



طراحی کنترل کننده غیرخطی پهباد چهارروتور به کمک روش ترکیبی گرادیان ازدحام ذرات

حامد شهبازی*، وحید تیکنی

دانشکده مهندسی مکترونیک، دانشگاه اصفهان، اصفهان، ایران

تاریخچه داوری:

دریافت: ۳ مرداد ۱۳۹۵
بازنگری: ۴ آبان ۱۳۹۵
پذیرش: ۷ آذر ۱۳۹۵
ارائه آنلاین: ۹ آذر ۱۳۹۵

کلمات کلیدی:

چهارروتور
کنترل بهینه
گرادیان سیاست
بهینه‌سازی ازدحام ذرات

چکیده: در این مقاله با ترکیب ایده‌هایی از یادگیری تقویتی گرادیان سیاست و روش ازدحام ذرات یک روش ترکیبی بهینه‌سازی برای کنترل یک سامانه پیچیده غیرخطی ارائه شده است که کاربردهای فراوانی در جهان واقعی خواهد داشت. این سامانه ترکیبی بر روی یک پرند هدایت‌پذیر از دور چهارروتور نصب شده است که با هدف کنترل جهت‌گیری و موقعیت پهباد عمل می‌کند. در این روش با گرفتن ایده از روش‌های تقویتی، گرادیان سیاست در کنترل کننده مشتق‌گیر تناسبی یک چهارروتور محاسبه می‌شود و در روابط بهینه‌سازی ازدحام ذرات وارد می‌گردد تا بهینه‌سازی علاوه بر فاکتورهای لحاظ شده در روش‌های ازدحامی در جهت گرادیان سیاست کنترلی نیز انجام شود. برای انجام بهینه‌سازی ورودی‌های کنترلی و مشخصه‌های سامانه از جمله: زمان پاسخ‌دهی سامانه، خطای ماندگار سامانه، فراجش و زمان نشست سامانه در تابع هزینه برای بهینه‌سازی منظور شده‌اند. روش ارائه شده بر روی بستر عملی آزمون چهارروتور پیاده‌سازی و با تعدادی از روش‌های مرسوم مقایسه شده است.

۱- مقدمه

کنترل بهینه یک سامانه غیرخطی از جمله مسائل بسیار پیچیده در حوزه مهندسی کنترل به شمار می‌رود که توجه بسیاری از محققین حوزه کنترل را به خود معطوف ساخته است. از آنجایی که بسیاری از دستگاه‌های کنترل جهان واقعی غیرخطی هستند و کنترل آن‌ها با صرف کمترین هزینه و زمان یک نیاز واقعی است، اهمیت این حوزه هرروز در حال افزایش می‌باشد. در این مقاله با ترکیب ایده‌هایی از یادگیری تقویتی گرادیان سیاست و روش ازدحام ذرات یک روش ترکیبی برای کنترل یک سامانه پیچیده غیرخطی ارائه می‌دهیم که کاربردهای فراوانی در کاربردهای واقعی خواهد داشت. یادگیری تقویتی روشی از یادگیری ماشین جهت حل کردن مسائل تصمیم‌گیری پشت سرهم با استفاده از نوعی سازوکار تنبیه و تشویق می‌باشد. یادگیری تقویتی می‌تواند به دو صورت مستقیم و یا غیرمستقیم به مسئله اعمال شود. در دسته یادگیری مستقیم، یادگیری در ابتدا سعی می‌کند تا یک تخمینی از تابع مقدار را به دست آورد و سپس براساس آن سیاست بهینه را استخراج نماید. در روش‌های مستقیم، سعی یادگیر بر آن است که در فضای سیاست به طور مستقیم به جستجوی بهترین سیاست ممکن بپردازد؛ بنابراین دیگر نیازی به استخراج تابع مقدار نیست. این دسته از یادگیری معمولاً به دو شیوه گرادیان سیاست و تکاملی (یا تکاملی-عصبی) تقسیم می‌شود. در فرآیندهای تصمیم‌گیری پشت سر هم، یک عامل همواره سعی

می‌کند با استفاده از ارتباط با محیط پیرامون خویش رفتاری را که توسط عملگرهای خود انجام می‌دهد، بهینه نماید. عامل با استفاده از حسگرهای خود سعی می‌کند تا از محیط و حالت فعلی خودش درون محیط، مدلی را استخراج نماید. در مسأله یادگیری تقویتی محیط به عامل در ازای انجام هر عملی پاداشی را اعطا می‌نماید و درنهایت این وظیفه عامل است که سری اعمال خود در محیط را برای بیشینه کردن میزان پاداش انباشتی، بهینه نماید؛ بنابراین به‌طور کلی می‌توان یادگیری تقویتی را نوعی از یادگیری دانست که در آن عامل با استفاده از سعی و خطا و دریافت پاداش به‌صورت مثبت و یا منفی از طریق محاوره و تعامل با محیط، سودمندترین روش برای رسیدن به هدف خود را می‌آموزد. این نوع از یادگیری علاوه بر کاربرد در محیط‌های گسسته دارای کاربردهای دیگری در زمینه‌های پیوسته نظیر مسأله کنترل پاندول معکوس، مسأله گاری-میله می‌باشد. روش‌های یادگیری تقویتی در کل به دو مجموعه روش‌های یادگیری تابع مقدار و یادگیری تابع سیاست تقسیم می‌شوند.

روش‌های مبتنی بر فضای سیاست، سیاست را به‌صورت یک تابع تخمینی با یک سری پارامتر تعریف می‌کنند و معمولاً از روش‌های گرادیانی برای یافتن و تنظیم این پارامترها استفاده می‌شود. اگرچه روش‌های مبتنی بر گرادیان می‌توانند هزینه‌بر باشند، چنین روش‌هایی در چند سال اخیر بسیار مورد توجه بوده‌اند و همچنان در کانون توجه قرار دارند. در کنار روش‌های گرادیانی، روش‌های بهینه‌سازی دیگری نظیر سرد شدن

شبکه‌سازی شده استفاده می‌شوند. همچنین روش‌های بهینه‌سازی دیگری نظیر روش‌های مبتنی بر محاسبات تکاملی نیز مورد استفاده قرار گرفته‌اند که منجر به زمینه جدید رباتیک تکاملی گردیده است. در این مقاله بر روی کنترل بهینه ربات پرنده چهارروتور تمرکز شده است.

چهارروتور یا کوادروتور یک وسیله پرنده با شش درجه آزادی حرکت می‌باشد که قابلیت پرواز عمودی و انجام مانورهای پیچیده را داراست. این سازه دارای ساختار صلیب‌شکل است که چندین ملخ در گوشه‌های آن قرار دارد و با استفاده از تغییر سرعت ملخ‌ها می‌تواند حرکات و مانورهای مختلف را انجام دهد. این وسیله به سبب قابلیت نشست و برخاست عمودی در دسته عمودپروازها قرار می‌گیرد. ظرفیت حمل بار، سادگی ساختاری، قابلیت مانور پذیری بالا، داشتن قیود حرکتی کم و هزینه پایین تعمیر و نگهداری از جمله ویژگی‌ها است که موجب شده این وسیله مورد توجه قرار بگیرد. از دسته مشکلاتی که بر سر راه طراحی چنین پرنده‌هایی وجود دارد، مسأله کنترل پایدار این نوع پرنده‌ها می‌باشد که کار بسیار پیچیده‌ای است. از مشکلاتی که پایداری پرنده را مختل می‌کند می‌توان به تغییر دور ناخوابسته ملخ‌ها، عدم تعادل در سرعت چرخش پره‌ها، عدم تعادل وزنی پرنده و عدم پایداری در هنگام وزش باد اشاره کرد.

طراحی کنترل برای یک پهپاد با تعداد زیادی از پارامترهای طراحی به هم وابسته درگیر است. در مقاله پیش‌رو یک کنترلر مشتق‌گیر-تناسبی در ساخت‌افزار کنترلی آن پیاده‌سازی شده است و بهینه‌سازی پارامترهای کنترلی در آن با استفاده از الگوریتم جدید ارائه شده اتفاق می‌افتد و به‌طور مرتب رفتار سامانه را بهبود می‌دهد. در بخش بعدی به مروری از کارهای گذشته پیرامون طراحی کنترل‌کننده برای پرنده کوادروتور پرداخته شده است. بخش سوم مقاله به تشریح دینامیک کنترلی مسأله اختصاص دارد. بخش چهارم به مدل‌سازی روش بهینه‌سازی ترکیبی ارائه شده اختصاص داده شده است. در بخش پنجم این مقاله نتایج پیاده‌سازی‌ها و آزمایش‌های انجام‌شده روی بستر سخت‌افزاری، مورد بحث قرار گرفته است. در آخرین بخش، تحلیل نهایی مقاله ارائه خواهد شد.

۲- مرور کارهای گذشته

طراحی کنترلر بهینه و بهینه‌سازی عملکرد کنترل‌کننده‌ها یکی از موضوعات مهم تحقیق در حوزه مهندسی کنترل و هوش مصنوعی به حساب می‌آید. تاکنون روش‌های بهینه‌سازی مختلفی برای کنترل‌کننده‌ها ارائه شده است که از جمله آن‌ها می‌توان به روش الگوریتم ژنتیک، ازدحام ذرات و رقابت استعماری اشاره نمود.

در پرنده کوادروتور، عدم قطعیت‌ها در مدل دینامیکی و ساختار باعث پیچیدگی می‌شود و در طول پرواز کیفیت پرواز کوادروتور را با زمان تغییر می‌دهد که این عدم قطعیت جزئی از ذات مدل می‌باشد و مشکلات طراحی کنترل را افزایش می‌دهد. کنترلرهای غیر بهینه نمی‌توانند کارایی مورد نظر را برای مدت طولانی داشته باشند، درحالی‌که کنترلرهای بهینه تا حدودی

می‌توانند مقاوم بودن کنترل و تغییرات خطا را تضمین کنند. روش‌های کنترلی تناسبی [۳-۱] مشتق‌گیر انتگرالی هرکدام به دلایلی مورد توجه است. روش تناسبی-مشتق‌گیر به دلیل خاصیت همگرایی نمایشی‌اش با جبران ترم‌های کوریولیس و ژيروسکوپی و روش تناسبی مشتق‌گیر انتگرالی به دلیل عدم احتیاج به پارامترهای خاص مدل و سادگی کنترلرهای مناسبی محسوب می‌شوند. در راستای پیشرفت در نظریه‌های کنترل هوشمند، کنترلرهای تخصصی فازی، فازی کنترل و کنترلر عصبی ارائه شده است که می‌تواند با ترکیب شدن با کنترلر PID کلاسیک [۷-۴] تأثیرات مثبتی روی کنترلر کلاسیک بگذارد.

روش linear quadratic regulator (LQR) هم به سبب مزیت در ارائه سیگنال ورودی بهینه از بازخورد متغیرها موضوع برخی از کارها قرار گرفته است. مزیت این روش این است که می‌تواند سامانه را در محدوده مجاز طراحی قرار دهد و چون مربوط به نظریه خطی است به راحتی می‌توان کنترل بهینه حلقه بسته را تشکیل داد [۱۰-۸]. تخمین موقعیت زاویه‌ای کوادروتور یکی از چالش‌های طراحی کنترلر در محیط آزمایشگاهی است که بررسی شده است [۱۱]. روش‌های ترکیبی فازی نیز برای پایدارسازی کوادروتور مورد استفاده قرار گرفته است [۱۲].

روش اسلایدینگ مود برای عدم قطعیت‌های غیرخطی پاسخ مناسبی دارد، مزیت‌های این روش از جمله پاسخ سریع، مقاوم در برابر عدم قطعیت‌ها و اغتشاشات خارجی، سازگار با سامانه غیرخطی و ساده بودن پیاده‌سازی می‌باشد. این روش نیز مانند روش‌های دیگر نقطه‌ضعفی دارد که از جمله آن پدیده نوسان در اطراف صفحه لغزش است [۱۶-۱۳].

خطی‌سازی فیدبک، روشی است که سامانه غیرخطی را با استفاده از فیدبک‌های غیرخطی، به سامانه خطی تبدیل می‌کند [۲۱-۱۷]. کنترلر تطبیقی، یک روش کنترلی است که می‌تواند به‌طور اتوماتیک پارامترها را تنظیم و تولید کند و ارتباط کنترلی برای ساختن بهترین حالت کنترلی را فراهم کند [۲۴-۲۲].

اساس روش بک‌استپینگ (پسخوراند)، تجزیه کردن کنترلر به چندلایه می‌باشد. ما برای هر قسمت می‌توانیم با انتخاب یک کنترلر مجازی سامانه را پایدار کنیم. به تدریج کنترلر را اصلاح می‌کنیم تا برای تعقیب یا تنظیم پایدار شود. مزیت این روش این است که این کنترلر نرخ همگرایی سریعی دارد. از معایب این روش می‌توان به مقاوم نبودن آن اشاره کرد؛ اما می‌توان با استفاده از روش‌های دیگر این عیب را جبران کرد [۲۷-۲۵].

در برخی مقالات به روش‌های ترکیبی کنترلی پرداخته شده است. کنترلر عصبی با خطی‌سازی فیدبک، کنترلر PID و اسلایدینگ مود [۲۸]، اسلایدینگ مود و پسخوراند [۲۹] از این دست مقالات می‌باشند.

۳- بررسی دینامیک پهپاد

مطالعه دینامیک سازه کوادروتور به شناخت هرچه بیشتر فیزیک این سازه و رفتار آن کمک می‌کند. حرکت کوادروتور با استفاده از تغییر سرعت چرخش روتورها انجام می‌گیرد.

$$\ddot{y} = (\cos \phi \sin \theta \sin \psi - \sin \phi \cos \psi) \frac{U_1}{m} \quad (ب-۴)$$

$$\ddot{z} = -g + (\cos(\phi) \cos(\theta)) \frac{U_1}{m} \quad (پ-۴)$$

$$\ddot{\phi} = \dot{\theta} \dot{\psi} \left[\frac{I_{yy} - I_{zz}}{I_{xx}} \right] + \frac{J_r}{I_{xx}} \dot{\theta} \Omega_d + \frac{1}{I_{xx}} U_2 \quad (ت-۴)$$

$$\ddot{\theta} = \dot{\phi} \dot{\psi} \left[\frac{I_{zz} - I_{xx}}{I_{yy}} \right] - \frac{J_r}{I_{yy}} \dot{\phi} \Omega_d + \frac{1}{I_{yy}} U_3 \quad (ث-۴)$$

$$\ddot{\psi} = \dot{\theta} \dot{\phi} \left[\frac{I_{xx} - I_{yy}}{I_{zz}} \right] + \frac{1}{I_{yy}} U_4 \quad (ج-۴)$$

که x ، y و z موقعیت مرکز جرم در سازه و ϕ و θ و ψ زوایای اویلر هستند. همچنین m ، I_x ، I_y و I_z جرم و ممان اینرسی کوادروتور است. L فاصله بین روتورها و مرکز جرم است، J_r و Ω_R ممان اینرسی و سرعت زاویه‌ای تیغه‌های پرواز هستند. U_1 ، U_2 ، U_3 و U_4 ورودی‌های کنترلی سامانه و به ترتیب نیروی مجموع، مومنت‌های چرخش حول محور y ، چرخش حول محور x و چرخش حول محور z تولید شده توسط تیغه‌های پروازی می‌باشند.

۳-۲- طراحی کنترلر PD

روش کنترل کلاسیک PD ارائه شده در [۳۲ و ۳۳] معادلات زیر را به منظور کنترل PD یک کوادروتور ارائه می‌کنند:

$$U_1 = \frac{m(g + K_{pz} E_z + K_{dz} (-\dot{z}))}{\cos \phi \cos \theta} \quad (الف-۵)$$

$$U_2 = (K_{pp} E_p + K_{dp} (-\dot{\phi})) \quad (ب-۵)$$

$$U_3 = (K_{pt} E_t + K_{dt} (-\dot{\theta})) \quad (پ-۵)$$

$$U_4 = (K_{ps} E_s + K_{ds} (-\dot{\psi})) \quad (ت-۵)$$

در این معادلات m جرم کلی پرنده K_{pi} و K_{di} ها به ترتیب بهره‌های کنترلی نسی و مشتق‌گیر هستند. در بهینه‌سازی یک کنترلر PD این پارامترهای متغیر بهینه‌سازی می‌شوند تا رفتار کنترل‌کننده بهینه شود.

در این مقاله قصد داریم ضرایب کنترل‌کننده PD را همان‌گونه که در شکل ۲ نشان داده شده، به کمک روش ارائه شده (روش ترکیبی گرادیان ازدحام ذرات) بهینه‌سازی با سایر روش‌های موجود مقایسه کنیم.

با توجه به ساختار ساده و عملکرد مقاوم این نوع از کنترلر، طراحی و پیاده‌سازی آن در صنایع بسیار رواج دارد. معادله کنترلر PD به صورت تابع تبدیل زیر است:

$$U = [U_1 \ U_2 \ U_3 \ U_4] \quad (۳)$$

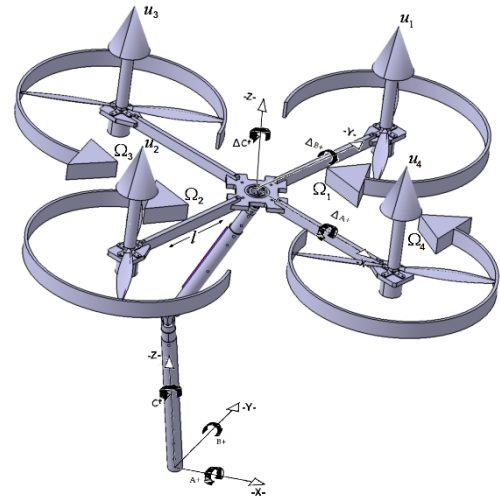


Fig. 1. Quadrotor Schematic
شکل ۱: بیکربندی کوادروتور

بیکربندی کوادروتور در شکل ۱ ملاحظه می‌شود که متشکل از دو بال می‌باشد و هرکدام دارای موتورهایی در انتهای خود هستند. موتور شماره ۱ و ۳ در راستای ساعت‌گرد می‌چرخند و موتور شماره ۲ و ۴ در راستای پادساعت‌گرد گردش می‌کنند.

۳-۱- مدل ریاضی کوادروتور

به منظور به دست آوردن مدل دینامیکی سامانه به صورت مدل فضای حالت متغیرهای حالت را به صورت $\dot{X} = f(X, U)$ منظور می‌کنیم که U بردار ورودی و X بردار متغیرهای حالت است و به صورت زیر تعریف می‌شوند [۳۰]:

$$X = \left[x \ \dot{x} \ y \ \dot{y} \ z \ \dot{z} \ \phi \ \dot{\phi} \ \theta \ \dot{\theta} \ \psi \ \dot{\psi} \right] \quad (۱)$$

$$U = [U_1 \ U_2 \ U_3 \ U_4] \quad (۲)$$

که ورودی‌های کنترلی به شکل زیر تعریف می‌شوند:

$$U_1 = \sum_{i=1}^4 T_i = b(\Omega_1^2 + \Omega_2^2 + \Omega_3^2 + \Omega_4^2) \quad (الف-۳)$$

$$U_2 = (-T_2 + T_4) = b(-\Omega_2^2 + \Omega_4^2) \quad (الف-۳)$$

$$U_3 = (T_1 + T_3) = b(\Omega_1^2 - \Omega_3^2) \quad (الف-۳)$$

$$U_4 = (-1)^i \sum_{i=1}^4 M_{D_i} = d(-\Omega_1^2 + \Omega_2^2 - \Omega_3^2 + \Omega_4^2) \quad (الف-۳)$$

در معادلات بالا b ضریب رانشی و d ضریب پسا است. معادلات حرکتی برگرفته شده از مدل دینامیکی کوادروتور به صورت زیر هستند [۳۰ و ۳۱]:

$$\ddot{x} = (\cos \phi \sin \theta \cos \psi + \sin \phi \sin \psi) \frac{U_1}{m} \quad (الف-۴)$$

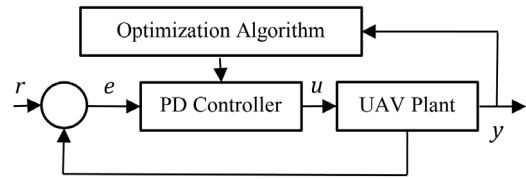


Fig. 2. System block diagram
شکل ۲: طرح کلی کنترلی سامانه

که بهره کنترلی تناسبی و T_d ثابت زمانی مشتق‌گیر است؛ بنابراین مقدار بهره کنترلی مشتق‌گیر $K_d = K_p T_d$ است.

به منظور محاسبه بهره‌های بهینه برای سامانه از الگوریتم‌های بهینه‌سازی استفاده شده است، به این ترتیب که ۸ بهره کنترلی $k_{dz}, k_{pt}, k_{dt}, k_{ps}, k_{ds}, k_{pp}, k_{dp}$ با استفاده از الگوریتم بهینه‌ساز محاسبه و مقداردهی می‌شوند. بدین منظور ورودی‌های کنترلی و انتگرال مرتبه دوم به علاوه مشخصه‌های سامانه از جمله: زمان پاسخ‌دهی سامانه، خطای ماندگار سامانه، فراجش و زمان نشست سامانه در تابع هزینه برای بهینه‌سازی منظور شده‌اند. به این ترتیب تابع هزینه ما شامل تابع هزینه کنترل و مجموع ضرایبی از موارد ذکر شده می‌باشد.

$$J_i = e^{-\beta} (M_{pi} + e_{ssi}) + (1 - e^{-\beta})(t_{si} + t_{ri}) \quad (7)$$

در تابع هزینه t_s زمان نشست سامانه، t_r زمان فراز، e_{ss} خطای ماندگار و M_p بیشینه فراجش پاسخ سامانه است. مقدار β مطابق با خواسته‌های طراح از مسأله قابل تغییر است، برای یکسان در نظر گرفتن تأثیر مشخصه‌های منظور شده در تابع هزینه این مقدار برابر $0/69$ منظور شده است. الگوریتم‌های بهینه‌ساز با تنظیم بهره‌های کنترلی و کاهش تابع هزینه به کنترل بهتر سامانه کمک می‌کنند.

۴- طراحی بهینه‌ساز ترکیبی

در بخش دوم روش‌های اصلی بهینه‌سازی کنترلی معرفی شد و در انتهای آن مشکلات موجود در این روش‌ها بررسی گردید. الگوریتم معرفی شده در این بخش بر مبنای ترکیب و بهبود روش‌های گذشته شکل گرفته‌اند و سعی دارند تا ایرادات اساسی روش‌های قبلی را بر طرف کنند. در اینجا یک روش جدید برای بهینه‌سازی کنترلی که ایده‌های یادگیری گرادین سیاست و ایده تکاملی ازدحام ذرات (جستجوی موازی، تأثیر تکامل ذرات و جهش) را با هم ترکیب کرده است تا راهکار کامل‌تری از تکامل کنترل ارائه نماید. این روش بسیاری از مشکلات موجود در طراحی خودکار سامانه‌های کنترل بهینه را حل نموده است و مزایای روش‌های گذشته را نیز داراست.

در روش ارائه‌شده از ترکیب یادگیری تقویتی گرادین سیاست با الگوریتم تکاملی ازدحام ذرات برای رسیدن به یک بهینه‌ساز بهتر استفاده می‌شود. روش‌های گرادین سیاست، نوعی از تکنیک‌های یادگیری تقویتی هستند که سعی می‌کنند یک سیاست پارامتری شده را براساس امید بازگشت (پاداش) انباشته‌شده در درازمدت با استفاده از گرادین کاهنده، بهینه‌سازی کنند. این

روش‌ها فاقد مشکلات نهفته در روش‌های سنتی یادگیری تقویتی همانند عدم ضمانت وجود یک تابع مقدار مناسب، مشکلات مرتبط با عدم قطعیت در حالات سامانه و پیچیدگی ناشی از حالات و اعمال پیوسته سامانه، می‌باشند. برای حل مسأله یادگیری پارامترهای سامانه کنترلی در ابتدا رفتار دینامیکی سامانه کنترلی را در قالب یک مسأله یادگیری تقویتی (عامل-نقاد) تعریف می‌کنیم. برای این کار لازم است تا مفاهیم تابع سیاست، پارامترها، تابع پاداش-تنبیه، تابع ارزش و گرادین سیاست را تعریف کنیم. در اینجا کل یا بخشی از سامانه کنترلی را می‌توان در قالب یک تابع سیاست نگاهت از حالت و مقدار به عمل خروجی توصیف نمود. سیاست توسط مجموعه پارامترهای پارامتربندی شده است که این پارامترها وزن‌ها و بایاس‌های سامانه کنترلی نوسان‌کننده هستند. از آنجایی که خروجی سامانه کنترلی همواره دارای مقداری خطا است رفتار این تابع سیاست را به شکل تصادفی تعریف می‌کنیم:

$$\pi_{\theta}(x, u, y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y - g_{\theta}(x, u))^2\right) \quad (8)$$

در این معادله سیاست پارامتربندی شده ورودی‌های حالت x و ورودی u و عمل خروجی y را به یک‌میزان احتمال نسبت می‌دهد که درون یک توزیع احتمال گوسی با مرکز $g_{\theta}(x, u)$ و واریانس σ قرار گرفته است. این شکل از تعریف تابع سیاست احتمالی در بسیاری از کاربردها به کار گرفته شده‌اند که ازجمله آن‌ها به روش‌های به کار رفته برای ورودی سامانه‌های کنترل بهینه در مراجع [۳۴-۳۶] می‌توان اشاره نمود. تابع $g_{\theta}(x, u)$ رفتار متوسط کل یا بخشی از سامانه کنترلی موردنظر را با ورودی‌های x, u و پارامتر θ بیان می‌نماید که درون هر یک از الگوریتم‌های این بخش تعریف شده‌اند. در تعریف این تابع حالت ورودی x حالت سامانه و عمل u مقدار ورودی سامانه است.

تابع پاداش-تنبیه در چنین سامانه‌ای خروجی واقعی سامانه کنترلی را با خروجی مورد انتظار آن مقایسه می‌کند و بر مبنای این مقایسه مقداری از پاداش یا تنبیه برای آن تولید می‌کند. یک شکل از تعریف آن با نادیده گرفتن مقدار ورودی در زمان ورودی به شکل زیر است:

$$r(x, y) = (z_{desired}(x, u) - y)^2 \quad (9)$$

در این معادله پاداش داده شده به عمل y در حالت x برابر توان دوم اختلاف سیگنال ورودی موردنظر در حالت x ($z_{desired}(x, u)$) با مقدار عمل y تعریف می‌شود. به دلیل آن که سیگنال ورودی $z_{desired}$ با در نظر گرفتن شرایط ورودی تولید شده است (مقدار ورودی در زمان تولید $z_{desired}$ دیده می‌شود) اثری از مقدار u در این معادله مشاهده می‌شود.

تابع ارزش سیاست به صورت جمع کاهش‌یافته مقادیر پاداش در طول یک دوره آزمایش تعریف می‌شود:

$$V^{\pi}(x(t)) = E \int_t^{\infty} e^{-\frac{s-t}{\tau}} r(x(s), y(s)) ds \quad (10)$$

که در آن نماد E مقدار مورد انتظار و ضریب کاهش در حالت پیوسته به

تا به مقدار بهینه پارامترها دست یابیم. در اینجا $\Delta\theta$ به کمک بردار گرادین سیاست برآورد می‌شود. برای برآورد $\Delta\theta$ در این روش از تخمین به‌دست‌آمده از تابع پاداش $r(x(s), y(s))$ استفاده می‌شود.

$$\pi_{\theta+\Delta\theta}(x, u) \approx \pi_{\theta}(x, u) + \nabla J_{\theta} \cdot \Delta\theta \quad (14)$$

که در آن بردار گرادین (مشتق) تابع ارزش نسبت به θ است. با مشتق‌گیری از J نسبت به θ در معادله (۱۳) خواهیم داشت:

$$\nabla J_{\theta} = 2 \cdot \sum_{i=1}^m (z_{desired}(x, u) - g_{\theta}(x, u)) \cdot \nabla(g_{\theta}(x, u)) \quad (15)$$

که $\nabla(g_{\theta}(x, u))$ یا به‌طور نسبی $\frac{\partial}{\partial \theta_j}(g_{\theta}(x, u))$ از رابطه زیر محاسبه می‌شود:

$$\nabla(g_{\theta}(x, u)) = dG_{\theta} = [dG1 \ dG2 \ dG3 \ dG4 \ dG5 \ dG6 \ dG7 \ dG8] \quad (16)$$

به عنوان مثال $\frac{\partial}{\partial \theta_1}(g_{\theta}(x, u))$ در رابطه (۱۷) محاسبه شده است.

$$\frac{\partial}{\partial \theta_1}(g_{\theta}(x, u)) = dG1 = \begin{pmatrix} 0 \\ dt \cdot \left((\cos \phi \sin \theta \cos \psi + \sin \phi \sin \psi) \left(\frac{E_z}{\cos \phi \cos \theta} \right) \right) \\ 0 \\ dt \cdot \left((\cos \phi \sin \theta \sin \psi - \sin \phi \cos \psi) \left(\frac{E_z}{\cos \phi \cos \theta} \right) \right) \\ 0 \\ dt \cdot \left(\frac{E_z}{\cos \phi \cos \theta} \right) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (17)$$

در حالت کلی ماتریس dG_{θ} به شکل زیر به دست می‌آید که محاسبه مقادیر dG_{ij} در پیوست مقاله آمده است.

در روش عامل-نقاد طبیعی، بروز رسانی‌های گرادین عامل از راهکار محاسبه گرادین طبیعی آماری^۱ به دست می‌آید. در حالی که نقاد مقادیرهای سیاست طبیعی و پارامترهای اضافی را همزمان توسط یک تابع ارزش و با استفاده از رگرسیون خطی محاسبه می‌کند. در مرجع [۳۷] نشان داده شده است که بهبود روی سیاست با کمک گردبان طبیعی بسیار مناسب‌تر از انواع مشابه است؛ زیرا که این گردبان مستقل از دستگاه مختصات سیاست

شکل ضرب عبارت $e^{-\frac{s-t}{\tau}}$ در معادله وارد شده است. در این معادله τ ثابت کاهش نامیده می‌شود. در عمل برای محاسبه این مقدار از جمع نمونه‌های به‌دست‌آمده از تابع پاداش $r(x(s), y(s))$ استفاده می‌شود.

برای مدل‌سازی مسأله کنترل چهارروتور باید تابع سیاست پارامتری شده $\pi_{\theta}(x, u, y)$ را با بردار پارامتر $\theta = [K_{pz} \ K_{dz} \ K_{pp} \ K_{dp} \ K_{pt} \ K_{dt} \ K_{ps} \ K_{ds}]$ در آن K_{di} و K_{pi} ها به ترتیب بهره‌های کنترلی نسبی و مشتق‌گیر هستند. در صورتی که رفتار دینامیکی کل سامانه کنترل‌کننده PD و دینامیک پرنده چهارروتور را که در معادلات ۳ تا ۵ توصیف شده‌اند، باهم ترکیب کنیم رفتار میانگین توصیف‌کننده سامانه در قالب تابع $g_{\theta}(x, u)$ به شکل زیر به دست می‌آید:

$$g_{\theta}(x, u) = \begin{pmatrix} x(1) + dt \cdot x(2) \\ x(2) + dt \cdot \left((\cos \phi \sin \theta \cos \psi + \sin \phi \sin \psi) \left(\frac{g + K_{pz} E_z + K_{dz} (-\dot{Z})}{\cos \phi \cos \theta} \right) \right) \\ x(3) + dt \cdot x(4) \\ x(4) + dt \cdot \left((\cos \phi \sin \theta \sin \psi - \sin \phi \cos \psi) \left(\frac{g + K_{pz} E_z + K_{dz} (-\dot{Z})}{\cos \phi \cos \theta} \right) \right) \\ x(5) + dt \cdot x(6) \\ x(6) + dt \cdot \left(-g + \left(\frac{g + K_{pz} E_z + K_{dz} (-\dot{Z})}{\cos \phi \cos \theta} \right) \right) \\ x(7) + dt \cdot x(8) \\ x(8) + dt \cdot \left(\dot{\theta} \psi \left[\frac{I_{yy} - I_{zz}}{I_{xx}} \right] + \frac{J_r}{I_{xx}} \dot{\Omega}_d + \frac{1}{I_{xx}} \left((K_{pp} E_p + K_{dp} (-\dot{\theta})) \right) \right) \\ x(9) + dt \cdot x(10) \\ x(10) + dt \cdot \left(\dot{\phi} \psi \left[\frac{I_{zz} - I_{xx}}{I_{yy}} \right] - \frac{J_r}{I_{yy}} \dot{\Omega}_d + \frac{1}{I_{yy}} \left((K_{ps} E_s + K_{ds} (-\dot{\theta})) \right) \right) \\ x(11) + dt \cdot x(12) \\ x(12) + dt \cdot \left(\dot{\theta} \dot{\phi} \left[\frac{I_{xx} - I_{yy}}{I_{zz}} \right] + \frac{1}{I_{zz}} \left((K_{ps} E_s + K_{ds} (-\dot{\psi})) \right) \right) \end{pmatrix} \quad (11)$$

این تابع خروجی کلی کنترلر در حالت x با ورودی u را در شکل متوسط تولید می‌نماید:

$$y = \pi_{\theta}(x, u) = g_{\theta}(x, u) + \mathcal{N}(0, \sigma) \quad (12)$$

که در اینجا $\mathcal{N}(0, \sigma)$ نویز گوسی به مرکز صفر و واریانس σ می‌باشد. این معادله به‌طور مستقیم از معادله (۸) به دست می‌آید. تابع ارزش بردار θ را در شکل ساده ($\gamma = 1$) بدون در نظر گرفتن نویز با معادله زیر تعریف می‌کنیم:

$$J(\dot{\theta}) = \sum_{i=1}^m r(x, y) = \sum_{i=1}^m (z_{desired}(x, u) - g_{\theta}(x, u))^2 \quad (13)$$

فرآیند یادگیری گرادین سیاست، یک روند تکراری است که بردار θ را با بهترین مقادیر اولیه به دست آمده در مرحله اول الگوریتم مقادیردهی اولیه می‌کند و هر بار بردار θ با بردار جدید $\theta + \Delta\theta$ جایگزین می‌شود

$$VG(i)^{it} = VG(i)^{it-1} + \alpha \cdot (BPos(it) - Pos(i)^{it-1}) + \beta \cdot \nabla J(Pos(i)^{it-1}) \quad (20)$$

در این معادله α و β دو ضریب وزن‌دهی هستند که مقدار بیشینه ثابت اولیه دارند و در طول اجرا به‌طور تصادفی از مقدار بیشینه تولید می‌شوند. مقدار بیشینه α با افزایش دوره‌های الگوریتم به‌تدریج کاهش می‌یابد تا الگوریتم بهتر همگرا گردد. $BPos(it)$ بهترین بردار در مجموعه بردارهای دور it است که کمترین مقدار خطا را تولید کرده است و $\nabla J(Pos(i)^{it-1})$ مقدار گرادیان سیاست محاسبه شده به ازای $Pos(i)$ در دور $it-1$ است. این گرادیان توسط روش ارائه‌شده در رابطه (۱۹) محاسبه می‌گردد.

الگوریتم پس از محاسبه $\nabla J(Pos(i)^{it-1})$ و $VG(i)$ و بروز رسانی هر بردار $Pos(i)$ ، آن‌ها را ارزیابی کرده و خطای برآورد را با محاسبه تابع ارزش به دست می‌آورد. پس از آن بهترین بردار آن دور که کمترین خطا را تولید کرده است، استخراج شده و در $BPos$ قرار می‌گیرد. پس از پایان یک دور مقدار بیشینه α بروز رسانی می‌شود و بهترین بردار پارامتر θ_{best} انتخاب می‌شود. در پایان حلقه اصلی الگوریتم، بهترین بردار پارامتر به خروجی تابع ارسال می‌شود و یک کنترلر به‌طور کامل ورودی می‌بیند.

۵- نتایج آزمایش‌های عملی

طراحی یک سامانه کنترلی برای کوادروتور نیازمند طراحی یک بستر آزمایشگاهی برای پیاده‌سازی و آزمون عملی کنترلر است. به منظور کم کردن پیچیدگی‌های کنترل و جلوگیری از آسیب دیدن سازه، یک پایه برای محدود کردن درجات آزادی سامانه طراحی و ساخته شد.

به همین منظور یک پایه برای طراحی کنترلر طراحی شده است به این ترتیب که پایه نشان داده شده در شکل ۳ حرکت و چرخش در راستاهای رول، پیچ و یاو را برای کوادروتور عملی می‌کند. مدل حرکت در یک بعد برای آزمون و پیاده‌سازی عملی در ادامه مورد استفاده قرار گرفته شده است. کوادروتور یک سامانه شش درجه آزادی است که از چهارروتور به‌عنوان پیشران استفاده می‌کند. پیشران‌ها موتورهای DC هستند که جریان و ولتاژ موردنیاز آن‌ها توسط باتری و برای چرخش آن‌ها توسط یک میکروکنترلر PWM تولید می‌شود. به‌عنوان هسته کنترلی کوادروتور از یک میکروکنترلر آردوینو استفاده شده است. آردوینو یک بستر محاسباتی فیزیکی متن‌باز است



Fig. 3. Quadrotor on the base with three degrees of freedom

شکل ۳: نحوه قرارگیری ربات روی پایه با سه درجه آزادی

انتخاب شده است و می‌تواند تا حد بالایی کارآمدتر از گرادیان‌های سیاست معمولی عمل کند. این روش مورد توجه فراوانی در ادبیات حوزه یادگیری تقویتی قرار گرفته است و در مراجع [۳۸-۴۰] ایده‌های نوینی بر روی آن توسعه داده شده است.

$$\nabla(g_\theta(x, u)) = dG_\theta = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ dg_{2,1} & dg_{2,2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ dg_{4,1} & dg_{4,2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ dg_{6,1} & dg_{6,2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & dg_{8,3} & dg_{8,4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & dg_{10,5} & dg_{10,6} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & dg_{12,7} & dg_{12,8} \end{bmatrix} \quad (18)$$

برای استفاده از سیاست تصادفی تعریف شده در معادله (۱۸) در روش عامل-نقاد طبیعی $\nabla_\theta \log(\pi_\theta(x, u, y))$ باید محاسبه شود:

$$\nabla_\theta \log(\pi_\theta(x, u, y)) = \frac{(y - g_\theta(x, u))}{\sigma^2} \nabla(g_\theta(x, u)) \quad (19)$$

که $\nabla(g_\theta(x, u))$ از رابطه (۱۸) محاسبه می‌شود. در ادامه بهینه‌سازی ازدحام ذرات را با روش گرادیان سیاست ترکیب می‌کنیم. در این روش، مجموعه ذراتی در فضای پاسخ فرض می‌شوند و یک سرعت ابتدایی به آن‌ها اختصاص داده می‌شود، همچنین کانال‌های ارتباطی بین ذرات در نظر گرفته می‌شود. سپس این ذرات در فضای پاسخ حرکت می‌کنند و نتایج حاصل بر مبنای یک ملاک شایستگی پس از هر بازه زمانی محاسبه می‌شود. با گذشت زمان، ذرات به سمت ذراتی که دارای ملاک شایستگی بالاتری هستند و در گروه ارتباطی یکسانی قرار دارند، شتاب می‌گیرند.

فرآیند بهینه‌سازی ترکیبی پارامترهای سامانه کنترلی، یک فرآیند تکراری است که از یک مجموعه N تایی مقادیر اولیه تصادفی شروع به کار می‌کند و پارامترهای بهینه فضای حالت را جستجو می‌کند. هر مقدار اولیه داخل یک بردار θ تایی $(Pos(i))$ قرار می‌گیرد که بردار پارامتر θ را مقداردهی می‌کند. این بردار Pos در طول اجرای الگوریتم مرتباً بروز رسانی می‌شود. در حلقه اصلی الگوریتم، تکرار تا زمانی که خطای برآورد زیاد است و یا تعداد مشخصی تکرار به پایان نرسیده است، ادامه می‌یابد. در هر دور هر بردار $Vec(i)$ با بردار جدید $Pos(i) + VG(i)$ جایگزین می‌شود که در اینجا $VG(i)$ بردار سرعت کلی $Pos(i)$ می‌باشد. بردار سرعت کلی $Pos(i)$ در دور it از رابطه زیر محاسبه می‌شود:

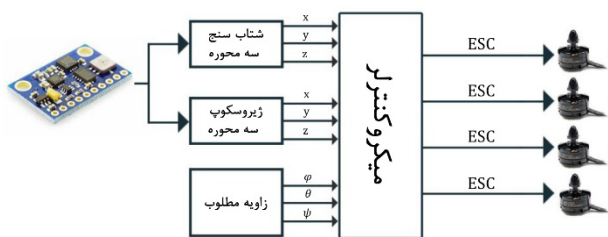


Fig. 4. Quadrotor Hardware Control Block Diagram

شکل ۴: بلوک دیاگرام سخت‌افزار کنترلی چهارروتور

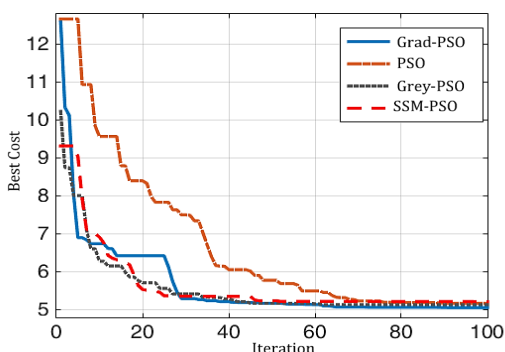


Fig. 5. Comparison of optimization curves of different methods

شکل ۵: مقایسه تابع هزینه طراحی کنترل‌کننده با الگوریتم‌های فراابتکاری

تابع هزینه اقدام می‌کنند. در کنترلر مشتق‌گیر تناسبی تعداد پارامترهای کنترلی عبارتند از k_{pz} ، k_{dz} ، k_{dt} ، k_{ps} ، k_{ds} ، k_{pp} ، k_{dp} ، k_{dp} که هشت جزء مورد بررسی ما برای بهینه‌سازی و کمینه‌سازی تابع هزینه به شمار می‌روند. بهینه‌ساز در طول زمان با اندازه‌گیری هزینه کنترلی سامانه آن را ذخیره‌سازی می‌کند و به تدریج با اجرای روش‌های بهینه‌سازی مشخص شده سعی می‌کند تا ضرایب کنترل‌کننده مشتق‌گیر-تناسبی را بهبود بخشد. شکل ۵ تابع هزینه را برحسب تعداد تکرار عملیات بهینه‌سازی نمایش می‌دهد، همانطور که مشخص است در طول روند بهینه‌سازی تابع هزینه که تابعی از زمان فراز سامانه، خطای ماندگار سامانه، فراجش و زمان نشست سامانه است در طول تکرارهای متعدد روند نزولی داشته است. در میان روش‌های آزمایش شده روش ارائه شده (Grad-PSO) نسبت به سایر روش‌ها سرعت بهینه‌سازی بسیار بیشتری دارد. بالایی سرعت بهینه‌سازی به دلیل حرکت در جهت گرادینت سیاست می‌باشد که بهینه‌سازی سریع‌تری را شکل داده است. شکل‌های ۶ و ۷ پاسخ سامانه به ورودی‌های پله برای درجات آزادی پیچ^۱ و یاو^۲ است که به ترتیب در جدول‌های ۱ و ۲ تحلیل شده است. در این جدول‌ها t_s زمان نشست سامانه، t_r زمان فراز و M_p بیشینه فراجش پاسخ سامانه با هم مقایسه شده‌اند. پارامترهای طراحی شده در هر کنترل‌کننده نیز در این جدول‌ها آورده شده است. همان‌گونه که دیده می‌شود روش ارائه شده زمان نشست و فراز کمتری نسبت به سایر روش‌های مقایسه شده دارد و از این‌رو پارامترهای کنترلی طراحی شده توسط این روش پارامترهای بهتری نسبت به روش‌های دیگر داراست.

- 1 Roll
- 2 Pitch

که براساس یک برد میکروکنترلر ساده تهیه شده است.

همان‌گونه که در شکل ۴ نشان داده شده است، میکروکنترلر سیگنال آنالوگ را از شتاب‌سنج و ژيروسکوپ می‌خواند. این سیگنال به مقادیر دیجیتال تبدیل می‌شود تا بتواند در میکروکنترلر مورد استفاده قرار بگیرد. میکروکنترلر با توجه به مقادیر خوانده شده و زاویه مطلوب منظور شده برای کوادروتور برای چهار موتور سیگنال Pulse Width Modulation (PWM) تولید می‌کند و به این ترتیب سرعت چرخش هر پیشران را معین می‌کند. همان‌گونه که در شکل هم مشخص است سیگنال PWM تولید شده توسط میکروکنترلر وارد کنترلرهای سرعت (ESC) می‌شود و در خروجی سرعت موتورها را تعیین می‌کند. مقادیر شتاب‌سنج و ژيروسکوپ از یک inertial measurement unit (IMU) مبتنی بر حسگرهای ابعاد کوچک به دست می‌آیند. حسگر GY80 در شکل ۴ برای محاسبه موقعیت زاویه‌ای با حسگرهای شتاب‌سنج و ژيروسکوپ مورد استفاده قرار گرفته است.

به صورت کلی سه استراتژی برای ترکیب داده‌ها موجود است. اولین روش با عنوان روش اصلاح شناخته می‌شود؛ برای مثال اطلاعات خروجی از یک حسگر برای اصلاح دیگری استفاده می‌شود. استراتژی دوم اختلاط نام دارد که با ترکیب و اختلاط بخش‌های یک حسگر اقدام به تولید داده می‌کند. روش سوم و یا بهترین روش برای ترکیب اطلاعات حسگرها ادغام است. در این روش مقادیر هر حسگر با منظور کردن ضرایب وزنی و آماری مختلف با همدیگر ترکیب می‌شوند تا بهینه‌ترین پاسخ ممکن را در خروجی تولید کند. به صورت کلی الگوریتم‌های ادغام از فیلتر کالمن برای تشخیص سیگنال نویزی در طول زمان استفاده می‌کند تا سیگنالی نزدیک به مقدار واقعی تولید کند. حسگرهای IMU به دو نوع دیجیتال و آنالوگ تقسیم‌بندی می‌شوند. حسگر آنالوگ خروجی را به صورت ولتاژ به میکروکنترلر می‌دهد، میکروکنترلر با استفاده از آنالوگ به دیجیتال (ADC) این مقدار را به مقادیر دیجیتال تبدیل می‌کند. برای تخمین دقیق موقعیت کوادروتور خروجی حسگرهای شتاب‌سنج و ژيروسکوپ با ضرایب وزنی مشخص ادغام می‌شوند. فیلتر کالمن این ضرایب وزنی در طی زمان و با تغییر به نویز داده‌های خوانده شده از شتاب‌سنج تغییر می‌کند. بردار به‌عنوان بردار نویزی بردار محاسبه شده، هر دو کسینوس زاویه بردار شتاب را در سازه مشخص می‌کنند. با ترکیب این دو مقدار با ضرایب وزنی مشخص مقدار نهایی در هر جهت به دست می‌آید. شکل ۴ بلوک دیاگرام سخت‌افزاری سامانه را نشان می‌دهد.

در این بخش به ارزیابی و مقایسه روش ارائه شده در این مقاله پرداخته می‌شود. بدین منظور الگوریتم ازدحام ذرات عادی (PSO)، ازدحام ذرات خاکستری (Grey-PSO) [۴۱]، ازدحام ذرات دانشجو راهنما (SSM-PSO) [۴۲] و روش ارائه شده گرادینت ازدحام ذرات (Grad-PSO) روی یک کنترلر PD پیاده‌سازی شدند تا به بررسی و بهبود عملکرد کنترلر پس از تنظیم بهره‌های کنترلی توسط این دو روش فراابتکاری و مقایسه عملکرد آن‌ها پرداخته شود. تعداد تکرار در هر دو الگوریتم برابر منظور شده است و الگوریتم‌ها به صورت مستقل به محاسبه بهره‌های کنترلی به منظور کاهش

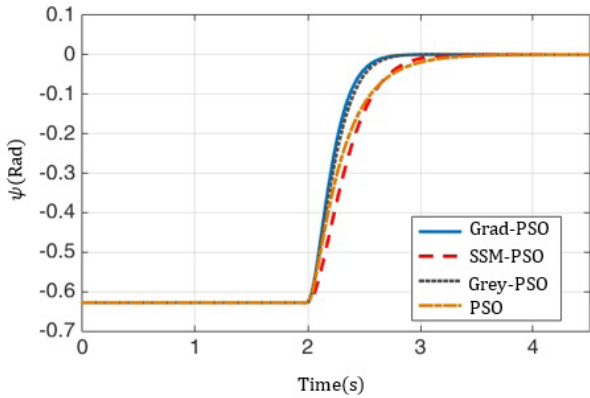


Fig. 7. The controller response to the step input in yaw control ψ

شکل ۷: پاسخ کنترلر به ورودی پله برای کنترل ψ

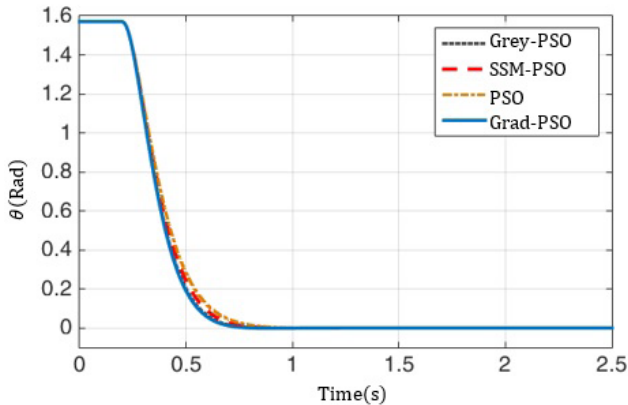


Fig. 8. The controller response to the step input in pitch control θ

شکل ۸: پاسخ کنترلر به ورودی پله برای کنترل θ

جدول ۳: مقایسه کنترلر رول

Table 3. Comparison of different methods

K_p	K_d	t_r	t_s	کنترلر
-/۹۷	-/۲۲	-/۵۵	-/۴۱	Gradient PSO
۱	-/۱۶	-/۶۳	-/۴۶	SSM PSO
-/۸۷	-/۱۵	-/۵۸	-/۴۳	Grey PSO
-/۸۲	-/۱۵	-/۷۱	-/۴۹	PSO

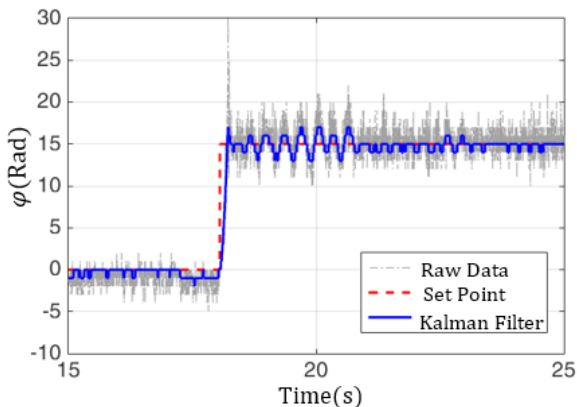


Fig. 9. Experimental result in roll control using PD controller

شکل ۹: نتایج تجربی کنترلر رول با استفاده از کنترلر PD

به صورت کلی گرادیان ازدحام ذرات به خاطر جستجوی بهتر که در فضای پارامترها انجام می‌دهد به تابع هزینه کمتری منجر می‌شود. این روش قادر است به‌طور گسترده در بهینه‌سازی کنترل‌کننده‌ها به کار گرفته شود. سوق دادن جهت حرکت ذره‌ها در راستای گرادیان تابع سیاست باعث می‌شود تا بهینه‌سازی با نرخ بیشتر و به پارامترهای بهتری منجر گردد.

در شکل ۹ نتایج آزمون عملی کنترل‌کننده روی بستر آزمایش ساخته‌شده به تصویر درآمده است. شکل‌های ۹ و ۱۰ پاسخ تغییرات زاویه رول و پیچ به ورودی پله را نشان می‌دهد.

فیلم نمایش این تست از لینک زیر قابل دریافت است:

<http://goo.gl/jYolmU>

شکل ۱۱ دنباله تصویر آزمایش کنترلر پیچ را نشان می‌دهد.

جدول ۱: مقایسه کنترلر رول

Table 1. Comparison of different methods

K_p	K_d	$(\%)M_p$	t_r	t_s	کنترلر
-/۹۸	-/۱۶	-/۰۱	-/۵۹۰۰	-/۴۰	Gradient PSO
-/۷۴	-/۱۵	-/۰۲	-/۸۳	-/۴۸	SSM PSO
-/۶۹	-/۱۳	-/۰۱۲	-/۶۶	-/۴۵	Grey PSO
-/۶۶	-/۱۴	-/۰۳۳	-/۷۳	-/۵۲	PSO

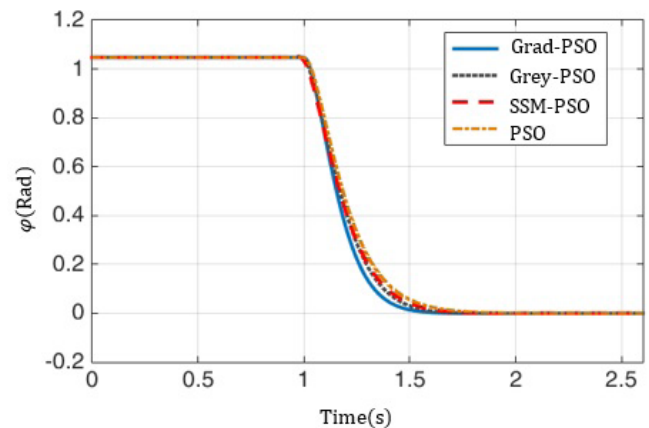


Fig. 6. The controller response to the step input in roll control ϕ

شکل ۶: پاسخ کنترلر به ورودی پله برای کنترل ϕ

جدول ۲: مقایسه کنترلر یاو

Table 2. Comparison of different methods

K_p	K_d	$(\%)M_p$	t_r	t_s	کنترلر
-/۹۷	-/۲۲	-/۰۱	-/۵۵	-/۴۱	Gradient PSO
۱	-/۱۶	-/۰۱۸	-/۶۳	-/۴۶	SSM PSO
-/۸۷	-/۱۵	-/۰۱۳	-/۵۸	-/۴۳	Grey PSO
-/۸۲	-/۱۵	-/۰۴	-/۷۱	-/۴۹	PSO

Autopilots for small unmanned aerial vehicles: a survey, 8(1) (2010) 36-44.

- [6] D.-J. Lee, I. Kammer, V. Dobrokhodov, K.J.I.J.o.C. Jones, Automation, Systems, Autonomous feature following for visual surveillance using a small unmanned aerial vehicle with gimbaled camera system, 8(5) (2010) 957-966.
- [7] D.-i. Han, J.-h. Kim, C.-o. Min, S.-j. Jo, J.-h. Kim, D.-w.J.I.J.o.C. Lee, Automation, Systems, *Development of Unmanned Aerial Vehicle (UAV) system with waypoint tracking and vision-based reconnaissance*, 8(5) (2010) 1091-1099.
- [8] E. Altug, J.P. Ostrowski, C.J. Taylor, Quadrotor control using dual camera visual feedback, in: *Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference on, IEEE, 2003*, pp. 4294-4299.
- [9] D. Suter, T. Hamel, R. Mahony, Visual servo control using homography estimation for the stabilization of an x4-flyer, in: *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on, IEEE, 2002*, pp. 2872-2877.
- [10] J. Dunfield, M. Tarbouchi, G. Labonte, Neural network based control of a four rotor helicopter, in: *Industrial Technology, 2004. IEEE ICIT'04. 2004 IEEE International Conference on, IEEE, 2004*, pp. 1543-1548.
- [11] M.G. Earl, R. D'Andrea, Real-time attitude estimation techniques applied to a four rotor helicopter, in: *Decision and Control, 2004. CDC. 43rd IEEE Conference on, IEEE, 2004*, pp. 3956-3961.
- [12] D. Lee, H.J. Kim, S.J.I.J.o.c. Sastry, Automation, systems, *Feedback linearization vs. adaptive sliding mode control for a quadrotor helicopter*, 7(3) (2009) 419-428.
- [13] B. Erginer, E.J.I.J.o.C. Altuğ, Automation, Systems, *Design and implementation of a hybrid fuzzy logic controller for a quadrotor VTOL vehicle*, 10(1) (2012) 61-70.
- [14] A. Sharma, A.J.I.J.E.E.C.E. Barve, *Controlling of Quad-rotor UAV using pld controller and Fuzzy logic controller*, 1(2) (2012) 38-41.
- [15] L.-X. Wang, A course in fuzzy systems and control: Desing of Fuzzy Systems from Input-Output Data, in: *Prentice-Hall, Upper Saddle River, NJ, 1997*.
- [16] C.-C.J.I.T.o.s. Lee, man., cybernetics, *Fuzzy logic in control systems: fuzzy logic controller. I*, 20(2) (1990) 404-418.
- [17] L. Reznik, *Fuzzy controllers handbook: how to design them, how they work*, Elsevier, 1997.
- [18] E.H.-K. Fung, Y.-K. Wong, Y. Ma, C.-W.M. Yuen, W.-K.J.I.J.o.C. Wong, Automation, Systems, *Smart hanger dynamic modeling and fuzzy controller design*, 9(4) (2011) 691.

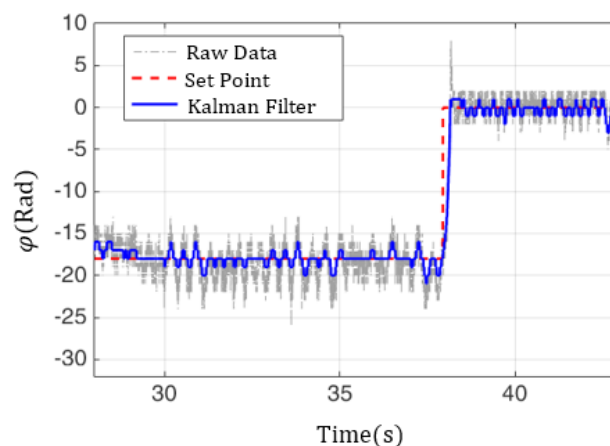


Fig. 10. Experimental result in pitch control using PD controller
 شکل ۱۰: نتایج تجربی کنترل پیچ با استفاده از کنترلر PD

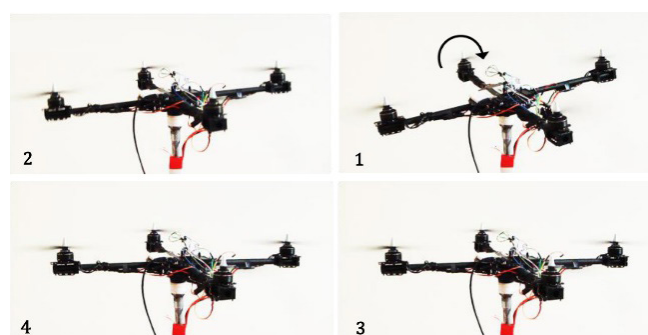


Fig. 11. Image sequence of pitch control using PD controller
 شکل ۱۱: دنباله تصویر کنترل پیچ با استفاده از کنترلر PD

۶- نتیجه گیری و جمع بندی

در این مقاله یک روش جدید ترکیبی با کمک گرادیان سیاست و ازدحام ذرات ارائه شد. این روش با بهره‌گیری از ایده‌های روش‌های یادگیری تقویتی گرادیان بردار سیاست کنترل‌کننده را محاسبه و از این گرادیان در بهینه‌سازی ازدحام ذرات برای یافتن بهترین پارامترهای کنترلی سامانه استفاده می‌کند. بهینه‌ساز در طول زمان با اندازه‌گیری هزینه کنترلی سامانه، آن را ذخیره‌سازی و به تدریج با اجرای روش‌های بهینه‌سازی مشخص شده سعی می‌کند تا ضرایب کنترل‌کننده مشتق‌گیر-تناسبی را بهبود بخشد. نتایج این مقاله بر روی یک زیرساخت سخت‌افزاری کوادروتور مورد آزمایش قرار گرفته تا توانایی‌های کنترل‌کننده نهایی به‌خوبی آشکار گردد.

مراجع

- [1] A. Gessow, G.C. Myers, *Aerodynamics of the Helicopter*, Macmillan New York, 1952.
- [2] J.J.P. Leishman, New, *Principles of Helicopter Aerodynamics*, Cambridge Univ, (2000).
- [3] M.J. Hirschberg, *The American Helicopter*, (2000).
- [4] P. Castillo, R. Lozano, A.E. Dzul, *Modelling and control of mini-flying machines*, Physica-Verlag, 2006.
- [5] H. Chao, Y. Cao, Y.J.I.J.o.C. Chen, Automation, Systems,

- [30] S.J.F.P.C.J.J.o.B.U.o.A. Kaiyuan, *Astronautics, Attitude control of quadrotor aircraft via nonlinear PID [J]*, 9 (2011) 003.
- [31] S. Bouabdallah, *Design and control of quadrotors with application to autonomous flying*, Epfl, 2007.
- [32] H. Bolandi, M. Rezaei, R. Mohsenipour, H. Nemati, S.M.J.I.C. Smailzadeh, *Automation, Attitude control of a quadrotor with optimized PID controller*, 4(03) (2013) 335.
- [33] A.L. Salih, M. Moghavvemi, H.A. Mohamed, K.S.J.S.r. Gaeid, *essays, Flight PID controller design for a UAV quadrotor*, 5(23) (2010) 3660-3667.
- [34] G. Endo, J. Morimoto, T. Matsubara, J. Nakanishi, G.J.T.I.J.o.R.R. Cheng, *Learning CPG-based biped locomotion with a policy gradient method: Application to a humanoid robot*, 27(2) (2008) 213-228.
- [35] T. Matsubara, J. Morimoto, J. Nakanishi, M.-a. Sato, K.J.R. Doya, A. Systems, *Learning CPG-based biped locomotion with a policy gradient method*, 54(11) (2006) 911-920.
- [36] Y. Nakamura, T. Mori, M.-a. Sato, S.J.N.N. Ishii, *Reinforcement learning for a biped robot based on a CPG-actor-critic method*, 20(6) (2007) 723-735.
- [37] S. Bhatnagar, R.S. Sutton, M. Ghavamzadeh, M.J.A. Lee, *Natural actor-critic algorithms*, 45(11) (2009) 2471-2482.
- [38] S. Bhatnagar, M. Ghavamzadeh, M. Lee, R.S. Sutton, *Incremental natural actor-critic algorithms*, in: *Advances in neural information processing systems*, 2008, pp. 105-112.
- [39] F.S. Melo, M. Lopes, *Fitted natural actor-critic: A new algorithm for continuous state-action MDPs*, in: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, Springer, 2008, pp. 66-81.
- [40] T. Degris, P.M. Pilarski, R.S. Sutton, *Model-free reinforcement learning with continuous action in practice*, in: *American Control Conference (ACC), 2012, IEEE, 2012*, pp. 2177-2182.
- [41] M.-S. Leu, M.-F.J.A.S.C. Yeh, *Grey particle swarm optimization*, 12(9) (2012) 2985-2996.
- [42] Y. Liu, Z. Qin, X. He, *Supervisor-student model in particle swarm optimization*, in: *Evolutionary Computation, 2004. CEC2004. Congress on, IEEE, 2004*, pp. 542-547.
- [19] A. Hafaifa, F. Laouad, K.J.I.J.o.C. Laroussi, *Automation, Systems, A numerical structural approach to surge detection and isolation in compression systems using fuzzy logic controller*, 9(1) (2011) 69-79.
- [20] M. Sugeno, I. Hirano, S. Nakamura, S. Kotsu, *Development of an intelligent unmanned helicopter*, in: *Fuzzy Systems, 1995. International Joint Conference of the Fourth IEEE International Conference on Fuzzy Systems and The Second International Fuzzy Engineering Symposium., Proceedings of 1995 IEEE Int, IEEE, 1995*, pp. 33-34.
- [21] B. Kadmiry, D. Driankov, *Fuzzy control of an autonomous helicopter*, in: *IFSA World Congress and 20th NAFIPS International Conference, 2001. Joint 9th, IEEE, 2001*, pp. 2797-2802.
- [22] C. Cavalcante, J. Cardoso, J.J. Ramos, O.R. Neves, *Design and tuning of a helicopter fuzzy controller*, in: *Fuzzy Systems, 1995. International Joint Conference of the Fourth IEEE International Conference on Fuzzy Systems and The Second International Fuzzy Engineering Symposium., Proceedings of 1995 IEEE Int, IEEE, 1995*, pp. 1549-1554.
- [23] N.I. Vitzilaios, N.C.J.J.o.I. Tsourveloudis, R. Systems, *An experimental test bed for small unmanned helicopters*, 54(5) (2009) 769-794.
- [24] R. Garcia, K.P. Valavanis, *The implementation of an autonomous helicopter testbed*, in: *Unmanned Aircraft Systems*, Springer, 2008, pp. 423-454.
- [25] Y.J.H.R.-W. Wu, Μάρτιος, *Development and Implementation of a Control System for a quadrotor UAV*, (2009).
- [26] R.J.U.P.D. Barclay, University of Sheffield, United Kingdom, *A Generic Simulator for Quad-Rotor Unmanned Aerial Vehicles*, (2005).
- [27] M.B. Srikanth, Z.T. Dydek, A.M. Annaswamy, E. Lavretsky, *A robust environment for simulation and testing of adaptive control for mini-UAVs*, in: *American Control Conference, 2009. ACC'09., IEEE, 2009*, pp. 5398-5403.
- [28] Z.-Y. Zhao, M. Tomizuka, S.J.I.t.o.s. Isaka, man., *cybernetics, Fuzzy gain scheduling of PID controllers*, 23(5) (1993) 1392-1398.
- [29] J.G.J.t.A. Ziegler, *Optimum settings for automatic controllers*, 65 (1943) 433-444.

برای ارجاع به این مقاله از عبارت زیر استفاده کنید:

Please cite this article using:

H.Shahbazi, V. Tikani, *Design of a Nonlinear Controller on Quadrotor Drone Using Combined Method of Gradient Particle Swarm Optimization*, *Amirkabir J. Mech. Eng.*, 50(5) (2018) 989-998.

DOI: 10.22060/mej.2016.859

