

Integrated Fault-Tolerant Attitude Control of Quadrotor Using Reinforcement Learning

Sajad Roshanravan¹, Saeed Shamaghdari²

¹ Ph.D. Student, Department of Control Systems, Faculty of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran

s_roshanravan@elec.iust.ac.ir

² Associate Professor, Department of Control Systems, Faculty of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran

shamaghdari@iust.ac.ir

Abstract:

This paper deals with the optimal attitude control problem for quadrotor unmanned air vehicles (UAV) with unknown nonlinear dynamics subject to component and actuator faults. The proposed integrated optimal fault tolerant control (FTC) scheme is based on reinforcement learning (RL) algorithm, without requiring prior knowledge of the system dynamics. To solve the Hamilton-Jacobi-Bellman (HJB) equation, an identifier-critic-based online RL strategy is employed with a dual neural network (NN) approximation structure. The forgetting factor in the proposed identifier update law is variable and dependent on the state estimation errors and measurement noise estimation. Choosing this variable forgetting factor increases the convergence speed and decreases the estimation error of identifier NN weights compared to the constant one while maintaining its robustness. When a fault occurs, the system continues to operate under the former control policy until the fault is detected. On the other hand, the optimal control design in the RL framework requires the initial stabilizing control condition. In order to make it possible to initiate the control learning process from the former applied FTC, this condition is relaxed by leveraging a stabilizing term in the critic update law. The Uniformly Ultimately Boundedness (UUB) of identifier and critic NN weight errors and, as a result, the convergence of the control input to the neighborhood of the optimal solution are all proved by Lyapunov theory. In the proposed method, changes in the values of faults are detected by comparing the HJB error to a predefined threshold. Finally, the simulation results are given to validate the effectiveness of the developed method.

Keywords: Quadrature attitude control, Component and actuator faults, Fault-Tolerant optimal control, Fault detection, Reinforcement learning.

Article Type: Research

Received: 19. 06. 2023

Revised: 11. 11. 2023

Accepted: 09. 01. 2024

Corresponding author: Saeed Shamaghdari

Corresponding author's address: Department of Control Systems, Faculty of Electrical Engineering, Iran University of Science and Technology, Farjam St., Narmak, Tehran, Iran



1. Motivation of the work

In the last two decades, the use of quadrotor UAVs has received much attention due to their high maneuverability, the possibility of vertical flight and low construction costs in high-risk environments. These types of UAVs have relatively complex non-linear dynamics, which are very vulnerable to faults due to the presence of external obstacles in their environments. Maintaining system performance when a fault occurs is a key requirement in critical safety systems. Therefore, in addition to guaranteeing stability, the control system must be able to meet the desired optimal performance. One of the most important challenges in the design of the fault tolerance optimal control (FTOC) is the lack of knowledge of the exact model of the system dynamics. According to the knowledge of the authors, there are few model-free active FTOC methods in literature that have investigated the integrated design of fault detection within an FTOC system [1]. This article seeks to present a new model-free integrated fault-tolerant attitude control method based on reinforcement learning for quadrotor UAVs, which is able to maintain its optimal performance in the face of both component and actuator faults.

2. Contributions

Compared to previous related researches, the most important achievements and advantages of the proposed method can be summarized as follows:

- 1) The first model-free integrated FTOC method for attitude tracking of quadrotor that is able to deal with both loss of effectiveness actuator and component faults with two sudden and intermittent behaviors.
- 2) Compared to the supervisory control approach, the proposed structure does not require the primary stabilizing controller or any switching in the control law when the controller adapts to the faulty system.
- 3) Due to the lack of need to prior knowledge of quadrotor dynamics and the initial admissible control policy, it can be used in a wide range of UAVs with similar dynamic structure.
- 4) The forgetting factor in the update law of the identifier neural network (NN) is considered as a variable, which improves the behavior of the identifier NN weights in converging to the real values, compared to the fixed one. Such that, it increases the convergence speed and decreases the estimation error of identifier NN weights while maintaining its robustness against the measurement noise.
- 5) Detecting the occurrence of a new fault is done without the need for any filter bank or secondary model and only based on the residual error value of the HJB equation.

3. Procedures

In this method, the fault-tolerant optimal tracking control (FTOTC) problem for the quadrotor is equated to an optimal stabilizing control problem for a nonlinear augmented system with affine form, where a discounted cost function is minimized. The augmented system

consists of quadrotor dynamics and reference trajectory dynamics. Using this formulation, solving the FTOTC problem for the quadrotor situation is equivalent to solving the Hamilton-Jacobi-Bellman (HJB) equation for the augmented system. Solving the HJB equation requires knowledge of system dynamics. Therefore, in order to solve the HJB equation online without the need to know the dynamics of the quadrotor, an identifier-critic NN approximation structure has been used. In this structure, the detector and critic networks are responsible for approximating the unknown dynamics of the quadrotor and the optimal value of the cost function, respectively. Then the FTOTC policy is calculated based on the approximations obtained from the system dynamics and the optimal value of the cost function. The training process in this method begins with identifying the occurrence of a new fault or its value change, which is detected by the fault detection unit solely based on the residual value of the HJB equation.

4. Findings

To assess the performance and efficiency of the proposed method in controlling the quadrotor attitude, the simulation results on a quadrotor nonlinear model using MATLAB software are presented in Section 5. Examining the behavior of the identifier NN weights reveals that variable considering the forgetting factor in the identifier update law has improved its convergence rate and reduced approximation errors. Furthermore, in comparison with the data-driven approach [2], the quadrotor roll angle under the proposed method, with the same training termination conditions, has converged to the desired trajectory in shorter times and was able to track it in both healthy and faulty modes.

5. Conclusion

In this article, a new FTOTC method based on reinforcement learning algorithm is presented for the design of quadrotor attitude control system. The proposed method has been developed using a dual approximation structure including identifier and critic NNs. In comparison to triple approximation structures including identifier, actor and critic NNs, this structure exhibits lower computational burden and higher convergence rate. The forgetting factor in the identifier update law has been considered variable, which, compared to the fixed one, has resulted in an increase in convergence rate and a reduction in the residual estimation error. In order to eliminate the need to know the initial admissible controller, a stabilizing term is employed in the critic NN update law, which provides the compensation of the fault effect without any prior knowledge of the fault value and system dynamics. Using the Lyapunov stability theory, it has been demonstrated that the tracking error and the identifier and critic weight errors will be ultimately uniformly bounded. The simulation results of the proposed method show its efficiency in the attitude-loop of the quadrotor autopilot.

کنترل وضعیت تحمل‌پذیر عیب یکپارچه کوادروتور با استفاده از یادگیری تقویتی

سجاد روشن‌روان^۱، سعید شمع‌داری^۲

۱- دانشجوی دکتری - دانشکده مهندسی برق - دانشگاه علم و صنعت ایران - تهران - ایران

s_roshanravan@elec.iust.ac.ir

۲- دانشیار - دانشکده مهندسی برق - دانشگاه علم و صنعت ایران - تهران - ایران

shamaghdari@iust.ac.ir

چکیده: در این مقاله به مسئله طراحی کنترل وضعیت بهینه برای پرنده کوادروتور با دینامیک غیرخطی که در معرض عیوب عملکرد و اجزا قرار دارد پرداخته شده است. سیستم کنترل وضعیت بهینه تحمل‌پذیر عیب پیشنهادی مبتنی بر یادگیری تقویتی است و بدون نیاز به شناخت قبلی از دینامیک پرنده و به صورت یکپارچه طراحی می‌شود. بدین معنی که آشکارسازی عیب و بازطراحی کنترل‌کننده را به طور همزمان مورد بحث قرار می‌دهد. به منظور حل برخط معادله همیلتون-ژاکوبی-بلمن (HJB) بدون نیاز به شناخت دینامیک پرنده، از یک ساختار تخمین با دو شبکه عصبی شناساگر-نقاد استفاده شده است. ضریب فراموشی در قانون بهروزرسانی شبکه شناساگر متغیر و تابعی از خطای تخمین حالات و تخمین نویز اندازه‌گیری می‌باشد که باعث بهبود مشخصه‌های حالت گذرا و پایدار آن می‌شود. از سوی دیگر به منظور حذف نیازمندی فرآیند آموزش به کنترل‌کننده پایدارساز اولیه، از یک جمله پایدارساز در قانون بهروزرسانی شبکه نقاد استفاده شده است که امکان شروع فرآیند آموزش را از کنترل‌کننده بهینه پیشین فراهم می‌کند که لزوماً پایداری سیستم معیوب جدید را تضمین نمی‌کند. همچنین در این ساختار، آشکارسازی عیب بدون نیاز به هیچگونه بانکی از مدل و صرفاً مبتنی بر مقدار باقی مانده معادله HJB انجام می‌پذیرد. پایداری فراگیر یکنواخت وزن‌های هر دو شبکه و در نتیجه همگرایی قانون کنترل به پاسخ بهینه با استفاده از قضیه لیاپانوف اثبات شده و کارایی آن با استفاده از شبیه‌سازی مورد ارزیابی قرار گرفته است.

کلمات کلیدی: کنترل وضعیت کوادروتور، عیب اجزا و عملکرد، کنترل بهینه تحمل‌پذیر عیب، آشکارسازی وقوع عیب، یادگیری تقویتی

نوع مقاله: پژوهشی

دریافت: ۱۴۰۲/۰۳/۲۹

بازنگری: ۱۴۰۲/۰۸/۲۰

پذیرش: ۱۴۰۲/۱۰/۱۹

نام نویسنده‌ی مسئول: دکتر سعید شمع‌داری

نشانی نویسنده‌ی مسئول: ایران - تهران - میدان رسالت - خیابان هنگام - خیابان دانشگاه - دانشگاه علم و صنعت ایران - دانشکده

مهندسی برق

۱- مقدمه

پیشرفت‌های چشمگیر صورت گرفته در حوزه عملگرهای دوار پربازده و منابع تغذیه جریان بالا در دهه‌های اخیر باعث گردید ساخت ربات‌های زمینی و هوایی بدون سرنشین بسیار مورد توجه قرار گیرند. در این بین، پرنده‌های مولتی روتور به دلیل هزینه‌های ساخت پایین، قابلیت مانور بالا و کم بودن هزینه‌های پروازی در هر دو حوزه صنعتی و نظامی جهت انجام ماموریت‌های پروازی مختلف از قبیل تصویر برداری، شناسایی و نظارت و ... به صورت فراگیر مورد استفاده قرار گرفتند [۱-۳].

در مقایسه با سایر پرنده‌های بدون سرنشین، پرنده‌های عمود پرواز با توجه به امکان شناور بودن در یک مکان ثابت، نداشتن محدودیت حداقل سرعت جابه‌جایی، عدم نیاز به باند فرود و قابلیت مانور بالا جهت انجام ماموریت در محیط‌های شهری و سر پوشیده بیشتر مورد استفاده قرار می‌گیرند. لذا با توجه به حضور موانع در این محیط‌ها، این پرنده‌ها بسیار در معرض آسیب‌های ناشی از برخورد از قبیل شکستگی پره‌ها قرار دارند. از این رو لازم است سیستم کنترل در آن‌ها به نحوی طراحی گردد که بتواند عملکرد مطلوب را در مقابل این دسته از عیوب احتمالی حفظ نماید [۴].

به منظور جبران سازی اثر عیب روش‌های کنترلی بسیار وسیعی تحت عنوان کنترل تحمل‌پذیر عیب^۱ ارائه شده است که به طور کلی به دو دسته فعال^۲ و غیرفعال^۳ تقسیم می‌شوند [۵]. در دسته غیرفعال یک قانون کنترل ثابت مقاوم^۴ به نحوی طراحی می‌شود که بتواند پایداری و عملکرد سیستم معیوب را همانند حالت بدون عیب حفظ نماید. بزرگترین مشکل این دسته از روش‌ها، محدود بودن کارایی آن‌ها به عیوبی است که از قبل در فرآیند طراحی در نظر گرفته شده‌اند [۶]. بر خلاف دسته غیرفعال، در روش‌های فعال ساختار کنترل‌کننده ثابت نمی‌باشد و از اطلاعات بر خط عیب که توسط قسمت آشکارسازی و تشخیص^۵ بدست می‌آید جهت بازطراحی کنترل‌کننده استفاده می‌شود [۷-۹].

در بعضی موارد علاوه بر پایداری لازم است عملکرد بهینه یا از پیش تعیین شده سیستم در تمامی شرایط عملیاتی تضمین داده شود [۱۰-۱۲]. طبق دانش نویسندگان، تعداد کمی از مقالات موضوع طراحی کنترل وضعیت بهینه تحمل‌پذیر عیب یا به صورت کلی‌تر کنترل ردیاب بهینه تحمل‌پذیر عیب^۶ (FTOTC) را مورد بررسی قرار داده‌اند که انگیزه اصلی این پژوهش را باعث شده است. الگوریتم یادگیری تقویتی^۷ یک ابزار قدرتمند جهت محاسبه قانون کنترل بهینه و زیر بهینه محسوب می‌شود [۱۳]. از این رو در این مقاله، الگوریتم یادگیری تقویتی به عنوان رویکردی کارآمد جهت طراحی کنترل وضعیت بهینه تحمل‌پذیر عیب مورد توجه قرار گرفته است.

به منظور حل مسئله کنترل تحمل‌پذیر عیب بهینه، تنها تعداد کمی روش طراحی مبتنی بر یادگیری تقویتی ارائه شده است که تحت

بعضی شرایط محدود کننده توسعه داده شده‌اند. از اولین مقالات ارائه شده در این زمینه می‌توان به [۱۴-۱۶] اشاره کرد که برای دسته‌های نسبتاً کوچکی از سیستم‌های غیرخطی ارائه شده‌اند و استفاده از آن‌ها را در سیستم‌های واقعی با محدودیت مواجهه می‌کند. در [۱۴] سیستم‌های غیرخطی چند ورودی-چند خروجی به فرم مرتبه- m مورد بحث قرار گرفت و روش پیشنهادی مبتنی بر ساختار شبکه عملگر-نقاد^۸ ارائه شد. پس از آن در [۱۵]، نویسندگان با کاهش پارامترهای یادگیری از طریق جایگزینی آموزش نرم بردار وزن‌های شبکه به جای بردار وزن‌ها، توانستند سرعت همگرایی الگوریتم بهبود بخشند. سپس در [۱۶]، روش ارائه شده در [۱۵] به دسته بزرگتری از سیستم‌های غیرخطی تک‌ورودی-تک‌خروجی گسسته زمان به فرم پسخور صریح^۹ توسعه داده شد. تضمین پایداری سیستم در هر سه این روش‌ها منوط به شروع فرآیند آموزش از کنترل‌کننده پایدارساز اولیه است که برای سیستم‌های در معرض عیب نیازمندی بسیار سخت گیرانه‌ای محسوب می‌شود. انتشار این مقالات باعث توجه حداکثری محققین به حوزه کنترل‌کننده‌های تحمل‌پذیر عیب بهینه مبتنی بر یادگیری تقویتی گردید و در نتیجه دسته‌های بزرگ‌تری از سیستم‌های غیرخطی از قبیل ورودی افاین^{۱۱} که دینامیک پرنده کوادروتور را نیز شامل می‌شوند مورد بررسی قرار گرفتند [۱۷-۲۵]. البته هیچ یک از این مقالات نیز، تمامی اهداف مورد نظر این مقاله را پوشش نمی‌دهند؛ بطوریکه در همگی آن‌ها طراحی قانون کنترل بهینه با هدف پایدارسازی ارائه شده بود و به جزء [۲۲-۲۴]، همگی به شناخت کامل دینامیک سیستم احتیاج داشتند. البته روش‌های ارائه شده در [۲۲-۲۴]، نیز به شناخت اولیه از دینامیک داخلی سیستم احتیاج داشتند. قابل توجه است که در [۲۲] و [۲۴] روش‌های پیشنهادی مبتنی بر ساختار تخمین سه‌گانه، شامل شبکه‌های شناساگر، عملگر و نقاد ارائه شده بودند. درحالیکه در [۲۳] روش پیشنهادی مبتنی بر ساختار تخمین دوگانه شامل شبکه‌های شناساگر و نقاد بنا شده بود. همچنین بررسی‌های انجام شده نشان می‌دهد که مسئله FTOTC برای سیستم‌های غیرخطی افاین بسیار محدود مورد بحث قرار گرفته است که از این میان می‌توان به [۲۶] اشاره نمود که مبتنی بر شناخت دینامیک سیستم و به روش تکرار سیاست^{۱۲} ارائه شده است. قابل توجه است که نیازمندی روش تکرار سیاست به شناخت دینامیک سیستم و همچنین کنترل‌کننده پایدارساز اولیه استفاده از آن را در سیستم‌های کنترل تحمل‌پذیر عیب فعال باچالش‌های جدی مواجه می‌کند.

طراحی همزمان بخش‌های کنترل و شناسایی عیب در سیستم‌های فعال، به گونه‌ای که هر بخش نیازمندی بخش دیگر را برآورده کند، با توجه به وجود نامعینی‌های متقابل بین آن‌ها، هنوز نیز مسئله‌ای صعب و پر چالش به شمار می‌آید [۲۷]. وجود این نیازمندی باعث توجه ویژه محققین به طراحی همزمان بخش‌های کنترل و شناسایی عیب در

همچنین با استفاده از قضیه لیاپانوف^{۲۰} نشان داده می‌شود که نه تنها متغیرهای حالت تخمین زده توسط شناساگر به متغیرهای حالت سیستم عملی همگرا می‌شوند بلکه وزن‌های شبکه شناساگر نیز به مقادیر واقعی خود همگرا می‌شوند. سپس به منظور تخمین برخط تابع هزینه عملکرد کنترل کننده از یک شبکه نقاد استفاده شده است. در انتها قانون کنترل بهینه با استفاده از تخمین‌های بدست آمده از دینامیک سیستم و شاخص عملکرد بهینه محاسبه می‌شود.

بعد از وقوع عیب در سیستم، ساختار کنترل تحمل‌پذیر عیب لازم است با وضعیت جدید انطباق پیدا کند. از سوی دیگر، شروع فرآیند آموزش از کنترل کننده پایدارساز اولیه یک شرط ضروری در همگرایی روش‌های مبتنی بر یادگیری تقویتی به قانون کنترل بهینه و تضمین پایداری سیستم در طول دوره آموزش محسوب می‌شود. از این رو پیدا کردن یک کنترل کننده پایدارساز اولیه برای سیستم معیوب ناشناخته یک نیازمندی کلیدی به شمار می‌آید که در عین حال کاری بسیار دشوار است [۳۰]. اخیراً به منظور کاهش این محدودیت، یک قانون به‌روزرسانی^{۲۱} اصلاح شده برای شبکه نقاد در [۳۱] ارائه شده است که در آن از یک جمله پایدارساز استفاده گردیده است. این جمله زمانی فعال می‌شود که تابع لیاپانوف کاندید برای سیستم حلقه بسته بهینه در طول مسیر سیستم غیر کاهشی باشد. سپس در [۳۲]، به منظور مقابله با تغییرات در دینامیک داخلی سیستم، یک جمله مقاوم به قانون به‌روزرسانی شبکه نقاد اضافه شده است. در هر دو روش ارائه شده در [۳۱] و [۳۲]، قانون کنترل بهینه با استفاده از ساختاری متشکل از تنها یک شبکه نقاد محاسبه می‌شود که به شناخت کامل از دینامیک سیستم احتیاج دارد. به منظور رفع این نیازمندی، روشی مبتنی بر یادگیری تقویتی انتگرالی^{۲۲} با ساختاری متشکل از دو شبکه عصبی در [۲۹] ارائه گردیده است.

قانون به‌روزرسانی شبکه نقاد ارائه شده در این مقاله اگر چه از ساختاری مشابه با [۳۱] و [۳۲] برخوردار است، ولی از تخمین دینامیک سیستم استفاده می‌کند که توسط شناساگر بدست آمده است. با استفاده از قضیه لیاپانوف اثبات می‌شود که قانون به‌روزرسانی پیشنهادی برای شبکه نقاد علاوه بر همگرایی به قانون کنترل بهینه، پایداری سیستم را در طول دوره آموزش حفظ می‌نماید. در روش فعال پیشنهادی، آشکارسازی تغییر مقدار عیب بدون نیاز به هیچگونه بانک فیلتر و یا مدل ثانویه و تنها با مقایسه مقدار خطای باقی مانده معادله HJB نسبت یک آستانه از پیش تعیین شده انجام می‌شود. این مقدار آستانه سطح مجاز خطای معادله HJB را تعیین می‌کند و مبتنی بر شرط توقف فرآیند یادگیری کنترل کننده انتخاب می‌شود.

در مقایسه با پژوهش‌های مرتبط پیشین، مهم‌ترین دستاوردها و مزایای روش پیشنهادی را می‌توان به صورت زیر خلاصه نمود:

سال‌های اخیر شده است که به عنوان روش‌های کنترل تحمل‌پذیر عیب یکپارچه^{۱۳} شناخته می‌شوند [۲۸]. در میان تمامی مقالات اشاره شده، تنها [۲۵] مسئله آشکارسازی وقوع عیب را مورد بحث قرار داده است؛ در حالیکه این بخش یک نیازمندی کلیدی در روش‌های فعال محسوب می‌شود. در این روش که مبتنی بر ساختار کنترل نظارتی^{۱۴} ارائه شده است از یک توالی کلیدزنی و انطباق به منظور جبران سازی اثر عیب بهره گرفته شده است و علاوه بر آشکارسازی به جداسازی عیب نیز احتیاج دارد. همچنین مقالات مذکور تنها به جبران سازی عیوب عملگر پرداخته‌اند، در حالیکه که عیوب اجزا مانند شکستگی در بدنه و پره‌ها نیز یکی از عیوب غیر قابل امتناع در سیستم‌های پروازی محسوب می‌شوند. از سوی دیگر، تمامی این روش‌ها به منظور همگرایی کنترل کننده به قانون کنترل بهینه، به قانون پایدارساز اولیه احتیاج دارند که برای سیستم معیوب ناشناخته، شرطی بسیار دشوار محسوب می‌شود. در حالیکه روش پیشنهادی این مقاله، به شناخت قبلی از دینامیک پرنده و کنترل کننده پایدارساز اولیه احتیاج ندارد و قادر است ضمن پایدارسازی، ردیابی سیگنال مرجع مطلوب را در حضور عیوب اجزاء و عملگر با دو رفتار زمانی ناگهانی^{۱۵} و تکرار شونده^{۱۶} تضمین دهد.

در این مقاله، با استفاده از الگوریتم یادگیری تقویتی و رویکرد شناسایی عیب مبتنی بر آزمایش معقولیت^{۱۷}، روشی جدید جهت طراحی کنترل وضعیت بهینه تحمل‌پذیر عیب یکپارچه برای پرنده کوادروتور پیشنهاد شده است که به هیچ گونه شناخت قبلی از دینامیک پرنده احتیاج ندارد. دینامیک پرنده کوادروتور به صورت غیرخطی، چند ورودی-چند خروجی، زیر فعال می‌باشد. از این رو کنترل کننده متناظر نیز به فرم غیرخطی بوده و علاوه بر پایداری، ردیابی مقادیر حالات مطلوب را نیز در حضور عیوب اجزاء و عملگر با دو رفتار زمانی ناگهانی و تکرار شونده تضمین می‌نماید.

در این روش، مسئله طراحی کنترل وضعیت بهینه برای پرنده کوادروتور با مسئله کنترل پایدارساز بهینه برای یک سیستم غیرخطی افزوده که شامل دینامیک خطای وضعیت پرنده و مسیر مرجع مطلوب است معادل می‌شود. با استفاده از این معادل سازی، حل مسئله کنترل وضعیت بهینه تحمل‌پذیر عیب برای پرنده کوادروتور با حل معادله همیلتن-ژاکوبی-بلمن^{۱۸} (HJB) برای سیستم افزوده برابر خواهد بود. حل معادله HJB به شناخت دینامیک سیستم احتیاج دارد. از این رو در این مقاله، به منظور رفع نیازمندی به شناخت دینامیک سیستم، از یک شناساگر عصبی تطبیقی^{۱۹} جهت تخمین دینامیک پرنده استفاده شده است. برخلاف شناساگر عصبی تطبیقی ارائه شده در [۲۹]، ضریب فراموشی در قانون به‌روزرسانی ارائه شده متغیر و تابعی از خطای تخمین حالات و تخمین نویز اندازه‌گیری در نظر گرفته شده است. نشان داده می‌شود که استفاده از ضریب فراموشی متغیر در مقایسه با حالت ثابت، ضمن حفظ مقاومت نسبت به نویز اندازه‌گیری، باعث افزایش سرعت همگرایی و کاهش خطای تخمین می‌شود.

۲-۱- نیروها و گشتاورهای ناشی از هر روتور

هر یک از روتورهای پرنده شامل یک نیروی بالابرنده عمود بر صفحه چرخش روتورها و یک لختی القایی در خلاف جهت چرخش روتورها است.

نیروی بالابرنده هر یک از روتورها طبق رابطه زیر قابل محاسبه است:

$$T_i = b\Omega_i^2, i=1,2,3,4 \quad (1)$$

که در آن b ضریب ثابت نیروی بالابرنده و Ω_i سرعت زاویه‌ای ملخ هر یک از روتورها است. ممان آیرودینامیکی ناشی از روتورها به صورت زیر محاسبه می‌شود:

$$\tau_i = d\Omega_i^2, i=1,2,3,4 \quad (2)$$

که در آن d ضریب ثابت نیروی پسا^{۲۴} ملخ است.

۲-۲- دینامیک دورانی

معادلات دینامیک دورانی به صورت زیر بدست می‌آید [۳۲]:

$$\ddot{\phi} = \frac{I_y - I_z}{I_x} \dot{\theta}(t)\dot{\psi}(t) + \frac{l}{I_x} u_1 - \frac{J_r}{I_x} \Omega \dot{\theta} \quad (3)$$

$$\ddot{\theta} = \frac{I_z - I_x}{I_y} \dot{\psi}(t)\dot{\phi}(t) + \frac{l}{I_y} u_2 - \frac{J_r}{I_y} \Omega \dot{\phi} \quad (4)$$

$$\ddot{\psi} = \frac{I_x - I_y}{I_z} \dot{\phi}(t)\dot{\theta}(t) - \frac{l}{I_z} u_3 \quad (5)$$

که در آن ϕ ، θ و ψ به ترتیب زوایای غلت، پیچ و چرخ می‌باشد. l فاصله دوران هر روتور از مرکز ثقل کوادروتور است. $(i = x, y, z)$ I_i مرکز اینرسی کوادروتور در راستای محورهای مختصات و J_r اینرسی هر ملخ است و u_i ها به ازای $i = 1, 2, 3$ ورودی‌های کنترلی مجازی سیستم می‌باشند. مقدار u_i ها و Ω به صورت زیر بدست می‌آیند:

$$\Omega = \Omega_1 + \Omega_3 - \Omega_2 - \Omega_4 \quad (6)$$

$$\begin{aligned} u_1 &= b(\Omega_4^2 - \Omega_2^2) \\ u_2 &= b(\Omega_3^2 - \Omega_1^2) \\ u_3 &= d(\Omega_2^2 + \Omega_4^2 - \Omega_1^2 - \Omega_3^2) \end{aligned} \quad (7)$$

که در آن u_1 ، u_2 و u_3 به ترتیب گشتاورهای زوایای غلت، پیچ و چرخ می‌باشند.

(۱) اولین روش کنترل وضعیت بهینه تحمل‌پذیر عیب یکپارچه که هر دو عیب کاهش کارایی عملگر و اجزا با دو رفتار زمانی ناگهانی و تکرار شونده در آن در نظر گرفته شده است.

(۲) در مقایسه با رویکرد کنترل نظارتی، ساختار پیشنهادی در هنگام انطباق کنترل‌کننده با سیستم معیوب به کنترل‌کننده پایدارساز اولیه و یا هیچگونه کلیدزنی در قانون کنترل احتیاج ندارد.

(۳) با توجه به عدم نیاز به شناخت قبلی از دینامیک پرنده و کنترل‌کننده پایدارساز اولیه در دسته وسیعی از پرنده‌های با ساختار دینامیکی مشابه قابل استفاده خواهد بود.

(۴) ضریب فراموشی^{۲۳} در قانون به‌روزرسانی شبکه شناساگر به صورت متغیر در نظر گرفته شده است که در مقایسه با حالت ثابت باعث بهبود رفتار شبکه شناساگر در همگرایی به پاسخ واقعی می‌شود. بطوریکه باعث کاهش خطای تخمین وزن‌ها، افزایش سرعت همگرایی شبکه و مقاومت به نویز اندازه‌گیری شده است.

(۵) شناسایی وقوع عیب و تغییر اندازه آن بدون نیاز به هیچگونه بانک فیلتر و یا مدل ثانویه و تنها مبتنی بر مقدار خطای باقی مانده معادله HJB انجام می‌شود.

در ادامه مقاله، ابتدا در قسمت ۲ مدل ریاضیاتی دینامیک کوادروتور بیان می‌شود. سپس در قسمت ۳ انواع عیوب احتمالی و هدف از طراحی FTOTC بیان و در قسمت ۴ روش پیشنهادی ارائه می‌گردد. در انتها نیز نتایج حاصل از شبیه‌سازی روش بیان شده و جمع‌بندی مقاله به ترتیب در قسمت‌های ۵ و ۶ آورده شده است.

۲- مدل دینامیکی کوادروتور

پرنده‌های کوادروتور پرکاربردترین دسته از پرنده‌های بدون سرنشین عمود پرواز محسوب می‌شوند که دارای دینامیک شش درجه آزادی غیرخطی می‌باشند. این پرنده‌ها دینامیک نسبتاً پیچیده‌ای را شامل هستند بطوریکه پارامترهای فراوانی، رفتار حرکتی آن‌ها را تحت تاثیر قرار می‌دهد. به منظور ساده‌سازی تحلیل رفتاری و طراحی سیستم کنترل، مدل‌سازی دینامیک آن‌ها تحت فرضیات ساده‌کننده‌ای انجام می‌پذیرد. این فرضیات فاصله چشم‌گیری با واقعیت نداشته و با آن سازگار است. مدل‌سازی دینامیک کوادروتور با مفروضات ساده‌ساز زیر انجام می‌گیرد:

(۱) بدنه کوادروتور و پروانه‌ی موتورها صلب در نظر گرفته شده است.

(۲) ساختار هندسی و جرمی متقارن است.

(۳) ماتریس لختی دورانی متقارن است.

(۴) مرکز جرم و مرکز مختصات محلی بر هم منطبق و در وسط بدنه پرنده قرار دارد.

معادلات دینامیکی کوادروتور با استفاده از روش لاگرانژ قابل محاسبه است. به این منظور بایستی نیروها و گشتاورهای اعمالی به پرنده مشخص شود که در ادامه به آن پرداخته شده است.

که در آن $\rho(t) = \text{diag} \{ \rho_1(t), \rho_2(t), \rho_3(t) \}$ است و $f_F(x(t), t)$ و $g_F(x(t), t)$ به ترتیب دینامیک های داخلی و ورودی سیستم معیوب می باشند.

فرض ۲: عیب کاهش کارایی عملگر محدود است به این معنی که برای

$$0 < \rho_i(t) \leq 1, i = 1, 2, 3$$

توجه ۱: عیوب با توجه رفتار زمانی شان، به سه دسته ناگهانی، افزایشی و تکرار شونده تقسیم می‌شوند که در سیستم معیوب (۱۰) دو حالت ناگهانی و تکرار شونده در نظر گرفته شده است [۳۴].

فرض ۳: عیب کاهش کارایی عملگر $\rho_i(t)$ به صورت تکه‌ای ثابت و دینامیک‌های $f_F(x(t), t)$ و $f_G(x(t), t)$ به صورت تکه‌ای پیوسته از سمت راست هستند که نقاط ناپیوستگی در لحظات وقوع عیب ایجاد می‌شوند.

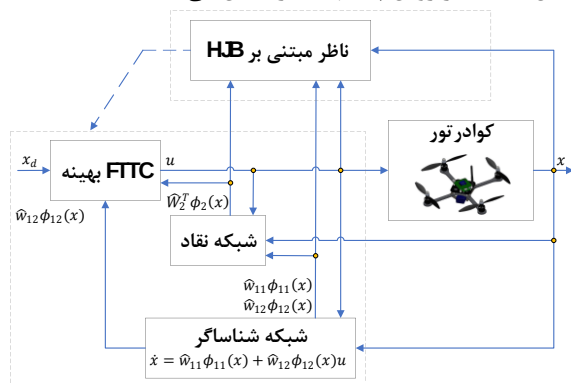
بر t_{F_k} را k امین لحظه تغییر میزان عیب در سیستم در نظر بگیرید. بر طبق فرض ۳، سیستم (۱۰) را در طول بازه زمانی $[t_{F_k}, t_{F_{k+1}}]$ می‌توان به صورت مستقل از زمان زیر نوشت:

$$\dot{x}(t) = f_F(x(t)) + g_F(x(t))(\rho u(t)) \quad (12)$$

بدون از دست دادن کلیت روابط، مسئله طراحی FTOTC برای سیستم (۱۱) در طول باز زمانی $[t_{F_k}, t_{F_{k+1}}]$ انجام می‌شود و سپس نتایج حاصل، به تمامی $t \geq 0$ توسعه داده می‌شود.

٤- طراحی FTOTC

طراحی کنترل‌کننده در این روش مبتنی بر یک ساختار تخمین شبکه عصبی دوگانه متشکل از شبکه‌های شناساگر و نقاد صورت می‌پذیرد. شکل ۱ ساختار روش پیشنهادی را نشان می‌دهد.



شکل (۱): ساختار روش FTOTC پیشنهادی

در این ساختار، شبکه شناساگر وظیفه تخمین دینامیک ناشناخته سیستم و شبکه نقاد وظیفه تخمین مقدار تابع هزینه عملکرد بهینه را بر عهده دارد. سپس قانون FOTC مبتنی بر تخمین بدست آمده از دینامیک سیستم و مقدار تابع هزینه محاسبه می‌شود. در اینجا، هر دو شبکه عصبی در یک زمان به‌روزرسانی می‌شوند که در مقایسه با روش ترتیبی، یک الگوریتم آموزش هم‌زمان را نتیجه می‌دهند. شروع فرآیند

۳- تعریف مسئلہ

در این بخش مدل دینامیکی کوادروتور در معرض عیوب عملکرد و اجزاء، به صورت یک سیستم غیرخطی با فرم افاین نسبت به ورودی بیان می‌شود و همچنین فرضیات لازم در دینامیک سیستم و عیوب جهت حل مسئله FTOTC ارائه می‌گردد.

۳-۱- بیان مدل کوادروتور به صورت سیستم غیرخطی افاین

بردار حالت x_1 و x_2 به صورت $x_1 = [\phi \ \theta \ \psi]^T$ و $x_2 = [\dot{\phi} \ \dot{\theta} \ \dot{\psi}]^T$ را در نظر بگیرید. با صرف نظر از Ω با توجه به کوچک بودن مقدار آن، مدل دینامیکی (۳) تا (۵) را می‌توان به فرم غیرخطی افاین نسبت به ورودی به صورت زیر بازنویسی کرد:

$$\dot{x}_1(t) = x_2(t) \quad (\Lambda)$$

$$\dot{x}_2(t) = f(x(t)) + g(x(t))u(t)$$

که در آن $g = \text{diag} [l / I_x, l / I_y, l / I_z]$ ، $u = [u_1, u_2, u_3]$

$f(x)$ به صورت زیر است:

$$f(x) = \left[\left((I_y - I_z) / I_x \right) x_{22} x_{23}, \left((I_z - I_x) / I_y \right) x_{23} x_{21}, \right. \\ \left. \left((I_x - I_y) / I_z \right) x_{21} x_{22} \right]^T$$

در اینجا فرض شده است که بردار حالت $x(t)$ قابل اندازه‌گیری است و $f(x)$ و $g(x)$ ناشناخته بوده به ترتیب دینامیک‌های داخلی^{۲۵} و ورودی^{۲۶} سیستم هستند.

فرض ۱: $f(x(t)) + g(x(t))u(t)$ در طول هر مجموعه بسته $\mathbb{N} \in \mathbb{R}^n$ پیوسته و لپشیتز، $f(0) = 0$ و سیستم پایدار پذیر است. همچنین $f(x)$ نیز در طول $\mathbb{N}$ پیوسته و لپشیتز به صورت $f(x) \leq b_f x$ و $g(x)$ محدود و به صورت $g(x) \leq b_g$ است که b_o و b_f اعداد مثبت مشخصی هستند.

۳-۲- عیب اجزا و عملگر

عیوب اجزا باعث تغییرات ناخواسته در دینامیک داخلی و ورودی می‌شوند و عیوب عملگر باعث تغییر دینامیک عملگر می‌شوند. عیب کاهش کارایی برای عملگر i ام به فرم یک عیب ضریب شونده متغیر با زمان به صورت زیر مدل‌سازی می‌شود:

$$u_i^F(t) = \rho_i u_i(t) \quad i = 1, 2, 3 \quad (9)$$

که $u_i^F(t)$ خروجی عملگر معیوب $\dot{I}m$ و $\rho_i(t)$ مقدار کاهش کارایی آن عملگر نشان می‌دهد. سپس، دینامیک سیستم در معرض عیب اجزا و عملگر به صورت زیر قابل بازنویسی است:

$$\dot{x}(t) = f_E(x(t), t) + g_E(x(t), t)(\rho(t)u(t)) \quad (1 \cdot)$$

که در آن $\Gamma \in R^{(k_{wf}+k_{wg}) \times (k_{wf}+k_{wg})}$ یک ماتریس بهره یادگیری مثبت معین است و سرعت همگرایی وزن‌ها را به مقادیر واقعی شان را مشخص می‌کند.

توجه ۲: مقدار k در رابطه (۱۴)، پهنای باند فیلتر را مشخص می‌کند که به منظور حفظ مقاومت نسبت به نویز اندازه‌گیری بایستی کمینه مقدار در نظر گرفته شود. همچنین $(1+e_2^{k_2})/(1+e_1^{k_1})$ در رابطه (۱۵)، مقدار ضریب فراموشی و بهره حالت ماندگار را مشخص می‌کند. بنابراین انتخاب مقادیر مناسب برای l ، k_1 و k_2 به یک مصالحه بین خطای تخمین، سرعت همگرایی و مقاومت احتیاج دارد.

تعریف ۱: تابع $\phi_1(x(t), u(t))$ را برانگیخته کامل^{۲۷} گوئیم در صورتی که اعداد مثبت ثابت α_1 و α_2 برای تمامی $t \geq 0$ به نحوی وجود داشته باشد که:

$$\alpha_1 I \leq \int_t^{t+T_1} \phi_1(x(\tau), u(\tau)) \phi_1(x(\tau), u(\tau)) d\tau \leq \alpha_2 I \quad (18)$$

می‌توان نشان داد که مثبت معین بودن ماتریس P در (۱۵) یک شرط ضروری برای همگرایی خطای تخمین $\tilde{W}_1$ است که صورت $\tilde{W}_1 = W_1 - \hat{W}_1$ تعریف می‌شود. $\lambda_{\min}(\cdot)$ و $\lambda_{\max}(\cdot)$ را به ترتیب به عنوان بزرگترین و کوچکترین مقدار ویژه ماتریس‌های متناظر با آن در نظر بگیرید. لم ۱ رابطه بین برانگیخته کامل بودن بردار ϕ_1 و مثبت معین بودن P را نشان می‌دهند.

لم ۱ [۲۵]: در صورتی که بردار $\phi_1(x, u)$ در (۱۴)، برانگیخته کامل باشد، ماتریس P در (۱۵) مثبت معین خواهد بود؛ به این معنی که عدد مثبت $\sigma > 0$ به نحوی وجود دارد که $\lambda_{\min}(P) > \sigma$.

قضیه ۱: سیستم (۱۳) با قانون به‌روزرسانی (۱۷) که رابطه (۱۱) را تخمین می‌زند در نظر بگیرید. اگر بردار ϕ_1 برانگیخته کامل باشد، سپس خطای تخمین $\tilde{W}_1$ زمانی که خطای تخمین ε_1 صفر باشد پایدار مجانبی نمایی خواهد بود و زمانی که خطای تخمین ε_1 صفر نباشد پایدار فراگیر یکنواخت^{۲۸} خواهد بود.

اثبات: با توجه به رابطه (۱۲) و (۱۴) می‌توان نوشت:

$$\dot{x}_f = \frac{x - x_f}{k} = W_1^T \phi_{1f} + \varepsilon_{1f} \quad (19)$$

که در آن سیگنال فیلتر شده ε_{1f} است و به صورت $k \dot{\varepsilon}_{1f} + \varepsilon_{1f} = \varepsilon_1$ تعریف می‌شود. با جایگذاری (۱۹) در (۱۵) و سپس انتگرال‌گیری از آن، می‌توان دریافت که:

$$D = PW_1 - v \quad (20)$$

که در آن $v = -\int_0^t e^{-\frac{l+e_1^{k_1}}{1+e_2^{k_2}}(t-\tau)} \phi_{1f}^T(\tau) \varepsilon_{1f}^T(\tau) d\tau$ است. از آنجایی که $e_1, \varepsilon_{1f}, \phi_{1f}$ و e_2 همگی محدود هستند می‌توان نتیجه گرفت که متغیر v نیز محدود است؛ یا عبارت دیگر یک عدد $\varepsilon_v > 0$ وجود دارد که $v \leq \varepsilon_v$.

آموزش در این روش با شناسایی وقوع عیب جدید صورت می‌پذیرد که توسط واحد آشکارسازی عیب و صرفاً مبتنی بر مقدار باقی مانده معادله HJB تشخیص داده می‌شود.

۴-۱- شناساگر عصبی تطبیقی

در این بخش، به منظور تخمین دینامیک ناشناخته کوادروتور، یک شناساگر عصبی تطبیقی مبتنی بر مقادیر ورودی-حالت ارائه می‌شود. بر اساس فرض ۱، سیستم ناشناخته (۱۱) را می‌توان به صورت زیر تخمین زد:

$$\dot{x} = W_1^T \phi_1(x, u) + \varepsilon_1 \quad (12)$$

که در آن $W_1 = [w_{11}^T, w_{12}^T] \in R^{(k_{w11}+k_{w12}) \times n}$ و $\phi_1(x, u) = [\phi_{11}^T(x), u^T \phi_{12}^T(x)]^T \in R^{k_{w11}+k_{w12}}$ و $\phi_{11} \in R^{k_{w11}}$ و $\phi_{12} \in R^{k_{w2} \times m}$ توابع پایه شبکه عصبی و ε_1 خطای تخمین هستند. با توجه به اینکه بردار وزن‌های ایده‌آل W_1 ناشناخته است، رابطه (۱۲) می‌تواند به صورت زیر تخمین زده شود:

$$\dot{x} = \hat{W}_1^T \phi_1(x, u) \quad (13)$$

به منظور ارائه قانون به‌روزرسانی برای تخمین $\hat{W}_1$ ، متغیرهای فیلتر شده x_f و ϕ_{1f} به صورت زیر تعریف شده‌اند:

$$\begin{cases} k\dot{x}_f + x_f = x \\ k\dot{\phi}_{1f} + \phi_{1f} = \phi_1 \end{cases} \quad (14)$$

که $k > 0$ یک عدد مثبت ثابت است. برای هر $l, k_1, k_2 > 0$ ماتریس‌های پایه فیلتر شده $D \in R^{(k_{wf}+k_{wg}) \times n}$ به صورت زیر تعریف می‌شوند:

$$\begin{cases} \dot{P} = -\left(\frac{l+e_1^{k_1}}{1+e_2^{k_2}}\right)P + \phi_{1f} \phi_{1f}^T \\ \dot{D} = -\left(\frac{l+e_1^{k_1}}{1+e_2^{k_2}}\right)D + \phi_{1f} \left[\frac{x - x_f}{k}\right]^T \end{cases} \quad (15)$$

که در آن $P(0) = 0$ و $D(0) = 0$ ، فیلتر شده بردار خطای تخمین حالت شناساگر به صورت $e_1 = x_f - \hat{x}_f$ و $e_2 = x - \hat{x}_f$ بردار خطای حالت فیلتر شده به صورت $e_2 = x - x_f$ است. بردار خطای حالت فیلتر شده $\hat{x}_f$ به صورت $k\dot{\hat{x}}_f + \hat{x}_f = \hat{x}$ تعریف می‌شود که در آن $\hat{x}$ بردار تخمین حالت x است. همچنین ماتریس کمکی $M \in R^{(k_{wf}+k_{wg}) \times n}$ که تابعی از P و D است را به صورت زیر نظر بگیرید:

$$M = P\hat{W}_1 - D \quad (16)$$

سپس قانون به‌روزرسانی برای $\hat{W}_1$ را می‌توان به صورت زیر طراحی کرد:

$$\dot{\hat{W}}_1 = -\Gamma M \quad (17)$$

$$\dot{x}_d(t) = H(x_d(t)) \quad (27)$$

که در آن $H(x_d(t))$ لپشیتز و $H(0) = 0$ است.

لذا خطای ردیابی به صورت زیر قابل تعریف است:

$$e_d(t) \triangleq x(t) - x_d(t) \quad (28)$$

لذا با استفاده از (۲۸) و با تعریف بردار حالت $z = [e_d^T \ x_d^T]^T$ برای سیستم افزوده، خواهیم داشت:

$$\dot{z} = \hat{F}(z) + \hat{G}(z)u \quad (29)$$

که در آن $\hat{F}: R^{2n} \rightarrow R^{2n}$ ، $\hat{G}: R^{2n} \rightarrow R^{2n \times m}$ ، $u \in R^m$ به صورت زیر هستند:

$$\hat{F}(z) = \begin{pmatrix} \hat{f}_F(x_d + e_d) + \hat{g}_F(x_d + e_d)\hat{v} - H(x_d) \\ H(x_d) \end{pmatrix} = \begin{pmatrix} \hat{w}_{11}\phi_{11}(x_d + e_d) - H(x_d) \\ H(x_d) \end{pmatrix}$$

$$\hat{G}(z) = \begin{pmatrix} \hat{g}_F(x_d + e_d)\hat{\rho} \\ 0 \end{pmatrix} = \begin{pmatrix} \hat{w}_{12}\phi_{12}(x_d + e_d) \\ 0 \end{pmatrix}$$

فرض ۵: با توجه به فرض ۱ و ۴، می توان نتیجه گرفت که دینامیک داخلی و ورودی سیستم افزوده به ترتیب لپشیتز و محدود هستند؛ بطوریکه $\hat{G}(z) \leq G_M$ و $\hat{F}(z) \leq L_F z$

می توان نشان داد که با تعریف یک تابع عملکرد تضعیف شده مناسب، مسئله ردیابی بهینه برای سیستم (۱۲) معادل با مسئله پایداری بهینه برای سیستم (۲۹) خواهد بود [۳۶]. برای این منظور تابع هزینه تضعیف شده با افق بینهایت زیر را در نظر بگیرید:

$$V(z(t)) = \int_t^\infty e^{-\gamma(\tau-t)} [z^T(\tau)Q_T z(\tau) + u^T(\tau)Ru(\tau)] d\tau \quad (30)$$

که در آن $Q_T \in R^{2n \times 2n}$ ماتریس مثبت معین به صورت زیر است:

$$Q_T = \begin{pmatrix} Q_{n \times n} & 0_{n \times n} \\ 0_{n \times n} & 0_{n \times n} \end{pmatrix}$$

با مشتق گیری از (۳۰) نسبت به زمان، همیلتونین متناظر با آن به صورت زیر بدست می آید:

$$H(z, u, \nabla V) = \nabla V(\hat{F}(z) + \hat{G}(z)u) - \gamma V(z) \quad (31)$$

$$+ z^T Q_T z + u^T R u = 0$$

که $\nabla V = \frac{\partial V(z)}{\partial z}$ با بهره گیری از قید ایستایی^{۲۹}

$$\frac{\partial H(z, u, \nabla V)}{\partial u} = 0, [10], \text{ قانون کنترل FTOTC بهینه } u^*(z) \text{ به}$$

صورت زیر بدست می آید:

$$u^*(z) = -\frac{1}{2} R^{-1} \hat{G}^T(z) \nabla V^*(z) \quad (32)$$

سپس با جایگذاری (۲۰) در (۱۶) خواهیم داشت:

$$M = P\hat{W}_1 - P\hat{W}_1 - v = -P\hat{W}_1 + v \quad (21)$$

رابطه (۲۱) نشان می دهد که ماتریس M حاوی اطلاعات خطای تخمین وزن های شبکه است. حال تابع لیاپانوف کاندید V_1 را به صورت زیر نظر بگیرید:

$$V_1 = \frac{1}{2} tr(\tilde{W}_1^T \Gamma^{-1} \tilde{W}_1) \quad (22)$$

بنابراین مشتق زمانی آن به صورت زیر قابل محاسبه خواهد بود:

$$\dot{V}_1 = tr(\tilde{W}_1^T \Gamma^{-1} \dot{\tilde{W}}_1) = -tr(\tilde{W}_1^T P \tilde{W}_1) + tr(\tilde{W}_1^T v) \quad (23)$$

الف) زمانی $\varepsilon_1 = 0$ باشد؛ در این حالت $v = 0$ است و در نتیجه (۲۳) به صورت زیر قابل بازنویسی است:

$$\dot{V}_1 = -tr(\tilde{W}_1^T P \tilde{W}_1) < -\lambda_{\min}(P) \tilde{W}_1^2 \leq -\beta_1 V_1 \quad (24)$$

که با توجه مثبت معین بودن P ، $\beta_1 = 2\lambda_{\min}(P) / \lambda_{\max}(\Gamma^{-1})$ یک عدد مثبت است. لذا طبق قضیه لیاپانوف خطای تخمین به صورت نمایی به صفر همگرا می شود.

ب) زمانی $\varepsilon_1 \neq 0$ باشد، رابطه (۲۳) را می توان به صورت زیر بازنویسی کرد:

$$\dot{V}_1 = -tr(\tilde{W}_1^T P \tilde{W}_1) + tr(\tilde{W}_1^T v) \leq -\tilde{W}_1(\lambda_{\min}(P) \tilde{W}_1 - \varepsilon_v) \quad (25)$$

با توجه به (۲۵)، برای اینکه $\dot{V}_1$ منفی باشد باید نامساوی زیر برآورد شود:

$$\tilde{W}_1 > \frac{\varepsilon_v}{\lambda_{\min}(P)} \quad (26)$$

سپس با توجه به قضیه لیاپانوف توسعه داده شده، خطای تخمین پایدار UBB خواهد بود.

۴-۲- FTOTC مبتنی بر یادگیری تقویتی

تا زمانی که وقوع عیب جدید در سیستم آشکارسازی نشود سیستم تحت کنترل کننده قبلی عمل می کند که لزوما پایداری سیستم را در وضعیت جدید تضمین نمی کند. در این مقاله، FTOTC مبتنی بر الگوریتم یادگیری تقویتی طراحی و محاسبه می شود که به صورت کلی به کنترل کننده پایداری اولیه احتیاج دارد. از این رو به منظور کاهش این محدودیت، از یک جمله پایداری ساز در قانون به روزرسانی شبکه نقاد استفاده می شود. نشان داده خواهد شد که فرآیند یادگیری تحت قانون به روزرسانی پیشنهادی نه تنها به قانون کنترل بهینه همگرا می شود بلکه پایداری سیستم را نیز در طول دوره آموزش تضمین می نماید.

فرض ۴: فرض کنید $x_d(t)$ مسیر مرجع محدود دارای دینامیک $H(x_d(t)) \in R^n$ به صورت زیر باشد:

وجود خواهد داشت. همچنین، ماتریس مثبت معین $\bar{Q}(z) \in R^{n \times n}$ وجود دارد که:

$$\nabla J^T (\hat{F}(z) + \hat{G}(z)u^*) = -\nabla J_1^T \bar{Q}(z) \nabla J_1 \quad (39)$$

که $\bar{Q}(z) = 0$ اگر و تنها اگر $z = 0$. همچنین $\lim_{z \rightarrow \infty} \bar{Q}(z) = \infty$ روی مجموعه پیوسته Ω داریم $\bar{Q}_{min} \leq \bar{Q}(z) \leq \bar{Q}_{max}$ که $\bar{Q}_{min}$ و $\bar{Q}_{max}$ اعداد مثبت هستند.

قابل اثبات است که پیدا کردن مقدار تابع هزینه $\hat{V}(z)$ معادل پیدا کردن وزن‌های شبکه نقاد به نحوی است که $E = \frac{1}{2}e^2$ را کمینه سازد [۳۷]. اگرچه این کمینه سازی لزوماً پایداری سیستم را در طول فرآیند آموزش تضمین نمی‌دهد. بنابراین، قانون به‌روزرسانی شبکه نقاد بایستی علاوه بر همگرایی به تابع هزینه بهینه، پایداری سیستم را نیز در طول دوره آموزش تضمین نماید. قانون به‌روزرسانی زیر را در نظر بگیرید:

$$\begin{aligned} \dot{\hat{W}}_2 = & -\alpha_1 \left(\frac{\bar{\sigma}}{m_s} \right) \hat{e} + \alpha_1 \Xi(z, \hat{u}) \left(\frac{1}{2} \nabla \phi_2(z) \wedge \nabla J_1 \right) \\ & + \alpha_1 \left((K_1 \bar{\sigma}^T - K_2) \hat{W}_2 + \frac{1}{4} \nabla \phi_2(z) \wedge \nabla \phi_2^T(z) \hat{W}_2 \left(\frac{\bar{\sigma}^T}{m_s} \right) \hat{W}_2 \right) \end{aligned} \quad (40)$$

که در آن $\bar{\sigma} = \hat{\sigma}^T / (\hat{\sigma}^T \hat{\sigma} + 1)$ ، $\hat{\sigma} = \nabla \phi_2(z) \hat{F}(z) - \gamma \phi_2(z) - \frac{1}{2} \nabla \phi_2(z) \wedge \nabla \phi_2^T(z) \hat{W}_2$ ، $m_s = \hat{\sigma}^T \hat{\sigma} + 1$ و $J_1(z)$ در فرض ۷ تشریح شده است. همچنین $K_2 \in R^{N_2 \times N_2}$ و $K_1 \in R^{N_2}$ ، $\alpha_1 > 0$ و $\Xi(z, \hat{u})$ تابع نشان دهنده وضعیت پایداری سیستم حلقه بسته است و به صورت زیر تعریف می‌شود:

$$\Xi(z, \hat{u}) = \begin{cases} 0 & \text{if } \nabla J_1 z < 0 \\ 1 & \text{else} \end{cases} \quad (41)$$

بر اساس (۳۵)، داریم:

$$\begin{aligned} z^T Q_T z = & \gamma \hat{W}_2^T \phi_2(z) - \hat{W}_2^T \nabla \phi_2(z) \hat{F}(z) \\ & + \frac{1}{4} \hat{W}_2^T \nabla \phi_2(z) \wedge \nabla \phi_2^T(z) \hat{W}_2 - \varepsilon_{HJB} \end{aligned} \quad (42)$$

در صورتی که خطای تخمین شبکه نقاد را به صورت $\tilde{W}_2 = W_2 - \hat{W}_2$ در نظر بگیریم، با استفاده از (۴۰) و (۴۲) دینامیک خطای وزن‌های شبکه نقاد به صورت زیر قابل محاسبه است:

$$\begin{aligned} \dot{\tilde{W}}_2 = & \alpha_1 \frac{\bar{\sigma}}{m_s} \left[-\tilde{W}_2^T \hat{\sigma} + \frac{1}{4} \tilde{W}_2^T \nabla \phi_2(z) \wedge \nabla \phi_2^T(z) \tilde{W}_2 - \varepsilon_{HJB} \right] \\ & - \alpha_1 \Xi(z, \hat{u}) \left(\frac{1}{2} \nabla \phi_2(z) \wedge \nabla J_1 \right) \\ & - \alpha_1 \left((K_1 \bar{\sigma}^T - K_2) \tilde{W}_2 + \frac{1}{4} \nabla \phi_2(z) \wedge \nabla \phi_2^T(z) \tilde{W}_2 \left(\frac{\bar{\sigma}^T}{m_s} \right) \tilde{W}_2 \right) \end{aligned} \quad (43)$$

قضیه ۲: سیستم غیرخطی (۲۹) را به همراه معادله HJB ردیاب (۳۳) و قانون کنترل FTOTC تخمینی (۳۷) در نظر بگیرید. همچنین قانون به‌روزرسانی (۱۷) و (۴۰) را به ترتیب جهت آموزش وزن‌های

در آن $V^*(z)$ تابع هزینه تضعیف شده بهینه است که در $H(z, u^*, \nabla V^*) = 0$ صدق می‌کند. سپس با جایگذاری (۳۲) در (۳۱)، معادله HJB ردیاب به صورت زیر بدست می‌آید:

$$\begin{aligned} z^T Q_T z - \gamma \hat{W}_2^T \phi_2(z) + \nabla V^T \hat{F}(z) \\ - \frac{1}{4} \nabla V^T(z) \hat{G}(z) R^{-1} \hat{G}^T(z) \nabla V^*(z) = 0 \end{aligned} \quad (33)$$

بر طبق قضیه تخمین مرتبه بالای وایرستراس، می‌توان تابع هزینه بهینه (۳۰) را با استفاده از یک شبکه عصبی تک لایه با تعداد مناسب توابع پایه به صورت زیر تخمین زده شود:

$$V^*(z) = W_2^T \phi_2(z) + \varepsilon_2(z) \quad (34)$$

که $W_2 \in R^{N_2}$ ماتریس ناشناخته از وزن‌های ایده‌آل شبکه عصبی، $\phi_2 \in R^{N_2}$ بردار توابع پایه شبکه نقاد و $\varepsilon_2 \in R$ خطای تخمین است.

فرض ۶: بردار وزن‌های ایده‌آل شبکه نقاد، بردار توابع پایه شبکه نقاد و گرادین آن محدود هستند به صورتی که $\phi_2 < \phi_{2M}$ ، $W_2 < W_{2M}$ و $\nabla \phi_2 < \phi_{2dM}$. همچنین خطای تخمین شبکه نقاد و گرادین آن محدود در نظر گرفته شده است به نحوی که $\varepsilon_2 < \varepsilon_{2M}$ و $\nabla \varepsilon_2 < \varepsilon_{2dM}$. با جایگذاری رابطه (۳۴) و تخمین آن در (۳۳)، معادله HJB ردیاب (۳۳) به صورت زیر قابل بازنویسی است:

$$\begin{aligned} z^T Q_T z - \gamma \hat{W}_2^T \phi_2(z) + \hat{W}_2^T \nabla \phi_2(z) \hat{F}(z) \\ - \frac{1}{4} \hat{W}_2^T \nabla \phi_2(z) \wedge \nabla \phi_2^T(z) \hat{W}_2 + \varepsilon_{HJB} = 0 \end{aligned} \quad (35)$$

که $\varepsilon_{HJB} = \nabla \varepsilon_2^T (\hat{F}(z) + \hat{G}(z)u^*) + \gamma \varepsilon_2(z) + \frac{1}{4} \nabla \varepsilon_2^T \wedge \nabla \varepsilon_2$ و $\Lambda = \hat{G}(z) R^{-1} \hat{G}^T(z)$ است. در عمل بردار وزن‌های ایده‌آل W_2 ناشناخته است. بنابراین از تخمین تابع هزینه بهینه به صورت زیر استفاده می‌شود:

$$\hat{V}(z) = \hat{W}_2^T \phi_2(z) \quad (36)$$

که تخمین ماتریس وزن‌های ایده‌آل W_2 است. بنابراین قانون کنترل بهینه تخمینی $\hat{u}$ به صورت زیر قابل بیان است:

$$\hat{u} = -\frac{1}{2} R^{-1} \hat{G}^T(z) \nabla \phi_2^T(z) \hat{W}_2 \quad (37)$$

با جایگذاری (۳۶) و (۳۷) در (۳۱)، تخمین همیلتونین به صورت زیر قابل محاسبه خواهد بود:

$$\begin{aligned} \hat{H}(z, \hat{W}_2) = & z^T Q_T z + \hat{W}_2^T \nabla \phi_2(z) \hat{F}(z) - \gamma \hat{W}_2^T \phi_2(z) \\ & - \frac{1}{4} \hat{W}_2^T \nabla \phi_2(z) \wedge \nabla \phi_2^T(z) \hat{W}_2 \triangleq \hat{e}(z(t)) \end{aligned} \quad (38)$$

فرض ۷: برای سیستم (۲۹) با تابع هزینه عملکرد (۳۰) و قانون FTOTC (۳۲)، تابع لیاپانوف کاندید پیوسته شعاعی نامحدود $J_1(z)$ به طوری که $J_1(z) = \nabla J_1^T z = \nabla J_1^T (\hat{F}(z) + \hat{G}(z)u^*) \leq 0$ باشد

$$\delta_2 = -\frac{1}{4m_s} \nabla \phi_2(z) \Lambda \nabla \phi_2^T(z) W_2 \bar{\sigma}^T W_2 + K_2 \bar{\sigma}^T W_2 - K_1 W_2$$

با توجه به فرض ۵ و ۶ و همچنین $\bar{\sigma} < 1$ ، می‌توان نتیجه گرفت که δ محدود به مقدار δ_M است. در صورتی که K_1 و K_2 به نحوی انتخاب شوند که $N > 0$ باشد، سپس حد بالای رابطه (۴۸) به صورت زیر قابل محاسبه خواهد بود:

$$j \leq \nabla J_1^T \left(\hat{F}(z) - \frac{\Lambda \nabla \phi_2^T(z) \hat{W}_2}{2} \right) - \zeta^2 \sigma_{min}(N) + \zeta \delta_M - \Xi(z, \hat{u}) \left(\frac{\tilde{W}_2^T \nabla \phi_2(z) \Lambda \nabla J_1}{2} - \tilde{W}_1 (\lambda_{min}(P) \tilde{W}_1 - \varepsilon_v) \right) \quad (49)$$

بر اساس رفتار مشتق زمانی تابع لیاپانوف دو حالت زیر قابل رخداد است:

الف) $\Xi(z, \hat{u}) = 0$ ، با توجه به شرط برانگیخته کامل بودن و رابطه (۴۱) می‌توان نتیجه گرفت یک مقدار مثبت ثابت $\hat{z}_{min}$ وجود دارد که $0 < \hat{z}_{min} < \hat{z}$. سپس از رابطه (۴۹) داریم:

$$\begin{aligned} j &\leq -\hat{z}_{min} \nabla J_1 - \zeta^2 \lambda_{min}(N) + \zeta \delta_M \\ &= -\hat{z}_{min} \nabla J_1 - \lambda_{min}(N) \left(\zeta - \frac{\delta_M}{2\lambda_{min}(M)} \right)^2 \end{aligned} \quad (50)$$

$$-\lambda_{min}(P) \left(\tilde{W}_1 - \frac{\varepsilon_v}{2\lambda_{min}(P)} \right)^2 + \Pi$$

که در آن $\Pi = \frac{\delta_M^2}{4\lambda_{min}(M)} + \frac{\varepsilon_v^2}{4\lambda_{min}(P)}$. در نتیجه برای اینکه j منفی باشد بایستی نامساوی‌های زیر برقرار باشد:

$$\nabla J_1 \geq \frac{\Pi}{\hat{z}_{min}} \triangleq B_{\nabla J_1} \quad (51)$$

$$-\lambda_{min}(N) \left(\zeta - \frac{\delta_M}{2\lambda_{min}(N)} \right)^2 + \Pi < 0 \quad (52)$$

$$\Rightarrow \zeta \geq \sqrt{\frac{\Pi}{\lambda_{min}(N)}} + \frac{\delta_M}{2\lambda_{min}(N)} \triangleq B_\zeta$$

$$-\lambda_{min}(P) \left(\tilde{W}_1 - \frac{\varepsilon_v}{2\lambda_{min}(P)} \right)^2 + \Pi < 0 \quad (53)$$

$$\Rightarrow \tilde{W}_1 \geq \sqrt{\frac{\Pi}{\lambda_{min}(P)}} + \frac{\varepsilon_v}{2\lambda_{min}(P)} \triangleq B_{\tilde{W}_1}$$

در نتیجه با برقرار بودن حداقل یکی از نامساوی‌های (۵۱)، (۵۲) و (۵۳) می‌توان نتیجه گرفت که ∇J_1 ، ζ و $\tilde{W}_1$ پایدار فراگیر یکنواخت هستند.

ب) $\Xi(z, \hat{u}) = 1$ ، با اضافه و کم نمودن جمله $\frac{1}{2} \nabla J_1 \Lambda \nabla \varepsilon_2$ به سمت راست (۴۹) و با استفاده از (۳۲)، می‌توان دریافت که:

$$j \leq \nabla J_1^T \left(\hat{F}(z) + \hat{G}(z) u^* \right) + \frac{1}{2} \nabla J_1 \Lambda \nabla \varepsilon_2 \quad (54)$$

شبکه شناساگر و نقاد در نظر بگیرید. با توجه به فرض ۵ و ۶، اگر بردار توابع پایه ϕ_1 برانگیخته کامل باشد، سپس بردار حالت z ، خطای تخمین $\tilde{W}_1$ و خطای تخمین $\tilde{W}_2$ همگی پایدار UBB هستند. اثبات: تابع لیاپانوف کاندید زیر را در نظر بگیرید:

$$J = J_1(z) + \frac{1}{2} \text{tr}(\tilde{W}_1^T \Gamma^{-1} \tilde{W}_1) + \frac{1}{2\alpha_1} \tilde{W}_2^T \tilde{W}_2 \quad (44)$$

که Γ ماتریس بهره یادگیری برای قانون به‌روزرسانی شبکه شناساگر (۱۷) و $\alpha_1 > 0$ بهره یادگیری شبکه نقاد است.

مشتق زمانی تابع (۴۴) به صورت زیر محاسبه می‌شود:

$$j = \nabla J_1^T \left(\hat{F}(z) - \frac{\Lambda \nabla \phi_2^T(z) \hat{W}_2}{2} \right) + \text{tr} \left(\tilde{W}_1^T \Gamma^{-1} \tilde{W}_1 \right) + \tilde{W}_2^T \alpha_1^{-1} \tilde{W}_2 \quad (45)$$

با جایگذاری (۴۳) در (۴۵)، داریم:

$$\begin{aligned} j &= \nabla J_1^T \left(\hat{F}(z) - \frac{\Lambda \nabla \phi_2^T(z) \hat{W}_2}{2} \right) + \tilde{W}_2^T \bar{\sigma} \left(-\bar{\sigma}^T \tilde{W}_2 - \frac{\varepsilon_{HBB}}{m_s} \right) \\ &\quad + \frac{1}{4} \tilde{W}_2^T \nabla \phi_2(z) \Lambda \nabla \phi_2^T(z) W_2 \frac{\bar{\sigma}^T}{m_s} \tilde{W}_2 \\ &\quad - \frac{1}{4} \tilde{W}_2^T \nabla \phi_2(z) \Lambda \nabla \phi_2^T(z) W_2 \frac{\bar{\sigma}^T}{m_s} W_2 \\ &\quad + \frac{1}{4} \tilde{W}_2^T \nabla \phi_2(z) \Lambda \nabla \phi_2^T(z) \tilde{W}_2 \frac{\bar{\sigma}^T}{m_s} W_2 \\ &\quad - \Xi(z, \hat{u}) \left(\frac{\tilde{W}_2^T \nabla \phi_2(z) \Lambda \nabla J_1}{2} \right) \\ &\quad + \text{tr} \left(\tilde{W}_1^T \Gamma^{-1} \tilde{W}_1 \right) - \tilde{W}_2^T K_2 \hat{W}_2 + \tilde{W}_2^T K_1 \bar{\sigma}^T \hat{W}_2 \end{aligned} \quad (46)$$

دو جمله آخر (۴۶) را می‌توان به صورت زیر بازنویسی کرد:

$$\begin{aligned} \tilde{W}_2^T K_1 \bar{\sigma}^T \hat{W}_2 - \tilde{W}_2^T K_2 \hat{W}_2 &= \tilde{W}_2^T K_2 \tilde{W}_2 - \tilde{W}_2^T K_2 W_2 \\ -\tilde{W}_2^T K_1 \bar{\sigma}^T \tilde{W}_2 + \tilde{W}_2^T K_1 \bar{\sigma}^T W_2 & \end{aligned} \quad (47)$$

با تعریف کردن $\zeta = [\bar{\sigma}^T \tilde{W}_2, \tilde{W}_2]^T$ ، برای رابطه (۴۶) داریم:

$$\begin{aligned} j &= -\zeta^T \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \zeta + \zeta^T \delta \\ &\quad + \nabla J_1^T \left(\hat{F}(z) - \frac{\Lambda \nabla \phi_2^T(z) \hat{W}_2}{2} \right) \\ &\quad - \Xi(z, \hat{u}) \left(\frac{\tilde{W}_2^T \nabla \phi_2(z) \Lambda \nabla J_1}{2} \right) + \text{tr} \left(\tilde{W}_1^T \Gamma^{-1} \tilde{W}_1 \right) \end{aligned} \quad (48)$$

که درایه‌های ماتریس N به صورت زیر است:

$$\begin{aligned} N_{11} &= I \\ N_{21} = N_{12}^T &= -\frac{1}{4m_s} \nabla \phi_2(z) \Lambda \nabla \phi_2^T(z) W_2 + \frac{K_1}{2} \\ N_{22} &= -K_2 \end{aligned}$$

و درایه‌های بردار $\delta = [\delta_1 \quad \delta_2]^T$ به صورت زیر است:

$$\delta_1 = \frac{\varepsilon_{HBB}}{m_s}$$

۳-۴- آشکارسازی مبتنی بر HJB

پیش از اینکه عیب در سیستم اتفاق بیافتد، قانون کنترل با سیستم در انطباق^{۳۰} است. به عبارت دیگر سیستم در شرایط انطباق عمل می‌کند. در این مقاله، کنترل‌کننده FTOTC برای سیستم (۲۹) طراحی می‌شود که تابع هزینه عملکرد (۳۰) را کمینه می‌کند. بر اساس روش ارائه شده برای طراحی کنترل‌کننده، می‌توان دریافت که خطای باقی مانده معادله HJB در شرایط انطباق مقداری نزدیک به صفر دارد. لذا فرض کنید که نامساوی زیر در شرایط انطباق برای خطای باقی مانده معادله HJB برقرار باشد:

$$\hat{e}(z(t)) \leq \xi \quad (59)$$

که ξ یک مقدار از پیش تعیین شده بر اساس شرط توقف فرآیند یادگیری FTOTC است. بنابراین اولین زمانی که نامساوی (۵۹) نقض شود به عنوان زمان آشکارسازی عیب در نظر گرفته می‌شود.

توجه ۳: حساسیت آشکارسازی عیب رابطه مستقیمی با اثرگذاری عیوب عملکرد و اجزا بر مقدار معادله HJB دارد. در این رویکرد، تنها عیوبی قابل آشکارسازی هستند که باعث افزایش مقدار خطای معادله HJB یا به عبارت دیگر کاهش عملکرد سیستم شوند. از سوی دیگر بایستی توجه کرد که $\hat{e}(z(t)) \geq \xi$ لزوماً به معنی ناپایداری سیستم نخواهد بود.

۵- شبیه‌سازی

در این قسمت، به منظور بررسی عملکرد و کارایی روش پیشنهادی نتایج شبیه‌سازی حاصل از بکارگیری آن در کنترل زاویه چرخش پرنده کوادروتور آورده شده است. مدل دینامیکی پرنده کوادروتور مطابق رابطه (۸) و پارامترهای آن مطابق با جدول ۱ در نظر گرفته شده است [۳۸].

جدول (۱): پارامترهای مدل دینامیکی پرنده

پارامتر	مقدار	پارامتر	مقدار
$I_x \text{ (Kg.m}^2\text{)}$	$7.78e^{-3}$	۱	۰.۰۳۳
$I_y \text{ (Kg.m}^2\text{)}$	$7.78e^{-3}$	$J_r \text{ (Kg.m}^2\text{)}$	$6e^{-3}$
$I_z \text{ (Kg.m}^2\text{)}$	$1.44e^{-2}$		

دینامیک پرنده (۸) در معرض هر دو عیب عملکرد و اجزا به صورت زیر در نظر گرفته شده است:

$$\dot{x}_1(t) = x_2(t) \quad (60)$$

$$\dot{x}_2(t) = v(t)f(x(t)) + \rho(t)g(x(t))u(t)$$

که $\rho(t)$ و $v(t)$ نسبت به زمان رفتاری به صورت زیر دارند:

$$v(t) = \begin{cases} 1 & 0 < t < 36 \\ 0.5 & t \geq 36 \end{cases} \quad (61)$$

$$-\lambda_{\min}(N) \left(\zeta - \frac{\delta_M}{2\lambda_{\min}(N)} \right)^2 - \lambda_{\min}(P) \left(\tilde{W}_1 - \frac{\varepsilon_v}{2\lambda_{\min}(P)} \right)^2 + \Pi$$

با توجه به فرض ۵، Λ محدود به مقدار مثبت Λ_M خواهد بود به نحوی که $0 < \Lambda < \Lambda_M$. با استفاده از فرض ۶ و ۷، رابطه (۵۴) را می‌توان به صورت زیر بازنویسی کرد:

$$J \leq -\bar{Q}_{1\min} \nabla J_1^2 - \lambda_{\min}(N) \left(\zeta - \frac{\delta_M}{2\lambda_{\min}(N)} \right)^2 - \lambda_{\min}(P) \left(\tilde{W}_1 - \frac{\varepsilon_v}{2\lambda_{\min}(P)} \right)^2 + \frac{1}{2} \nabla J_1 \Lambda_M \varepsilon_{2dM} + \Pi \leq -\bar{Q}_{1\min} \nabla J_1^2 - \lambda_{\min}(N) \left(\zeta - \frac{\delta_M}{2\lambda_{\min}(N)} \right)^2 - \lambda_{\min}(P) \left(\tilde{W}_1 - \frac{\varepsilon_v}{2\lambda_{\min}(P)} \right)^2 + Y \quad (55)$$

که در آن $Y = \frac{\Lambda_M^2 \varepsilon_{2dM}^2}{4Q_{1\min}} + \Pi$. بنابراین در صورتی که حداقل یکی از نامساوی‌های زیر برقرار باشد، J منفی خواهد بود.

$$\nabla J_1 > \sqrt{\frac{2Y}{Q_{1\min}}} \triangleq B'_{\nabla J_1} \quad (56)$$

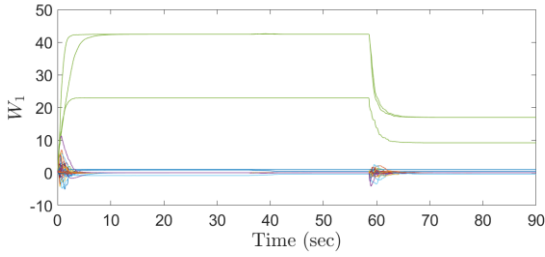
$$-\lambda_{\min}(N) \left(\zeta - \frac{\varepsilon_M}{2\lambda_{\min}(N)} \right)^2 + \Pi < 0 \quad (57)$$

$$\Rightarrow \zeta \geq \sqrt{\frac{Y}{\lambda_{\min}(N)}} + \frac{\delta_M}{2\lambda_{\min}(N)} \triangleq B'_\zeta$$

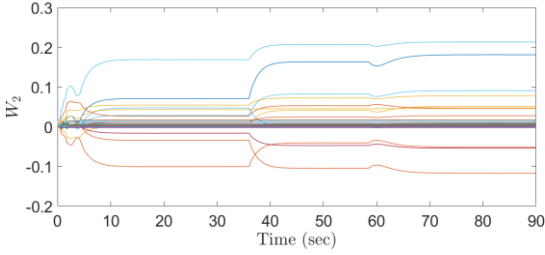
$$-\lambda_{\min}(P) \left(\tilde{W}_1 - \frac{\varepsilon_v}{2\lambda_{\min}(P)} \right)^2 + \Pi < 0 \quad (58)$$

$$\Rightarrow \tilde{W}_1 \geq \sqrt{\frac{Y}{\lambda_{\min}(P)}} + \frac{\varepsilon_v}{2\lambda_{\min}(P)} \triangleq B'_{\tilde{W}_1}$$

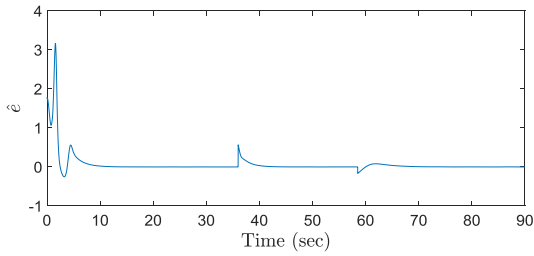
در نتیجه ∇J_1 ، ζ و $\tilde{W}_1$ پایدار فراگیر یکنواخت خواهد بود. در نتیجه برای هر دو حالت $\Xi(z, \hat{u}) = 1$ و $\Xi(z, \hat{u}) = 0$ ، در صورتی که $\zeta > \max(B'_\zeta, B'_\zeta) \triangleq \bar{B}_\zeta$ ، $\nabla J_1 > \max(B_{\nabla J_1}, B'_{\nabla J_1}) \triangleq \bar{B}_{\nabla J_1}$ و $\tilde{W}_1 > \max(B_{\tilde{W}_1}, B'_{\tilde{W}_1}) \triangleq \bar{B}_{\tilde{W}_1}$ باشد، سپس $\dot{J} < 0$ خواهد بود. با توجه به فرض ۷، محدود بودن ∇J_1 محدود بودن Z را نتیجه می‌دهد. می‌توان نشان داد که Z محدود به مقدار $\bar{B}_z$ است که توسط $\bar{B}_{\nabla J_1}$ قابل تعیین است. همچنین محدود بودن δ ، محدود بودن خطای تخمین وزن‌های $\tilde{W}_2$ به مقدار $\bar{B}_z$ را نتیجه می‌دهد.



(ج)



(د)



(و)

شکل (۲): خروجی سیستم حلقه بسته تحت روش پیشنهادی پیش از وقوع عیوب و پس از آن. (الف) ردیابی سیگنال مرجع مطلوب زاویه چرخش، (ب) سیگنال کنترل ورودی، (ج) وزن‌های شبکه شناساگر، (د) وزن‌های شبکه نقاد، (و) مقدار باقی مانده معادله HJB

همانطور که در شکل ۲ نشان داده شده است کنترل ردیاب زاویه چرخش در روش پیشنهادی در عدم حضور عیب در 14.32s به همگرایی رسیده است و وزن‌های شبکه شناساگر نیز در 9.41s به وزن‌های واقعی خود همگرا شده است. با توجه به قسمت (و) می‌توان دریافت تغییر میزان عیب باعث افزایش خطای معادله HJB گردیده است که طبق (۵۹)، با در نظر گرفتن $\xi = 0.05$ ، آشکارسازی رخداد عیوب به ترتیب در زمان‌های 36.04s و 58.01s صورت گرفته است. همچنین قابل توجه است که با در نظر گرفتن سه شرط فوق، فرآیند آموزش به ترتیب در زمان‌های 44.34s و 68.21s متوقف شده است. قسمت (ج) و (د) این شکل به ترتیب همگرایی وزن‌های شبکه شناساگر و نقاد را در طول فرآیند آموزش نشان می‌دهد که نرم خطای آن‌ها یعنی $\|\tilde{W}_1\|$ و $\|\tilde{W}_2\|$ در انتهای آموزش، در حالت بدون عیب و همچنین بعد از رخداد عیوب، کمتر از 0.011 بوده است.

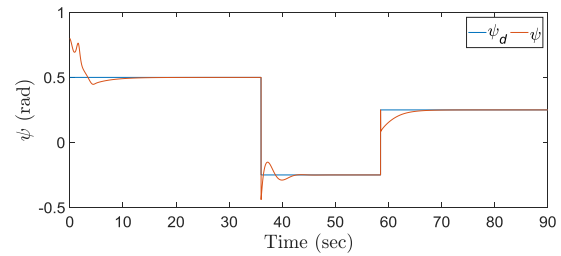
$$\rho(t) = \begin{cases} 1 & 0 < t < 58 \\ 0.4 & t \geq 58 \end{cases} \quad (62)$$

پارامترهای تنظیم سیستم شناساگر به صورت $l=1, k_1=1, k_2=0.1, k=0.05, \Gamma=40$ بردار توابع پایه به صورت $\phi_1(x, u) = [x_{21} \ x_{22} \ x_{23} \ x_{21}x_{22} \ x_{21}x_{23} \ x_{22}x_{23} \ u^T]^T$ و $W_1(0)$ به صورت یک بردار تصادفی با ابعاد مناسب در نظر گرفته شده است. همچنین در شبکه نقاد $\alpha_1=10$ و بردار توابع پایه شامل تمامی تک‌جمله‌ای‌های از درجه ۲ تا ۴ در نظر گرفته شده است. مسیر مرجع مطلوب زاویه چرخش به صورت تکه‌ای ثابت زیر در نظر گرفته شده است:

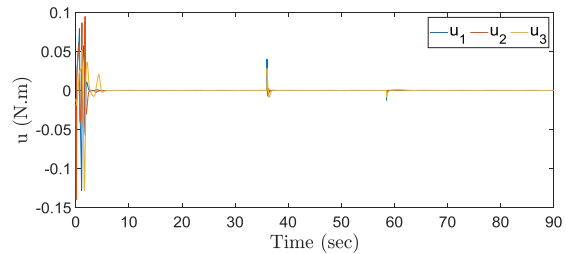
$$\psi_d(t) = \begin{cases} 0.5 & 0 \leq t < 36 \\ -0.25 & 36 \leq t < 58 \\ 0.25 & 58 \leq t \end{cases} \quad (63)$$

وضعیت اولیه پرنده در شبیه‌سازی به صورت $x_1(0) = [-1, 0.5, 0.8]^T$ و $x_2(0) = [0, 0, 0]^T$ در نظر گرفته شده است.

به منظور بررسی کارایی روش پیشنهادی در آموزش کنترل‌کننده وضعیت بهینه کوادروتور و همچنین جبران سازی اثر عیب، نتایج شبیه‌سازی حاصل از شبیه‌سازی روش مذکور با روش ارائه شده در [۳۶] مقایسه شده است که ترتیب در شکل‌های ۲ الی ۳ نشان داده شده است. شرط توقف آموزش برآورده شدن هر سه نامساوی $\dot{e}(z) < \infty$ ، $\hat{e}(z) < 0.2$ و $\hat{W}_2 < 0.2$ در نظر گرفته شده است. قسمت (الف) هر یک از این دو شکل سیگنال مرجع زاویه چرخش و خروجی پرنده و قسمت (ب) ورودی کنترلی را نشان می‌دهند.



(الف)



(ب)

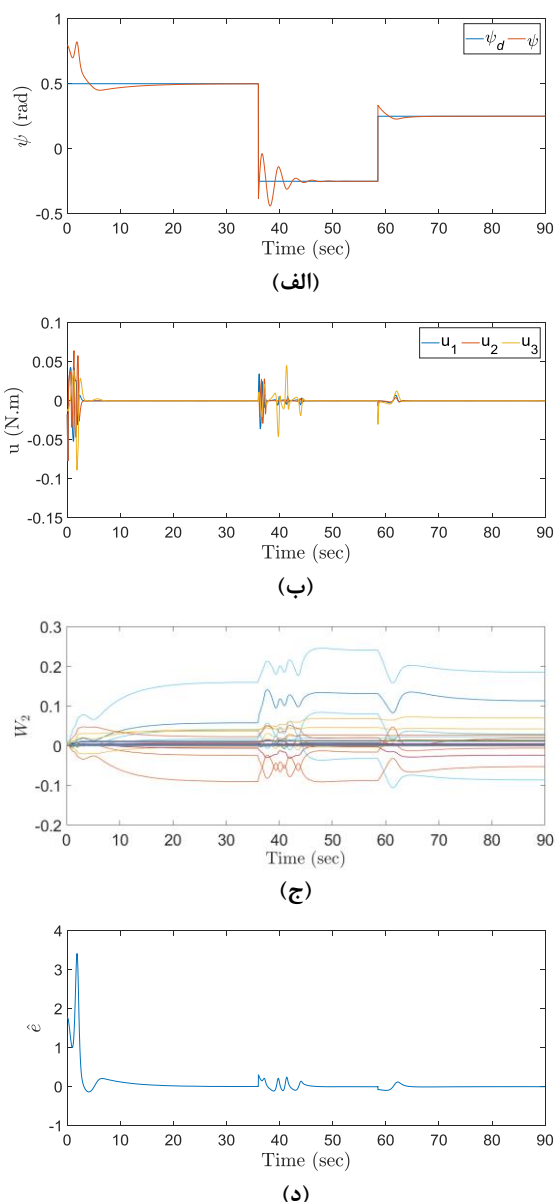
همگرا شود و قادر است وضعیت مرجع مطلوب را در هر دو حالت سالم و معیوب دنبال نماید.

۶- نتیجه گیری

در این مقاله یک روش جدید FTOTC مبتنی بر الگوریتم یادگیری تقویتی جهت طراحی سیستم کنترل وضعیت پرنده کوادروتور ارائه شده است. روش پیشنهادی مبتنی بر یک ساختار تخمین با دو شبکه عصبی شناساگر و نقاد توسعه داده شده است و در مقایسه با ساختارهای تخمین سه گانه شامل شبکه های شناساگر، عملگر و نقاد، باعث کاهش بار محاسباتی و افزایش سرعت همگرایی FTOTC شده است. ضریب فراموش در شبکه شناساگر متغیر در نظر گرفته شده است که در مقایسه با حالت ثابت، باعث افزایش سرعت همگرایی و کاهش خطای حالت ماندگار گردیده است. به منظور حذف نیازمندی به شناخت کنترل کننده پایدار ساز اولیه، در قانون به روزرسانی شبکه نقاد از یک جمله پایدار ساز استفاده شده است که جبران سازی اثر عیب را بدون نیاز به هیچ گونه شناخت قبلی از مقدار عیب و دینامیک سیستم فراهم می کند. همچنین با استفاده از قضیه لیاپانوف نشان داده شد که خطای ردیابی و خطای وزن های شبکه شناساگر و نقاد پایدار فراگیر یکنواخت خواهد بود. نتایج حاصل از بکارگیری روش پیشنهادی، کارایی آن را در حلقه کنترل سیستم خلبان خودکار کوادروتور نشان می دهد.

مراجع

- [1] D. K. Villa, A. S. Brandao, M. Sarcinelli-Filho, "A survey on load transportation using multirotor UAVs", Journal of Intelligent & Robotic Systems, vol. 98, no. 2, pp. 267-296, 2020.
- [2] Yousefi S A, Kazemi M H. Control of Quadcopter with Suspended Load by Internal Model Control and Input Shaping for Reducing the Load Oscillation. Journal of Iranian Association of Electrical and Electronics Engineers 2022; 19 (3): 215-226.
- [3] Barshooi A H, Amirkhani A. Cooperative Control of Mobile Robots in Creating a Runway Platform for Quadrotor Landing. Journal of Iranian Association of Electrical and Electronics Engineers 2023; 20 (1): 153-162.
- [4] W. Zhao, H. Liu, F. L. Lewis, "Data-driven fault-tolerant control for attitude synchronization of nonlinear quadrotors", IEEE Transactions on Automatic Control, vol. 66, no. 11, pp. 5584-5591, 2021.
- [5] A. A. Amin, K. M. Hasan, "A review of fault tolerant control systems: advancements and applications", Measurement, vol. 143, pp. 58-68, 2019.
- [6] C. Liu, B. Jiang, R. J. Patton, K. Zhang, "Decentralized output sliding-mode fault-tolerant control for heterogeneous multiagent systems", IEEE transactions on cybernetics, vol. 50, no. 12, pp. 4934-4945, 2019.
- [7] S. Roshanravan, Sobhani B. Gendeshmin, S. Shamaghdari, "Design of an actuator fault-tolerant controller for an air vehicle with nonlinear dynamics", Proceedings of the Institution of Mechanical Engineers,



شکل (۳): خروجی سیستم حلقه بسته تحت روش مبتنی بر داده پیش از وقوع عیوب و پس از آن. (الف) ردیابی سیگنال مرجع مطلوب زاویه چرخش، (ب) سیگنال کنترل ورودی، (ج) وزن های شبکه شناساگر، (د) وزن های شبکه نقاد، (و) مقدار b باقی مانده

معادله HJB

همانطور که در شکل ۳ نشان داده شده است، کنترل زاویه چرخش تحت روش ارائه شده در [۳۶] در مقایسه با روش ارائه شده در این مقاله، با در نظر گرفتن شروط توقف آموزش یکسان، در زمان های طولانی تری به همگرایی رسیده است. قابل توجه است این روش نیز به صورت مستقل از مدل و داده محور ارائه شده است که منجر به عدم نیاز این روش به شبکه شناساگر شده است. اگر چه در عوض برای آموزش کنترل کننده بهینه در هر گام، به مرحله جمع آوری داده احتیاج دارد که باعث کاهش سرعت همگرایی آن شده است. از این رو با مقایسه شکل ۲ و ۳ می توان دریافت که روش پیشنهاد شده در این مقاله توانسته است با سرعت بالاتری به کنترل کننده بهینه

- control for nonlinear systems with actuator faults", ISA transactions, vol. 120, pp. 99-109, 2022.
- [22] P. Deptula, Z. I. Bell, E. A. Doucette, J. W. Curtis, W. E. Dixon, "Data-based reinforcement learning approximate optimal control for an uncertain nonlinear system with control effectiveness faults", *Automatica*, vol. 116, pp. 108922, 2020.
- [23] H. Lin, B. Zhao, D. Liu, C. Alippi, "Data-based fault tolerant control for affine nonlinear systems through particle swarm optimized neural networks", *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 4, pp. 954-964, 2020.
- [24] K. Li, Y. Li, "Adaptive nn optimal consensus fault-tolerant control for stochastic nonlinear multiagent systems", *IEEE Transactions on Neural Networks and Learning Systems*, 2021.
- [25] H.-J. Ma, L.-X. Xu, G.-H. Yang, "Multiple environment integral reinforcement learning-based fault-tolerant control for affine nonlinear systems", *IEEE transactions on cybernetics*, vol. 51, no. 4, pp. 1913-1928, 2019.
- [26] K. Zhang, H. Zhang, Z. Gao, H. Su, "Online adaptive policy iteration based fault-tolerant control algorithm for continuous-time nonlinear tracking systems with actuator failures", *Journal of the Franklin Institute*, vol. 355, no. 15, pp. 6947-6968, 2018.
- [27] J. Lan, R. J. Patton, "A new strategy for integration of fault estimation within fault-tolerant control", *Automatica*, vol. 69, pp. 48-59, 2016.
- [28] F. Sabbaghian-Bidgoli, M. Farrokhi, "Polynomial fuzzy observer-based integrated fault estimation and fault-tolerant control with uncertainty and disturbance", *IEEE Transactions on Fuzzy Systems*, vol. 30, no. 3, pp. 741-754, 2020.
- [29] A. Mishra, S. Ghosh, "Simultaneous identification and optimal tracking control of unknown continuous-time systems with actuator constraints", *International Journal of Control*, vol. 95, no. 8, pp. 2005-2023, 2022.
- [30] S. Pakkhesal, S. Shamaghdari, "Sum- of- squares- based policy iteration for suboptimal control of polynomial time- varying systems", *Asian Journal of Control*, vol. 24, no. 6, pp. 3022-3031, 2022.
- [31] T. Dierks, S. Jagannathan, "Optimal control of affine nonlinear continuous-time systems", In *Proceedings of the 2010 American control conference*, pp. 1568-1573, IEEE, 2010.
- [32] X. Yang, D. Liu, Q. Wei, "Robust tracking control of uncertain nonlinear systems using adaptive dynamic programming", In *International conference on neural information processing*, pp. 9-16. Springer, Cham, 2015.
- [33] Y.-C. Choi, H.-S. Ahn, "Nonlinear control of quadrotor for point tracking: Actual implementation and experimental tests", *IEEE/ASME transactions on mechatronics*, vol. 20, no. 3, pp. 1179-1192, 2014.
- [34] C. Edwards, T. Lombaerts, and H. Smaili, "Fault tolerant flight control", *Lecture notes in control and information sciences*, vol. 399, pp. 1-560, 2010.
- [35] J. Na, M. N. Mahyuddin, G. Herrmann, X. Ren, P. Barber, "Robust adaptive finite time parameter estimation and control for robotic systems", *International Journal of Robust and Nonlinear Control*, vol. 25, no. 16, pp. 3045-3071, 2015.
- [36] H. Modares, F. L. Lewis, Z.-P. Jiang, " $\mathcal{H}_\infty$ Tracking control of completely unknown continuous-time systems via off-policy reinforcement learning", *IEEE transactions on neural networks and learning systems*, vol. 26, no. 10, pp. 2550-2562, 2015.
- Part G: *Journal of Aerospace Engineering*, vol. 233, no. 10, pp. 3534-3546, 2019.
- [8] S. Roshanravan, S. Shamaghdari, "Simultaneous fault detection and isolation and fault-tolerant control using supervisory control technique: asynchronous switching approach", *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 234, no. 8, pp. 900-911, 2020.
- [9] S. Roshanravan, S. Shamaghdari, "Adaptive Fault-Tolerant Tracking Control for Affine Nonlinear Systems With Unknown Dynamics via Reinforcement Learning", *IEEE Transactions on Automation Science and Engineering*, vol. 21, no. 1, pp. 569-580, 2024.
- [10] J.-X. Zhang, G.-H. Yang, "Prescribed performance fault-tolerant control of uncertain nonlinear systems with unknown control directions", *IEEE Transactions on Automatic Control*, vol. 62, no. 12, pp. 6529-6535, 2017.
- [11] G. Liu, N. Sun, T. Yang, Y. Fang, "Reinforcement Learning-Based Prescribed Performance Motion Control of Pneumatic Muscle Actuated Robotic Arms With Measurement Noises", *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2022.
- [12] J.-X. Zhang, G.-H. Yang, "Fault-tolerant output-constrained control of unknown Euler-Lagrange systems with prescribed tracking accuracy", *Automatica*, vol. 111, pp. 108606, 2020.
- [13] H. Zhang, L. Cui, Y. Luo, "Near-optimal control for nonzero-sum differential games of continuous-time nonlinear systems using single-network ADP", *IEEE transactions on cybernetics*, vol. 43, no. 1, pp. 206-216, 2012.
- [14] Z. Wang, L. Liu, H. Zhang, G. Xiao, "Fault-tolerant controller design for a class of nonlinear MIMO discrete-time systems via online reinforcement learning algorithm", *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 46, no. 5, pp. 611-622, 2015.
- [15] L. Liu, Z. Wang, H. Zhang, "Adaptive fault-tolerant tracking control for MIMO discrete-time systems via reinforcement learning algorithm with less learning parameters", *IEEE Transactions on Automation Science and Engineering*, vol. 14, no. 1, pp. 299-313, 2016.
- [16] Z. Wang, L. Liu, Y. Wu, H. Zhang, "Optimal fault-tolerant control for discrete-time nonlinear strict-feedback systems based on adaptive critic design", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 6, pp. 2179-2191, 2018.
- [17] Q.-Y. Fan, G.-H. Yang, "Active complementary control for affine nonlinear control systems with actuator faults", *IEEE Transactions on Cybernetics*, vol. 47, no. 11, pp. 3542-3553, 2016.
- [18] Q. Y. Fan, G. H. Yang, "Adaptive fault- tolerant control for affine non- linear systems based on approximate dynamic programming", *IET Control Theory & Applications*, vol. 10, no. 6, pp. 655-663, 2016.
- [19] H. Jiang, H. Zhang, Y. Liu, J. Han, "Neural-network-based control scheme for a class of nonlinear systems with actuator faults via data-driven reinforcement learning method", *Neurocomputing*, vol. 239, pp. 1-8, 2017.
- [20] M. E. Dehshalie, M. B. Menhaj, M. Karrari, "Fault tolerant cooperative control for affine multi-agent systems: An optimal control approach", *Journal of the Franklin Institute*, vol. 356, no. 3, pp. 1360-1378, 2019.
- [21] S. Zhang, C. Huang, K. Ji, H. Zhang, "Prescribed performance incremental adaptive optimal fault-tolerant

- [37] H. Modares, F. L. Lewis, "Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning", *Automatica*, vol. 50, no. 7, pp. 1780-1792, 2014.
- [38] X. Wang, Q. Wang, C. Sun, "Prescribed performance fault-tolerant control for uncertain nonlinear MIMO system using actor-critic learning structure", *IEEE Transactions on Neural Networks and Learning Systems*, 2021.

زیرنویس‌ها

-
- ¹ Fault-Tolerant Control
² Active
³ Passive
⁴ Robust
⁵ Fault Detection and Diagnosis
⁶ Fault-Tolerant Optimal Tacking Control
⁷ Reinforcement Learning
⁸ mn-Order
⁹ Actor-Critic
¹⁰ Strict Feedback
¹¹ Affine
¹² Policy Iteration
¹³ Integrated Fault-Tolerant Control
¹⁴ Supervisory
¹⁵ Abrupt
¹⁶ Intermittent
¹⁷ Plausibility Test
¹⁸ Hamilton-Jacobi-Bellman
¹⁹ Adaptive Neural Network Identifier
²⁰ Lyapunov
²¹ Update Law
²² Integral Reinforcement Learning
²³ Forgetting Factor
²⁴ Drag
²⁵ Drift
²⁶ Input
²⁷ Persistently Excited
²⁸ Uniformly Ultimately Boundedness
²⁹ Stationary condition
³⁰ Matched