطرناک که از بلایای طبیعی از قبیل سیل، زمین لرزه، آتشفشان، طوفان و یا یک خطر هسته‌ای رنج می‌برند را ندارند.

از کاربردهای این پرنده‌ها می‌توان به بازرسی و کنترل مرزها، ماموریت‌های نظامی، ماموریت‌های شناسایی، حمل بار، عکس‌برداری، عملیات جستجو، کشف آتش‌سوزی‌ها، استفاده در محیط‌های خطرناک و غیرقابل دسترس، بازدید و بازرسی از خطوط انتقال نفت و خطوط فشار قوی اشاره کرد. نیاز به پایداری پرنده در شرایط محیطی مختلف و در تمامی کاربردهای مذکور سبب ایجاد تیم‌های تحقیقاتی و اجرای مطالعات در این

* نویسنده عهده‌دار مکاتبات: k.torabi@eng.ui.ac.ir



سرنشین پرداخته اند[۳].

کننده تناسبی انتگرالی استفاده نموده اند[۸]. زقلاچ و همکاران با به کارگیری کنترل کننده‌های مد لغزشی، فازی به طراحی یک کنترل کننده‌ی مقاوم در برابر آسیب دیدگی یکی از پروانه‌های یک اوتاروتور با جفت پروانه‌های هم محور پرداخته اند[۹]. هو و همکاران با تبدیل مسئله کنترل سیستم چند ورودی و جند خروجی به مسئله کنترل چند ورودی و یک خروجی و استفاده از تئوری فیدبک کمی^۷، به طراحی کنترل کننده‌ی مربوط به یک کوادکوپتر پرداخته است[۱۰].

کوئنتینا و همکاران روشی برای طراحی هگزا روتور با قابلیت رد اغتشاشات با استفاده از پروانه‌های زاویه داده شده ارائه داده اند. در این روش استفاده شاخص مقاوم بودن به عنوان اندازه گیری توانایی رد اغتشاشات خارجی و شاخص انرژی به عنوان اندازه گیری انرژی مصرف شده توسط هگزا روتور معلق در هوا پیشنهاد می‌شود. ایشان یک مسئله بهینه سازی چند هدفه ارائه می‌دهند که در آن توابع هدف، شاخص قدرت و شاخص انرژی هستند. این مسئله با کمک الگوریتم تکاملی با رویکرد پارتو^۸ حل می‌شود[۱۱].

اراسموس و همکاران یک مولتی روتور با یک محموله معلق ناشناخته (جرم بار و طول کابل ناشناخته است و زوایای تاب برای اندازه گیری در دسترس نیست) که اجازه حرکت در یک محور را دارد، در نظر گرفته‌اند. ایشان یک روش کنترل تطبیقی برای به حداقل رساندن اثر بار معلق بر روی پرده طراحی و پیاده سازی کرده‌اند[۱۲]. چن و همکاران به بررسی یک سیستم حمل و نقل هوایی دارای چهار روتور پرداخته‌اند که محموله را با چهار کابل حمل می‌کند. در این پژوهش تنظیمات پایدار برای نشان دادن مزایای سیستم چهار کابل مورد بحث قرار گرفته و یک کنترل کننده تطبیقی برای حمل بار توسط چهار روتور طراحی شده است[۱۳].

کیس و همکاران به طراحی و پیاده‌سازی کنترل کننده‌ی تطبیقی در حضور اغتشاشات خارجی پرداخته‌اند[۱۴]. پی و همکاران یک مدل ساده شده از کوادروتور را در نظر گرفته‌اند و به منظور کنترل آن از یادگیری تقویتی^۹ استفاده نموده‌اند. در این مقاله آموزش پرده در دو حالت هاور و تعقیب مسیر پیشنهادی (مسیر دایره‌ای) انجام گرفته است و در ادامه آموزش بر روی یک کواد روتور واقعی نیز اجرا شده است[۱۵].

سلطانی و همکاران کنترل کوادروتور حامل مایع را مورد مطالعه قرار داده و کنترل کننده‌ای طراحی نمودند که بر خلاف مطالعات پیشین، بدون نیاز به اندازه‌گیری یا تخمین حالات مایع می‌تواند پایداری کوادروتور را در انتقال

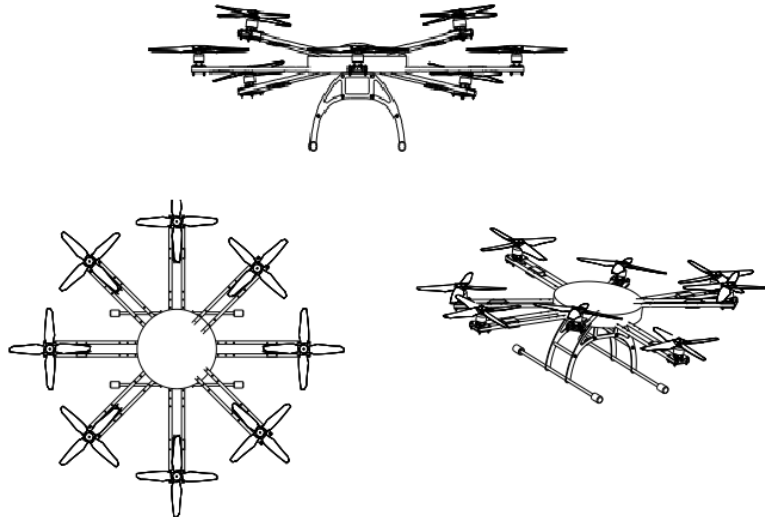
اینورنیزی و همکاران از یک مدلسازی مقید باهدف بهبود عملکرد کنترل کننده کوادکوپتر با زوایای روتور متغیر^۱ و اجرای فرایند تعقیب استفاده کرده اند[۴]. همچنین ایشان در سال ۲۰۱۸، مدلسازی و کنترل یک کوادکوپتر با زاویه بازوی متغیر را مورد بررسی قرار داده و اعتبارسنجی آن را با ساخت و تست مدل واقعی به انجام رسانده‌اند. در این پژوهش از یک کنترل کننده تناسبی-انتگرالی-مشتقی استفاده شده و از توانایی فعال‌سازی سیستم به منظور برطرف کردن مشکل ردیابی مسیر بهره برده شده است[۵]. آواز و نینگ به بررسی قابلیت روش ذرات گرداب چسبناک^۲ برای مدل سازی برهمکنش‌های آیرودینامیکی مولتی روتورها با در نظر گرفتن حجم محاسبات آن می‌پردازند، در واقع یک مدل مولتی روتور مبتنی بر روش ذرات گرداب چسبناک همراه با توصیه‌هایی برای ثبات عددی و کارایی محاسباتی معرفی شده است[۶].

ژو و همکاران برای رفع مشکلات اوتاروتورهای معمولی با ساختار تک لایه و کواکسیال، یک پیکربندی هشت روتوره جدید بر اساس بررسی اولیه آیرودینامیکی پیشنهاد داده‌اند. در این پژوهش برای دستیابی به یک تحلیل آیرودینامیکی دقیق، تحلیل با پنج مدل مختلف اوتاروتور با سه پیکربندی متفاوت (از جمله پیکربندی معمولی، پیکربندی کواکسیال و پیکربندی جدید) انجام گرفته است؛ ایشان شبیه سازی دینامیک سیالاتی را برای بررسی تأثیر اندازه تیغه روتور، باد و تداخل بین روتورها بر عملکرد و کارایی پرده اجرا کرده‌اند[۷].

علاوه بر مطالعات انجام گرفته در زمینه‌ی ساختار و چیدمان پروانه‌ها، تحقیقات بر طراحی کنترل کننده‌ی مناسب نیز مورد توجه پژوهشگران بوده است. با توجه به کارایی مورد انتظار از پرده در پژوهش‌های انجام گرفته انواع مختلف کنترل کننده‌ها مورد بررسی قرار گرفته است. کنترل کننده‌ی مقاوم در برابر نامعینی‌ها یکی از رایج‌ترین موضوعاتی است که به آن پرداخته شده است و در ادامه برخی از این پژوهش‌ها به عنوان نمونه آورده شده است. اسریدهار و همکاران به منظور کنترل حرکات زوایای رول^۳، پیچ^۴ و یاو^۵ یک کوادروتور تیلت روتور از روش کنترل مود لغزشی^۶ بهره برده‌اند. ایشان همچنین برای کنترل زوایای تیلت روتورها به صورت مستقل از کنترل

- 1 Tilt arm
- 2 viscous vortex particle method (VPM)
- 3 roll
- 4 pitch
- 5 yaw
- 6 Sliding Mode Control

7 Quantitative feedback theory
8 evolutionary algorithm with a Pareto approach
9 Reinforcement Learning



شکل ۱. شماتیک مدل اکتاروتور مورد نظر

Fig. 1. Schematic of the octorotor model

پرداخته شده و معادلات دینامیکی سیستم استخراج شده است. سپس مختصری در خصوص روش یادگیری تقویتی و الگوریتم یادگیری گرادیان سیاست گرایی قطعی با تأخیر دوقلو توضیحاتی داده شده و ساختار شبکه‌ی عصبی و نحوه‌ی طراحی کنترل کننده تشریح شده است. در نهایت نتایج شبیه‌سازی عددی و قابلیت مقاوم بودن در برابر تاثیر اختلالات احتمالی در عملکرد پرنده بررسی شده و نتیجه‌های گرفته شده تشریح شده است.

۲- ساختار و پیکره‌بندی اکتاروتور پیشنهادی

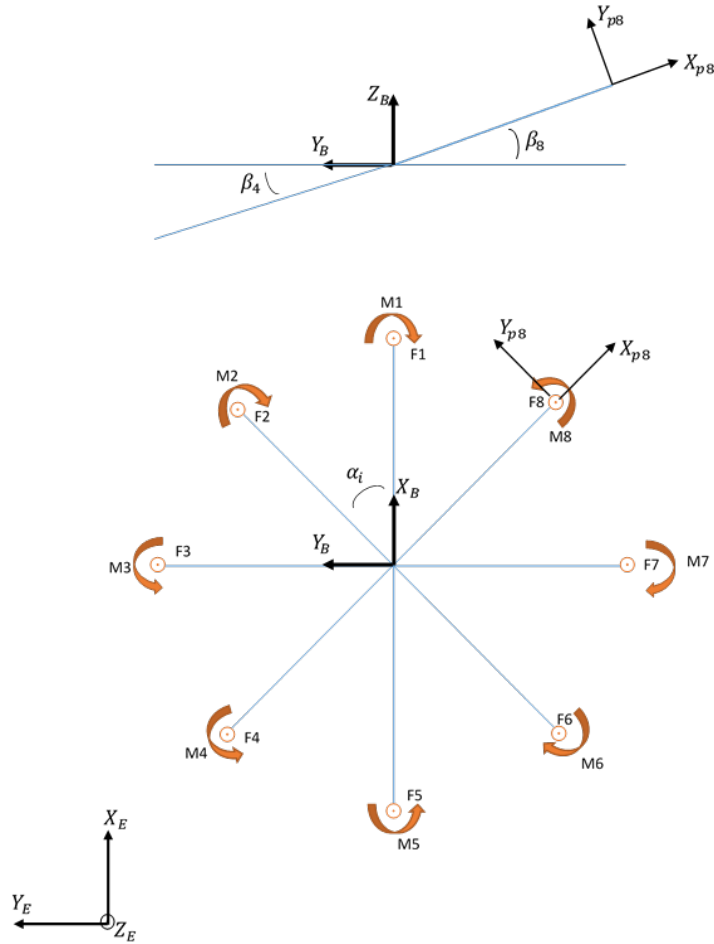
پیکره بندی یک اکتاروتور متداول را در ذهن داشته باشید که در آن تمامی روتورها در یک صفحه قرار گرفته و نیروی تراست حاصل از آن‌ها در یک راستا و به موازات یکدیگر هستند، در اینگونه مدل‌ها تنها قابلیت کنترل ۴ درجه آزادی به صورت مستقل (سه درجه‌ی آزادی موقعیت مرکز جرم و یک درجه‌ی آزادی زاویه‌ی یاو) وجود دارد [۱۹]. در این پژوهش به منظور دستیابی به شش درجه‌ی آزادی مستقل، ساختار اکتاروتور مطابق شکل ۱ به نحوی در نظر گرفته می‌شود که چهار روتور آن همانند یک اکتاروتور متداول جانمایی شده و چهار روتور دیگر بر روی چهار بازو با زاویه‌ی متغیر قرار خواهند گرفت. هر یک از چهار بازوی زاویه‌دار به نحوی جانمایی می‌گردند که یک طرف آن‌ها بر مرکز جرم اکتاروتور قرار گرفته و با صفحه‌ی چهار بازوی ثابت دارای زاویه‌ای متغیر هستند. با این آرایش، محور دوران چهار روتور قرار گرفته در یک صفحه با یکدیگر موازی بوده و محور هر یک از چهار

نقطه به نقطه فراهم نماید. بدین منظور، ابتدا یک کنترل کننده با خطی‌سازی معادلات حرکت کوادروتور حامل مایع و فرض صلب بودن مایع از طریق جایگذاری مناسب قطب‌ها طراحی شده است [۱۶].

شمس‌الهی و همکاران به سیستمی مرکب از یک کوادروتور و یک بازوی رباتیک سری پرداخته‌اند. ایشان یک الگوریتم تخمین کنترل برای دستیابی به ردیابی مسیر برای کوادروتور ارائه نمودند [۱۷]. لی و همکاران (۲۰۲۱) یک روش مبتنی بر یادگیری تقویتی را برای کنترل یک مولتی روتور در کاربردهای دنیای واقعی پیشنهاد داده‌اند. برای انجام این کار، یک تابع پاداش جدید برای یک مدل شبکه عصبی پیشنهاد شده که بازده انرژی را در نظر می‌گیرد [۱۸].

با توجه به پیشینه‌ی بیان شده در بالا، عمده‌ی پژوهش‌های انجام گرفته در زمینه‌ی مولتی روتورها به دو بخش ارائه‌ی ساختار و چیدمان جدید پروانه‌های پرنده و طراحی کنترل کننده‌ی متناسب با کارایی مورد انتظار از آن، قابل تقسیم می‌باشد. با توجه به افزایش مطالعات در این زمینه در سال‌های اخیر می‌توان به اهمیت این موضوع و لزوم ارتقای سیستم‌های موجود پی برد، بنابراین در این پژوهش نیز به هردو بخش پرداخته شده است. در این پژوهش سعی بر آن است تا با ایجاد ساختاری جدید در پیکره‌بندی روتورهای یک اکتاروتور و طراحی کنترل کننده‌ی مقاوم به روش یادگیری تقویتی، قابلیت کنترل موقعیت و زاویه‌ی اکتاروتور به صورت مجزا در پرنده ایجاد شود.

در ادامه، ابتدا به معرفی ساختار و پیکره‌بندی جدید برای یک اکتاروتور



شکل ۲. آرایش روتورها، نیروی پیشران و گشتاور تولیدی توسط روتورهای اوکتاروتور و نمایش دستگاه‌های مختصات

Fig. 2. Rotor configuration, thrust, and torque generated by octorotor rotors and reference frame display

روتور دیگر راستای متفاوتی با یکدیگر دارند. به منظور ثابت نگه داشتن مرکز

جرم در حین تغییر زوایا، زوایای دو بازوی رو در روی یکدیگر با هم یکسان

در نظر گرفته می‌شود. انتظار می‌رود با استفاده از تغییر در سرعت روتورها و زوایای بازوها امکان ایجاد حرکت مستقل در ۶ درجه‌ی آزادی فراهم گردد.

فرضیات:

- طول هر ۸ بازو (L_i) با یکدیگر برابر می‌باشد.

- چیدمان بازوها در صفحه $X-Y$ مطابق شکل ۲ به صورت متقارن

و با زوایای $\alpha_i = (i - 1) \pi / 4$ می‌باشد.

- چهار بازوی اول، سوم، پنجم و هفتم ثابت هستند

$\beta_r = \beta_f = \beta_s = \beta_x = 0$

- بازوهای بازو بهی متغیر که به صورت متقارن مقابل یکدیگر قرار گرفته‌اند، دوه‌دو دارای سرعت‌های دورانی و زوایای یکسانی

هستند. در واقع مطابق شکل ۲ $\beta_1 = \beta_5 = \beta_x$ ، $\omega_1 = \omega_5 = \omega_x$ ،

$\omega_r = \omega_v = \omega_y$ و $\beta_r = \beta_v = \beta_y$ می‌باشد.

- مرکز جرم کل پرنده ثابت و در مرکز هندسی پرنده قرار دارد.

۲-۱- دستگاه‌های مختصات مرجع

همانطور که در شکل ۲ نمایش داده شده است، سه نوع دستگاه مختصات

در نظر گرفته شده است:

- دستگاه مختصات F_E : به عنوان دستگاه مختصات اینرسی در نظر

گرفته شده است.

- دستگاه مختصات F_B : به عنوان دستگاه مختصات بدنی در نظر گرفته

شده است که مرکز آن بر مرکز جرم پرنده منطبق است.

- دستگاه مختصات F_{P_i} : ۸ دستگاه مختصات که مرکز آن‌ها بر مرکز

جرم روتورها منطبق است.

به منظور انتقال یک بردار از دستگاه مختصات بدنی به دستگاه مختصات

اینرسی می‌توان از ماتریس انتقال زیر استفاده نمود [۲۰].

۳-۱- دینامیک دورانی:

دینامیک دورانی سیستم در دستگاه بدنی به صورت زیر بیان می‌گردد [۲۱]:

$$I \dot{\omega} = -\omega \times I \omega + \tau + \tau_{ext} \quad (4)$$

در این رابطه I ممان اینرسی، τ گشتاورهای داخلی، τ_{ext} گشتاورهای ناشی نیروهای خارجی و ω سرعت زاویه‌ای F_B نسبت به F_E است که در دستگاه F_B نوشته شده است.

گشتاورهای داخلی τ به صورت زیر محاسبه می‌گردد:

$$\tau = \tau_{thrust} + \tau_{drag} + \tau_{gyroscopic} \quad (5)$$

که در آن τ_{thrust} گشتاور نیروی پیشران، τ_{drag} گشتاور پسا و $\tau_{gyroscopic}$ گشتاور ناشی از اثرات ژيروسکوپی هستند.

نیروی پیشران در دستگاه مختصات مستقر در مرکز جرم روتور به صورت زیر محاسبه می‌گردد.

$${}^{P_i} F_{thrust_i} = \begin{bmatrix} 0 & 0 & k_f \omega_i^2 \end{bmatrix}^T \quad (6)$$

در این رابطه k_f ضریب نیروی پیشران^۱ و ω_i سرعت چرخشی روتور i ام می‌باشد. بنابراین گشتاور نیروهای پیشران در دستگاه بدنی به صورت زیر محاسبه می‌گردد

$${}^B \tau_{thrust} = \sum_{i=1}^8 \left({}^B P_i \times {}^B R_{P_i} {}^{P_i} F_{thrust_i} \right) \quad (7)$$

گشتاور پسا^۲ در دستگاه مختصات مستقر در مرکز جرم روتور به صورت زیر محاسبه می‌گردد.

$${}^{P_i} \tau_{drag_i} = \begin{bmatrix} 0 & 0 & \mu_i k_m \omega_i^2 \end{bmatrix}^T \quad (8)$$

$${}^E R_B = R_{z,\psi} R_{y,\theta} R_{x,\varphi} = \begin{bmatrix} c_\psi s_\theta & c_\psi s_\theta s_\varphi - s_\psi c_\varphi & c_\psi s_\theta c_\varphi + s_\psi s_\varphi \\ c_\theta s_\psi & s_\psi s_\theta s_\varphi - c_\psi s_\varphi & s_\psi s_\theta c_\varphi - c_\psi s_\varphi \\ -s_\theta & c_\theta s_\varphi & c_\theta c_\varphi \end{bmatrix} \quad (1)$$

در رابطه‌ی بالا نماد های S و C به ترتیب برابر sin و cos بوده و زوایای φ ، θ و ψ زوایای رول، پیچ و یاء هستند و جهت مثبت آن‌ها براساس قانون دست راست تعیین می‌گردد.

همچنین به منظور انتقال یک بردار از دستگاه مختصات روتورها به دستگاه مختصات بدنی می‌توان از ماتریس انتقال زیر استفاده نمود. لازم به ذکر است در اوکتاروتور پیشنهادی مقدار β_i برای چهار روتور برابر صفر و برای چهار روتور دیگر مقادیر متغیری می‌باشد.

$${}^B R_{P_i} = R_{z,\alpha_i} R_{y,\beta_i} = \begin{bmatrix} c_{\alpha_i} & -s_{\alpha_i} & 0 \\ s_{\alpha_i} & c_{\alpha_i} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_{\beta_i} & 0 & s_{\beta_i} \\ 0 & 1 & 0 \\ -s_{\beta_i} & 0 & c_{\beta_i} \end{bmatrix} \quad (2)$$

باتوجه به رابطه‌ی (۲)، مختصات مرکز جرم روتور i ام (P_i) در دستگاه F_B به صورت زیر محاسبه می‌گردد.

$${}^B P_i = {}^B R_{P_i} \begin{bmatrix} L_i \\ 0 \\ 0 \end{bmatrix} \quad (3)$$

که در این رابطه L_i طول بازوی i ام می‌باشد.

۳- معادلات دینامیکی

با استفاده از روش استاندارد نیوتن اولیبر برای سیستم‌های دینامیکی، می‌توان معادلات دینامیکی مدل پیشنهادی اوکتاروتور را با در نظر گرفتن نیروها و گشتاورهایی که توسط هر روتور ایجاد می‌شود همراه با اثرات ژيروسکوپی و اینرسی به دست آورد.

1 propeller thrust coefficient

2 drag torque

$$m\ddot{X} = m \begin{bmatrix} 0 \\ 0 \\ -g \end{bmatrix} + {}^E R_B {}^B F_{thrust} + f_{ext} \quad (13)$$

در این رابطه توابع m به عنوان مجموع جرم پرنده، f_{ext} به عنوان نیروی خارجی و ${}^B F_{thrust}$ به عنوان مجموع نیروی پیشران در دستگاه بدنی به صورت زیر تعریف می‌گردند

$${}^B F_{thrust} = \sum_{i=1}^8 {}^B R_{P_i} {}^{P_i} F_{thrust_i} \quad (14)$$

۴- طراحی کنترلر به روش یادگیری تقویتی

یادگیری تقویتی یکی از روش‌های یادگیری در سیستم‌های هوشمند است که بر اساس رابطه‌ی علت و معلولی عمل می‌کند. در این روش یادگیری، عامل هوشمند با توجه به وضعیتی که در محیط دارد، عملی را بر روی محیط انجام می‌دهد و منتظر نتیجه‌ی عملش می‌ماند، این نتیجه می‌تواند در قالب یک پاداش یا تنبیه باشد. اگر نتیجه در قالب پاداش باشد، عمل انجام شده مطلوب بوده و عامل به هدفی که در آن محیط دارد نزدیک شده است، ولی اگر نتیجه در قالب تنبیه باشد، عمل انجام شده نامطلوب بوده و عامل از هدفش دور شده است. عامل باید یاد بگیرد که چه اعمالی را انجام دهد تا پاداش بیشتری را کسب کند و در نهایت به هدفش برسد.

در واقع هدف از یادگیری تقویتی آموزش یک عامل برای تکمیل یک کار در یک محیط نامشخص است؛ عامل مشاهدات و پاداش را از محیط دریافت می‌کند و اقداماتی را برای محیط ارسال می‌کند و پاداش معیاری است که نشان می‌دهد یک عمل تا چه حد در تکمیل هدف موفق است [۲۳]. دو بخش مهم در روش یادگیری تقویتی شامل خط مشی^۲ و الگوریتم یادگیری است. خط مشی اقدامات را بر اساس مشاهدات از محیط^۳ انتخاب می‌کند و به طور معمول، یک تابع تقریب با پارامترهای قابل تنظیم یا یک شبکه عصبی عمیق است. همچنین الگوریتم یادگیری به طور مداوم پارامترهای خط مشی را بر اساس اقدامات، مشاهدات و پاداش‌ها به روز می‌کند. هدف الگوریتم یادگیری یافتن خط مشی بهینه‌ای است که پاداش انباشته مورد انتظار در درازمدت را به حداکثر برساند.

در این رابطه k_m ضریب پسای روتورها^۱ و μ_i تعیین کننده‌ی جهت اعمال گشتاورپسا می‌باشد که برای هر روتور در جهت عکس دوران روتور هست. با توجه به ساختار در نظر گرفته شده برای اوکتاروتور در این پژوهش (شکل ۲) بردار μ_i به صورت زیر می‌باشد:

$$\mu = [-1 \quad -1 \quad 1 \quad 1 \quad 1 \quad -1 \quad -1 \quad 1] \quad (9)$$

در نتیجه مجموع گشتاورهای پسا در دستگاه بدنی به صورت زیر قابل محاسبه می‌باشد:

$${}^B \tau_{drag} = \sum_{i=1}^8 {}^B R_{P_i} {}^{P_i} \tau_{drag_i} \quad (10)$$

گشتاور ناشی از اثرات ژيروسکوپیی در دستگاه بدنی به صورت زیر محاسبه می‌گردد [۲۲]:

$${}^B \tau_{gyroscopic} = -\sum_{i=1}^8 \omega \times J_i \bar{\omega}_i \quad (11)$$

که در آن J_i ممان اینرسی روتور^۱ ام می‌باشد. با جایگذاری روابط (۷) و (۹) و (۱۱) در رابطه‌ی (۵)، گشتاورهای داخلی τ به صورت زیر محاسبه می‌گردد:

$$\begin{aligned} \tau &= \tau_{thrust} + \tau_{drag} + \tau_{gyroscopic} = \\ & \sum_{i=1}^8 \left({}^B P_i \times {}^B R_{P_i} {}^{P_i} F_{thrust_i} \right) + \\ & \sum_{i=1}^8 {}^B R_{P_i} {}^{P_i} \tau_{drag_i} - \sum_{i=1}^8 \omega \times J_i \bar{\omega}_i \end{aligned} \quad (12)$$

۳-۲- دینامیک انتقالی

دینامیک انتقالی سیستم در دستگاه اینرسی به صورت زیر بیان می‌گردد [۲۱]:

2 policy
3 environment

1 propeller drag coefficient

جدول ۱. الگوریتم آموزش به روش TD3
Table 1. TD3-based training algorithm

الگوریتم TD3
* مقاردهی اولیه برای شبکه‌های منتقد ^۱ Q_{μ_1} ، Q_{μ_2} و شبکه‌ی عملگر ^۲ π_{η} با پارامترهای تصادفی μ_1 ، μ_2 و η
* مقاردهی اولیه به شبکه‌های هدف $\mu'_1 \leftarrow \mu_1$ ، $\mu'_2 \leftarrow \mu_2$ ، $\eta' \leftarrow \eta$
* آماده‌سازی فضای ذخیره سازی تجربه‌ها B
* شروع حلقه‌ی تکرار برای $t = 1$ to T
- تعیین اکشن به همراه نویز اکتشافی ^۳ $\epsilon \sim N(0, \sigma)$ که $a \sim \pi_{\eta}(s) + \epsilon$ در حالی که ϵ می‌باشد و به دست آوردن پاداش r و موقعیت جدید s'
- ذخیره سازی مجموعه‌ی (s, a, r, s') در B
- تشکیل یک نمونه‌ی N تایی از تجربه‌های قبلی ذخیره شده در B
- $\tilde{a} \leftarrow \pi_{\eta}(s') + \epsilon$ ، $\epsilon \sim clip(N(0, \bar{\sigma}), -c, c)$
- $h \leftarrow r + \gamma \min_{i=1,2} Q_{\mu'_i}(s', \tilde{a})$
- به روز رسانی پارامترهای شبکه‌ی منتقد: $\mu_i \leftarrow arg \min_{\mu_i} N^{-1} \Sigma(h - Q_{\mu_i}(s, a))$ ^۴
- اگر شماره‌ی اپیزود t در مجموعه‌ی b قرار داشته باشد:
- به روز رسانی پارامترهای شبکه‌ی عملگر (η) بوسیله‌ی گرادینت خط مشی قطعی ^۴
$\nabla_{\eta} J(\eta) = N^{-1} \Sigma \nabla_a Q_{\mu_1}(s, a) \Big _{a=\pi_{\eta}(s)} \nabla_{\eta} \pi_{\eta}(s)$
- به روز رسانی پارامترهای شبکه‌های هدف
$\mu'_i \leftarrow \chi \mu_i + (1 - \chi) \mu'_i$
$\eta' \leftarrow \chi \eta + (1 - \chi) \eta'$
- پایان شرط
* پایان حلقه‌ی تکرار

¹ critic networks

² actor network

³ exploration noise

⁴ deterministic policy gradient

۴-۱ الگوریتم آموزش TD3

می‌توانند توابع ارزش^۶ را بیش از حد تخمین بزنند و سبب ایجاد خط مشی غیر بهینه شوند. برای کاهش تخمین بیش از حد تابع ارزش، تغییراتی (نظیر آموزش دو تابع ارزش Q ^۷ و استفاده از مقدار حداقل تابع ارزش در طول به‌روزرسانی خط‌مشی، به‌روزرسانی خط‌مشی با تناوبی کمتر از توابع Q و اضافه کردن نویز به عامل هدف^۸) در الگوریتم گرادینت سیاست گرای قطعی داده شده و الگوریتم گرادینت سیاست گرای قطعی با تأخیر دوقلو ایجاد شده است. الگوریتم آموزش گرادینت سیاست گرای قطعی با تأخیر دوقلو در جدول ۱ بیان شده است [۲۴].

در جدول ۱ پارامترهای Q_{μ_1} ، Q_{μ_2} و π_{η} به ترتیب بیانگر شبکه‌های منتقد اول و دوم و شبکه‌ی عملگر می‌باشد. پارامترهای μ_1 ، μ_2 و η بیانگر پارامترهای قابل تنظیم شبکه‌های منتقد و شبکه‌ی عملگر می‌باشد و μ'_1 ، μ'_2

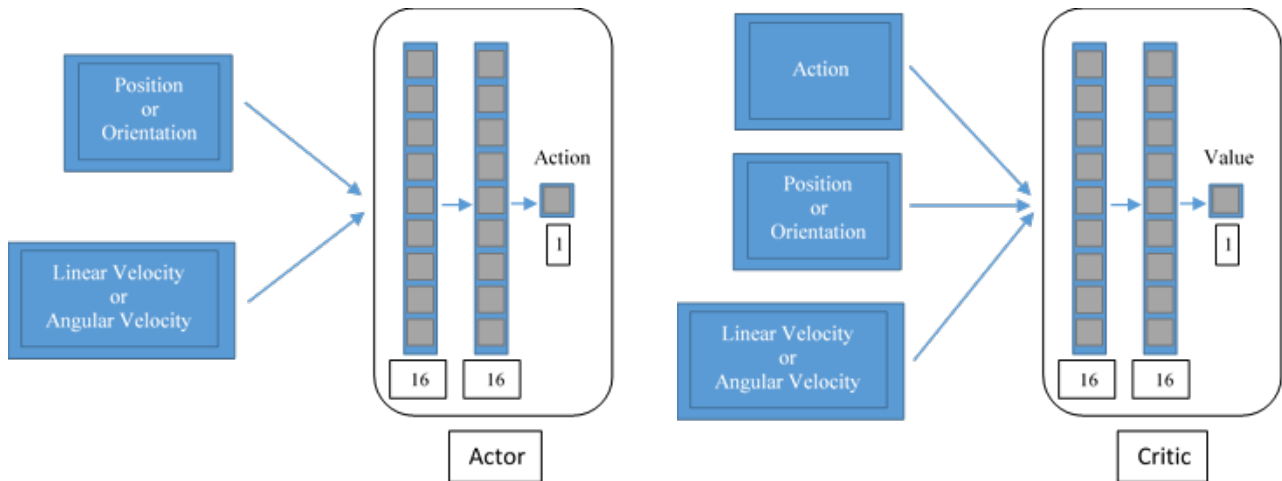
در روش یادگیری تقویتی از الگوریتم‌های آموزشی متعددی برای حل مسئله استفاده شده است. در این پژوهش از روش یادگیری گرادینت سیاست گرای قطعی با تأخیر دوقلو^۱ به منظور آموزش شبکه‌ی عصبی عمیق استفاده شده است. این الگوریتم یک روش یادگیری تقویتی بدون مدل^۲، برخط^۳ و بدون سیاست^۴ است و یک عامل آن یک عامل یادگیری تقویتی عملگر-منتقد^۵ است که به دنبال یک خط مشی بهینه برای به حداکثر رساندن پاداش انباشته درازمدت می‌باشد. در واقع این الگوریتم توسعه‌ای از الگوریتم گرادینت سیاست گرای قطعی است. عوامل گرادینت سیاست گرای قطعی

- 1 Twin-Delayed Deep Deterministic Policy Gradient
- 2 model-free
- 3 online
- 4 off-policy
- 5 actor-critic

6 value functions

7 Q-value

8 target action



شکل ۳. ساختار شبکه های عملگر و منتقد

Fig. 3. Actor-critic network configuration

و پارامترهای τ_x ، τ_y ، τ_z به ترتیب بیانگر گشتاور مورد نیاز حول محورهای X، Y و Z به منظور رسیدن به وضعیت مطلوب می باشد. به منظور طراحی شبکه‌ی عصبی و آموزش آن به روش یادگیری تقویتی، برای هر درجه‌ی آزادی نیاز به یک شبکه‌ی عملگر و دو شبکه‌ی منتقد می باشد که مطابق شکل ۳ در نظر گرفته شده است. ساختار هر سه شبکه مشابه هم و به صورت شبکه‌ای با لایه‌ی ورودی دارای ۲ نورون شامل دو پارامتر موقعیت در یک راستا و سرعت خطی در همان راستا (یا وضعیت و سرعت دورانی)، دو لایه‌ی پنهان با ۱۶ نورون و تابع فعال‌سازی یکسوساز^۱ و لایه‌ی خروجی با ۱ نورون در نظر گرفته شده است. در شبکه‌ی عملگر از تابع فعال‌سازی تانژانت هذلولوی^۲ استفاده شده است و خروجی آن به منظور تعیین نیروی (گشتاور) مورد نیاز در راستای مورد نظر به منظور رسیدن به موقعیت (یا وضعیت) هدف استفاده می گردد؛ همچنین از خروجی شبکه‌های منتقد مطابق الگوریتم بیان شده در جدول ۱ به منظور تخمین تابع ارزش استفاده می شود.

آموزش شبکه به نحوی انجام می گیرد که با فرض قرارگیری پرنده در هر موقعیت و وضعیتی و با هر سرعتی به سمت صفر یا به عبارتی قرارگیری در مبدا دستگاه مختصات هدایت شود و در آن محل به صورت شناور بماند. در هنگام آموزش موقعیت در بازه‌ای با شعاع ۲ متر و وضعیت در بازه‌ای با شعاع ۶۰ درجه به صورت تصادفی انتخاب می شوند. نکته‌ی قابل ذکر در مورد

μ' و η' مقادیر تصادفی اختصاص داده شده به آن‌ها می باشد. S بیانگر فضای حالت فعلی، S' بیانگر فضای حالت جدید، I بیانگر میزان پاداش در هر اپیزود، h بیانگر میزان تابع ارزش، ϵ بیانگر میزان نویز در نظر گرفته شده، a بیانگر میزان اکشن و γ بیانگر ضریب به روزرسانی پارامترهای شبکه‌های هدف می باشد.

۴-۲- ساختار کنترلی پیشنهادی

هدف از این پژوهش طراحی یک کنترل کننده شبکه‌ی عصبی به روش یادگیری تقویتی به منظور کنترل پرنده در ۶ درجه‌ی آزادی به صورت مستقل است. برای این منظور در مرحله‌ی آموزش به شبکه‌های عصبی از روابط کلی زیر به عنوان ۶ رابطه‌ی دینامیکی مستقل برای در نظر گرفتن محیط آموزش استفاده شده است و ۶ شبکه به صورت مجزا برای کنترل آن‌ها آموزش داده شده است.

$$m\ddot{X} = [F_x \quad F_y \quad F_z] \quad (15)$$

$$I\dot{\omega} + \omega \times I\omega = [\tau_x \quad \tau_y \quad \tau_z] \quad (16)$$

در این روابط، پارامترهای F_x ، F_y و F_z به ترتیب بیانگر نیروی مورد نیاز در راستاهای X، Y و Z به منظور رسیدن به موقعیت مطلوب

1 ReLU
2 Hyperbolic Tangent

بخش ۲ و با محاسبه‌ی برآیند نیروها در راستای Z (سطر سوم معادلات (۱۸)) خواهیم داشت:

$$z : \underbrace{\varpi_x^2 (-2k_f \sin \beta_x \sin \theta + 2k_f \cos \beta_x \cos \theta \cos \varphi)}_{p.1} + \underbrace{\varpi_y^2 (2k_f \sin \beta_y \cos \theta \sin \varphi + 2k_f \cos \beta_y \cos \theta \cos \varphi)}_{p.2} + \underbrace{(\varpi_2^2 + \varpi_4^2 + \varpi_6^2 + \varpi_8^2) k_f \cos \theta \cos \varphi}_{p.3} = mg + F_z \quad (19)$$

با فرض این که κ درصد از نیروهای راستای Z توسط روتورهای با زاویه‌ی ثابت و $(1-\kappa)$ درصد توسط روتورهای ثابت تعیین گردد از سه بخش رابطه‌ی (۱۹) خواهیم داشت:

$$p.1: \varpi_x^2 = \frac{\kappa(mg + F_z)}{-k_f \sin \beta_x \sin \theta + k_f \cos \beta_x \cos \theta \cos \varphi} \quad (20)$$

$$p.2: \varpi_y^2 = \frac{\kappa(mg + F_z)}{k_f \sin \beta_y \cos \theta \sin \varphi + k_f \cos \beta_y \cos \theta \cos \varphi} \quad (21)$$

$$p.3: (\varpi_2^2 + \varpi_4^2 + \varpi_6^2 + \varpi_8^2) = \frac{(1-\kappa)(mg + F_z)}{k_f \cos \theta \cos \varphi} \quad (22)$$

با محاسبه‌ی برآیند نیروها در راستای X و Y در معادلات (۱۸) و جایگذاری روابط (۲۰) و (۲۱) خواهیم داشت:

$$\tan \beta_x = \frac{(F_x \cos \theta \cos \psi + F_y \cos \theta \sin \psi - (1-\kappa)(mg + F_z)) + (F_x \cos \theta \cos \psi + F_y \cos \theta \sin \psi)}{(1-\kappa)(mg + F_z)} \quad (23)$$

$$\frac{mg \sin \theta - F_z \sin \theta \cos \theta \cos \varphi}{\cos \theta \sin \psi - mg \sin \theta - F_z \sin \theta \sin \theta}$$

این نحوه‌ی آموزش به شبکه این است که این یادگیری سبب ایجاد کنترلر مقاوم در پرنده نیز خواهد شد چراکه عوامل خارجی سبب انحراف پرنده از مسیر هدف شده و شبکه آموزش دیده است از هر موقعیت و وضعیتی و با هر سرعتی به مسیرهدف بازگردد.

تابع پاداش در آموزش نیز به نحوی در نظر گرفته می‌شود که علاوه بر حرکت به سمت هدف با کمترین تلاش کنترلی به هدف دست یابد. به عنوان نمونه برای آموزش در راستای X تابع پاداش به صورت زیر تعریف شده است:

$$r = -\rho(x - x_d)^2 - (1-\rho)u^2 \quad (17)$$

که در آن $x_d = 0$ بوده و بیانگر موقعیت هدف در راستای X، پارامتر u بیانگر تلاش کنترلی در راستای X و ρ بیانگر ضریب تاثیر می‌باشد. همچنین قابل ذکر است برای موقعیت هدف دیگری به جز مبدا مختصات، نیاز به آموزش مجدد شبکه نیست و تنها با ایجاد یک تغییر متغیر در دینامیک مسئله می‌توان پرنده را به سمت اهداف مورد نظر هدایت نمود.

مزیت دیگر این آموزش این است که با توجه به مشابهت ساختار شبکه، رابطه‌ی دینامیکی و هدف کنترلی در ۳ درجه‌ی آزادی انتقالی، می‌توان از شبکه‌ی آموزش داده شده در راستای X برای راستاهای Y و Z نیز استفاده نمود و نیازی به آموزش مجدد در این راستاها نیست. همچنین از شبکه‌ی آموزش داده شده برای دوران حول محور X می‌توان برای دوران حول محورهای Y و Z نیز استفاده کرد.

۴-۲-۱- محاسبه‌ی سرعت روتورها وزوایای بازوها

در مدل اوکتاروتور پیشنهادی ۶ خروجی نیرو و گشتاور گرفته شده از شبکه‌ها باید مبنای محاسبه‌ی سرعت ۸ روتور اصلی پرنده و زوایای چهار بازو قرار بگیرد. برای این منظور به صورت زیر عمل می‌شود.

با توجه به روابط (۱۳) و (۱۵) خواهیم داشت:

$$m \begin{bmatrix} 0 \\ 0 \\ -g \end{bmatrix} + {}^E R_B \left(\sum_{i=1}^8 {}^B R_{P_i} {}^{P_i} F_{thrust_i} \right) + f_{ext} = \begin{bmatrix} F_x \\ F_y \\ F_z \end{bmatrix} \quad (18)$$

با در نظر گرفتن فرضیات مدل در نظر گرفته شده برای اوکتاروتور در

$$\tan \beta_y = \frac{(F_x (\cos \psi \sin \theta \sin \varphi - \sin \psi \cos \varphi) + F_y (\sin \psi \sin \theta \sin \varphi + \cos \psi \cos \varphi)) \dots}{(1-\kappa)(mg + F_z) - (F_x (\cos \psi \sin \theta \sin \varphi - \sin \psi \cos \varphi) + (mg + F_z) \cos \theta \sin \varphi) \cos \theta \cos \varphi} \dots \quad (24)$$

$$+ F_y (\sin \psi \sin \theta \sin \varphi + \cos \psi \cos \varphi) + (mg + F_z) \cos \theta \sin \varphi \cos \theta \sin \varphi$$

$$\omega_8^2 = \left(\begin{array}{cc} -\frac{\tau_y}{k_f L \cos \frac{\pi}{4}} & -\frac{\tau_x}{k_f L \cos \frac{\pi}{4}} \\ +\frac{\tau_z}{k_m} + \frac{(1-\kappa)(mg + F_z)}{k_f \cos \theta \cos \varphi} \end{array} \right) / 4 \quad (29)$$

در نهایت با استفاده از روابط (۲۰)، (۲۱)، (۲۳)، (۲۴) و (۲۶) تا (۲۹) سرعت‌های دورانی ۸ روتور و زوایای بازوهای غیرثابت براساس خروجی شبکه‌های عصبی قابل محاسبه می‌باشد.

۵- شبیه سازی و نتایج

۵-۱- شبیه‌سازی عددی

به منظور اجرای شبیه‌سازی عددی پارامترهای مربوط به اوکتاروتور مطابق جدول ۲ در نظر گرفته شده است.

همانطور که بیان شد آموزش شبکه به نحوی انجام می‌گیرد که با فرض قرارگیری پرنده در هر موقعیت و وضعیتی و با هر سرعتی به سمت صفر یا به عبارتی قرارگیری در مبدا دستگاه مختصات هدایت شود و در آن محل به صورت شناور بماند. هر اپیزود مربوط به آموزش شبکه در ۴ ثانیه که شامل ۴۰۰ گام زمانی^۱ است انجام گرفته است. در این پژوهش آموزش کامل شبکه (به منظور پایداری در حالت شناوری) در کمتر از ۲۰۰ هزار گام زمانی برای هر درجه آزادی انجام گرفته است، در حالی که در سایر پژوهش‌ها که از یادگیری تقویتی برای کنترل کوادروتور استفاده کرده اند، نظیر هانگبو و همکاران [۲۵] حدود ۲۱۵۰ میلیون گام زمانی و پی و همکاران [۱۵] حدود ۱۰ میلیون گام زمانی برای دستیابی به پایداری مناسب در حالت شناوری نیاز داشته اند. در واقع مدل پیشنهادی و مستقل بودن درجات آزادی امکان کاهش ابعاد شبکه نسبت به پژوهش‌های قبلی و در نتیجه کاهش چشمگیر زمان آموزش را فراهم کرده است که از مزیت‌های این پژوهش محسوب می‌گردد.

شکل ۴ و ۵ بیانگر هدایت پرنده به سمت قرارگیری در مرکز مختصات

همچنین با توجه به روابط (۴) و (۱۲) برای دینامیک اوکتاروتور و رابطه‌ی (۱۶) خواهیم داشت:

$$\sum_{i=1}^8 ({}^B P_i \times {}^B R_{P_i} {}^{P_i} F_{thrust_i}) + \sum_{i=1}^8 {}^B R_{P_i} {}^{P_i} \tau_{drag_i}^B \quad (25)$$

$$- \sum_{i=1}^8 \omega \times J_i \bar{\omega}_i + \tau_{ext} = \begin{bmatrix} \tau_x \\ \tau_y \\ \tau_z \end{bmatrix}$$

با استفاده از سه معادله‌ی (۲۵) و معادله‌ی (۲۲) یک دستگاه چهار معادله و چهار مجهول خواهیم داشت که نتیجه‌ی آن به صورت زیر خواهد بود:

$$\omega_2^2 = \left(\begin{array}{cc} -\frac{\tau_y}{k_f L \cos \frac{\pi}{4}} + \frac{\tau_x}{k_f L \cos \frac{\pi}{4}} \\ -\frac{\tau_z}{k_m} + \frac{(1-\kappa)(mg + F_z)}{k_f \cos \theta \cos \varphi} \end{array} \right) / 4 \quad (26)$$

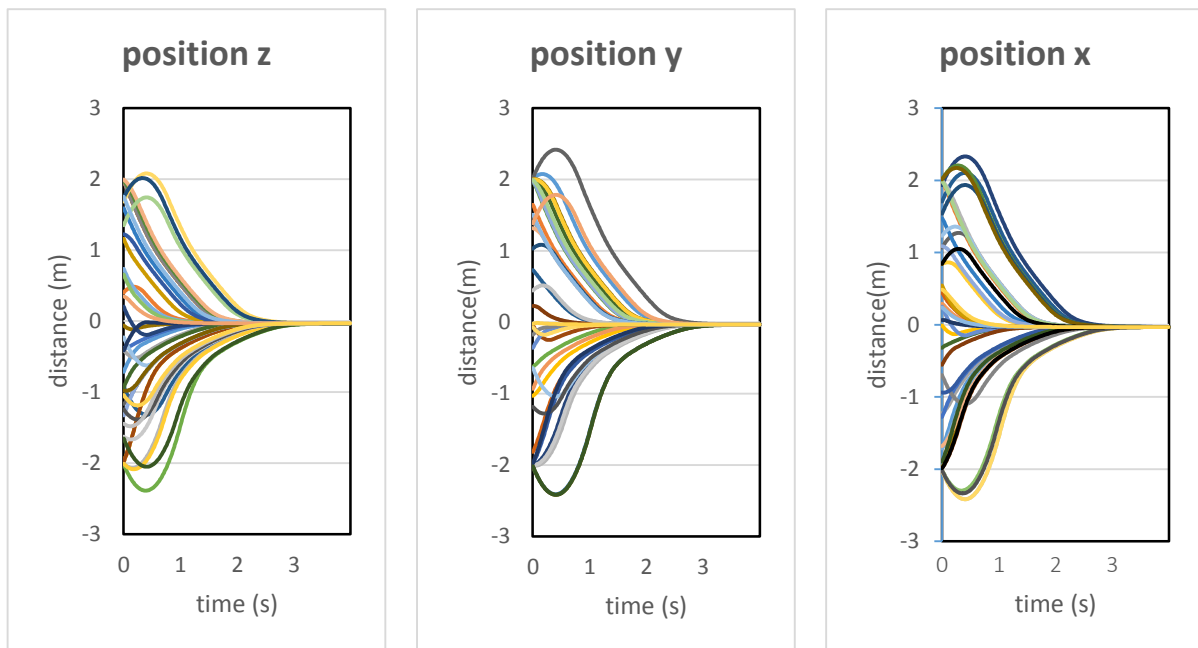
$$\omega_4^2 = \left(\begin{array}{cc} -\frac{\tau_y}{k_f L \cos \frac{\pi}{4}} + \frac{\tau_x}{k_f L \cos \frac{\pi}{4}} \\ +\frac{\tau_z}{k_m} + \frac{(1-\kappa)(mg + F_z)}{k_f \cos \theta \cos \varphi} \end{array} \right) / 4 \quad (27)$$

$$\omega_6^2 = \left(\begin{array}{cc} -\frac{\tau_y}{k_f L \cos \frac{\pi}{4}} - \frac{\tau_x}{k_f L \cos \frac{\pi}{4}} \\ -\frac{\tau_z}{k_m} + \frac{(1-\kappa)(mg + F_z)}{k_f \cos \theta \cos \varphi} \end{array} \right) / 4 \quad (28)$$

جدول ۲. پارامترهای اوکتاروتور

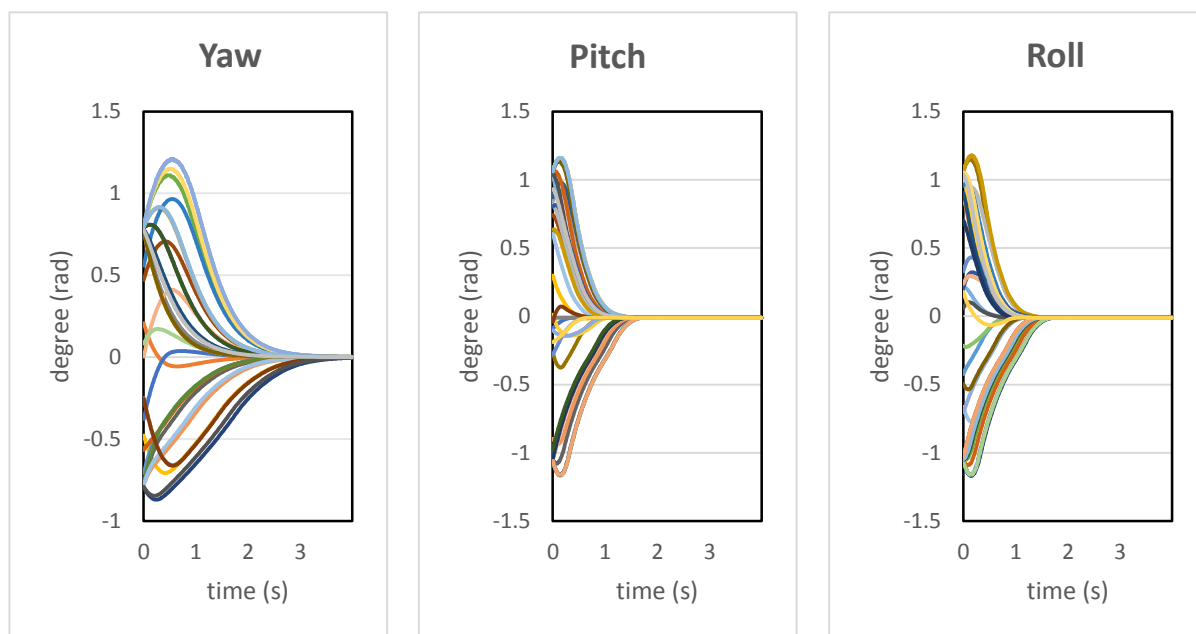
Table 2. Octorotor parameters

مقادیر	پارامترها
۱ (kg)	جرم
۰/۲۳ (m)	طول بازوها
$۳/۱۳ \times ۱۰^{-۵}$ (N.s ^۲)	ضریب نیروی پیشران
$۷/۵ \times ۱۰^{-۷}$ (N.ms ^۲)	ضریب پسای روتورها
$۷/۵ \times ۱۰^{-۳}$ (kg.m ^۲)	ممان اینرسی حول محور x
$۷/۵ \times ۱۰^{-۳}$ (kg.m ^۲)	ممان اینرسی حول محور y
$۱/۳ \times ۱۰^{-۲}$ (kg.m ^۲)	ممان اینرسی حول محور z



شکل ۴. فاصله‌ی محورهای x، y و z اوکتاروتور از مرکز [۰،۰،۰] در مانور معلق ماندن در ۴ ثانیه. نقاط شروع از ۴۰ حالت مختلف با موقعیت و سرعت خطی تصادفی انتخاب شده است.

Fig. 4. The distance of the x, y, and z axes of the octorotor from the origin [0,0,0] in the hovering maneuver within four seconds. Starting points are selected from forty modes with random position and linear velocity



شکل ۵. میزان رول، پیچ و یاو در مانور معلق ماندن در ۴ ثانیه. نقاط شروع از ۴۰ حالت مختلف با وضعیت و سرعت زاویه ای تصادفی انتخاب شده است.

Fig. 5. The roll, pitch, and yaw rates in the hovering maneuver within four seconds. The starting points are selected from forty modes with random position and angular velocity.

هدایت سریع پرنده به مسیر هدف از پیش تعیین شده می‌باشد. موارد مطرح شده در بالا و نتایج ترسیم شده در شکل‌های ۴، ۵ و ۶ بیانگر عملکرد مناسب کنترلر در محیط شبیه‌سازی می‌باشد در حالی که عملکرد پرنده در محیط واقعی می‌تواند تحت تاثیر عوامل مختلفی قرار گرفته و عملکرد مناسبی نداشته باشد. از جمله این عوامل می‌توان به عملکرد ضعیف باتری، اختلال در سیگنال‌های ارسالی و یا عملکرد معیوب موتورها اشاره نمود، بنابراین برای عملکرد مناسب کنترلر در محیط واقعی باید کنترلر نسبت به این تغییرات مقاوم باشد. با توجه به این که ایجاد این اختلالات را می‌توان باعث عدم اعمال نیروی کافی برای تعقیب مسیر از پیش تعیین شده و در نتیجه انحراف از مسیر دانست و توجه به این نکته که آموزش شبکه به نحوی انجام گرفته است که در هر شرایطی (موقعیت و سرعت دلخواه) به سمت هدف هدایت می‌گردد، بنابراین کنترلر پیشنهادی از قابلیت مقاوم بودن در برابر اختلالات بیان شده برخوردار هست.

اگر نتیجه‌ی ایجاد اختلال در پرنده را به صورت ضعف در اعمال نیروی لازم تعریف نماییم یا به عبارت دیگر تنها درصدی از نیروهای تعیین شده توسط شبکه در محیط واقعی اعمال شود، نیروی اعمال شده توسط موتورها

و برقراری پایداری در حالت شناوری با شروع از ۴۰ حالت تصادفی می‌باشد. لازم به ذکر است هر شش درجه‌ی آزادی به صورت همزمان و مستقل عمل می‌کند.

به منظور بررسی قابلیت کنترل اکتاروتور در مسیرها و جهت‌گیری‌های از پیش تعیین شده، مطابق جدول ۳ برای جابه‌جایی در راستای محور X از تابع مرتبه دوم، برای جابه‌جایی در راستای محورهای Y و Z از توابع خطی و برای رول و پیچ از توابع پله و برای یاو از تابع ثابت برای تعیین مسیر از پیش تعیین شده استفاده شده است که هر شش درجه‌ی آزادی باید به صورت همزمان و مستقل آن‌ها را تعقیب نمایند.

نمودارهای ترسیم شده در شکل ۶ بیانگر قابلیت اکتاروتور پیشنهادی در کنترل مستقل شش درجه‌ی آزادی در مسیرها و جهت‌گیری‌های مطرح شده در جدول ۳ در طی ۱۰ ثانیه می‌باشد که در واقع ثابت‌کننده‌ی قابلیت‌های ساختار و کنترلر پیشنهادی برای اکتاروتور هست. بیشترین میزان انحراف در شکل ۶ (الف) مربوط به انحراف پرنده در راستای X از مسیر تعیین شده در ثانیه‌ی ۸ می‌باشد که به دلیل تغییر ناگهانی سرعت از پیش تعیین شده از $1/33$ - به صفر (مطابق جدول ۳) می‌باشد، نمودار ترسیم شده بیانگر قابلیت کنترلر در

جدول ۳. مسیرها و جهت گیری های از پیش تعیین شده برای اوکتاروتور

Table 3. Preset trajectories and orientations for the octorotor

جابه جایی در راستای محور Z:	جابه جایی در راستای محور Y:	جابه جایی در راستای محور X:
$\begin{cases} \begin{cases} z = 0 \\ \dot{z} = 0 \end{cases} & t < 1 \\ \begin{cases} z = 0.5t - 0.5 \\ \dot{z} = 0.5 \end{cases} & 1 \leq t < 3 \\ \begin{cases} z = 1 \\ \dot{z} = 0 \end{cases} & 3 \leq t < 6 \\ \begin{cases} z = -0.5t + 4 \\ \dot{z} = -0.5 \end{cases} & 6 \leq t < 8 \\ \begin{cases} z = 0 \\ \dot{z} = 0 \end{cases} & t \geq 8 \end{cases}$	$\begin{cases} \begin{cases} y = 0 \\ \dot{y} = 0 \end{cases} & t < 4 \\ \begin{cases} y = 0.25t - 1 \\ \dot{y} = 0.25 \end{cases} & 4 \leq t < 6 \\ \begin{cases} y = -0.25t + 2 \\ \dot{y} = -0.25 \end{cases} & 6 \leq t < 8 \\ \begin{cases} y = 0 \\ \dot{y} = 0 \end{cases} & t \geq 8 \end{cases}$	$\begin{cases} \begin{cases} x = 0 \\ \dot{x} = 0 \end{cases} & t < 2 \\ \begin{cases} x = \frac{-t^2 + 10t - 16}{4/5} \\ \dot{x} = \frac{-2t + 10}{4/5} \end{cases} & 2 \leq t < 8 \\ \begin{cases} x = 0 \\ \dot{x} = 0 \end{cases} & t \geq 8 \end{cases}$
دوران حول محور Z:	دوران حول محور Y:	دوران حول محور X:
$\begin{cases} \psi = 0 \\ \dot{\psi} = 0 \end{cases} \quad 0 \leq t \leq 10$	$\begin{cases} \begin{cases} \theta = 0 \\ \dot{\theta} = 0 \end{cases} & t < 4 \\ \begin{cases} \theta = \frac{15^* \pi}{18} \\ \dot{\theta} = 0 \end{cases} & 4 \leq t < 7 \\ \begin{cases} \theta = 0 \\ \dot{\theta} = 0 \end{cases} & t \geq 7 \end{cases}$	$\begin{cases} \begin{cases} \varphi = 0 \\ \dot{\varphi} = 0 \end{cases} & t < 2 \\ \begin{cases} \varphi = \frac{30^* \pi}{18} \\ \dot{\varphi} = 0 \end{cases} & 2 \leq t < 5 \\ \begin{cases} \varphi = 0 \\ \dot{\varphi} = 0 \end{cases} & t \geq 5 \end{cases}$

۵-۲- شبیه سازی سخت افزار در حلقه:

ایده اصلی شبیه سازی سخت افزار در حلقه، به قراردادن یک یا چند زیر سیستم از یک سامانه به صورت سخت افزار واقعی در حلقه شبیه سازی مربوط می شود. مزیت عمده این روش این است که بدون نیاز به ساخت تمام اجزای سیستم، می توان سیستم کنترل را در شرایطی تا حد امکان واقعی مورد آزمایش قرار داد و چون به جای برخی از زیر سیستم ها، سخت افزار واقعی قرار گرفته است، نتایج شبیه سازی به عملکرد واقعی سیستم نزدیک تر خواهد بود و در نتیجه می توان از ریسک و هزینه های آزمایش هایی که ممکن است ناموفق باشند، جلوگیری کرد.

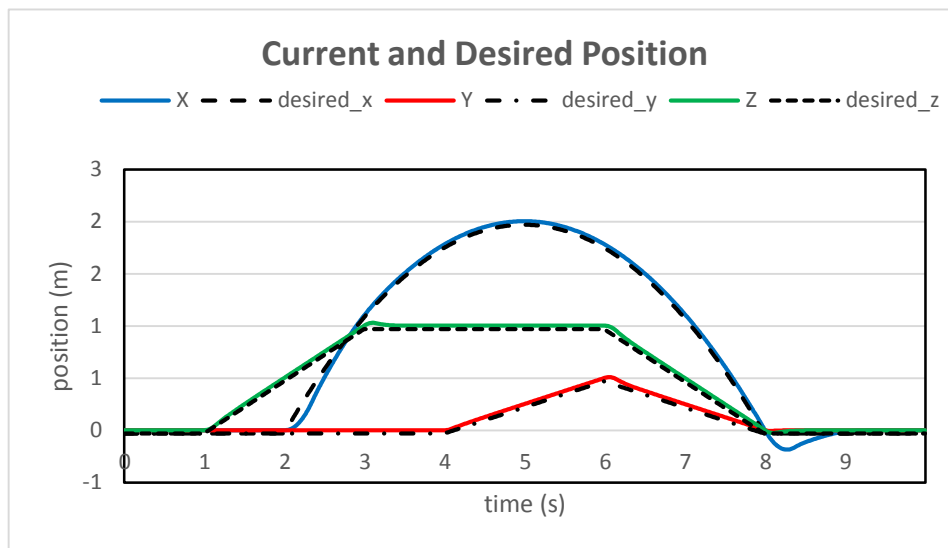
در این پژوهش به منظور اجرای یک تست ساده ی سخت افزار در حلقه، دینامیک پرنده با در نظر گرفتن دینامیک موتورها به صورت روابط ریاضی

به صورت رابطه ی زیر تعیین می گردد [۱۵].

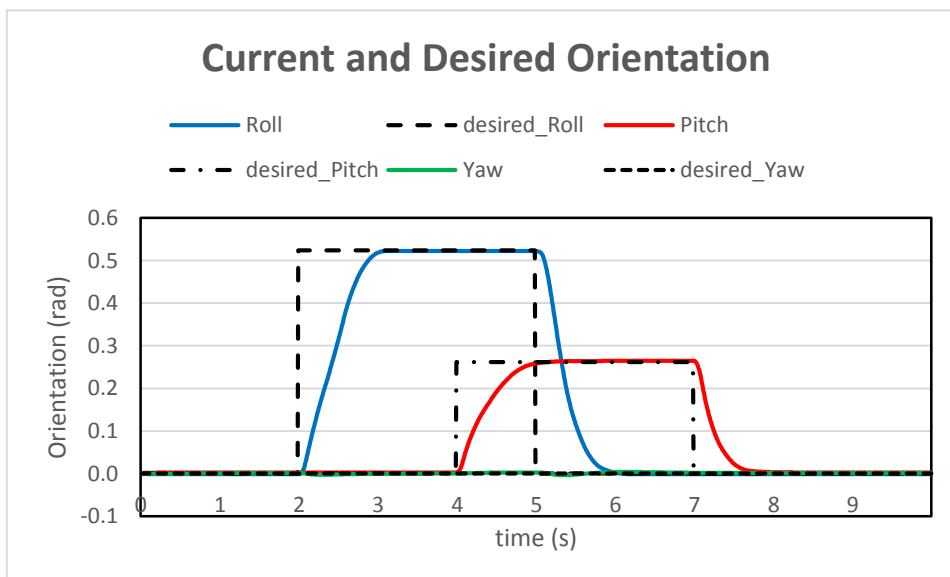
$$\begin{bmatrix} F_{x_{real}} \\ F_{y_{real}} \\ F_{z_{real}} \end{bmatrix} = C \begin{bmatrix} F_x \\ F_y \\ F_z \end{bmatrix} \quad (30)$$

که در آن $F_{x_{real}}$ ، $F_{y_{real}}$ و $F_{z_{real}}$ بیانگر نیروهای اعمال شده در محیط واقعی و C بیانگر درصد انتقال نیرو می باشد.

شکل ۷ نمودار مربوط به حالت انتقال پرنده در راستای محور Z به میزان ۲ متر در ۴ ثانیه به ازای مقادیر مختلف C می باشد که بیانگر مقاوم بودن کنترلر طراحی شده در برابر اختلالات ایجاد شده در محیط واقعی هست.



(الف)



(ب)

شکل ۶. نمودارهای تعقیب مسیر (الف) و جهت‌گیری (ب) اوکتاروتور در شش درجه‌ی آزادی مستقل

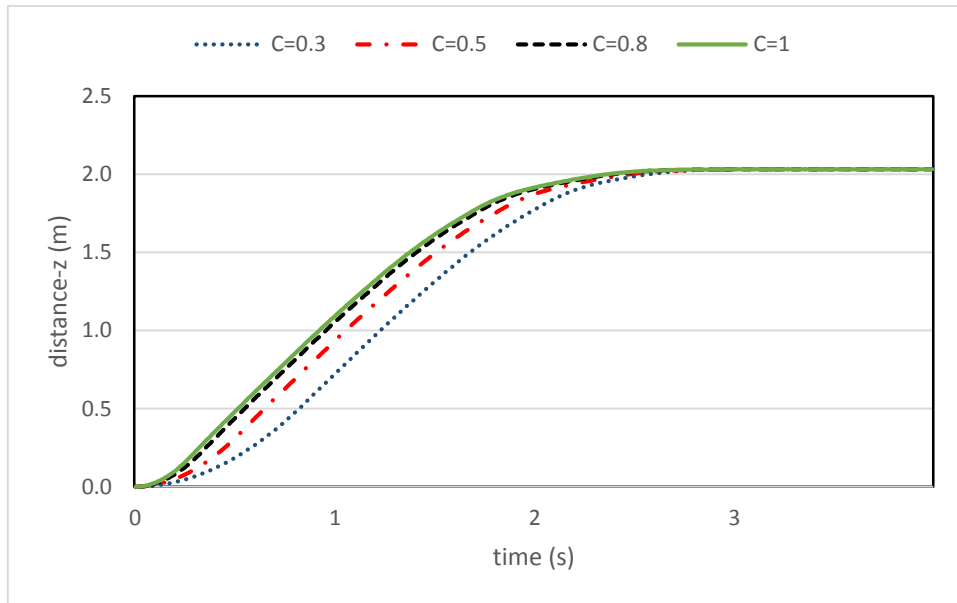
Fig. 6. (a) Trajectory tracking and (b) orientation graphs of the octorotor in independent 6-DoF

می‌گردد (شکل ۸). همچنین به منظور شبیه‌سازی عملکرد موتورها از رابطه‌ی (۳۱) به عنوان دینامیک موتورها استفاده شده است.

$$\frac{d\omega_i}{dt} = K_{\omega}(\omega_i^{des.} - \omega_i) \quad (31)$$

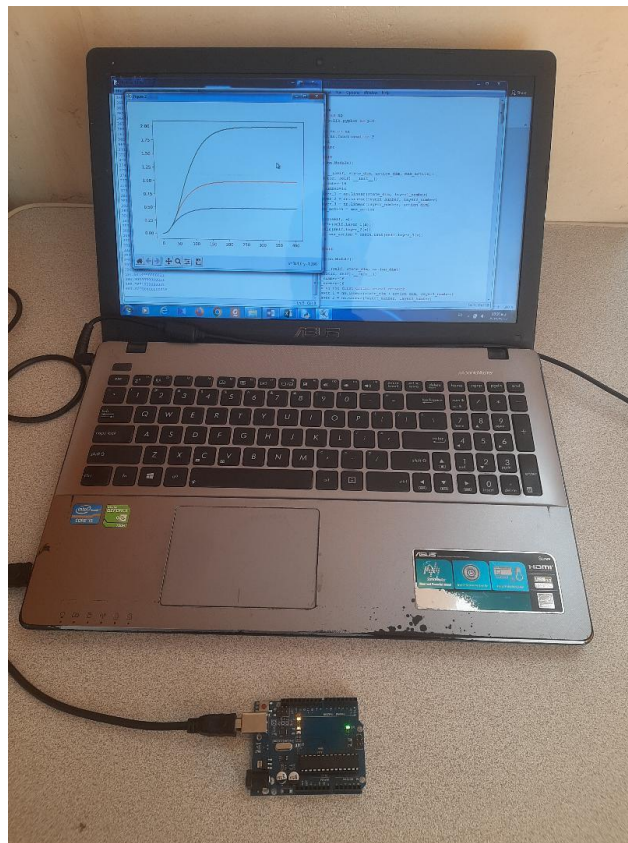
شبیه‌سازی شده و بر روی یک برد الکترونیکی آپلود شده است؛ از طرفی کنترل‌کننده‌ی طراحی شده در پایتون بر روی یک لپ‌تاپ قرار دارد و تست رهگیری مسیر از پیش‌تعیین شده بوسیله‌ی برقراری ارتباط سریال بین پایتون (کنترل‌کننده) و آردوینو (دینامیک پرنده) انجام گرفته است.

برد الکترونیکی مورد استفاده در این پژوهش از نوع آردوینو با تراشه‌ی ATmega328P می‌باشد که از طریق یک کابل سریال به لپ‌تاپ متصل



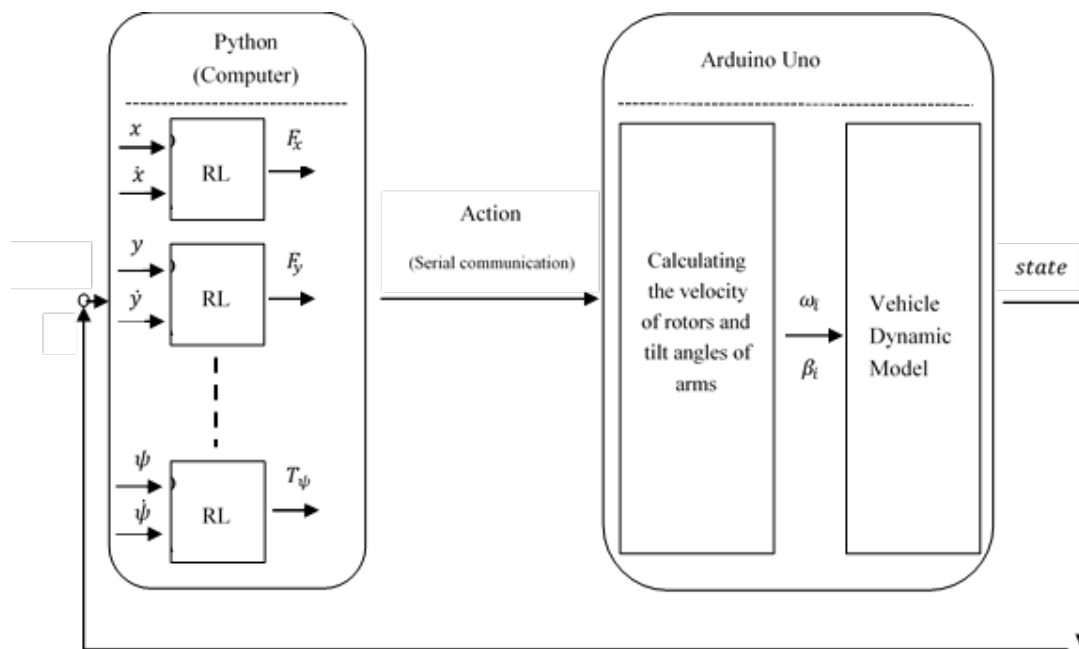
شکل ۷. میز تاثیر مقادیر مختلف در جابه‌جایی اوکتاروتور در راستای محور z

Fig. 7. The effect of different values of C on octorotor transmission along the z-axis



شکل ۸. سخت افزار در حلقه

Fig. 8. Hardware in the loop



شکل ۹. بلوک دیاگرام سخت افزار در حلقه

Fig. 9. Block diagram of hardware in the loop

دینامیک پرنده می‌باشد.

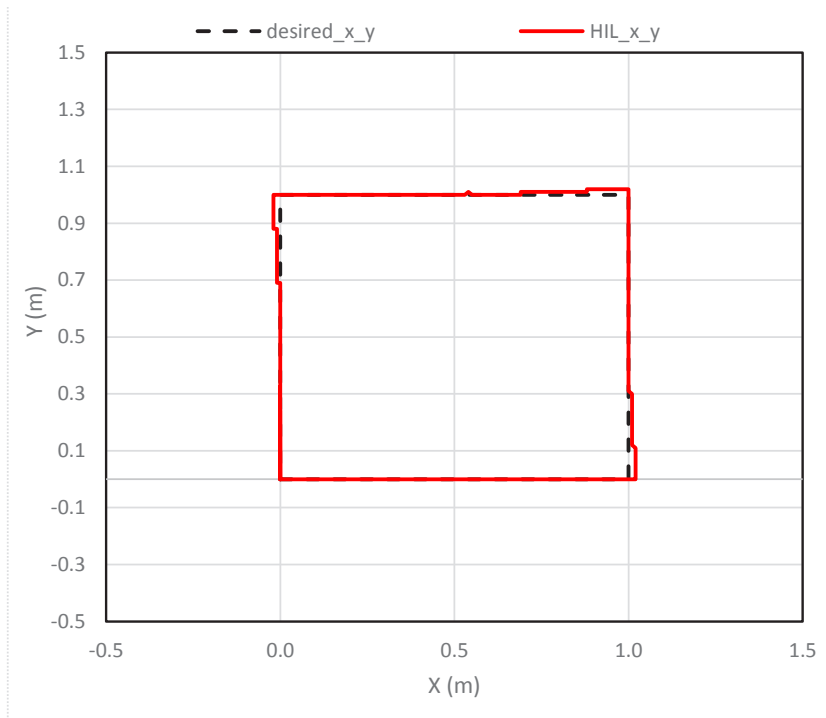
۶- نتیجه‌گیری

در این پژوهش یک ساختار و پیکره‌بندی جدید برای یک اوکتاروتور ارائه شده است به نحوی که چهار روتور آن همانند یک اوکتاروتور متداول جانمایی شده و چهار روتور دیگر بر روی چهار بازو با زاویه‌ی متغیر قرار خواهند گرفت. از قابلیت‌های این پیکره‌بندی افزودن قابلیت دنبال‌کردن جهت‌گیری به اهداف پرنده علاوه بر دنبال‌کردن موقعیت می‌باشد. همچنین کنترل‌کننده‌ای بوسیله‌ی شبکه‌های عصبی و روش یادگیری تقویتی برای این پرنده طراحی شده که علاوه بر قابلیت کنترل اوکتاروتور در شش درجه‌ی آزادی مستقل از قابلیت مقاوم بودن در برابر اختلالات یک یا چند موتور نیز برخوردار است. بواسطه‌ی مستقل بودن درجات آزادی، ابعاد شبکه‌ی در نظر گرفته شده برای این کنترلر نسبت به پژوهش‌های مشابه کوچکتر بوده و زمان آموزش نیز به صورت چشمگیری کاهش یافته است؛ بدون این‌که از قابلیت‌ها و دقت کنترلر طراحی شده کاسته شود. نمودارهای ترسیم شده در این مقاله به خوبی بیانگر قابلیت‌های پیکره‌بندی و کنترلر طراحی شده می‌باشد و می‌توان از آن به منظور استفاده در پرنده‌های واقعی بهره‌گرفت.

که در این رابطه، پارامتر $K_{\omega} = 20 \text{ s}^{-1}$ بیانگر تاخیر زمانی موتورها، ω_i^{des} بیانگر سرعت زاویه‌ای مطلوب روتور i ام و ω_i^{des} بیانگر سرعت زاویه‌ای ایجاد شده توسط موتورها می‌باشد.

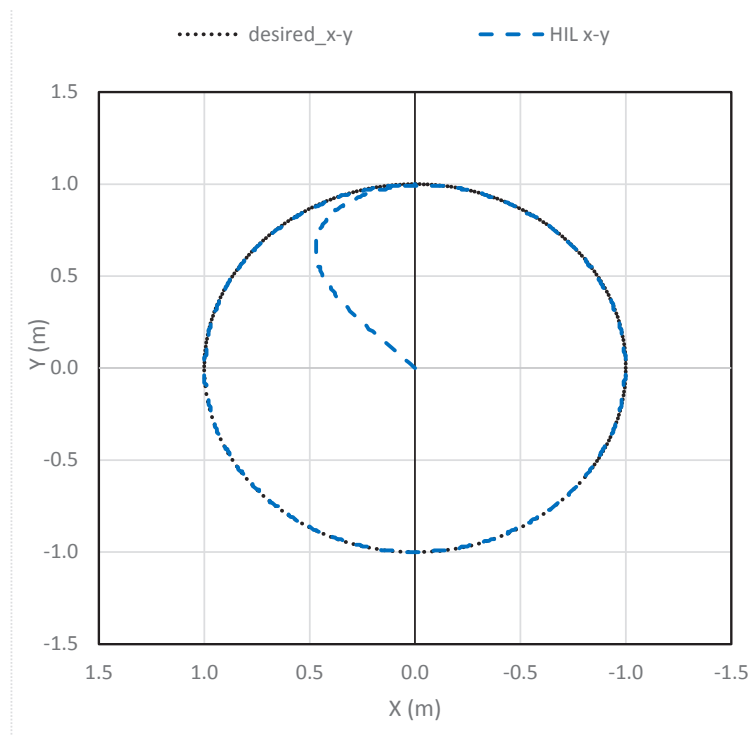
همان‌طور که در شکل ۹ نمایش داده شده است، نیروهای مورد نیاز در شش درجه‌ی آزادی برای رهگیری مسیر توسط کامپیوتر و پایتون تعیین شده و به صورت یکجا به عنوان یک رشته از طریق پورت سریال به برد آردوینو ارسال می‌گردد. نیروهای دریافت شده توسط برد الکترونیکی مبنای محاسبه‌ی سرعت روتورها و زاویه بازوهای متغیر قرار گرفته و سپس به دینامیک پرنده اعمال شده و موقعیت جدید پرنده تخمین زده می‌شود و در نهایت ۱۲ پارامتر شامل موقعیت، وضعیت، سرعت‌های انتقالی و دورانی پرنده به صورت یکجا به عنوان یک رشته به پایتون ارسال می‌گردد. این فرایند به صورت یک حلقه ادامه خواهد داشت و رهگیری مسیر از پیش تعیین شده مورد ارزیابی قرار می‌گیرد.

شکل ۱۰ و شکل ۱۱ به ترتیب نمودار تعقیب میسر مربعی‌شکل و دایره‌ای شکل را نمایش می‌دهد که بیانگر صحت عملکرد کنترل‌کننده در تست سخت‌افزار در حلقه با وجود تاخیرات ارسال و دریافت اطلاعات از طریق پورت سریال و در نظر گرفتن دینامیک موتورها در شبیه‌سازی



شکل ۱۰. تعقیب مسیر مربعی شکل در تست سخت‌افزار در حلقه

Fig. 10. Square-shaped trajectory tracking in HIL testing



شکل ۱۱. تعقیب مسیر دایره‌ای شکل در تست سخت‌افزار در حلقه

Fig. 11. Circular-shaped trajectory tracking in HIL testing

- 2018, pp. 281-286.
- [11] V.M. Arellano-Quintana, E.A. Portilla-Flores, E.A. Merchán-Cruz, Multi-objective design optimization of a hexa-rotor with disturbance rejection capability using an evolutionary algorithm, *IEEE Access*, 6 (2018) 69064-69074.
- [12] A. Erasmus, H. Jordaan, Robust adaptive control of a multicopter with an unknown suspended payload, *IFAC-PapersOnLine*, 53(2) (2020) 9432-9439.
- [13] T. Chen, J. Shan, A novel cable-suspended quadrotor transportation system: From theory to experiment, *Aerospace Science and Technology*, 104 (2020) 105974.
- [14] S. Kase, M. Oya, Adaptive tracking controller for hexacopters with a wind disturbance, *Artificial Life and Robotics*, 25(2) (2020) 322-327.
- [15] C.-H. Pi, K.-C. Hu, S. Cheng, I.-C. Wu, Low-level autonomous control and tracking of quadrotor using reinforcement learning, *Control Engineering Practice*, 95 (2020) 104222.
- [16] A. Soltani, A. H. Vahidi Bajestani, M. Goharkhah, Point to Point Control of a Liquid Carrying Quadrotor, *AUT Journal of Mechanical Engineering*, 54(4) (2022) 727-746. (In Persian)
- [17] H. Shamsollahi, F. Rekabi, F.A. Shirazi, M.J. Sadigh, Control of a Quadrotor Equipped with Robotic Arm Based on Disturbance Estimation, *AUT Journal of Mechanical Engineering*, 54 (2022) 768-747. (In Persian)
- [18] H. Lee, M. Jeong, C. Kim, H. Lim, C. Park, S. Hwang, H. Myung, Low-level Pose Control of Tilting Multicopter for Wall Perching Tasks Using Reinforcement Learning, in: 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, 2021, pp. 9669-9676.
- [19] S. Rajappa, M. Ryll, H.H. Bühlhoff, A. Franchi, Modeling, control and design optimization for a fully-actuated hexarotor aerial vehicle with tilted propellers, in: 2015 IEEE international conference on robotics and automation (ICRA), IEEE, 2015, pp. 4006-4013.
- [20] A. Nematy, M. Kumar, Modeling and control of a single axis tilting quadcopter, in: 2014 American Control
- [1] G. Barbaraci, Modeling and control of a quadrotor with variable geometry arms, *Journal of Unmanned Vehicle Systems*, 3(2) (2015) 35-57.
- [2] D. Brescianini, R. D'Andrea, Design, modeling and control of an omni-directional aerial vehicle, in: 2016 IEEE international conference on robotics and automation (ICRA), IEEE, 2016, pp. 3261-3266.
- [3] Y. Tadokoro, T. Ibuki, M. Sampei, Maneuverability analysis of a fully-actuated hexrotor UAV considering tilt angles and arrangement of rotors, *IFAC-PapersOnLine*, 50(1) (2017) 8981-8986.
- [4] D. Invernizzi, M. Lovera, Geometric tracking control of a quadcopter tiltrotor UAV, *IFAC-PapersOnLine*, 50(1) (2017) 11565-11570.
- [5] D. Invernizzi, M. Giurato, P. Gattazzo, M. Lovera, Full pose tracking for a tilt-arm quadrotor UAV, in: 2018 IEEE Conference on Control Technology and Applications (CCTA), IEEE, 2018, pp. 159-164.
- [6] E.J. Alvarez, A. Ning, High-fidelity modeling of multicopter aerodynamic interactions for aircraft design, *AIAA Journal*, 58(10) (2020) 4385-4400.
- [7] H. Zhu, H. Nie, L. Zhang, X. Wei, M. Zhang, Design and assessment of octocopter drones with improved aerodynamic efficiency and performance, *Aerospace Science and Technology*, 106 (2020) 106206.
- [8] S. Sridhar, R. Kumar, M. Radmanesh, M. Kumar, Non-linear sliding mode control of a tilting-rotor quadcopter, in: Dynamic Systems and Control Conference, American Society of Mechanical Engineers, 2017, pp. V001T009A007.
- [9] S. Zeglache, H. Mekki, A. Bouguerra, A. Djerioui, Actuator fault tolerant control using adaptive RBFNN fuzzy sliding mode controller for coaxial octorotor UAV, *ISA transactions*, 80 (2018) 267-278.
- [10] K. Hu, Q. Wu, Y. Li, W. Chen, Quadrotor QFT Control Based on Model Reduction Method, in: 2018 5th International Conference on Information, Cybernetics, and Computational Social Systems (ICCSS), IEEE,

- Advanced Intelligent Mechatronics (AIM), IEEE, 2020, pp. 1648-1653.
- [24] S. Fujimoto, H. Hoof, D. Meger, Addressing function approximation error in actor-critic methods, in: International conference on machine learning, PMLR, 2018, pp. 1587-1596.
- [25] J. Hwangbo, I. Sa, R. Siegwart, M. Hutter, Control of a quadrotor with reinforcement learning, IEEE Robotics and Automation Letters, 2(4) (2017) 2096-2103.
- Conference, IEEE, 2014, pp. 3077-3082.
- [21] N. Osmić, M. Kurić, I. Petrović, Detailed octorotor modeling and PD control, in: 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, 2016, pp. 002182-002189.
- [22] M. Mikkelsen, Development, modelling and control of a multicopter vehicle, in, 2015.
- [23] Y.-W. Dai, C.-H. Pi, K.-C. Hu, S. Cheng, Reinforcement learning control for multi-axis rotor configuration UAV, in: 2020 IEEE/ASME International Conference on

چگونه به این مقاله ارجاع دهیم

D. Sharifi, M. Irani Rahaghi, K. Torabi, H. Shahbazi, Reinforcement learning-based controller design for a proposed octorotor with tilt-arm angles, Amirkabir J. Mech Eng., 55(10) (2024) 1175-1194.

DOI: 10.22060/mej.2024.21904.7538



