

طراحی بهینه کنترل کننده‌ی گام به عقب چهارپره با روش اتوماتای یادگیری تقویتی

رضا عباس پور^۱، علی اکبر قره ویسی^۲، سید محمد علی محمدی^۳

چکیده

در چند سال اخیر استفاده از پرنده‌های بدون سرنشین در مأموریت‌های نظامی و غیرنظامی به دلیل توانایی شناور بودن، پرواز در ارتفاع‌های کم و فضاهای بسته افزایش فزاینده‌ای داشته است. در این مقاله به تحلیل، طراحی و شبیه‌سازی کنترل‌کننده-ی بهینه در مورد هواپیمای ۴ موتوره با ۶ درجه آزادی پرداخته می‌شود که به چهارپره معروف است. چهارپره یک وسیله‌ی عمومی در تحقیقات پرنده‌های بدون سرنشین است زیرا طراحی و نگهداری ساده‌ای دارد. سیستم چهارپره به علت وجود اثرات غیرخطی آیرودینامیکی در مدل آن بسیار ناپایدار است و کنترل آن دشوار خواهد بود. طراحی کنترل‌کننده‌ی سطح پایین به دلیل غیرخطی بودن و ناپایداری ذاتی مدل یک چالش است. کنترل‌کننده‌ی مورد استفاده کنترل‌کننده‌ی گام به عقب است که بر عملگرهای زوایا اعمال شده و سپس با استفاده از روش اتوماتای یادگیری تقویتی بهینه شده است. مقایسه این روش بهینه‌سازی با جعبه ابزار بهینه‌سازی MATLAB، نشان می‌دهد که نتایج بهتری بدست آمده است.

کلیدواژه

چهارپره، کنترل گام به عقب، اتوماتای یادگیری تقویتی، کنترل غیرخطی

مقدمه

بی‌نیاز است. تغییر سرعت موتورها موجب حرکت در جهت افق می‌شود و کاهش یا افزایش همزمان سرعت آنها حرکت عمودی را ایجاد می‌کند. بر خلاف چرخبال‌های معمول که زاویه پیچ متغییری دارند، چهارپره زاویه پیچ ثابتی دارد، بنابراین تمامی حرکات اساسی با کنترل سرعت پره‌ها تغییر می‌یابد [۱]. مهمترین مزیت پرنده‌های چند موتوره، افزایش مانورپذیری، افزایش توانایی حمل بار و قابلیت نشست و برخاست عمودی است. مهمترین عیب چهارپره، افزایش مصرف انرژی به دلیل افزایش تعداد موتورها است.

تحقیقات زیادی در زمینه‌ی طراحی کنترل‌کننده برای پایدارسازی چهارپره صورت گرفته است. بوعبدالله^۴ و همکارانش [۲] کنترل‌کننده‌های مختلفی نظیر LQ^۵ و PID^۶ را بر روی مدلی که توسط خودشان ارائه شده بود، مورد مطالعه قرار داده‌اند، در این تحقیقات مشخص شد که اثرات ژيروسکوپ^۷ در بدنه صلب نادیده گرفته شده است. روش‌های کنترل کلاسیک برای سیستم‌هایی با

پرنده‌های بدون سرنشین مزایا و کاربردهای فراوانی دارند که از جمله‌ی آنها می‌توان به امداد و نجات، نظارت از راه دور، جستجو، عملیات نظامی و عدم نیاز به حضور انسان‌ها در شرایط خطرناک را نام برد. اشیاء پرنده همواره جذبه‌ی خاصی در تحقیقات و توسعه‌ها داشته‌اند. از سوی دیگر پیشرفت‌های اخیر تکنولوژی در زمینه‌ی حساسه‌ها، منابع ذخیره انرژی، عملگرها و پردازنده‌ها، موجب جهش قابل توجهی در این حوزه شده‌اند.

چهارپره در کلاس هواپیمای بدون سرنشین قرار دارد و به علت ویژگی‌های خاص بسیار مورد توجه قرار گرفته است. این پرنده وسیله‌ای با ساختاری صلیبی است که جفت پره‌های آن در خلاف جهت یکدیگر می‌چرخند، با این کار اثر ممان اینرسی ایجاد شده حذف می‌گردد و به این صورت از داشتن پره دم

^۱ کارشناس ارشد برق-کنترل، دانشگاه شهید باهنر کرمان،

r.abbaspour1989@yahoo.com

^۲ استادیار بخش مهندسی برق، دانشگاه شهید باهنر کرمان

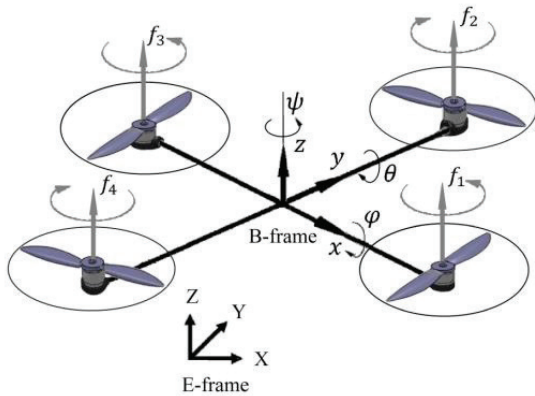
^۳ استادیار بخش مهندسی برق، دانشگاه شهید باهنر کرمان

تاریخ دریافت: ۱۹ آبان ۱۳۹۲ تاریخ پذیرش: ۱۱ دی ۱۳۹۲

Bouabdallah^۴
Linear Quadratic^۵
Proportional Integral Derivative^۶
Gyroscopic effects^۷

مدل سازی سیستم دینامیکی

معادلات دینامیکی سیستم غیرخطی، از مدل نیوتون-اولر توصیف شده در [۱۱] استفاده شده است. همچنین پارامترهای سیستم شبیه‌سازی شده برگرفته از یک مدل واقعی با نام "OS4"^۹ است تا توانایی پرواز موفق داشته باشد [۱۲].



شکل ۱. نمایش ساده‌ای از سیستم چهارپره

در مدل فیزیکی سیستم باید فرضیات زیر در نظر گرفته شوند [۱۳]:

- سازه و پره‌ها کاملاً صلب هستند.
- سازه متقارن فرض شده است.
- ضرایب تراست^۱ و دراگ^{۱۱} متناسب با مجذور سرعت پره‌ها هستند.
- مرکز چهارچوب بدنه ثابت و مرکز ثقل، منطبق بر هم فرض می‌شوند.

معادلات دینامیک این جسم صلب که تحت اثر نیروهای خارجی اعمال شده به مرکز جرم در چهارچوب بدنه ثابت بیان شده با استفاده از مدل‌سازی دینامیکی سیستم توسط معادله نیوتون-اولر بصورت رابطه (۱) می‌باشد [۱۳]:

$$[mI_{3 \times 3} \quad 0] \begin{bmatrix} \dot{V} \\ \dot{\omega} \end{bmatrix} + [\omega \times mV] = \begin{bmatrix} F + F_d \\ \tau + \tau_d \end{bmatrix} \quad (1)$$

که در این رابطه، m جرم پرنده، I ماتریس اینرسی، V بردار سرعت خطی و ω بردار سرعت زاویه‌ای بدنه است. F بردار نیروهای اعمال شده به سیستم، τ بردار گشتاورهای اعمال شده و اندیس d نیز نشان دهنده اغتشاشات وارد شده بر سیستم است. ماتریس اینرسی نیز به صورت قطری در نظر گرفته می‌شود. معادلات حرکت جسم صلب را می‌توان به صورت رابطه (۲) نوشت:

$$\begin{cases} \dot{\xi} = V \\ m\dot{V} = \mathcal{R}_1 F_b \\ \mathcal{R}_1 = \mathcal{R}_1 S(\omega) \\ I\dot{\omega} = -\omega \times I\omega + \tau_a \end{cases} \quad (2)$$

لحظه‌ای سرعت پره‌ها به صفر می‌رسید که این کار در عمل ناممکن است. مندرجاً و همکارانش [۵] از کنترل کننده‌ای مبتنی بر فیدبک بینایی استفاده کردند که برای این منظور، دوربینی بر بدنه چهارپره یا بر سطح زمین نصب می‌شود و با پردازش تصاویر و با استفاده از کنترل فازی از برخورد با مانع جلوگیری به عمل می‌آید اما تاخیر ایجاد شده در پردازش و ارسال اطلاعات به چهارپره توسط رایانه موجب ناهمواری در حرکت سیستم می‌شود. کنترل لغزشی یک روش قوی کنترلی دیگر است که ساده و مقاوم است اما نیازمند سوئیچینگ^۶ است که منجر به پدیده‌ی ضربه زدن می‌شود [۶]. روش‌های کنترل تطبیقی نیز بدلیل نیاز به شناسایی سیستم از سرعت پاسخ پایین‌تری نسبت به سایر روش‌ها برخوردارند [۱۶]. همچنین کنترل کننده‌های دیگری نیز با روش‌هایی نظیر فازی تطبیقی و کنترل لغزشی انتگرالی جهت کنترل چهارپره استفاده شده‌اند [۸، ۷]. کنترل موفق چهارپره در حال حمل بار با وزن ناشناخته، در حضور جریان‌های باد مساله‌ای به روز در تحقیقات در این زمینه است [۹]. پروژه‌های تجاری زیادی در زمینه چهارپره وجود دارد که از جمله آنها می‌توان به X-UFO، Draganflyer و MD4-200 اشاره کرد. بسیاری از محققان نیز ساخت این پرنده را ترجیح داده‌اند به عنوان مثال: OS4، X4-Flyer و STARMAC.

قبل از اینکه طراحی کنترل کننده صورت گیرد، ضروریست مجموعه مشخصات مربوط به نوع ماموریت و وظایف هواپیمای بدون سرنشین مورد نیاز تعریف گردد که از جمله‌ی آنها مانورپذیری، چابکی، دقت ردیابی و معیارهای بهینگی است که باید رضایت‌بخش باشند. مساله پایدارسازی و ردیابی چهارپره با استفاده از حلقه‌های کنترل تودرتو انجام می‌پذیرد. در کنترل حلقه داخلی از کنترل کننده‌ی غیرخطی گام به عقب^۸ استفاده می‌شود و حلقه‌ی کنترل بیرونی تراژکتوری لازم برای کنترل کننده‌ی حلقه داخلی به منظور دسترسی به نقاط مطلوب را فراهم می‌سازد [۱۰].

برای حل مساله، این مقاله یک کنترل کننده‌ی بهینه‌ی گام به عقب پیشنهاد می‌شود که نتایج آن با مقاله [۱۱] مقایسه شده است. در روش پیشنهادی پایداری سیستم حول نقاط تعادل با استفاده از توابع لیاپانوف بازگشتی تضمین می‌شود. روش گام به عقب به صورت بازگشتی به سیستم اعمال می‌شود تا کل سیستم پایدار شود. بررسی نتایج شبیه‌سازی نشان می‌دهند که روش پیشنهادی در این مقاله نتایج بهتری نسبت به [۱۱] ارائه می‌دهد.

^۹ Omnidirectional Stationary Flying Outstretched Robot
^{۱۰} thrust
^{۱۱} drag

^۶ Mendez
^۷ switching logic
^۸ Backstepping Control

منظور طراحی کنترل‌کننده‌های پایدارساز برای دسته‌ای خاص از سیستم‌های با دینامیک غیرخطی معرفی گردید. این سیستم‌ها از زیر سیستم‌های غیر قابل تقلیلی تشکیل می‌شوند که می‌توانند با استفاده از سایر روش‌ها پایدار شوند. بدلیل وجود این ساختار بازگشتی، طراح می‌تواند روند طراحی را از یک سیستم پایدار مشخص آغاز نماید و مابقی کنترل‌کننده‌ها نیز، به تدریج هر زیر سیستم بیرونی را پایدار می‌کنند. با رسیدن به آخرین زیر سیستم بیرونی طراحی پایان می‌یابد. از این رو این فرآید با نام "گام به عقب" شناخته شده است. طراحی ورودی کنترل واقعی برای اعمال به سیستم اصلی، در آخرین مرحله توسط یک تابع لیاپانوف نهایی انجام می‌پذیرد، این تابع لیاپانوف از مجموع توابع لیاپانوف مربوط به هر مرحله تشکیل می‌شود.

ایده‌ی اصلی روش گام به عقب، طراحی یک کنترل‌کننده بازگشتی با فرض تعدادی متغیر حالت به عنوان کنترل مجازی است، سپس قانون‌های کنترلی پایدارساز برای ورودی‌های کنترلی مجازی طراحی می‌شود. در آخرین مرحله ورودی کنترل واقعی برای پایدارسازی کل سیستم استفاده می‌شود. با استفاده از روش گام به عقب می‌توان با ترکیب قانون کنترل، سیستم را وادار به ردیابی مسیر مطلوب کرد. در اولین گام از طراحی با روش گام به عقب برای سیستم چهارپره خطای ردیابی زاویه رول اینگونه در نظر گرفته می‌شود:

$$z_1 = x_{1d} - x_1 \quad (6)$$

می‌توان با استفاده از قضیه لیاپانوف، تابع لیاپانوف معین مثبت را به صورت زیر تعریف کرد:

$$V(z_1) = \frac{1}{2} z_1^2 \quad (7)$$

که مشتق این تابع لیاپانوف به صورت رابطه (8) بدست می‌آید:

$$\dot{V}(z_1) = z_1(\dot{x}_{1d} - \dot{x}_2) \quad (8)$$

حال می‌توان پایداری z_1 را با تعریف ورودی کنترل مجازی x_2 بدست آورد.

$$x_2 = \dot{x}_{1d} + \alpha_1 z_1 \quad \text{with: } \alpha_1 > 0 \quad (9)$$

در نتیجه مشتق تابع لیاپانوف نیمه معین منفی خواهد بود:

$$\dot{V}(z_1) = -\alpha_1 z_1^2 \quad (10)$$

سپس با استفاده از تغییر متغیر زیر:

$$z_2 = x_2 - \dot{x}_{1d} - \alpha_1 z_1 \quad (11)$$

در دومین گام تابع لیاپانوف به صورت زیر تکمیل می‌گردد:

$$V(z_1, z_2) = \frac{1}{2} (z_1^2 + z_2^2) \quad (12)$$

و مشتق زمانی آن به صورت زیر بدست می‌آید:

به طوریکه S یک ماتریس پاد متقارن، R ماتریس دوران در چهارچوب اینرسی و $\xi = (x, y, z)$ موقعیت مرکز جرم پرنده در چهارچوب اینرسی E را مشخص می‌کند، همچنین τ_a بردار کنترلی اعمالی به چهارپره است. حرکات اساسی بر مبنای ورودی‌های کنترلی است. رابطه بین گشتاور ایجاد شده توسط پره‌ها و ورودی‌های کنترلی U به صورت رابطه (3) است:

$$\tau_a = \begin{bmatrix} (f_2 - f_4)l \\ (f_3 - f_1)l \\ \sum_{i=1}^4 \tau_{M_i} \end{bmatrix} = \begin{bmatrix} lb(\Omega_2^2 - \Omega_4^2) \\ lb(\Omega_3^2 - \Omega_1^2) \\ k_\tau(\Omega_1^2 - \Omega_2^2 + \Omega_3^2 - \Omega_4^2) \end{bmatrix} = \begin{bmatrix} l.U_2 \\ l.U_3 \\ U_4 \end{bmatrix} \quad (3)$$

f_i نیروی تولید شده توسط موتورهای M_i و Ω_i ها سرعت زاویه‌ای پره‌ها هستند. می‌توان متغیرهای حالت را به صورت زیر در نظر گرفت:

$$X = [\phi \ \dot{\phi} \ \theta \ \dot{\theta} \ \psi \ \dot{\psi} \ z \ \dot{z}]^T$$

در رابطه بالا، (ψ, θ, ϕ) ، زوایای اوپلر هستند که جهت‌گیری چهارپره را نشان می‌دهند. ψ زاویه یار حول محور z، θ زاویه پیچ حول محور y و ϕ زاویه رول حول محور x است. رابطه ورودی‌های کنترلی و سرعت زاویه‌ای پره‌ها به صورت رابطه (4) می‌باشند:

$$\begin{cases} U_1 = b(\Omega_1^2 + \Omega_2^2 + \Omega_3^2 + \Omega_4^2) \\ U_2 = bl(-\Omega_2^2 + \Omega_4^2) \\ U_3 = bl(\Omega_1^2 - \Omega_3^2) \\ U_4 = d(-\Omega_1^2 + \Omega_2^2 - \Omega_3^2 + \Omega_4^2) \end{cases} \quad (4)$$

می‌توان ماتریس انتقال بین سرعت زاویه‌ای در چهارچوب بدنه و چهارچوب زمین را به صورت ماتریس همانی در نظر گرفت، بنابراین معادلات به صورت رابطه (5) ساده می‌شوند:

$$\dot{X} = f(X, U) = \begin{pmatrix} x_2 \\ a_1 x_4 x_6 + a_2 x_4 \Omega + b_1 U_2 \\ x_4 \\ a_3 x_2 x_6 + a_4 x_2 \Omega + b_2 U_3 \\ x_6 \\ x_4 x_2 a_5 + b_3 U_4 \\ x_8 \\ -g + (\cos x_1 \cos x_3) \frac{1}{m} U_1 \end{pmatrix} \quad (5)$$

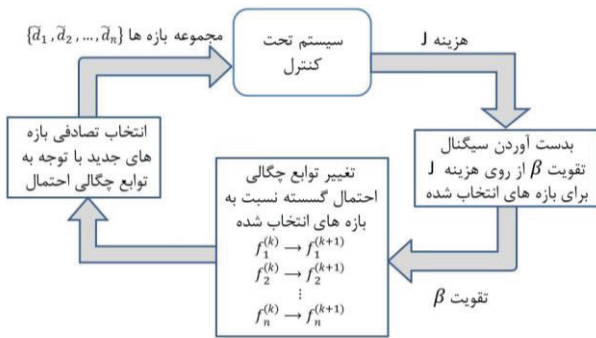
به طوری‌که:

$$\begin{aligned} a_1 &= (I_{yy} - I_{zz})/I_{xx} & b_1 &= l/I_{xx} \\ a_2 &= -J_R/I_{xx} & b_2 &= l/I_{yy} \\ a_3 &= (I_{zz} - I_{xx})/I_{yy} & b_3 &= l/I_{zz} \\ a_4 &= J_R/I_y & & \\ a_5 &= (I_{xx} - I_{yy})/I_{zz} & & \end{aligned}$$

مشاهده می‌شود که معادلات غیرخطی (5) شامل اثرات ژيروسکوپی بدنه و پره‌ها است.

کنترل گام به عقب

کنترل گام به عقب یک روش توسعه یافته در تئوری کنترل است که توسط کتوتویک^{۱۲} و همکارانش [۱۴] در دهه ۹۰ به



شکل ۲. الگوی روش DARLA

$$f_i^{(0)}(d) = \begin{cases} \frac{1}{N_i} & d = 1, 2, \dots, N_i \\ 0 & \text{other} \end{cases} \quad (18)$$

که N_i تعداد متغیرهای تصمیم λ و d اندیس بازه است. سپس در هر مرحله مجموعه جدیدی از متغیرهای تصمیم با توجه به توابع چگالی احتمال انتخاب شده و به سیستم کنترل اعمال می شود. این انتخاب با استفاده از احتمال تجمعی هر یک از متغیرهای تصمیم طبق رابطه (۱۹)، انجام می شود [۱۵]:

$$C_i^{(k)}(x) = \sum_{q=1}^x f_i^{(q)} \quad , d = 1, 2, \dots, N_i \quad (19)$$

که $C_i^{(k)}(x)$ احتمال تجمعی x از متغیر تصمیم λ ، در تکرار k ام است. پس از محاسبه احتمال تجمعی، توسط عدد تصادفی $r \in (0, 1]$ ، متغیر انتخاب می شود. بدیهی است هر مقداری که بیشترین مقدار تابع چگالی احتمال را به خود نسبت دهد، شانس بیشتری برای انتخاب شدن دارد. پس از انتخاب متغیرهای تصمیم در مرحله قبل، این مجموعه متغیرها به سیستم کنترل اعمال می شود. پس از مشخص شدن خروجی‌ها تابع هزینه محاسبه می شود. ساختار DARLA به گونه‌ای است که هدف می نیم کردن هزینه می باشد. پس از محاسبه اسکالر هزینه، این هزینه با هزینه‌های تکرارهای قبلی مقایسه شده و سیگنال تقویت $\beta(J)$ محاسبه می شود. سیگنال تقویت معیاری از عملکرد مناسب مجموعه متغیرهای انتخاب شده است؛ یعنی هر چه مجموعه انتخاب شده مناسب تر باشد مقدار این سیگنال بیشتر است. تابعی که برای محاسبه سیگنال تقویت از روی تابع هزینه استفاده می شود مطابق رابطه (۲۰) است:

$$\beta(J) = \min \left\{ 1, \max \left\{ 0, \frac{J_{mean} - J}{J_{mean} - J_{min}} \right\} \right\} \quad (20)$$

در این رابطه، J_{min} و J_{mean} به ترتیب میانگین و مینیمم هزینه‌های قبلی هستند. محدوده تغییرات β بین صفر و یک است و هزینه‌هایی که بالاتر از میانگین هزینه‌های قبلی باشد منجر به تقویت صفر و هزینه‌هایی که کمتر از مینیمم

$$\dot{V}(z_1 z_2) = z_2(a_1 x_4 x_6 + a_2 x_4 + b_1 U_2) - z_2(\dot{x}_{1d} - \alpha_1(z_2 + \alpha_1 z_1)) - z_1 z_2 - \alpha_1 z_1^2 \quad (13)$$

ورودی کنترل واقعی برای کنترل زاویه رول با فرض $(\dot{x}_{1,2,3d} = 0)$ و اینکه $\dot{V}(z_1 z_2) < 0$ را ارضا نماید به صورت زیر است:

$$U_2 = \frac{1}{b_1} (z_1 - a_1 x_4 x_6 - a_2 x_4 \Omega - \alpha_1(z_2 + \alpha_1 z_1) - \alpha_2 z_2) \quad (14)$$

قسمت $\alpha_2 z_2$ با $\alpha_2 > 0$ به منظور پایداری z_1 افزوده شده است. به طور مشابه ورودی‌های U_3 و U_4 با تغییر متغیرهای زیر بدست می آیند:

$$\begin{cases} z_3 = x_{3d} - x_3 \\ z_4 = x_4 - \dot{x}_{3d} - \alpha_3 z_3 \\ z_5 = x_{5d} - x_5 \\ z_6 = x_6 - \dot{x}_{5d} - \alpha_5 z_5 \end{cases} \quad (15)$$

$$U_3 = \frac{1}{b_2} (z_3 - a_3 x_2 x_6 - a_4 x_2 \Omega - \alpha_3(z_4 + \alpha_3 z_3) - \alpha_4 z_4) \quad (16)$$

$$U_4 = \frac{1}{b_3} (z_5 - a_5 x_2 x_4 - \alpha_5(z_6 + \alpha_5 z_5) - \alpha_6 z_6) \quad (17)$$

روش اتوماتای یادگیری تقویتی گسسته

روش اتوماتای یادگیری تقویتی^{۱۳} (RLA) اولین بار توسط هوول^{۱۴} و همکارانش در سال ۱۹۹۷ ارائه شد. این روش در واقع یکی از کلاس‌های یادگیری تقویتی است که از توابع چگالی احتمال برای یافتن مقدار بهینه متغیرهای تصمیم مساله استفاده می نماید. مبنای این روش همانند سایر روش های یادگیری تقویتی بر اساس اندرکنش با محیط بنا شده است و مکانیزم آن به طور ساده چنین است که: اگر مجموعه‌ای از متغیرهای تصمیم انتخاب شده عملکرد خوبی از سیستم را نتیجه دهند، باید احتمال انتخاب آنها افزایش یابد. الگوریتم این روش به این صورت است: ۱- انتخاب مجموعه‌ای از متغیرهای تصمیم به صورت تصادفی و با توجه احتمال تجمعی. ۲- اعمال این مجموعه متغیرها به مدل. ۳- مشاهده خروجی‌ها و محاسبه تابع هزینه. ۴- محاسبه سیگنال تقویت با توجه به تابع هزینه. ۵- تغییر دادن توابع چگالی احتمال با استفاده از سیگنال تقویت و بازگشت به مرحله ۱ [۱۷]. الگوی روش اتوماتای یادگیری تقویتی گسسته^{۱۵} در شکل (۲) رسم شده است.

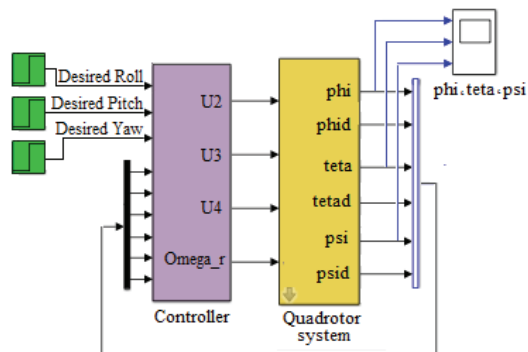
در این روش ابتدا مقادیر هر یک از متغیرهای تصمیم به بازه‌های ترجیحاً (و نه لزوماً) با طول مساوی تقسیم می شود. چنانچه سیستم تحت کنترل دارای n متغیر تصمیم باشد، n تابع چگالی احتمال گسسته وجود خواهد داشت. توابع چگالی اولیه f به صورت رابطه (۱۸) تعریف می شوند:

^{۱۳} Reinforcement Learning Automata (RLA)

^{۱۴} Howell

^{۱۵} (DARLA) Action Reinforcement Learning Discrete Automata

است. کنترل‌کننده طراحی شده روی زیر سیستم وضعیت اعمال شده است که مدل شبیه سازی سیستم مذکور و پارامترهای آن در شکل ۳ و جدول ۱ ارائه شده است.



شکل ۳. بلوک دیاگرام سیستم

پارامترها	مقادیر	واحد
I_{XX}	6.225e-3	$Kg.m^2$
I_{YY}	6.225e-3	$Kg.m^2$
I_{ZZ}	1.121e-2	$Kg.m^2$
J_r	6e-5	$Kg.m^2$
g	9.81	m/s^2
l	0.23	m
m	0.65	kg
d	7.5e-7	Nms^2

جدول ۱. پارامترهای سیستم چهارپره

به طوریکه I_{XX} ممان اینرسی حول محور x ، m جرم پرنده، d ضریب دراگ، l طول بازو و J_r نیز اینرسی موتور است. موتور مورد استفاده در چهارپره "OS4" از نوع بدون جاروبک^{۱۸} با توان مصرفی ۳۵ وات است و مدل آن LRK195.03 می‌باشد. البته نوع گیربکس و پرها نیز بسیار مهم هستند. بیشترین تراست ممکن ایجاد شده ۲/۴۳N می‌باشد، و ورودی‌های کنترلی به مقادیر زیر محدود می‌شوند:

$$-1.2 \leq U_{2,3} \leq 1.2 N$$

$$-0.05 \leq U_4 \leq 0.05 Nm$$

هدف، بدست آوردن ضرایب $(\alpha_1, \dots, \alpha_6)$ به گونه‌ای است که خروجی زوایا به صورت بهینه پایدار گردد. فرض بر این است که هر ضریب بتواند مقداری بین ۰ تا ۱۰۰ را اختیار نماید. در نظر گرفتن این بازه مثبت، پایداری سیستم به مفهوم لیاپانوف را تضمین می‌کند. بنابراین مقادیر بدست آمده از این روش بهینه ساز خللی در پایداری سیستم وارد نمی‌کند.

در ابتدا این محدوده، به ۲۰۰ بازه با طول ۰/۵ تقسیم می‌شود. در روش DARLA مرکز هر یک از این بازه ها بعنوان نماینده ضرایب در نظر گرفته می‌شود. تابع هزینه برای این بهینه‌سازی

هزینه‌های قبلی باشد تقویت یک را در بر خواهد داشت. با افزایش تکرارها همواره میانگین هزینه‌ها روندی نزولی دارد و با توجه به محدود و خطی بودن رابطه تقویت و هزینه، همگرایی روش تضمین می‌شود. پس از محاسبه سیگنال تقویت، توابع چگالی احتمال متناسب با آن و با مرکزیت متغیر انتخاب شده تغییر می‌یابد. اما برای تغییر در تابع چگالی احتمال از تابع نمایی معکوس (۲۱) استفاده می‌شود:

$$Q(d, r) = \lambda * 2^{-(d-r)^2} \quad (21)$$

که پارامتر λ یک عدد کاملاً مثبت است و در سرعت و دقت همگرایی روش موثر می‌باشد. توابع چگالی احتمال توسط رابطه زیر تغییر می‌یابند:

$$f_i^{(k+1)}(d) = \alpha_i^{(k)} \left(f_i^{(k)} + \beta(k)Q(d, \bar{d}_i) \right) \quad (22)$$

$$d = 1, 2, \dots, N_i; i = 1, 2, \dots, n$$

در این رابطه α به منظور نرمالیزه کردن تابع چگالی احتمال بکار رفته است و به صورت رابطه (۲۳) محاسبه می‌شود:

$$\alpha_i^{(k)} = \frac{1}{\sum_{q=1}^{N_i} f_i^{(k)} + \beta(k)Q(d, \bar{d}_i)} \quad (23)$$

پس از تغییر توابع چگالی احتمال، فرایندهای گذشته دوباره تکرار می‌شود. با تکرار این حلقه به تعداد کافی، توابع چگالی احتمال در نقاط بهینه ماکزیمم می‌شوند. بنابراین در تکرار انتهایی، در جایی که این توابع بیشترین مقدار خود را دارند، آن نقاط همان نقاط بهینه می‌باشند. با توجه به اینکه ساختار DARLA یک ساختار تکرار است، بنابراین برای خاتمه عملیات باید ضابطه‌ای را تعریف نمود. در این مقاله "تکرار الگوریتم به تعداد مشخص" بعنوان ضابطه پایان به کار رفته است.

یکی از نقاط قوت بسیار سودمند در یادگیری تقویتی توانایی کنترل فرآیندهایی است که دارای نامعینی^{۱۶} هستند. این نامعینی می‌تواند ناشی از دینامیک‌های مدل نشده^{۱۷} یا پارامترهای متغیر در ساختار سیستم باشد که بصورت بدون تغییر مدل شده‌اند. این توانایی روش یادگیری تقویتی از آنجا نشأت می‌گیرد که در طی فرآیند یادگیری به دینامیک‌های سیستم تحت کنترل اتکالی نشده است.

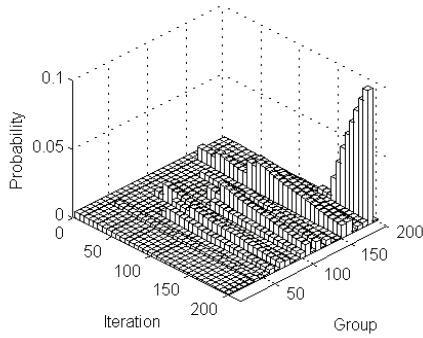
تعیین بهینه ضرایب کنترل پس گام با روش DARLA

در ابتدا با تعریف تمامی محدوده‌های ممکن برای مقادیر پارامترها و تقسیم این محدوده به بازه‌های کوچکتر با استفاده از روش DARLA در یک الگوریتم تکرار مناسب‌ترین بازه برای هر متغیر بدست می‌آید.

مدل مورد استفاده در این مقاله یک سیستم چهارپره واقعی با نام "OS4" می‌باشد که توسط بو عبدالله مورد مطالعه قرار گرفته

^{۱۸} brushless

^{۱۶} Uncertainty
^{۱۷} Un-modeled dynamics

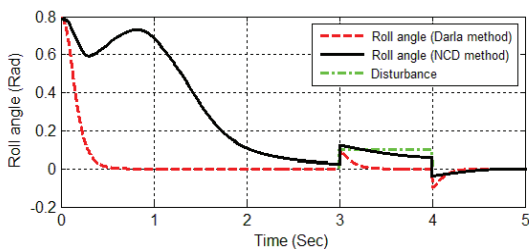


شکل ۵. روند تغییرات تابع چگالی احتمال گسسته α_2

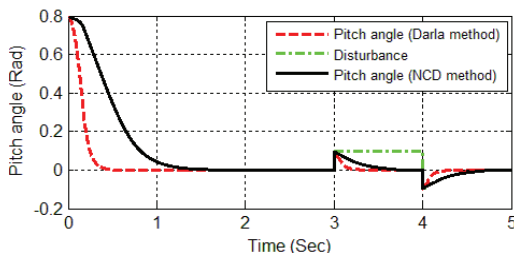
مشاهده می‌شود بازه‌ی ۱۸۹ در نمودار بالا بیشترین احتمال را کسب کرده است. یعنی مقدار بهینه α_2 بین اعداد ۹۴/۵ و ۹۴ قرار دارد. برای تعیین جواب دقیق‌تر می‌توان DARLA را مجدداً اجرا و بهترین مقادیر را بدست آورد.

پس از یافتن ضرایب بهینه، این ضرایب را در مدل جایگزین کرده و سپس شبیه‌سازی صورت می‌گیرد، در این شبیه‌سازی هدف پایدارسازی زوایای چهارپره می‌باشد به گونه‌ای که شرایط اولیه زوایا ۴۵ درجه در نظر گرفته شده است. جهت بررسی نیز نتایج این روش بهینه‌سازی با نتایج بدست آمده از بهینه‌سازی با استفاده از جعبه ابزار بهینه‌ساز MATLAB^{۱۹} مقایسه می‌گردد [۱۱].

همچنین به منظور بررسی دقیق‌تر در بازه زمانی ۳ تا ۴ ثانیه، اغتشاش پله به سیستم وارد شده است. همانگونه که مشاهده می‌شود، سیستم پس از این بازه زمانی به سرعت پایدار می‌شود.



شکل ۶. پایدارسازی زاویه حول محور x با شرایط اولیه ۴۵ درجه



شکل ۷. پایدارسازی زاویه حول محور y با شرایط اولیه ۴۵ درجه

به صورت رابطه (۲۳) در نظر گرفته شده است که در این تابع $e(t)$ سیگنال خطا و $u(t)$ سیگنال کنترل است:

$$J = G_e \int_0^{t_f} e^2(t) dt + G_u \int_0^{t_f} u_c^2(t) dt + G_M M_p + G_S E_{SS} + G_d \sup_t \left| \frac{de(t)}{dt} \right| \quad (23)$$

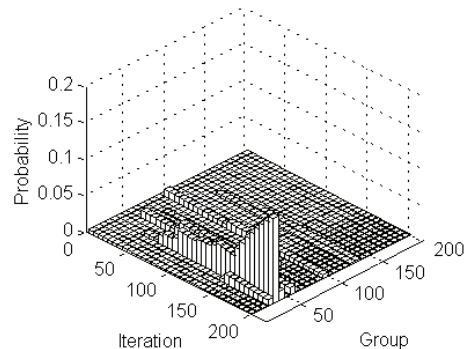
در این رابطه M_p فراجهش خروجی، E_{SS} خطای حالت دائمی زاویه و ضرایب G وزنه‌های المانهای هزینه می‌باشند. t_f زمان کل شبیه‌سازی است که باید به اندازه کافی بزرگ انتخاب شود. در این مدل $t_f = 5^{sec}$ مدت زمان مناسبی است.

هر یک از المانهای بکار رفته در تابع هزینه تاثیرات خاص خود را دارد؛ بخش اول تابع هزینه باعث کاهش خطای ردیابی در کمترین زمان می‌شود، انتگرال دوم انرژی کنترل را محدود می‌کند، M_p و E_{SS} به ترتیب فراجهش و خطای حالت دائمی را کاهش می‌دهد، و در نهایت $\sup_t \left| \frac{de(t)}{dt} \right|$ موجب ردیابی هموار ورودی مرجع می‌گردد. وزن های بکار رفته در این تابع هزینه به صورت

$$G_d = 340, G_u = 6e - 7, G_M = 3, G_e = 1600, G_S = 1e + 4$$

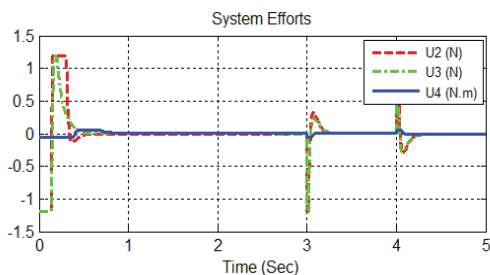
انتخاب شده‌اند. باید دقت کرد که اندازه وزنها نرمالیزه نیست و از این رو کوچکی و بزرگی یک وزن به تنهایی ارزش آن را معین نمی‌کند.

نمونه‌ای از طراحی پارامترهای کنترل توسط روش DARLA در شکل های (۵و۴) رسم شده است. در این شکل نحوه بدست آمدن α_1 پس از ۲۰۰ مرتبه اجرای الگوریتم DARLA دیده می‌شود. همانطور که مشهود است بعد از پایان بهینه‌سازی DARLA، بازه شماره ۲۸ بیشترین احتمال انتخاب شدن را دارد یعنی مقدار مناسب برای α_1 بین ۱۴ تا ۱۳/۵ می‌باشد.

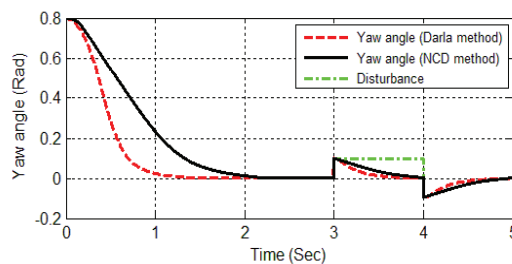


شکل ۴. روند تغییرات تابع چگالی احتمال گسسته α_1

و نیز برای α_2 مطابق شکل ۵ داریم:



شکل ۹. سیگنال‌های کنترلی ورودی



شکل ۸. پایدارسازی زاویه حول محور z با شرایط اولیه ۴۵ درجه

زمان خیز (ثانیه)			زمان نشست (ثانیه)			
زاویه رول	زاویه پیچ	زاویه یاو	زاویه رول	زاویه پیچ	زاویه یاو	
۰/۲۴	۰/۲۰	۰/۵۹	۰/۳۶	۰/۳۰	۰/۸۹	کنترل کننده گام به عقب بهینه شده با روش DARLA
۲/۰۹	۰/۶۸	۱/۲۰	۲/۶۰	۱/۰۴	۱/۶۵	کنترل کننده گام به عقب بهینه شده با جعبه ابزار بهینه ساز MATLAB

جدول ۲. بررسی و مقایسه عملکرد سیستم در روش بهینه سازی پیشنهادی

کنترلی گام به عقب بهینه‌سازی شده توسط DARLA پایداری بهتری در حضور اغتشاش نسبتاً زیاد را دارد. نوآوری استفاده شده در این مقاله، استفاده از روش بهینه سازی اتوماتای یادگیری تقویتی است که توانسته است مقدار زیادی عملکرد سیستم را بهبود بخشد. از سوی دیگر هدف از ردیابی و دفع اغتشاش کنترل کننده، هدایت و ایمنی حرکت چهارپره است. همانگونه که نشان داده شد، اعمال سریع و مناسب ورودی‌های کنترلی موجب ایجاد نتایج قابل قبول تری گردید. البته روش-های کنترلی متفاوتی را می‌توان در هدایت چهارپره به کار گرفت اما آنچه حائز اهمیت است مصالحه میان بار پردازشی و کارایی این روش‌ها است.

مراجع

[1] E. Altuğ, J. P. Ostrowski, and C. J. Taylor, "Control of a quadrotor helicopter using dual camera visual feedback". The International Journal of Robotics Research, 24(5), pp 329-341, 2005.

[2] S. Bouabdallah, A. Noth, and R. Siegwart. "PID vs LQ control techniques applied to an indoor micro quadrotor." Proceedings International Conference on Intelligent Robots and Systems, IROS, IEEE/RSJ. Vol. 3, pp. 2451-2456, 2004.

[3] C. Raimúndez, and J. L. Camaño. "Tracking in scale quad-rotors through adaptive augmentation." Neural Computing and Applications 23, no. 3-4, pp.635-643, 2013.

[4] L. C. Lai, C.C. Yang, and C. J. Wu, "Time-optimal control of a hovering quad-rotor

بحث و بررسی

شکل‌های ۶ تا ۸ خروجی زوایای پرنده با بهینه سازی توسط روش‌های اتوماتای تقویتی و جعبه‌ابزار MATLAB را نشان می‌دهند. با توجه به خروجی‌ها، برتری این روش بهینه‌سازی نمایان است. همچنین نتایج عملکرد سیستم در جدول (۲) ارائه شده است.

البته توجه به مقادیر ورودی کنترلی نیز حائز اهمیت است بنابراین به همین دلیل در شکل (۹) مقادیر ورودی کنترلی ارائه شده است. همان‌گونه که مشاهده می‌شود ورودی‌های کنترلی در بازه‌های مجاز قرار گرفته‌اند که این کار با وارد کردن یک بلوک اشباع کننده در مسیر ورودی‌ها صورت گرفته و سپس بهینه سازی انجام شده است. همان‌گونه که نتایج شبیه سازی نشان می‌دهند پایداری سیستم در حضور عنصر غیر-خطی سخت یعنی اشباع‌گر حاصل می‌شود. همچنین باید توجه نمود که بهینه سازی در حضور عنصر اشباع‌گر و به صورت آفلاین صورت گرفته است.

نتیجه گیری

در این مقاله با استفاده از روش اتوماتای یادگیری تقویتی، پارامترهای سیستم کنترلی چهارپره بطور بهینه تعیین گردید. روش‌های کنترلی کلاسیک موفقیت نسبی در پایداری زوایا در شرایط شناور و در حضور اغتشاش کم دارند اما برای سیستم‌های با دینامیک غیرخطی شدید و وابستگی شدید بین متغیرها مناسب نیستند. همچنین دیده می‌شود که روش

- In Robotics and Automation. ICRA, IEEE, pp. 2247-2252, April 2005.
- [12] S. Bouabdallah, "Design and control of quadrotors with application to autonomous flying". Ecole Polytechnique Federale de Lausanne, February 2007.
- [13] G. V. Raffo, "Modelado y Control de un Helicóptero QuadRotor", Doctoral dissertation, Master's thesis, Universidad de Sevilla, Dept. Ing. Sistemas y Automática, Sevilla, Spain, December 2007.
- [14] R. Sepulchre, M. Jankovic, P.V. Kokotovic, "constructive nonlinear control", Springer, August 1996.
- [15] M. Kashki, A. A. Gharaveisi, F. Kharaman, "Application of CDCARLA Technique in Designing Takagi-Sugeno Fuzzy Logic Power System Stabilizer (PSS)", In Power and Energy Conference, IEEE International. PECon'06, pp. 280-285, 2006.
- [۱۶] مردانی رضا، "بررسی امکان افزایش چابکی و کنترل پذیری یک چهارپره بدون سرنشین از طریق افزایش پارامترهای کنترلی"، پایان‌نامه کارشناسی ارشد دانشگاه صنعتی شریف، دی ماه ۱۳۹۰
- [۱۷] حیدری غلامعلی، قره ویسی علی اکبر، رشیدی نژاد مسعود، "طراحی بهینه کنترل PI در سیستم کنترل سرعت موتور القایی با روش هوشمند اتوماتای یادگیری تقویتی ترکیبی"، اولین کنگره مشترک سیستم‌های فازی و سیستم‌های هوشمند، دانشگاه فردوسی مشهد، شهریور ۱۳۸۶
- helicopter". Journal of Intelligent and Robotic Systems, 45(2), pp. 115-135, 2006.
- [5] M. A. Olivares-Mendez, P. Campoy, I. Mellado-Bataller, and L. Mejias. "See-and-avoid quadcopter using fuzzy control optimized by cross-entropy". International Conference In Fuzzy Systems, IEEE, pp. 1-7, June 2012.
- [6] Z. Fang, and W. Gao, "Adaptive integral backstepping control of a Micro-Quadrotor". 2nd International Conference In Intelligent Control and Information Processing (ICICIP), Vol. 2, pp. 910-915, July 2011.
- [7] S.L. Waslander, G. M. Hoffmann, J. S. Jang, and C. J. Tomlin, C. J. "Multi-agent quadrotor testbed control design: Integral sliding mode vs. reinforcement learning". International Conference In Intelligent Robots and Systems, IEEE/RSJ, pp. 3712-3717, August 2005.
- [8] C. Coza, and, C. J. B. Macnab, "A new robust adaptive-fuzzy control method applied to quadrotor helicopter stabilization". In Fuzzy Information Processing Society, NAFIPS, Annual meeting of the North American, IEEE, pp. 454-458, June 2006.
- [9] C. Nicol, C. J. B. Macnab, and A. Ramirez-Serrano, "Robust adaptive control of a quadrotor helicopter". Mechatronics, 21(6), pp. 927-938, 2011.
- [10] A. Nagaty, S. Saeedi, C. Thibault, M. Seto, and H. Li, "Control and Navigation Framework for Quadrotor Helicopters", Journal of Intelligent & Robotic Systems 70 ,no.1-4, 2013
- [11] S. Bouabdallah, and R. Siegwart, "Backstepping and sliding-mode techniques applied to an indoor micro quadrotor", Proceedings of the International Conference