مجله کنترل
ISSN (print) 2008-8345
ISSN (online) 2538-3752
مقاله کاربردی
جلد۱۶، شماره ۴، زمستان ۱۴۰۱، صفحه ۷۳-۵۷

کنترل نگاه به جلو یک ربات شبه اتومبیل از طریق یادگیری تقویتی

نسترن رنجکش[۱]، خوشنام شجاعی[۲،۳]

[۱] کارشناس ارشد، دانشکده مهندسی برق، واحد نجف آباد، دانشگاه آزاد اسلامی، نجف آباد، اصفهان، ایران، nastaran.ranjkesh@yahoo.com

[۲] دانشیار، دانشکده مهندسی برق، واحد نجف آباد، دانشگاه آزاد اسلامی، نجف آباد، اصفهان، ایران، shojaei@pel.iaun.ac.ir

[۳] مرکز تحقیقات پردازش دیجیتال و بینایی ماشین، واحد نجف آباد، دانشگاه آزاد اسلامی، نجف آباد، ایران

دریافت: ۱۴۰۱/۰۶/۱۸ ویرایش: ۱۴۰۱/۱۰/۰۳ پذیرش: ۱۴۰۱/۱۰/۲۸

چکیده: در این مقاله، مسأله بهبود عملکرد ردیابی خط سیر زمانی مرجع برای یک ربات شبه اتومبیل با قیود غیرهولونومیک در حضور اغتشاشات خارجی، غیر خطی‌ها و عدم قطعیت‌های پارامتری بررسی شده است. به همین منظور، ابتدا معادلات دینامیک و سینماتیک ربات شبه اتومبیل توصیف می شود و سپس برای ردیابی ربات شبه اتومبیل از روش نگاه به جلو در فضای دو بعدی استفاده می‌گردد. کنترل کننده پیشنهاد شده به کمک تکنیک کنترل سطح دینامیکی همراه با یادگیری تقویتی مبتنی بر شبکه عصبی بازیگر- منتقد طراحی خواهد شد و به منظور مقابله و جبران تأثیرات اغتشاشات خارجی، یک کنترل کننده مقاوم تطبیقی پیشنهاد می‌گردد. همچنین، برای بهبود عملکرد حالت گذرا و حالت ماندگار از روش عملکرد از پیش تعریف شده استفاده خواهد شد. یک شبکه عصبی بازیگر برای تخمین غیرخطی‌های نامعلوم و عدم قطعیت‌ها و یک شبکه عصبی منتقد برای ارزیابی عملکرد سیستم مورد استفاده قرار می‌گیرد. در ادامه، روش لیاپانوف مستقیم برای اثبات پایداری سیستم کنترل حلقه بسته و کرانداری نهایی یکنواخت خطاهای ردیابی، به کار گرفته خواهد شد. در انتها، اثر بخشی و کارآیی طرح کنترلی پیشنهادی با استفاده از نرم افزار متلب تایید می شود.

کلمات کلیدی: ربات شبه اتومبیل،کنترل نگاه به جلو، یادگیری تقویتی، کنترل کننده مقاوم تطبیقی، شبکه عصبی بازیگر، شبکه عصبی منتقد.

Look-ahead control of a car-like mobile robot via reinforcement learning

Nastaran Ranjkesh, Khoshnam Shojaei

Abstract: In this paper, the performance improvement problem of a reference trajectory tracking for a car-like mobile robot with nonholonomic constraints in the presence of external disturbances, nonlinearities and uncertain parameters is investigated. For this purpose, at first the dynamic and kinematic equations of the car-like mobile robot are expressed and then the look-ahead control method in two dimensional is used for the tracking of the car-like mobile robot. The purposed controller will be designed by using the dynamic surface control method with the reinforcement learning based on the actor-critic neural network and an adaptive robust controller is proposed to compensate the effects of external disturbances. Moreover, the prescribed performance control method will be utilized to improve the transient state and steady state. An actor neural network is used to estimate unknown nonlinearities and uncertainties and a critic neural network is employed to evaluate system performance. In the sequel, the direct Lyapunov method will be used to prove the closed-loop control system stability and uniform ultimate boundedness of tracking errors. Finally, the effectiveness and efficiency of the proposed control scheme are confirmed by using MATLAB software.

Keywords: Actor neural network, Adaptive robust controller, Car-like mobile robot, Critic neural network, Look-ahead control, Reinforcement learning

[Downloaded from joc.kntu.ac.ir on 2026-11-10]
[DOR: 20.1001.1.20088345.1401.16.4.5.3]
[DOI: 10.61186/joc.16.4.57]

۱- مقدمه

یادگیری ماشین به عنوان شاخه‌ای از هوش مصنوعی بر روی شناسایی خودکار الگوهای پیچیده و تصمیم گیری‌های هوشمند مبتنی بر داده‌های موجود تمرکز می‌کند. تکنولوژی یادگیری، در یادگیری ماشین به سه دسته تقسیم بندی می‌شود: ۱-یادگیری نظارت‌شده[1]، ۲-یادگیری بدون‌نظارت[2]، ۳-یادگیری تقویتی[3] [۱]. در یادگیری نظارت شده مجموعه داده‌های ورودی به دو گروه تست و آموزشی تقسیم می‌شوند. در این نوع یادگیری، جفت نمونه‌های ورودی - خروجی یا داده‌های برچسب گذاری شده برای آموزش مدل با هدف ایجاد تابعی می‌باشند که قادر به تخمین خروجی‌ها، در زمانی که ورودی‌های جدید به آن معرفی خواهد شد، استفاده می‌شوند. یادگیری نظارت شده جزء الگوریتم‌هایی می‌باشد که به ناظر نیاز خواهد داشت [۲-۳]. در یادگیری بدون نظارت برخلاف یادگیری نظارت شده، هیچ پاسخ صحیح و ناظری وجود ندارد. این یادگیری ویژگی‌های مختصری از داده‌ها را می‌آموزد وسعی می‌کند الگوها و ساختارها را از داده‌های بدون برچسب کشف کند و زمانی که داده جدیدی معرفی می‌شود، این یادگیری ویژگی‌های آموخته شده قبلی را برای شناسایی دسته‌ای از داده‌ها استفاده می‌کند [۴-۵]. در بعضی از موارد خروجی سیستم، توالی از اعمال[4] است. در این موارد، یک عمل به تنهایی مهم نیست و مسئله‌ای که اهمیت دارد آن سیاستی[5] است که توالی از اعمال صحیح را برای رسیدن به هدف مطلوب ایجاد کند. در این موارد، بیانی به عنوان بهترین عمل در هر حالت وجود ندارد و یک عمل در صورتی مطلوب است، اگر قسمتی از سیاست مطلوب باشد. دراین حالت، یادگیری ماشین باید قادر باشد به خوبی سیاست‌ها را ارزیابی کند و توالی از اعمال خوب گذشته را برای ایجاد سیاست مناسب یاد بگیرد. چنین یادگیری، یادگیری تقویتی نامیده می‌شود [۶]. این نوع یادگیری یک روش یادگیری برخط[6] است که با روش‌های یادگیری نظارت شده و بدون نظارت متفاوت است [۷]. یادگیری تقویتی برای تعیین اعمالی با هدف به حداکثر رساندن پاداش تجمعی پیش بینی شده در آینده، از مفهومی به نام عامل[7] استفاده می‌کند که قادر به انجام اعمال در محیط است [۸]. درمسئله یادگیری تقویتی به عامل این امکان داده می‌شود که ضمن تعامل با محیط و کسب تجربه، موقعیت‌ها را برای انجام اعمال به گونه‌ای جایگزین کند که سیگنال پاداش عددی ماکزیمم شود [۹]. در حقیقت به یادگیرنده[8] گفته نمی‌شود کدام عمل باید انجام شود، اما در عوض باید کشف کند که کدام اعمال ناشی از رفتار درست پاداشی مثبت و کدام اعمال ناشی از رفتار نادرست پاداش منفی را به همراه دارد و این نتایج با آزمایش به دست خواهند آمد [۸]. در واقع عامل باید به تجربیات خودش برای یادگیری متکی باشد و در هرعمل ارزیابی مناسب از محیط را به دست آورد و عمل را برای سازگاری با محیط اصلاح کند [۷].

امروزه به دلیل پیشرفت فراوان علوم به ویژه فناوری، یافتن ماشین‌هایی که به صورت خودمختار و هوشمند عمل می‌کنند آسان تر شده است. این ماشین‌ها به عنوان ربات شناخته می‌شوند. برخی از ربات‌ها به سیستم جابه‌جایی مجهز هستند که به آن‌ها این امکان را می‌دهد که از میانه یک محیط خاص حرکت کنند. این ربات‌ها به عنوان ربات‌های متحرک[9] شناخته می‌شوند [۱۰]. کنترل ربات‌های متحرک اهمیت قابل توجهی را در سال‌های اخیر به دست آورده است و تحقیقات زیادی در زمینه کنترل ردیابی ربات‌های متحرک چرخ‌دار انجام شده است [۱۱]. این گونه ربات‌ها می‌توانند از یک مکان به مکان دیگر و در محیط‌های مختلف اعم از شناخته شده و ناشناخته به صورت خود مختار و بدون کمک به اپراتور انسانی منتقل شوند و وظایف دلخواه را انجام دهند [۱۲]. ربات محیط اطراف خود را شناسایی کرده و عمل بعدی را مبتنی بر این شناخت برنامه ریزی می‌کند و سپس عمل مناسب را با استفاده از عملگرهای در دسترس انجام می‌دهد [۱۳]. ربات شبه اتومبیل ازجمله ربات‌های متحرک چرخدار می‌باشد. این ربات‌ها به یک کنترل کننده نیاز دارند تا در مسیر ایجاد شده حرکت کرده و از برخورد به موانع جلوگیری کنند [۱۴]. ربات‌های چرخ‌دار عمدتاً سیستم‌های تحریک ناقص مقید به قیود غیر هولونومیک[10] هستند [۱۵]. چنین سیستم‌هایی دارای معایبی از جمله پارامترهای غیرخطی و عدم قطعیت‌ها می‌باشند به همین دلیل نمی‌توانند از طریق بازخورد ثابت پایدار شوند. این بدان معناست که نمی‌توان نظریه کنترل خطی را برای حل مشکلات کنترل سیستم غیر هولونومیک به کار برد. به همین دلیل، این گونه مسائل کنترل غیرخطی، تعداد زیادی از پژوهش‌ها شامل تکنیک‌های مختلف کنترل خودکار را برانگیخته اند. به سبب وجود عدم قطعیت‌ها در دنیای واقعی خطای ردیابی خط سیر زمانی برای ربات‌های متحرک همیشه به وجود می‌آید و نمی‌تواند از بین رود [۱۶]. با وجود سادگی ظاهری مدل حرکتی ربات‌های متحرک چرخدار، طراحی قانون‌های کنترل پایدار برای این سیستم‌ها به دلیل وجود قیود غیرهولونومیک می‌تواند به عنوان یک چالش در نظر گرفته شود [۱۷]. به علاوه، به دلیل آن که مدل ربات‌ها به شدت غیرخطی است، برای طراحی کنترل کننده نیاز به تکنیک‌های کنترل غیرخطی از جمله کنترل تطبیقی [۱۸]،کنترل فازی [۱۹] و کنترل بهینه [۲۰] می‌باشد. در [۱۸]، بر طراحی یک کنترل کننده مدلغزشی تطبیقی برای ردیابی خط سیر زمانی تمرکز کرده است. در مرجع [۲۱]، برای ردیابی خط سیر زمانی ربات متحرک یک کنترل‌کننده پسگام پیشنهاد شده است. در [۱۹]،کنترل کننده فازی چند ورودی - چند خروجی برای ردیابی خط سیر زمانی مطلوب مختلف برای ربات متحرک طراحی شده است. در مرجع [۲۰]،

[1] Supervised learning
[2] Unsupervised learning
[3] Reinforcement learning
[4] Actions
[5] Policy
[6] Online
[7] Agent
[8] Learner
[9] Mobile robot
[10] Holonomic

[Downloaded from joc.kntu.ac.ir on 2026-11-10]
[DOR: 20.1001.1.20088345.1401.16.4.5.3]
[DOI: 10.61186/joc.16.4.57]

برای بررسی مسئله کنترل ردیابی خط سیرزمانی برای ربات‌های متحرک چرخدار غیرهولونومیک در حضور عدم قطعیت‌ها و اغتشاشات یک قانون کنترل بهینه پیشنهاد شده است. در [۲۲]، از روش کنترل بهینه و روش مقاوم برای کنترل ربات متحرک استفاده شده است که نه تنها سیستم غیرخطی را درمقابل اغتشاشات خارجی و عدم قطعیت‌ها مقاوم کرده است، بلکه تابع هزینه را بهینه می‌کند. در مرجع [۲۳]، یک طرح کنترلی شبکه عصبی تطبیقی برای یک ربات متحرک چرخدار با محدودیت‌های سرعت و قیود غیرهولونومیک بیان شده است. مسئله کنترل ردیابی برای دسته‌ای از سیستم‌های غیرخطی با غیرخطی‌های نامعلوم از طریق روش بازیگر- منتقد تطبیقی در مرجع [۲۴] بررسی شده است. یادگیری تقویتی برای یافتن الگوی مناسب جهت ردیابی مسیر ربات متحرک مبتنی بر یادگیری تقویتی در [۲۵] پیشنهاد شده است.

کنترل بهینه سیستم‌های غیرخطی موضوع مهمی در مهندسی کنترل می‌باشد. یافتن راه حل بهینه نیاز به حل معادله هامیلتن- جاکوبی- بلمن دارد که یک معادله دیفرانسیل غیر خطی است. حل صریح این معادله به طور کلی بسیار دشوار یا حتی غیر ممکن است. راه حل‌های کنترل بهینه برون خطی[1] هستند و به دانش کامل از دینامیک سیستم نیاز دارند. بنابراین، آن‌ها قادر نیستند با عدم قطعیت‌ها و تغییرات دینامیکی مقابله کنند. این امر محققان کنترل را برانگیخته است تا با توسعه کنترل کننده‌های مبتنی بر یادگیری تقویتی، خودمختاری تطبیقی را به شیوه‌ای بهینه اجرا کنند. برای تقریب راه حل معادله هامیلتن- جاکوبی- بلمن و رفع نیاز به دانش کامل در مورد دینامیک سیستم، ساختار بازیگر- منتقد به طور گسترده برای یافتن راه حل برخط برای این معادله ارائه شده است [۲۶]. در [۲۷]، از یک کنترل بهینه تطبیقی با استفاده از روش بازیگر – منتقد برای دسته‌ای از سیستم‌های غیرخطی در حضور دینامیک‌های مدل نشده پیشنهاد شده است. در [۲۸]، برای دسته‌ای از سیستم‌های پیوسته-زمان راه حل کنترل مدلغزشی مبتنی بر کنترل بهینه تطبیقی با استفاده از یادگیری تقویتی مبتنی بر روش بازیگر – منتقد پیشنهاد شده است. در [۲۹] کنترل ردیابی بهینه شده با استفاده از روش یادگیری تقویتی مبتنی بر بازیگر – منتقد برای دسته‌ای از سیستم‌های پیوسته زمان غیرخطی غیرافاین[2] بررسی شده است. همچنین، یک طرح کنترلی پسگام بهینه شده برای دسته‌ای از سیستم‌های فیدبک اکید با دینامیک‌های نامعلوم با استفاده از یادگیری تقویتی مبتنی بر بازیگر – منتقد در [۳۰] پیشنهاد شده است. در [۳۱]، یک طرح کنترل بهینه تطبیقی برای سیستم‌های غیرخطی فیدبک اکید[3] به کمک یادگیری تقویتی پیشنهاد شده است.

از آن جایی که عملکرد کنترل در حالت گذرا و حالت ماندگار در طراحی سیستم‌های کنترل از اهمیت بالایی برخوردار است، بسیاری از محققان از عملکرد از پیش تعریف شده[4] استفاده کرده‌اند. اما زمانی مطلوب برای رسیدن خطاهای سیستم کنترلی به درون یک باند تحمل از پیش تعریف شده و باقی ماندن در آن تنظیم نشده است. در مراجع [۳۲] و [۳۳] یک تابع عملکرد از پیش تعریف شده با زمان از پیش تنظیم شده برای آن که خطاهای سیستم به کران از پیش تعریف شده همگرا شوند، پیشنهاد شده است. اما زمان تنظیم در این مراجع به مشخصات پارامتری تابع عملکرد از پیش تعریف شده وابسته است [۳۴].

در این مقاله، ضمن آنکه برای تقریب راه حل معادله هامیلتن- جاکوبی- بلمن از تکنیک یادگیری تقویتی مبتنی بر شبکه عصبی بازیگر- منتقد استفاده شده است [۳۵]، نوآوری‌های زیر جهت کنترل نگاه به جلو ربات شبه اتومبیل ارائه خواهند شد:

۱- طراحی یک کنترل کننده سطح دینامیکی به منظور ردیابی خط سیر زمانی مطلوب مبتنی بر تکنیک کنترل نگاه به جلو برای یک ربات شبه اتومبیل در حضور عدم قطعیت های پارامتری و غیرپارامتری مدل های سینماتیکی و دینامیکی ربات؛

۲- تضمین یک عملکرد از پیش تعریف شده به گونه‌ای که فراجهش، نرخ همگرایی و زمان مطلوب برای رسیدن و باقی ماندن خطاهای ردیابی ربات شبه اتومبیل در کران از پیش تعریف شده که از قبل تنظیم شده است و به پارامترهای تابع عملکرد از پیش تعریف شده وابسته نمی‌باشد؛

۳- طراحی یک ساختار کنترل یادگیری تقویتی بازیگر-منتقد با یک تابع هزینه بلند مدت یا سیگنال تقویتی به منظور دستیابی به یک سیگنال کنترلی بهینه جهت تخمین و جبران دینامیک نامعین مدل ربات متحرک؛

۴- مقابله با تأثیر اغتشاشات خارجی با طراحی یک کنترل کننده مقاوم – تطبیقی کارآمد با کاهش میزان چترینگ؛

ساختار این مقاله به صورت زیر تنظیم شده است. در بخش ۲، مسأله کنترل نگاه به جلوی ربات شبه اتومبیل بیان می‌شود. بخش ۳ به طراحی کنترل کننده یادگیری تقویتی با استفاده از ساختار کنترلی سطح دینامیکی می پردازد. در بخش ۴، پایداری سیستم کنترل پیشنهادی به کمک روش لیاپانوف ارزیابی می‌گردد. نتایج شبیه سازی در بخش ۵ آورده می‌شود و نهایتاً بخش ۶ از مقاله نتیجه گیری می‌کند.

۲- فرمول بندی مسئله

۲-۱- مدل حرکت ربات شبه اتومبیل

یک ربات شبه اتومبیل با قیود غیر هولونومیک که معادلات حرکت آن به صورت زیر تعریف می‌شوند، را در نظر بگیرید:

$$M_1(p)\ddot{p} + C_1(p,\dot{p})\dot{p} + D_1\dot{p} + \tau_{d1}(t) = B_1(p)\tau_a - A^T(p)\lambda \qquad (۱)$$

که $p = [x, y, \psi, \gamma]$ نشان دهنده بردار مختصات تعمیم یافته است که (x, y) بیان کننده مختصات موقعیت و (ψ, γ) بیان کننده مختصات جهت گیری و فرمان می‌باشند. ماتریس اینرسی مثبت معین متقارن با $M_1(p) \in R^{4\times4}$، ماتریس کوریولیس و نیروی جانب مرکزگرا با

[1] Offline
[2] Affine
[3] Strict feedback
[4] Prescribed performance

[Downloaded from joc.kntu.ac.ir on 2026-11-10]
[DOR: 20.1001.1.20088345.1401.16.4.5.3]
[DOI: 10.61186/joc.16.4.57]

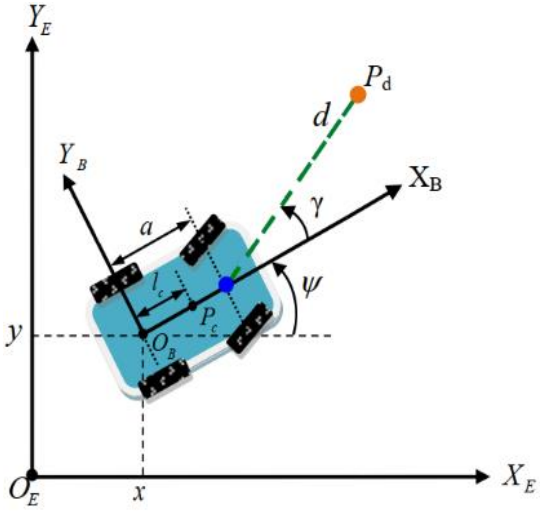
$C_1(p,\dot{p}) \in R^{4\times4}$، ماتریس میرایی و اصطکاک با $D_1 \in R^{4\times4}$، بردار اغتشاشات خارجی متغیر با زمان با $\tau_{d1} \in R^4$، ماتریس تبدیل ورودی با $B_1(p) \in R^{4\times2}$، و بردار گشتاورهای ورودی با $\tau_a \in R^2$ نمایش داده می‌شوند، $A(p) \in R^{2\times4}$ نشان دهنده ماتریس قیود مرتبه کامل و $\lambda \in R^2$ بردار ضرایب لاگرانژ است که نشان دهنده نیروهای محدودکننده می‌باشد [۳۶]. یک سیستم مکانیکی با n مختصات تعمیم یافته p که مشروط به m قید دو طرفه است و به صورت $G(p,\dot{p}) = 0$ نمایش داده می‌شوند، یک معادله هولونومیک نامیده می‌شود اگر به فرم $G(p) = 0$ در نظرگرفته شود، در غیر اینصورت یک قید سینماتیک است و معادله غیرهولونومیک را نشان می‌دهد. حال فرض برآن است که k قید هولونومیک و $m-k$ قید مستقل غیرهولونومیک موجود است که همه آن‌ها می‌توانند به فرم $A(p)\dot{p} = 0$ نوشته شوند که در رابطه ذکر شده $A(p)$ یک ماتریس مرتبه کامل $m \times n$ می‌باشد. حال اگر $S_1(p), \dots, S_{n-m}(p)$ یک مجموعه از میدان‌های برداری مستقل خطی در فضای تهی $A(p)$ باشد، یعنی $A(p)S_i = 0, i = 1, \dots, n-m$ برقرار باشد، آنگاه $S(p)$ ماتریس مرتبه کاملی است که از بردارهای $S(p) = [S_1(p) \quad S_{n-m}(p)]$ ساخته می‌شود[۳۷]. فرض می‌شود که $S(p) = [S_1(p),\ S_2(p)]^T$ ماتریس مرتبه کاملی است که ترکیبی از مجموعه‌ای از میدان‌های برداری هموار و مستقل خطی در فضای تهی $A(p)$ است، یا به عبارتی $A(p)S(p) = 0$ که در رابطه ذکر شده $S_1(p) = [cos\psi, sin\psi, \frac{1}{a}tan\gamma, 0]^T$ و $S_2(p) = [0,0,0,1]^T$ می‌باشد[۳۶]. مهم‌ترین ویژگی در معادلات سینماتیک ربات‌های متحرک چرخدار در مبحث سینماتیک حرکت وجود قیود غیرهولونومیکی می‌باشد. به عبارت دیگر، فرض می‌شود که حرکت ربات صفحه ای است و از لغزش جانبی و لغزش طولی چرخ‌های ربات در حرکت رو به جلو و رو به عقب صرف نظر می‌شود. قیود غیر هولونومیکی روابطی میان مختصات تعمیم یافته و سرعت‌های تعمیم یافته سیستم می‌باشند. از لحاظ مفهومی این قیود مربوط به عدم لغزش بین چرخ و سطحی است که ربات روی آن حرکت می‌کند. حضور قیود غیرهولونومیک یک پدیده مهم در سینماتیک این نوع سیستم‌ها می‌باشد[۳۸] و می‌توان آن را به صورت رابطه $A(p)\dot{p} = 0$ بیان کرد. مطابق با این رابطه، یک بردار شبه سرعت برای ربات شبه اتومبیل به صورت $v = [v(t), w(t)]^T$ وجود دارد که $w(t)$ و $v(t)$ به ترتیب نشان‌دهنده سرعت‌های خطی و زاویه‌ای هستند به گونه‌ای که

$$\dot{p}(t) = S(p)v(t) = S_1(p)v(t) + S_2(p)w(t) \qquad (۲)$$

با جایگذاری معادله (۲) و مشتق زمانی آن یعنی $\ddot{p}(t) = \dot{S}(p)v + S(p)\dot{v}$ در معادله (۱) و ضرب طرفین معادله در $S^T(p)$ معادله دینامیکی زیر به دست خواهد آمد:

$$M_2(p)\dot{v}(t) + C_2(p,\dot{p})v(t) + D_2(p)v + \tau_{d2}(t) = B_2(p)\tau_a \qquad (۳)$$

که بردارها و ماتریس‌های معادله (۳) به صورت $M_2(p) = S^T(p)M_1(p)S(p)$، $C_2(p,\dot{p}) = S^T(p)M_1(p)\dot{S}(p) + S^T(p)C_1(p,\dot{p})S(p)$، $D_2(p) = S^T(p)D_1S(p)$، $\tau_{d2}(t) = S^T(p)\tau_{d1}(t)$ و $B_2(p) = S^T(p)B_1(p)$ تعریف می‌شوند [۳۶].



شکل۱: پیکره بندی ربات شبه اتومبیل

۲-۲-نمایش فضای حالت

برای توسعه کنترل کننده نگاه به جلو، فرم فضای حالت معادله سینماتیک (۲) و دینامیکی (۳) را می‌توان به صورت زیر نشان داد:

(۴)

$$\dot{x} = \begin{bmatrix}\dot{p}\\ \dot{v}\end{bmatrix} = \underbrace{\begin{bmatrix}S(p)v(t)\\ 0\end{bmatrix}}_{f(x)} + \underbrace{\begin{bmatrix}0\\ M_2^{-1}B_2\end{bmatrix}}_{g(x)}\tau_a + \underbrace{\begin{bmatrix}0\\ M_2^{-1}(-C_2v - D_2v - \tau_{d2})\end{bmatrix}}_{q(x)}$$

که $x = [p, v] \in R^6$ بردار حالت را نشان می‌دهد.

تعریف۱: همچنین، معادله حالت خط سیر زمانی مطلوب هموار و کراندار $y_d(t): [0,\infty) \to R^2$ به صورت زیر تعریف می‌شود:

(۵)

$$\dot{x}_d = \begin{bmatrix}S(p_d)v(t)\\ 0\end{bmatrix} + \begin{bmatrix}0\\ I\end{bmatrix}\tau_{ad} \quad , y_d = h(p_d)$$

که $\dot{x}_d$ بردار حالت مطلوب، $S(p_d)v(t)$ معادله سینماتیک خط سیرزمانی مطلوب، τ_{ad} سیگنال کنترل حالت مطلوب، $y_d = h(p_d)$ معادله خط سیرزمانی مطلوب هموار و کراندار است.

۲-۳-هدف کنترلی

هدف نهایی کنترل در این مقاله پژوهشی، طراحی یک قانون کنترل برای ردیابی ربات شبه اتومبیل با پارامترهای مشخص شده در شکل ۱ با وجود اغتشاشات خارجی و عدم قطعیت‌ها می‌باشد، به گونه‌ای که خطای ردیابی $e(t) = y(t) - y_d(t)$ که در فضای دو بعدی ایجاد می‌شود، کراندار نهایی یکنواخت باشد. در رابطه ذکر شده $y(t) \in R^2$ نشان دهنده موقعیت یک نقطه کنترل مجازی است که نقطه نگاه به جلو نامیده می‌شود و در فاصله تعریف شده از محور چرخ‌های جلویی ربات شبه اتومبیل قرار

[Downloaded from joc.kntu.ac.ir on 2026-11-10]
[DOR: 20.1001.1.20088345.1401.16.4.5.3]
[DOI: 10.61186/joc.16.4.57]

گرفته است. به منظور دستیابی به هدف کنترلی ذکر شده لم‌ها، خاصیت‌ها و فرض‌های زیر در نظر گرفته می‌شوند:

لم ۱: نامساوی‌های زیر برای بر $x,y \in R^n$ و $N = N^T > 0$ برقرار است [۳۹]:

$$tr\{x^T y\} \le \frac{1}{2}\|x\|_F^2/k^2 + \frac{1}{2}k^2\|y\|_F^2 \quad (۶)$$

$$\lambda_{min}(N)\|x\|^2 \le x^T N x \le \lambda_{max}(N)\|x\|^2 \quad (۷)$$

که $\|x\|_F$ نشان دهنده نرم فروبنیوس می‌باشد و به صورت $\|x\|_F := \sqrt{tr\{x^T x\}}$ تعریف می‌شود.

لم ۲: نامساوی زیر برای هر $\gamma_d > 0$ و برای $\forall x \in R^n, h \in R$ برقرار است [۴۰] :

$$h\|x\| - x^T h tanh(0.2785hx/\gamma_d) \le n\gamma_d \quad (۸)$$

خاصیت ۱: برای ماتریس اینرسی $M_2(p)$، شرط $M_2(p) = M_2^T(p) > 0$ برقرار است و طبق نامساوی ریلی – ریتز که برای هر ماتریس متقارن برقرار است، داریم:

$$\lambda_{m2}\|x\|^2 \le x^T M_2(p) x \le \lambda_{M2}\|x\|^2, \forall x \in \Re^2, p \in \Re^4 \quad (۹)$$

که در این رابطه $\lambda_{M2} := \max_{\forall p \in \Re^4} \lambda_{max}(M_2(p))$ و $\lambda_{m2} := \min_{\forall p \in \Re^4} \lambda_{min} M_2(p)$ و $0 \le \lambda_{m2} \le \lambda_{M2} < \infty$ می‌باشد.

خاصیت ۲: ماتریس $\dot{M}_2(p) - 2C_2(p,v)$ پادمتقارن است. براین اساس، می‌توان اثبات نمود که تساوی زیر برقرار است:

$$x^T\left(\dot{M}_2(p) - 2C_2(p,v)\right)x = 0, \forall x,v \in \Re^2, p \in \Re^4 \quad (۱۰)$$

خاصیت ۳: ماتریس میرایی $D_2(p)$ یک ماتریس متقارن و مثبت معین است. پس شرط $D_2(p) = D_2^T(p) > 0$ برقرار است و طبق نامساوی ریلی – ریتز که برای هر ماتریس متقارن برقرار است، داریم:

$$\lambda_{d2}\|x\|^2 \le x^T D_2(p) x \le \lambda_{D2}\|x\|^2, \forall x \in \Re^2, p \in \Re^4 \quad (۱۱)$$

که در این رابطه $\lambda_{D2} := \max_{\forall p \in \Re^4} \lambda_{max}(D_2(p))$ و $\lambda_{d2} := \min_{\forall p \in \Re^4} \lambda_{min} D_2(p)$ و $0 \le \lambda_{d2} \le \lambda_{D2} < \infty$ می‌باشد.

فرض ۱: سیگنال‌های اغتشاش کراندار هستند به صورتی که رابطه $\|\tau_{d1}(t)\| \le \Delta_{\tau_d}$ برقرار است که $\Delta_{\tau_d} \in R^+$ اسکالر ثابت مثبت است.

فرض ۲: برای خط سیر زمانی مطلوب $y_d(t)$، مشتق مرتبه اول $\dot{y}_d(t)$ و مشتق مرتبه دوم آن $\ddot{y}_d(t)$ سیگنال‌هایی کراندار هستند به نحوی که روابط $sup_{t\ge 0}\|y_d(t)\| < B_y$، $sup_{t\ge 0}\|\dot{y}_d(t)\| < B_v$ و $sup_{t\ge 0}\|\ddot{y}_d(t)\| < B_a$ برقرار هستند و B_y، B_v و B_a ثابت‌های مثبت می‌باشند [۳۶].

۲-۴- تبدیل مختصات

به منظور دستیابی به هدف کنترلی ذکرشده برای ربات شبه اتومبیل از طریق روش کنترل نگاه به جلو مطابق با شکل ۱، نقطه P_d که در فاصله d از نقطه P_o که نقطه مرکزی محور چرخ‌های جلویی ربات شبه اتومبیل می‌باشد، تعریف می‌شود و مختصات آن در چارچوب زمین ثابت $\{O_E, X_E, Y_E\}$ به صورت زیر بیان می‌شود:

$$y = h(p) = \begin{bmatrix} x + acos\psi + dcos(\psi+\gamma) \\ y + asin\psi + dsin(\psi+\gamma) \end{bmatrix}. \quad (۱۲)$$

در رابطه فوق d یک پارامتر مثبت است و فاصله نگاه به جلو نامیده می‌شود. با مشتق گیری از رابطه (۱۲) و جایگذاری رابطه (۲) در آن، داریم:

$$\dot{y} = L_f h(x) + L_g h(x)\tau_a + L_q h(x) = J(p)\upsilon \quad (۱۳)$$

که $L_f h(x) = \nabla h f$، $L_g h(x) = \nabla h g$ و $L_q h(x) = \nabla h q$ مشتق لی h در جهت مسیر بردارهای f، g و q و ∇h گرادیان h را نشان می‌دهند [۴۱] و ماتریس $J(p)$ به صورت زیر به دست می‌آید:

$$J(p) = J_h(p)S(p) = \begin{bmatrix} cos\psi - (1/a)tan\gamma(asin\psi + dsin(\psi+\gamma)), -d\sin(\psi+\gamma) \\ sin\psi + (1/a)tan\gamma(acos\psi + d\cos(\psi+\gamma)), d\cos(\psi+\gamma) \end{bmatrix} \quad (۱۴)$$

که در رابطه ذکر شده $J_h(p) := \partial h(p)/\partial p$ بیانگر ماتریس ژاکوبین است.

فرض ۳: زاویه فرمان ربات شبه اتومبیل کراندار است به گونه‌ای که $|\gamma| \le \gamma_{max} \le \frac{\pi}{6}$ در نظر گرفته می‌شود تا از هرگونه تکینگی ممکن در تجزیه و تحلیل پایداری جلوگیری شود.

۲-۵- شبکه عصبی شعاعی پایه

شبکه عصبی تابع پایه شعاعی به عنوان یک تخمین‌گر قوی و خطی در پارامتر در کاربردهای کنترلی به دلیل توانایی تخمین کلی به طور گسترده استفاده می‌شود. اگر یک تابع پیوسته دلخواه به صورت $\xi(x):\Re^{n_o} \to \Re^{n_i}$ در نظر گرفته شود، آنگاه یک شبکه عصبی پایه شعاعی به صورت رابطه زیر وجود خواهد داشت:

$$\xi(x) = W^T\sigma(x) + e_w(x) \quad (۱۵)$$

که $x \in \Re^{n_i}$ نشان دهنده بردار ورودی شبکه عصبی، $W \in \Re^{n_h \times n_o}$ بیان کننده ماتریس‌های وزن شبکه عصبی که n_h تعداد گره‌های لایه مخفی و n_o تعداد گره‌های لایه خروجی می‌باشد، $e_w = [e_{w1}, e_{w2}, \dots e_{wn_o}]^T$ بردار خطای تخمین شبکه عصبی را نشان می‌دهد که در مجموعه محدب[1] $U \subset \Re^n$ کراندار است و $\sigma(x) =$

[1] Compact set

[Downloaded from joc.kntu.ac.ir on 2026-11-10]
[DOR: 20.1001.1.20088345.1401.16.4.5.3]
[DOI: 10.61186/joc.16.4.57]

$[\sigma_1(x), \dots, \sigma_{n_h}(x)]^T$ تابع پایه گوسی است که به صورت زیر نشان داده می‌شود:

(۱۶)

$$\sigma_k(x) = \exp\left(-(\mathrm{x} - \mu_k)^T (\mathrm{x} - \mu_k) / \lambda_k^2\right), \quad k = 1, 2, \cdots, n_h$$

که $\mu_k = [\mu_{k1}, \mu_{k2}, \dots, \mu_{kn}]^T$ بردار مراکز و λ_k انحراف استاندار می‌باشند. بر اساس ویژگی تقریب کلی شبکه عصبی شعاعی پایه، یک ماتریس وزنی ایده آل به صورت W^* وجود دارد به گونه‌ای که $\|e_w\|$ را برای همه $x \in U \subset \Re^n$ به صورت زیرکمینه خواهد کرد:

(۱۷)

$$\begin{cases} \xi(x) = W^*\sigma(x) + e_w^*(x), & \forall x \in U \subset \Re^n, \\ W^* := \arg\min \{sup_{x\epsilon U} \| f(x) - W\sigma(x) \|\} \end{cases}$$

که $e_w^*(x)$ خطای تقریب برای $W = W^*$ است. به دلیل آن که ماتریس وزنی شبکه عصبی W^* نامعلوم است، در این مقاله یک تخمین از آن به صورت $\widehat{W}$ در نظر گرفته می‌شود.

فرض ٤: ماتریس وزن ایده آل شبکه عصبی روی یک مجموعه محدب U کراندار است به نحوی که رابطه $\|W^*\|_F \leq B_w$ برقرار است و $B_w \in \Re^+$ نامعلوم است.

فرض ٥: خطای تابعی شبکه عصبی کراندار است به گونه‌ای $|e_w| \leq B_{ew}$ برقرار است که $i = 1,2, \dots, n_o$ و $B_{ew} \in \Re^+$ نامعلوم هستند.

فرض ٦: تابع پایه گوسی $\sigma(x)$ کراندار است به طوری که $\|\sigma(x)\| \leq \bar{\sigma}$ برقرار است که $\bar{\sigma} \in \Re^+$ کران بالای آن می‌باشد [۳۶].

۳- نتایج اصلی

۳-۱- طراحی کنترل کننده پیشنهادی

تکنیک کنترل پسگام یکی از روش‌های مشهور برای کنترل سیستم‌های غیرخطی فیدبک اکید با عدم قطعیت پارامتری است. اما این روش متحمل معضلی به نام انفجار پیچیدگی[1] می‌باشد. برای حل این مسئله، مرجع [۴۲] تکنیک کنترل سطح دینامیکی را از طریق معرفی یک فیلتر مرتبه اول در بین هر دو مرحله متوالی از روش طراحی کنترل پسگام پیشنهاد کرده است. در این بخش، به کمک معادلات دینامیک و سینماتیک ربات شبه اتومبیل که در بخش‌های قبلی به معرفی آن پرداخته شد، یک کنترل‌کننده سطح دینامیکی که از یک کنترل‌کننده مجازی Z_d برای جبران اثر سینماتیک حرکت ربات و یک کنترل کننده واقعی τ_a برای جبران اثر دینامیک تشکیل شده است، طراحی می‌شود. مطابق با روند طراحی کنترل‌کننده سطح دینامیکی، خطای سینماتیک و دینامیک ربات شبه اتومبیل به ترتیب به صورت روابط زیر تعریف می‌شوند:

$$Z_{1j} = e \quad (۱۸)$$

$$Z_2 = \upsilon - Z_f \quad (۱۹)$$

باتوجه به آن که ربات شبه اتومبیل یک سیستم مرتبه دوم است، پس دارای دو زیر سیستم مکانیکی می‌باشد که زیر سیستم مکانیکی اول با Z_{1j} و زیر سیستم مکانیکی دوم با Z_2 نمایش داده شده است. در زیر سیستم مکانیکی اول e بیان کننده خطاهای ردیابی می‌باشد و به صورت $e = [e_x, e_y]$ می‌باشد. در زیر سیستم مکانیکی دوم υ نشان دهنده بردار سرعت است و Z_f سیگنال فیلتر شده کنترل مجازی Z_d می‌باشدکه از یک فیلتر مرتبه اول عبورداده شده است. قبل از طراحی کنترل‌کننده، تکنیک عملکرد از پیش‌تعیین شده را به خطای ردیابی Z_{1j} در بخش بعدی اعمال می‌کنیم.

۳-۲- تابع عملکرد از پیش تعریف شده

ایده اصلی عملکرد از پیش تعریف شده که توسط مرجع [۴۳] در سال ۲۰۰۸ معرفی شده است، تحمیل قیودی کراندار بر روی متغیرهای حالت سیستم است. قیود کراندار از طریق یک تابع عملکرد از پیش تعریف شده توسط طراح معین می‌شوند. سپس، برای کاهش سختی طراحی کنترل‌کننده مناسب، سیستم کنترل مقید از طریق یک تابع نگاشت در فضای مقید به یک سیستم کنترل نامقید تبدیل می‌شود. مزیت اصلی تابع عملکرد از پیش تعریف شده آن است که تا زمانی که کرانداری تابع نگاشت تضمین شود، کرانداری قیود حتی اگر عدم قطعیت و اغتشاشات خارجی قوی در سیستم وجود داشته باشد، تضمین خواهد شد [۳۳]. عملکرد حالت گذرا و حالت ماندگار از پیش تعریف شده برای کنترل بردار خطای ردیابی ربات شبه اتومبیل $Z_{1j} = [z_x, z_y]$ به فرم قیود زیر تعریف می‌شود:

$$-\eta_{-j}(t) < Z_{1j}(t) < \bar{\eta}_j(t), j = (x, y). \quad (۲۰)$$

تابع عملکرد از پیش تعریف شده $-\eta_{-j}(t)$ و $\bar{\eta}_j(t)$ در این مقاله به صورت زیر بیان می‌شوند:

(۲۱)

$$-\eta_{-j}(t) = \begin{cases} -\left(\eta_{-j,0} - \eta_{-j,\infty}\right) \exp\left(-c_j \dfrac{T_j t}{T_j - t}\right) - \eta_{-j,\infty}, & t < T_j \\ -\eta_{-j,\infty}, & t \geq T_j \end{cases}$$

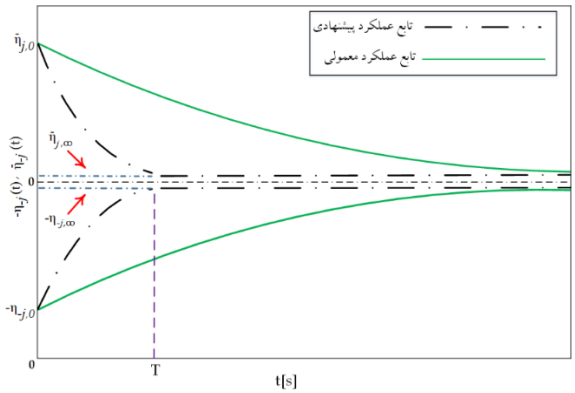
(۲۲)

$$\bar{\eta}_j(t) = \begin{cases} \left(\bar{\eta}_{j,0} - \bar{\eta}_{j,\infty}\right) \exp\left(-c_j \dfrac{T_j t}{T_j - t}\right) + \bar{\eta}_{j,\infty}(t), & t < T_j \\ \bar{\eta}_{j,\infty}, & t \geq T_j \end{cases}$$

که $\exp(\bullet)$ نشان دهنده تابع نمایی، $-\eta_{-j,0}$ و $\bar{\eta}_{j,0}$ مقدارهای اولیه تابع عملکرد از پیش تعریف شده، $-\eta_{-j,\infty}$ و $\bar{\eta}_{j,\infty}$ به ترتیب کران پایین و کران بالای باند نهایی از پیش تعریف شده $[-\eta_{-j,\infty}, \bar{\eta}_{j,\infty}]$، $c_j > 0$ نرخ همگرایی تابع عملکرد از پیش تعریف شده و T_j زمان تنظیم تابع عملکرد از پیش تعریف شده برای رسیدن به باند نهایی $[-\eta_{-j,\infty}, \bar{\eta}_{j,\infty}]$ می‌باشد.

[1] Explosion of complexity

[Downloaded from joc.kntu.ac.ir on 2026-11-10]
[DOR: 20.1001.1.20088345.1401.16.4.5.3]
[DOI: 10.61186/joc.16.4.57]



شکل۲: نمایش تابع عملکرد از پیش تعریف شده

تذکر ۱: همانطور که در شکل ۲ نشان داده شده است، تابع عملکرد از پیش تعریف شده سنتی که به صورت رابطه $-\eta_{-j}(t) = -\left(\eta_{-j,0}(t) - \eta_{-j,\infty}(t)\right)\exp(-c_j t) - \eta_{-j,\infty}$ و $\bar{\eta}_j(t) = \left(\bar{\eta}_{j,0}(t) - \bar{\eta}_{j,\infty}(t)\right)\exp(-c_j t) + \bar{\eta}_{j,\infty}(t)$ تعریف می‌شوند با گذشت زمان $(t \to \infty)$ به $-\eta_{-j,\infty}(t)$ و $\bar{\eta}_{j,\infty}(t)$ خواهند رسید. در حالی توابع عملکرد پیشنهاد شده در این مقاله که در روابط (۲۱) و (۲۲) تعریف شده است، در زمان از پیش تنظیم شده T_j می‌توانند به ترتیب به $-\eta_{-j,\infty}(t)$ و $\bar{\eta}_{j,\infty}(t)$ برسد. پارامترهای $-\eta_{-j,0}$، $\bar{\eta}_{j,0}$، $-\eta_{-j,\infty}$، $\bar{\eta}_{j,\infty}$ و T_j می‌توانند مطابق با نیازهای عملکردی تنظیم شوند [۳۴].

۳-۳- تابع نگاشت خطا

تبدیل خطای زیر برای به دست آوردن عملکرد از پیش تعریف شده برای ارزیابی خطای Z_{1j} معرفی می‌شود:

$$\varepsilon_j := T_{Z_{1j}}\left(\hat{Z}_{1j}(t)\right), j = x, y \qquad (۲۳)$$

که $\hat{Z}_{1j}(t)$ خطای ردیابی نرمالیزه شده می‌باشد و به صورت $\hat{Z}_{1j}(t) = Z_{1j}(t)/\eta_j$ تعریف می‌شود. تابع تبدیل $T_{Z_{1j}}\left(\hat{Z}_{1j}(t)\right): \boldsymbol{\Omega}_{Z_{1j}} \to \Re$ نگاشت اکیداً افزایشی هموار را نشان می‌دهد که $\boldsymbol{\Omega}_{Z_{1j}} := \{\hat{Z}_{1j}(t): \hat{Z}_{1j} \in (-a_{1j}, b_{1j})\}$ می‌باشد. اگر $T_{Z_{1j}}(0) = 0$ ، $-a_{1j}\eta_{-j}(0) < Z_{1j}(0) < b_{1j}\bar{\eta}_j(0)$ برقرار باشد و کنترل کننده $\varepsilon_j(t) \in L_\infty, \forall t \geq 0$ را تضمین کند، آنگاه $\hat{Z}_{1j}(t) \in \boldsymbol{\Omega}_{Z_{1j}}$ و شرایط رابطه (۲۰) برآورده می‌شود [۳۹]، [۴۳]، [۴۴]. همچنین، اگر کنترل‌کننده تضمین کند که ε_j با گذشت زمان به صفر میل می‌کند، آنگاه Z_{1j} نیز به صفر میل کرده و هدف کنترلی محقق می‌شود. در ادامه، با مشتق گیری از رابطه (۲۳) داریم:

(۲۴)

$$\dot{\varepsilon}_j := \frac{1}{\eta_j}\frac{\partial T_{Z_{1j}}\left(\hat{Z}_{1j}(t)\right)}{\partial \hat{Z}_{1j}(t)}\left(\dot{Z}_{1j}(t) - \frac{\dot{\eta}_j(t)}{\eta_j(t)}Z_{1j}(t)\right), j = x, y$$

رابطه (۲۴) را می‌توان به فرم برداری به صورت زیر بازنویسی کرد:

$$\dot{\varepsilon}_j(t) = R\left(\dot{Z}_{1j} + \Lambda Z_{1j}\right), j = x, y \qquad (۲۵)$$

که R و Λ به صورت رابطه (۲۶) و (۲۷) می‌باشند:

(۲۶)

$$R = diag\left[\frac{1}{\eta_x}\frac{\partial T_{Z_{1x}}(\hat{Z}_{1x})}{\partial \hat{Z}_{1x}(t)}, \frac{1}{\eta_y}\frac{\partial T_{Z_{1y}}(\hat{Z}_{1y})}{\partial \hat{Z}_{1y}(t)}\right]$$

(۲۷)

$$\Lambda = diag\left[-\frac{\dot{\eta}_x(t)}{\eta_x(t)}, -\frac{\dot{\eta}_y(t)}{\eta_y(t)}\right]$$

برای آن که خطای مقید $Z_{1j}(t)$ از طریق تابع عملکرد از پیش تعریف شده (۲۰) به خطای تبدیل شده غیر مقید ε_j تبدیل شود، از نگاشت خطای زیر که یک تابع اکیداً افزایشی است، استفاده می‌شود:

(۲۸)

$$\varepsilon_j = T_j\left(\hat{Z}_{1j}(t)\right) = \tan\left[\frac{\pi}{2} \times \frac{2\hat{Z}_{1j} - b_j + a_j}{b_j + a_j}\right], j = x, y$$

و $\frac{\partial T_{Z_{1j}}\left(\hat{Z}_{1j}(t)\right)}{\partial \hat{Z}_{1j}(t)}$ به صورت زیر به دست خواهد آمد:

(۲۹)

$$\frac{\partial T_{Z_{1j}}\left(\hat{Z}_{1j}(t)\right)}{\partial \hat{Z}_{1j}(t)} = \frac{\pi}{b_j + a_j} sec^2\left[\frac{\pi}{2} \times \frac{2\hat{Z}_{1j} - b_j + a_j}{b_j + a_j}\right] \quad j = x, y$$

با مشتق گیری از خطای رابطه (۱۸) و جایگذاری رابطه (۱۳) در آن داریم:

$$\dot{Z}_{1j} = Jv - \dot{y}_d \qquad (۳۰)$$

برای اعمال عملکرد از پیش تعریف شده به سیستم، $\dot{Z}_{1j}$ از رابطه (۳۰) در رابطه (۲۵) جایگذاری می‌شود:

$$\dot{\varepsilon}_j = R(Jv - \dot{y}_d) + R\Lambda Z_{1j}. \qquad (۳۱)$$

در ضمن، خطای خروجی فیلتر به صورت زیر تعریف می‌شود:

$$Z_e = Z_f(t) - Z_d(t). \qquad (۳۲)$$

با به دست آوردن v از رابطه (۱۹) و Z_f از رابطه (۳۲)، v به فرم $Z_2 + Z_e + Z_d$ به دست خواهد آمد که در رابطه (۳۱) جایگذاری می‌شود:

(۳۳)

$$\dot{\varepsilon}_j = RJ(Z_2 + Z_e + Z_d) - R\dot{y}_d + R\Lambda Z_1$$

حال، قانون کنترل مجازی برای پایداری رابطه (۳۳) به صورت زیر تعریف می‌شود:

[Downloaded from joc.kntu.ac.ir on 2026-11-10]
[DOR: 20.1001.1.20088345.1401.16.4.5.3]
[DOI: 10.61186/joc.16.4.57]

$$Z_d = (RJ)^{-1}(-K_1\varepsilon_j + R\dot{y}_d - R\Lambda Z_1) \qquad (۳۴)$$

که $K_1 \in R^{2\times 2}$ یک ماتریس ضریب قطری مثبت معین می‌باشد. سپس، کنترل کننده مجازی از یک فیلتر پایین گذر عبور داده می‌شود:

$$\tau_f \dot{Z}_f + Z_f = Z_d \quad , \quad Z_f(0) = Z_d(0) \qquad (۳۵)$$

که $\tau_f \in R^+$ ثابت طراحی است. معادله خطای سینماتیک حلقه بسته برای رابطه (۳۳) با جایگذاری قانون کنترل مجازی در آن به صورت زیر به دست می‌آید:

$$\dot{\varepsilon}_j = RJ(Z_2 + Z_e) - K_1\varepsilon_j \qquad (۳۶)$$

با مشتق گیری از رابطه (۳۲) و جایگذاری $\dot{Z}_f$ از معادله (۳۵) در آن رابطه $\dot{Z}_e = \dot{Z}_f - \dot{Z}_d = -Z_e/\tau_f - \dot{Z}_d$ به دست خواهد آمد که با توجه به مرجع [۴۲] می توان نتیجه گرفت که $\|\dot{Z}_e + Z_e/\tau_f\| \le \rho$ که ρ یک تابع پیوسته از سیگنال‌های حلقه بسته و مشتقات آن‌ها می‌باشد و در یک بازه محدود دارای یک مقدار ماکزیمم است و به صورت $\pi_2 = sup\{\rho(t)\}$ در نظر گرفته می‌شود. سپس، نامساوی زیر نتیجه گیری می‌گردد:

$$\dot{Z}_e^T Z_e \le -\left(\frac{1}{\tau_f} - 1\right)\|Z_e\|^2 + \frac{1}{4}\rho^2. \qquad (۳۷)$$

مطابق با روند طراحی روش کنترل سطح دینامیکی، تابع لیاپانوفی که برای زیر سیستم اول در نظر گرفته می‌شود، به صورت زیر تعریف می‌شود:

$$L_1 = \frac{1}{2}\varepsilon_j^T\varepsilon_j + \frac{1}{2}Z_e^T Z_e. \qquad (۳۸)$$

ضرایب کنترل‌کننده باید به نحوی انتخاب شوند که شرایط زیر برقرار شود:

$$\lambda_{min}\{K_1\} \ge 0.5\lambda_{max}\{RJ\},$$

با مشتق‌گیری از رابطه (۳۸) و جایگذاری روابط (۳۶) و (۳۷) در آن، می توان نوشت:

(۳۹)

$$\begin{aligned}\dot{L}_1 &= \varepsilon_j^T\dot{\varepsilon}_j + Z_e^T\dot{Z}_e \\ &\le -(\lambda_{min}\{K_1\} - 0.5\lambda_{max}\{RJ\})\|\varepsilon_j\|^2 + \varepsilon_j^T RJ Z_2 \\ &\quad -(1/\tau_f - 1 - 0.5\lambda_{max}\{RJ\})\|Z_e\|^2 + \frac{1}{4}\rho^2\end{aligned}$$

با مشتق زمانی از بردار خطای سرعت تعریف شده در رابطه (۱۹) و جایگذاری رابطه (۳) را در آن و ضرب طرفین رابطه در M_2، داریم:

(۴۰)

$$\begin{aligned}M_2(p)\dot{Z}_2 &= B_2(p)\tau_a - C_2(p,\upsilon)\upsilon(t) \\ &\quad -D_2(p)\upsilon(t) - \tau_{d2} - M_2(p)\dot{Z}_f.\end{aligned}$$

معادله فوق، معادله خطای حلقه باز در سطح دینامیک است.

۳-۴- شبکه عصبی بازیگر

در این بخش، شبکه عصبی بازیگر در ساختار کنترل تقویتی معرفی می‌شود. با جایگذاری $\upsilon = Z_2 + Z_f$ در رابطه (۴۰)، داریم:

(۴۱)

$$\begin{aligned}M_2(p)\dot{Z}_2 &= B_2(p)\,\tau_a - C_2(p,\upsilon)Z_2 - D_2(p)Z_2 \\ &\quad -\tau_d + \xi(p,\upsilon,\ Z_f,\ \dot{Z}_f)\end{aligned}$$

که $\xi(p,\upsilon,\ Z_f,\ \dot{Z}_f)$ شامل دینامیک‌های غیرخطی و عدم قطعیت ها می‌باشد و به صورت زیر تعریف می‌شود:

(۴۲)

$$\begin{aligned}\xi(p,\upsilon,\ Z_f,\ \dot{Z}_f) &= -C_2(p,\upsilon)\,Z_f - D_2(p)\,Z_f - \\ &\quad M_2\,\dot{Z}_f\end{aligned}$$

که می‌تواند توسط شبکه عصبی تابع شعاعی پایه به صورت زیر تخمین زده شود:

$$\xi(p,\upsilon,\ Z_f,\ \dot{Z}_f) = W_a^T\,\sigma_a(x) + e(x) \qquad (۴۳)$$

که $x = [p^T, \upsilon^T,\ Z_f{}^T,\ \dot{Z}_f{}^T]^T$ ورودی‌های شبکه عصبی بازیگر هستند. با جایگذاری رابطه (۴۳) در رابطه (۴۱) سیستم حلقه باز مبتنی بر شبکه عصبی بازیگر به صورت رابطه (۴۴) به دست می‌آید:

(۴۴)

$$\begin{aligned}M_2(p)\dot{Z}_2 &= B_2(p)\,\tau_a - C_2(p,\upsilon)Z_2 - D_2(p)Z_2 \\ &\quad -\tau_d + W_a^T\,\sigma_a(x) + e(x)\end{aligned}$$

که W_a یک ماتریس وزنی ایده آل برای شبکه عصبی بازیگر است و باید با تخمین خود که $\widehat{W}_a$ می‌باشد، جایگزین شود. در ادامه $\widehat{W}_a$ توسط یک قانون آموزش تطبیقی بر خط تخمین زده خواهد شد.

۳-۵- شبکه عصبی منتقد

در این بخش، شبکه عصبی منتقد در ساختار یادگیری تقویتی ارائه می شود. برای این منظور، تابع هزینه بلند مدت یا سیگنال یادگیری تقویتی شامل شبکه عصبی منتقد ($W_c^T\sigma_c(x)$) به صورت زیر پیشنهاد می‌شود:

$$S_c = \underbrace{\upsilon - Z_f}_{Z_2} + \left\|\underbrace{\upsilon - Z_f}_{Z_2}\right\| \underbrace{W_c^T\sigma_c(x)}_{CNN} \qquad (۴۵)$$

سیگنال یادگیری رابطه (۴۵) شامل سیگنال منتقد اولیه Z_2 و یک سیگنال منتقد ثانویه $\|Z_2\|\,W_c^T\sigma_c(x)$ می‌باشد. هدف از تعریف سیگنال یادگیری تقویتی در رابطه (۴۵) رسیدن به یک سیگنال کنترلی بهینه است که هزینه کنترل را با استفاده از شبکه عصبی بازیگر و شبکه عصبی منتقد کاهش دهد. از آن جایی که ماتریس وزن ایده آل W_c برای شبکه عصبی منتقد نامعلوم است، از تخمین آن $\widehat{W}_c$ استفاده می‌شود. بنابراین $\hat{S}_c$ به عنوان تخمین S_c مطابق با رابطه (۴۶) تعریف خواهد شد:

$$\hat{S}_c = \upsilon - Z_f + \|\upsilon - Z_f\|\,\widehat{W}_c^T\sigma_c(x) \qquad (۴۶)$$

[Downloaded from joc.kntu.ac.ir on 2026-11-10]
[DOR: 20.1001.1.20088345.1401.16.4.5.3]
[DOI: 10.61186/joc.16.4.57]

که در ادامه یک قانون تطبیق آنلاین برای به روز رسانی $\widehat{W}_c$ طراحی خواهد شد. از $\hat{S}_c$ برای ایجاد قانون به روز رسانی برخط وزن شبکه عصبی بازیگر استفاده خواهد شد. شبکه عصبی بازیگر نه تنها سیگنال کنترلی را برای ردیابی مسیر مطلوب تولید و غیر خطی‌های نامعلوم را جبران می‌کند، بلکه همچنین تابع هزینه بلند مدت $\hat{S}_c$ را به حداقل می‌رساند. در همین حال، شبکه عصبی منتقد تابع هزینه بلند مدت را به منظور تنظیم شبکه عصبی بازیگر تخمین می‌زند. بنابراین، توسط یادگیری تطبیقی، تخمین گر $\widehat{W}_a^T\sigma_a(x)$ به هزینه بهینه همگرا خواهد شد و یک کنترل‌کننده بهینه یا تقریباً بهینه به دست خواهد آمد [۴۴].

۳-۶- طراحی کنترل کننده

قانون کنترل برای ربات شبه اتومبیل به صورت رابطه (۴۷) پیشنهاد می‌شود:

(۴۷)

$$\tau_a = B_2^{-1}(-K_2Z_2 - \widehat{W}_a^T\sigma_a(x) - \mathrm{R}^TJ^T\varepsilon_j - u_r)$$

که $K_2 \in \mathcal{R}^{2\times2}$ بهره تناسبی می‌باشد. عبارت u_r یک کنترل کننده مقاوم تطبیقی است که به صورت زیر تعریف می‌شود:

$$u_r = \hat{h}Tanh\,(0.2785\ \hat{h}\ Z_2/C_n) \qquad (۴۸)$$

که C_n یک لایه مرزی برای کنترل کننده مقاوم تطبیقی است. عبارت $\hat{h}$ به صورت $\hat{h} = H_f\hat{r}_f$ بیان می‌شود که تخمین $h = H_fr_f$ می‌باشد. در این رابطه H_f بردار رگرسور برای ربات شبه اتومبیل را نشان می‌دهد که به صورت زیر تعریف می‌شود:

(۴۹)

$$H_{\mathrm{f}} = [1, -1, \|\widehat{W}_c\|_F, \|\widehat{W}_a\|_F \|\widehat{W}_c\|_F, , \|\widehat{W}_c\|_F \|\widehat{W}_a\|_F]$$

و $r_f = [r_1, r_2, r_3, r_4, r_5, r_6]^T$ بردار پارامترهای نامعلوم هستند و ثابت‌های کران بالای h را بیان می‌کنند که در بخش بعدی معرفی خواهند شد. قوانین تطبیق برای ماتریس‌های وزنی شبکه‌های عصبی بازیگر و منتقد با استفاده از اصلاح سیگما[1] به صورت زیر طراحی می‌شوند:

(۵۰)

$$\dot{\widehat{W}}_a = \Gamma_a\sigma_a(x)(Z_2 + \|Z_2\|\widehat{W}_c^T\sigma_c(x))^T - \Gamma_a\delta_a\widehat{W}_a$$

(۵۱)

$$\dot{\widehat{W}}_c = \Gamma_c\|Z_2\|\sigma_c(x)(\widehat{W}_a^T\sigma_a(x))^T - \Gamma_c\delta_c\widehat{W}_c$$

(۵۲)

$$\dot{\hat{r}}_{\mathrm{f}} = \Gamma_{r_f}\|Z_2\|H_f^T - \delta_{r_f}\Gamma_{r_f}\hat{r}_{\mathrm{f}}$$

که Γ_a ، Γ_c و Γ_{r_f} بهره‌های تطبیقی و δ_a ، δ_c و δ_{r_f} پارامترهای طراحی هستند. باید توجه شود که شبکه عصبی منتقد از طریق دریافت فیدبک از شبکه عصبی بازیگر در رابطه (۵۱) آموزش داده می‌شود و سپس از طریق تخمین تابع منتقد (۴۶) در شبکه عصبی بازیگر (۵۰) این شبکه عصبی مؤثرتر آموزش داده می‌شود [۴۴]. با جایگذاری قانون کنترل (۴۷) در (۴۴) دینامیک حلقه بسته برای ربات شبه اتومبیل به دست خواهد آمد:

(۵۳)

$$M_2\dot{Z}_2 = -K_2Z_2 - \hat{h}\,Tanh(\hat{h}\,Z_2/C_n) + \widetilde{W}_a^T\sigma_a(x) + (e(x) - \tau_{d2}) - \mathrm{R}^TJ^T\varepsilon_j - C_2Z_2 - D_2Z_2$$

۴- بررسی پایداری سیستم کنترل حلقه بسته

در این بخش، پایداری سیستم حلقه بسته از طریق روش مشهور لیاپانوف انجام می‌شود. پایداری سیستم حلقه بسته در قضیه زیر خلاصه می‌شود:

قضیه ۱: مدل دینامیکی و سینماتیکی ربات شبه اتومبیل که در معادلات (۱) و (۳) بیان شده است، را در نظر بگیرید. با در نظر گرفتن فرض ۱، کنترل‌کننده پیشنهادی (۴۷) با سیگنال‌های یادگیری تقویتی که به طور تطبیقی در (۴۶) طراحی می‌شود و قوانین تطبیق (۵۰)، (۵۱) و (۵۲) می‌توانند تضمین کنند که همه سیگنال‌ها در سیستم حلقه بسته کراندار می‌باشند و خطاهای ردیابی موقعیت کراندار نهایی یکنواخت هستند و به ناحیه کوچک اطراف مبدأ همگرا می‌شوند.

اثبات: برای اثبات پایداری سیستم حلقه بسته تابع لیاپانوف L_2 به صورت زیر پیشنهاد می‌شود:

(۵۴)

$$L_2 = L_1 + \frac{1}{2}Z_2^TMZ_2 + \frac{1}{2}tr\{\widetilde{W}_a^T\Gamma_a^{-1}\widetilde{W}_a\} + \frac{1}{2}tr\{\widetilde{W}_c^T\Gamma_c^{-1}\widetilde{W}_c\} + \frac{1}{2}\tilde{r}_f^T\Gamma_{r_f}^{-1}\tilde{r}_f$$

که $\widetilde{W}_a = W_a - \widehat{W}_a$ ، $\widetilde{W}_c = W_c - \widehat{W}_c$ و $\tilde{r}_f = r_f - \hat{r}_f$ می‌باشد. با مشتق‌گیری از رابطه (۵۴) و جایگذاری (۵۳) در آن و با درنظرگرفتن خاصیت ۲ خواهیم داشت:

(۵۵)

$$\begin{aligned}
&\dot{L}_2 = \dot{L}_1 - Z_2^TK_2Z_2 - Z_2^TH_{\mathrm{f}}\hat{r}_{\mathrm{f}}\mathrm{Tanh}(0.2785H_{\mathrm{f}}\hat{r}_{\mathrm{f}}Z_2/C_n) - Z_2^TD_2Z_2 + Z_2^T\widetilde{W}_a\sigma_a(x) - Z_2^T\mathrm{R}^TJ^T\varepsilon_j\\
&-tr\left\{\widetilde{W}_a^T\Gamma_a^{-1}\dot{\widehat{W}}_a\right\} - tr\left\{\widetilde{W}_c^T\Gamma_c^{-1}\dot{\widehat{W}}_c\right\} - \tilde{r}_f^T\Gamma_{r_f}^{-1}\dot{\hat{r}}_f + Z_2^T(e(x) - \tau_{d2})\\
&\le -(\lambda_{min}\{K_1\} - 0.5\lambda_{max}\{RJ\})\|\varepsilon_j\|^2 - (1/\tau_f(1/\tau_f - 1 - 0.5\lambda_{max}\{\mathrm{R}J\})\|Z_e\|^2 + \frac{1}{4}\rho^2\\
&-Z_2^TK_2Z_2 - -Z_2^TH_{\mathrm{f}}\hat{r}_{\mathrm{f}}\mathrm{Tanh}(0.2785\ H_{\mathrm{f}}\hat{r}_{\mathrm{f}}Z_2/C_n)\\
&-Z_2^TD_2Z_2 + Z_2^T\widetilde{W}_a\sigma_a(x) - tr\left\{\widetilde{W}_a^T\Gamma_a^{-1}\dot{\widehat{W}}_a\right\}\\
&-tr\left\{\widetilde{W}_c^T\Gamma_c^{-1}\dot{\widehat{W}}_c\right\} - \tilde{r}_f^T\Gamma_{r_f}^{-1}\dot{\hat{r}}_f + Z_2^T(e(x) - \tau_{d2})
\end{aligned}$$

با استفاده از $\widetilde{W}_a = W_a - \widehat{W}_a$ ،$\widetilde{W}_c = W_c - \widehat{W}_c$ و جایگذاری قوانین تطبیق (۵۰) و (۵۱) در تابع لیاپانوف، داریم:

(۵۶)

$$\begin{aligned}
&\dot{L}_2 \le -(\lambda_{min}\{K_1\} - 0.5\lambda_{max}\{RJ\})\|\varepsilon_j\|^2\\
&-(1/\tau_f - 1 - 0.5\lambda_{max}\{\mathrm{R}J\})\|Z_e\|^2 + \frac{1}{4}\rho^2\\
&-Z_2^TK_2Z_2 - Z_2^TH_{\mathrm{f}}\hat{r}_{\mathrm{f}}\mathrm{Tanh}(0.2785\ H_{\mathrm{f}}\hat{r}_{\mathrm{f}}Z_2/C_n)\\
&-Z_2^TD_2Z_2 + Z_2^T(e(x) - \tau_{d2})\\
&-tr\{W_a^T\sigma_a(x)\sigma_c^T(x)\|Z_2\|\widehat{W}_c\}
\end{aligned}$$

[1] Sigma-modification

[Downloaded from joc.kntu.ac.ir on 2026-11-10]
[DOR: 20.1001.1.20088345.1401.16.4.5.3]
[DOI: 10.61186/joc.16.4.57]

$$+tr\{\widehat{W}_a^T\sigma_a(x)\sigma_c^T(x)\|Z_2\|\widehat{W}_c\}+\delta_a tr\{\widetilde{W}_a^T\widehat{W}_a\}$$
$$-tr\{W_c^T\|Z_2\|\sigma_c(x)\sigma_a^T(x)\widehat{W}_a\}$$
$$+tr\{\widehat{W}_c^T\|Z_2\|\sigma_c(x)\sigma_a^T(x)\widehat{W}_a\}+\delta_c tr\{\widetilde{W}_c^T\widehat{W}_c\}$$
$$-\tilde{r}_f^T\Gamma_{r_f}^{-1}\dot{\hat{r}}_f$$

با در نظر گرفتن فرض ۱، ۴، ۵ و ۶ روابط کران بالای زیر به دست خواهد آمد:

(۵۷)

$$Z_2^T e(x)-Z_2^T\tau_d-tr\{W_a^T\sigma_a(x)\sigma_c^T(x)\|Z_2\|\widehat{W}_c\}$$
$$+tr\{\widehat{W}_a^T\sigma_a(x)\sigma_c^T(x)\|Z_2\|\widehat{W}_c\}-tr\{W_c^T\|Z_2\|$$
$$\sigma_c(x)\sigma_a^T(x)\widehat{W}_a\}+tr\{\widehat{W}_c^T\|Z_2\|\sigma_c(x)\sigma_a^T(x)\widehat{W}_a\}$$
$$\le\|Z_2\|\|e(x)\|-\|Z_2\|\|\tau_d\|+\|W_a\|_F\bar{\sigma}_a\bar{\sigma}_c\|Z_2\|$$
$$\|\widehat{W}_c\|_F+\|\widehat{W}_a\|_F\bar{\sigma}_a\bar{\sigma}_c\|Z_2\|\|\widehat{W}_c\|_F+\|W_c\|_F\|Z_2\|$$
$$\bar{\sigma}_c\bar{\sigma}_a\|\widehat{W}_a\|_F+\|\widehat{W}_c\|_F\|Z_2\|\bar{\sigma}_c\bar{\sigma}_a\|\widehat{W}_a\|_F$$
$$\le r_1\|Z_2\|-r_2\|Z_2\|+r_3\|Z_2\|\|\widehat{W}_c\|_F+r_4\|Z_2\|$$
$$\|\widehat{W}_a\|_F\|\widehat{W}_c\|_F+r_5\|Z_2\|\|\widehat{W}_a\|_F+r_6\|Z_2\|\|\widehat{W}_c\|_F$$
$$\|\widehat{W}_a\|_F:=h(Z_2,W_a,W_c,\widehat{W}_a,\widehat{W}_c)$$

که $h(Z_2,\widehat{W}_a,\widehat{W}_c)=H_f(Z_2,\widehat{W}_a,\widehat{W}_c)r_f$ می باشد. در این رابطه $r_f=[r_1,r_2,r_3,r_4,r_5,r_6]^T$ و $H_f(Z_2,W_a,W_c,\widehat{W}_a,\widehat{W}_c)$ در رابطه (۴۹) تعریف شده است. در ادامه، با استفاده از قانون تطبیق (۵۲) و به کارگیری لم ۲ تابع لیاپانوف به صورت زیر به دست می‌آید:

(۵۸)

$$\dot{L}_2\le-(\lambda_{min}\{K_1\}-0.5\lambda_{max}\{RJ\})\|\varepsilon_j\|^2$$
$$-(1/\tau_f-1-0.5\lambda_{max}\{RJ\})\|Z_e\|^2+\frac{1}{4}\rho^2$$
$$-Z_2^TK_2Z_2-Z_2^TD_2Z_2+\delta_a tr\{\widetilde{W}_a^T\widehat{W}_a\}$$
$$+\delta_c tr\{\widetilde{W}_c^T\widehat{W}_c\}+\delta_{r_f}\tilde{r}_f^T\hat{r}_{r_f}+2C_n$$

نامساوی‌های زیر از لم ۱ به دست می‌آید [۳۹]:

(۵۹)

$$\delta_a tr\{\widetilde{W}_a^T\widehat{W}_a\}\le-\delta_a\left(1-\frac{1}{2k^2}\right)\|\widetilde{W}_a\|_F^2$$
$$+\frac{1}{2}\delta_a k^2\|W_a\|_F^2,$$
$$\delta_c tr\{\widetilde{W}_c^T\widehat{W}_c\}\le-\delta_c\left(1-\frac{1}{2k^2}\right)\|\widetilde{W}_c\|_F^2$$
$$+\frac{1}{2}\delta_c k^2\|W_c\|_F^2$$
$$\delta_{r_f}\tilde{r}_f^T\hat{r}_f\le-\delta_{r_f}\left(1-\frac{1}{2k^2}\right)\|\tilde{r}_f\|^2$$
$$+\frac{1}{2}\delta_{r_f}k^2\|r_f\|^2$$

که در این روابط $K>2\sqrt{2}$ می‌باشد. تابع لیاپانوف (۵۴) به صورت زیر کراندار است:

(۶۰)

$$\lambda_{min}\|x_q(t)\|^2\le L_2(t)\le\lambda_{max}\|x_q(t)\|^2$$

که $x_q(t)$، λ_{min} و λ_{max} به صورت زیر تعریف خواهند شد:

$$x_q=\left[Z_1^T,Z_2^T,\widetilde{W}_{a11},\dots,\widetilde{W}_{an_hn_o},\widetilde{W}_{c11},\dots,\widetilde{W}_{cn_hn_o},\tilde{r}_f^T,Z_e^T\right]$$
$$\lambda_{min}=0.5\min\{1,\lambda_{min}\{M\},\lambda_{min}\{\Gamma_a^{-1}\},\lambda_{min}\{\Gamma_c^{-1}\},\lambda_{min}\left\{\Gamma_{r_f}^{-1}\right\}\}$$
$$\lambda_{max}=0.5\max\{1,\lambda_{max}\{M\},\lambda_{max}\{\Gamma_a^{-1}\},\lambda_{max}\{\Gamma_c^{-1}\},\lambda_{max}\left\{\Gamma_{r_f}^{-1}\right\}\}$$

با مشتق‌گیری از رابطه (۶۰) و جایگذاری (۵۸) در آن و استفاده از نامساوی‌های (۵۹)، نتیجه زیر به دست می‌آید:

(۶۱)

$$\dot{L}_2(t)\le-c_1\|\varepsilon_j\|^2-c_2\|Z_e\|^2-c_3\|Z_2\|^2$$
$$-c_4\|\widetilde{W}_a\|_F^2-c_5\|\widetilde{W}_c\|_F^2-c_6\|\tilde{r}_f\|^2+\Theta$$

که پارامترهای رابطه (۶۱) به صورت زیر تعریف خواهند شد:

(۶۲)

$$c_1=\lambda_{min}\{K_1\}-0.5\lambda_{max}\{RJ\},c_2=1/\tau_f-1-$$
$$0.5\lambda_{max}\{RJ\},c_3=\lambda_{min}\{K_2\}+\lambda_{min}\{D_2\}$$
$$c_4=\delta_a\left(1-\frac{1}{2k^2}\right),c_5=\delta_c\left(1-\frac{1}{2k^2}\right)$$
$$c_6=\delta_{r_f}\left(1-\frac{1}{2k^2}\right),$$

(۶۳)

$$\Theta=\frac{1}{4}\rho^2++\frac{1}{2}\delta_{wa}k^2\|W_a\|_F^2+\frac{1}{2}\delta_{wc}k^2\|W_c\|_F^2$$
$$+\frac{1}{2}\delta_{r_f}k^2\|r_f\|^2+2C_{n.}$$

در نهایت، رابطه (۶۱) به صورت رابطه زیر ساده سازی می شود:

$$\dot{L}_2\le-c_m\|x_q\|^2+\Theta \quad (۶۴)$$

که $c_m=min\{c_1,\dots,c_6\}$ می‌باشد. در رابطه (۶۴) c_m نشان دهنده مقادیر منفی و Θ بیان کننده مقادیر مثبت است. با توجه به رابطه (۶۰) مشتق تابع لیاپانوف به صورت زیر کراندار است:

(۶۵)

$$\dot{L}_2\le-c_m\|x_q\|^2+\Theta\le-c_mL_2/\lambda_{max}+\Theta$$

با حل نامساوی (۶۵) و سپس جایگذاری رابطه (۶۰) در آن خواهیم داشت:

(۶۶)

$$L_2\le\lambda_{max}\Theta/c_m+(L_2(t_0)-\lambda_{max}\Theta/c_m)$$
$$e^{-c_mt/\lambda_{max}}$$

با استفاده از رابطه (۶۶) به این نتیجه خواهیم رسید که با گذشت زمان($t\to\infty$) نامساوی زیر برای تابع لیاپانوف برقرار خواهد بود:

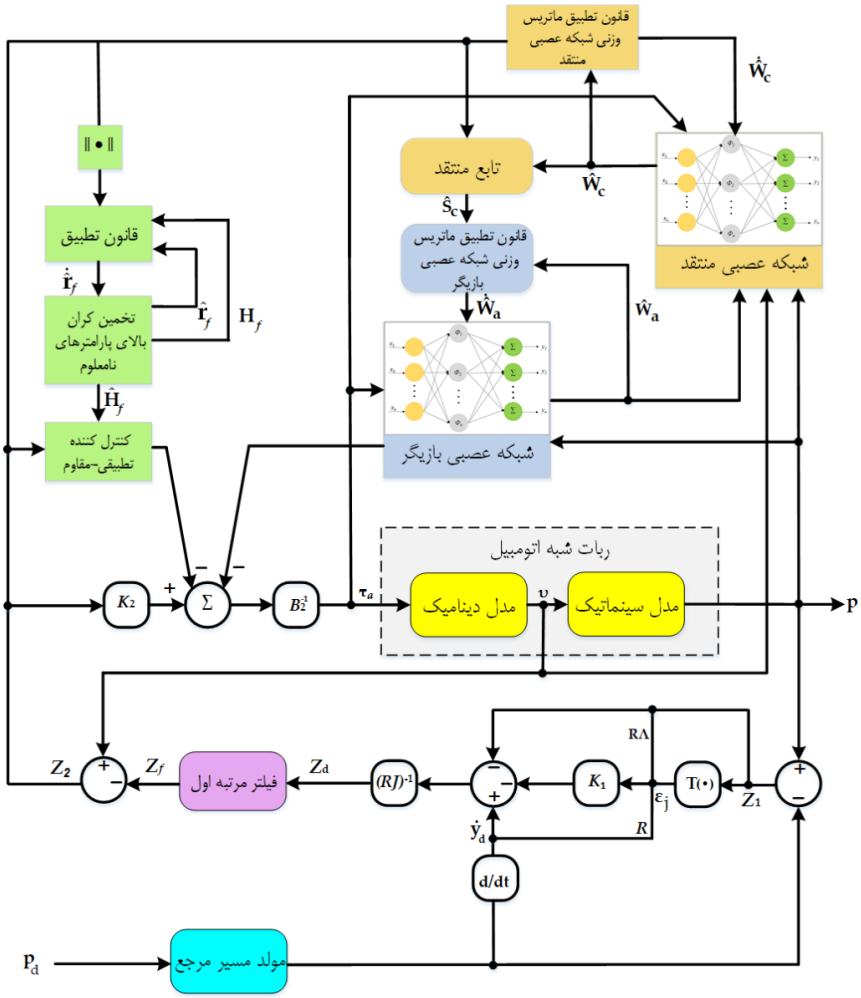
$$L_2\le max\{(L_2(t_0),\lambda_{max}\Theta/c_m\} \quad (۶۷)$$

در نتیجه، با درنظر گرفتن رابطه (۶۰) و (۶۷) مشتق تابع لیاپانوف $\dot{L}_2$ در خارج از مجموعه محدب زیر اکیداً منفی است:

(۶۸)

$$\Omega_q=\{x_q(t)|\ 0\le\|x_q(t)\|\le$$
$$\sqrt{max\{(L_2(t_0),\lambda_{max}\Theta/c_m\}/\lambda_{min}}$$

[Downloaded from joc.kntu.ac.ir on 2026-11-10]
[DOR: 20.1001.1.20088345.1401.16.4.5.3]
[DOI: 10.61186/joc.16.4.57]



شکل ۳: دیاگرام بلوکی الگوریتم کنترلی ربات شبه اتومبیل از طریق یادگیری تقویتی

از رابطه (۶۸) این نتیجه به عمل می‌آید که $\|x_q(t)\|$ خارج از ناحیه محدب Ω_q کاهش می‌یابد و $\|x_q(t)\|$ کران دار نهایی یکنواخت است. بنابراین، خطاهای ردیابی، خطاهای تخمین ماتریس‌های وزن شبکه عصبی بازیگر و منتقد و خطای تخمین کران بالای اغتشاشات خارجی کران دار نهایی یکنواخت هستند به عبارتی $x \in L_\infty$ می‌باشد. با استفاده

$\widehat{V}_{an_h n_o}, L_\infty$

آن جایی که

برقرار است و $\upsilon \prec \iota$, $\varepsilon_{1j} \subset L_\infty$ می‌باشد و ε_{1j} به ناحیه‌ای کوچک اطراف مبدأ همگرا می‌شود، پس $\hat{Z}_{1j} \in \Omega_{Z_{1j}}$ می‌باشد. بنابراین، $-a_{1j}\eta_{-j} < Z_{1j} < b_{1j}\bar{\eta}_j$ و خطاهای ردیابی Z_{1j} به ناحیه‌ای کوچکی اطراف مبدأ با عملکرد از پیش تعریف شده همگرا می‌شوند.

۵- نتایج شبیه سازی

در این بخش، اثر بخشی کنترل‌کننده یادگیری تقویتی پیشنهادی برای یک ربات شبه اتومبیل مورد بررسی قرار گرفته است. برای این منظور، به یک ربات شبه اتومبیل فرمان داده می‌شود تا در فاصله نگاه به جلو یک ربات مجازی مطلوب را ردیابی کند. تمام شبیه سازی‌ها در نرم افزارمتلب به کمک روش تقریب اویلر با زمان نمونه برداری ۲۰ میلی ثانیه انجام شده است. ماتریس‌های دینامیکی ربات شبه اتومبیل به صورت زیر تعریف می‌شوند:

(۶۹)

$$M_1(p) = \begin{bmatrix} m_c & 0 & -m_c l_c sin\psi & 0 \\ 0 & m_c & m_c l_c cos\psi & 0 \\ -m_c l_c sin\psi & m_c l_c cos\psi & I_c + I_f & I_f \\ 0 & 0 & I_f & I_f \end{bmatrix}$$

$$C_1(p,\dot{p})\dot{p} = \begin{bmatrix} -m_c l_c \dot{\psi}^2 cos\psi \\ -m_c l_c \dot{\psi}^2 sin\psi \\ 0 \\ 0 \end{bmatrix},$$

$$B_1(p) = \begin{bmatrix} cos\psi & 0 \\ sin\psi & 0 \\ asin\gamma cos\gamma & 0 \\ 0 & 1 \end{bmatrix}$$

در جدول ۱ پارامترهای ربات شبه اتومبیل مطابق با مرجع [۳۶] معرفی شده است:

[Downloaded from joc.kntu.ac.ir on 2026-11-10]
[DOR: 20.1001.1.20088345.1401.16.4.5.3]
[DOI: 10.61186/joc.16.4.57]

جدول (۱): پارامترهای ربات شبه اتومبیل

پارامتر	توضیحات	مقدار
r	شعاع چرخ‌های ربات	۰/۲m
m_c	جرم ربات شبه اتومبیل	$۳۰Kg$
I_c	لختی دورانی حول محور عمودی	$۵Kg.m^2$
I_f	لختی دورانی اتومبیل	$۱۰Kg.m^2$
a	فاصله بین چرخ‌های جلو و عقب	۱/۲m
L_c	فاصله بین مرکز جرم و نقطه O_B	۰/۵۵m
D_1	ضریب ماتریس میرایی	$۰/۲۵I_4$
d	فاصله P_o تا P_d	۰/۲۵m

در جدول ۲ پارامترهای شبکه عصبی و کنترل کننده بیان شده است:

جدول (۲): پارامترهای کنترل کننده و شبکه‌های عصبی بازیگر و منتقد

پارامتر	مقدار	پارامتر	مقدار	پارامتر	مقدار
تابع عملکرد					
$-\eta_{-j,0}$	-2.5 m	λ_c, λ_a	10[1,1,1,1, 1,1,1,1,1]^T	n_o	2
$\bar{\eta}_{j,0}$	2.5 m	K_2	10diag([10,10])	n_h	9
c	0.1	λ_c, λ_a	10[1,1,1,1, 1,1,1,1,1]^T	μ_c, μ_a	[-4,…,4]
a_1, b_1, a_2, b_2	2	δ_a	1	Γ_a	5
T	20m	δ_c	1	Γ_c	5
$-\eta_{-j,\infty}$	-0.1 m	τ_f	$\exp(-t)$ + 0.05	Γ_{r_f}	5
$\bar{\eta}_{j,\infty}$	0.1 m	C_n	0.3	δ_{r_f}	0.2

بردار موقعیت اولیه ربات شبه اتومبیل به صورت $[x, y, \psi, \gamma]^T$ = $[4, 0, 0, 0]^T$ می‌باشد. مدل زیر برای اغتشاشات خارجی و دینامیک‌های مدل نشده در نظر گرفته شده است:

(۷۰)

$$\tau_{d1} = 0.2[0.25(i) + 0.5sign(i) + 5\sin(2t)], \quad i = v, w$$

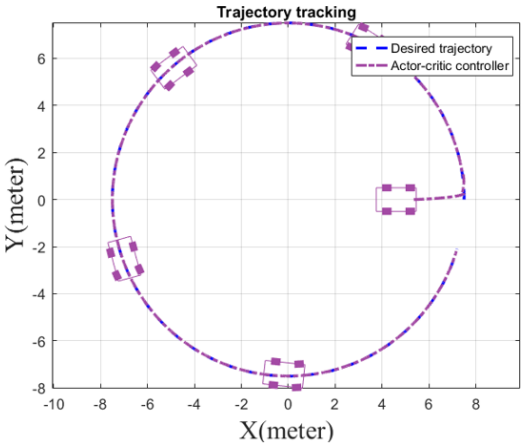
به منظور شبیه سازی اشباع ورودی عملگرهای واقعی، سیگنال‌های کنترل با $|\tau_a| \le 200, i = v(t), w(t)$ محدود گردیده‌اند. خط سیر زمانی مطلوب زیر برای ارزیابی عملکرد کنترل کننده پیشنهادی انتخاب شده است:

$$y_d = [y_{d1}(t), y_{d2}(t)]^T \quad (۷۱)$$
$$y_{d1} = x_g + R\cos(\omega_r t),$$
$$y_{d2} = y_g + R\sin(\omega_r t)$$

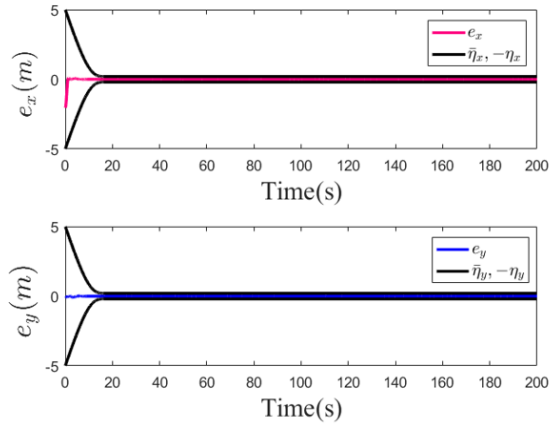
که $x_g = 0$ ، $y_g = 0$، $R = 7.5$ و $\omega_r = 0.03$ پارامترهای خط سیرزمانی مطلوب هستند.

مقدار اولیه پارامترهای نامعلوم به صورت $\hat{r}_f = [0.1,0.2,0.3,0.4,0.5,0.6]^T$ نظر گرفته شده است. پس از اعمال مقادیر ذکر شده در جدول‌های ۱ و ۲ شبیه سازی‌ها انجام شده است. شکل‌های حاصل از شبیه سازی شکل‌های ۴ تا ۱۱ می‌باشند که به ترتیب شامل ردیابی خط سیرزمانی مطلوب درصفحه دو بعدی $(x-y)$، خطای ردیابی موقعیت، ورودی‌های کنترل، خطای تخمین وزن‌های شبکه عصبی بازیگر و منتقد، تخمین پارامترها، خطای تخمین تابع منتقد و خطای خروجی فیلتر هستند که در ادامه به بررسی آن‌ها می‌پردازیم:

شکل ۴ نشان می‌دهد که ربات شبه اتومبیل خط سیر زمانی مطلوب را در حضور عدم قطعیت‌ها با موفقیت دنبال می کند. شکل ۵ نشان دهنده خطای ردیابی موقعیت است که به ناحیه کوچکی اطراف مبدأ همگرا شده است. شکل ۶ سیگنال کنترلی اعمال شده به ربات شبه اتومبیل از طریق کنترل کننده یادگیری تقویتی طراحی شده در حضور اغتشاشات خارجی را نشان داده است. در شکل‌های ۷، ۸ و ۹ به ترتیب اندازه نرم فروبنیوس تخمین وزن شبکه عصبی بازیگر، اندازه نرم فروبنیوس وزن شبکه عصبی منتقد و تخمین کران بالای عدم قطعیت‌های غیرپارامتری مشخص شده است که کراندار هستند. شکل ۱۰ سیگنال‌های تقویتی را نشان می‌دهد که به خوبی کراندار شده‌اند. شکل ۱۱ خطای خروجی فیلتر را نشان می‌دهد که در نزدیکی مبدأ همگرا شده است.

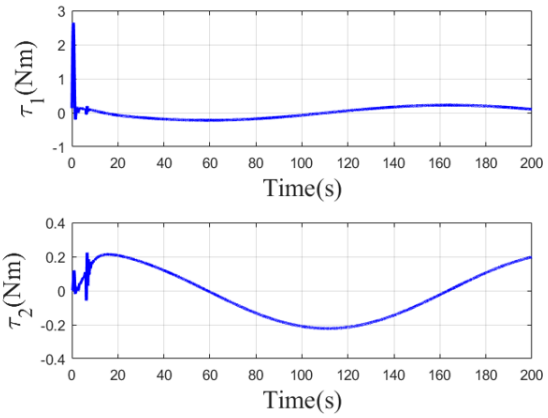


شکل (۴): ردیابی خط سیرزمانی مرجع در صفحه دوبعدی(x-y)

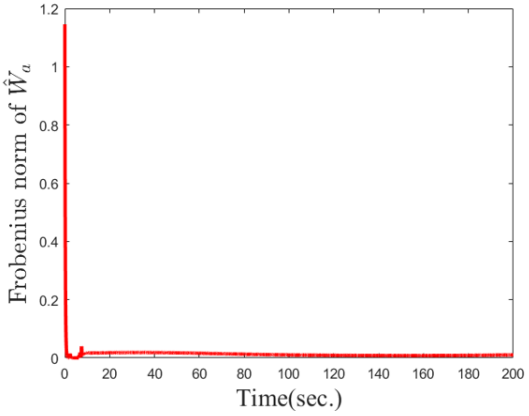


شکل (۵): خطای ردیابی موقعیت

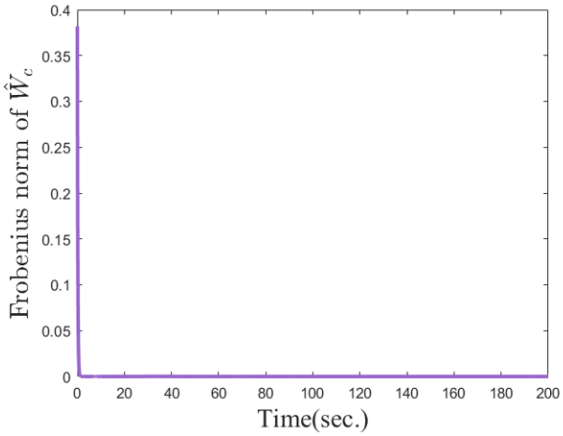
[Downloaded from joc.kntu.ac.ir on 2026-11-10]
[DOR: 20.1001.1.20088345.1401.16.4.5.3]
[DOI: 10.61186/joc.16.4.57]



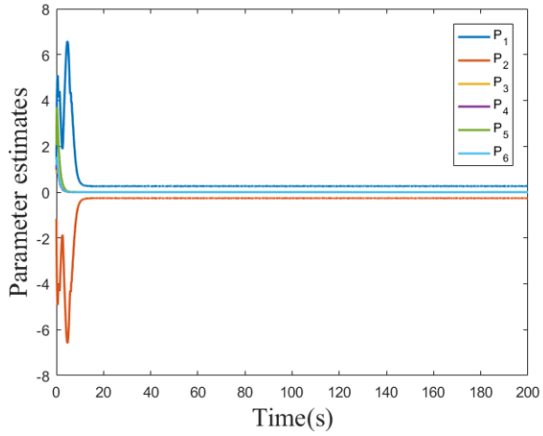
شکل (۶): گشتاورهای ورودی



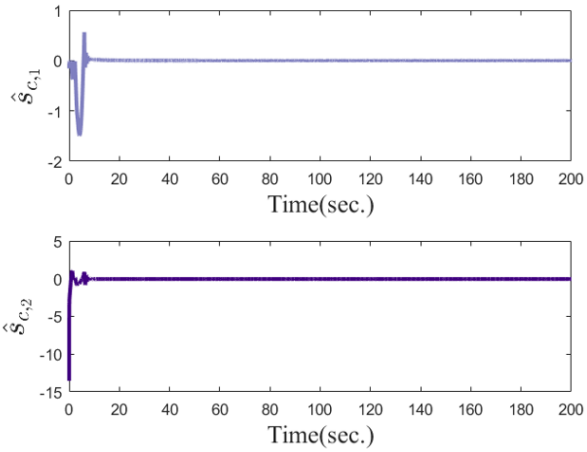
شکل (۷): نرم فروبنیوس ماتریس تخمین وزن شبکه عصبی بازیگر



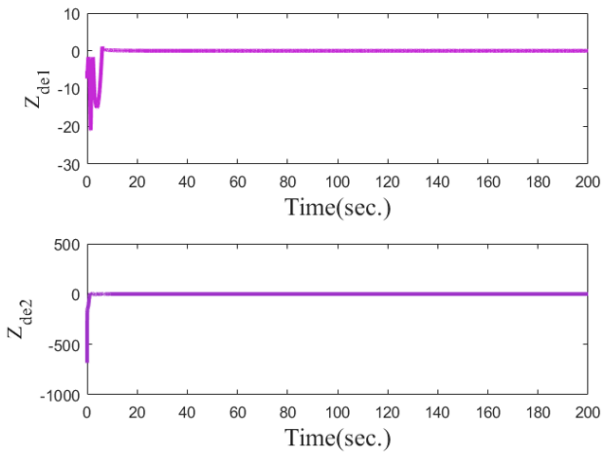
شکل (۸): نرم فروبنیوس ماتریس تخمین وزن شبکه عصبی منتقد



شکل (۹): تخمین پارامترها



شکل (۱۰): تخمین تابع منتقد



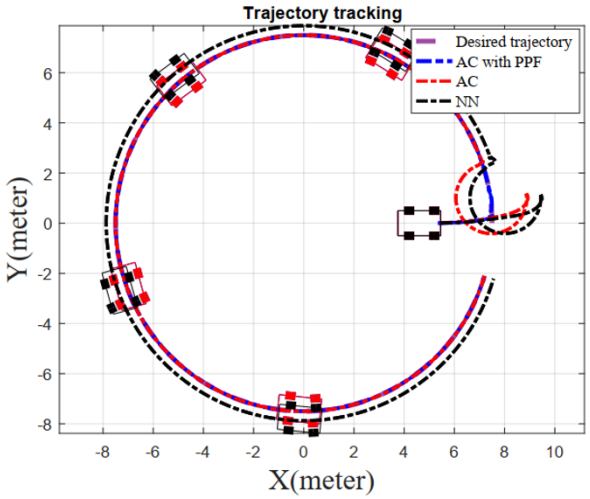
شکل (۱۱): خطای فیلتر

با توجه به شکل های ۴ تا ۱۱ به این نتیجه خواهیم رسید که تمام سیگنال-های سیستم حلقه بسته کران دار هستند. به منظور مقایسه و بررسی عملکرد کنترل کننده پیشنهادی در پاسخ گذرای ردیابی خط سیر زمانی ، زمان همگرایی خطاهای ردیابی و سیگنال کنترلی، مقایسه‌ای بر اساس روش کنترل سطح دینامیکی از طریق یادگیری تقویتی (AC with PPF) و روش کنترل سطح دینامیکی از طریق یادگیری تقویتی بدون عملکرد از

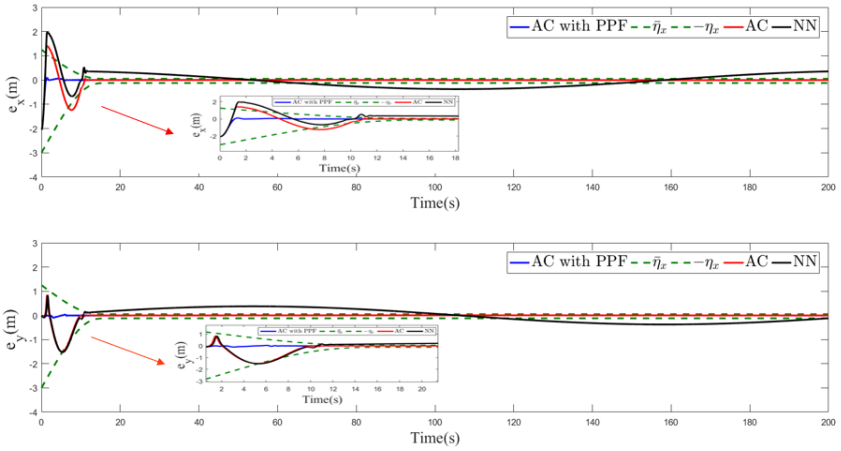
[Downloaded from joc.kntu.ac.ir on 2026-11-10]
[DOR: 20.1001.1.20088345.1401.16.4.5.3]
[DOI: 10.61186/joc.16.4.57]

پیش تعریف شده (AC) و کنترل سطح دینامیکی شامل یک شبکه عصبی (NN) انجام شده است که مقادیر آن‌ها مطابق با جدول ۱ و۲ انتخاب شده است با این تفاوت که بهره $K_1 = 2diag[10,10]$ در AC و NN در نظر گرفته شده است ونتایج آن در شکل‌های ۱۲ تا ۱۴ نشان داده شده است.

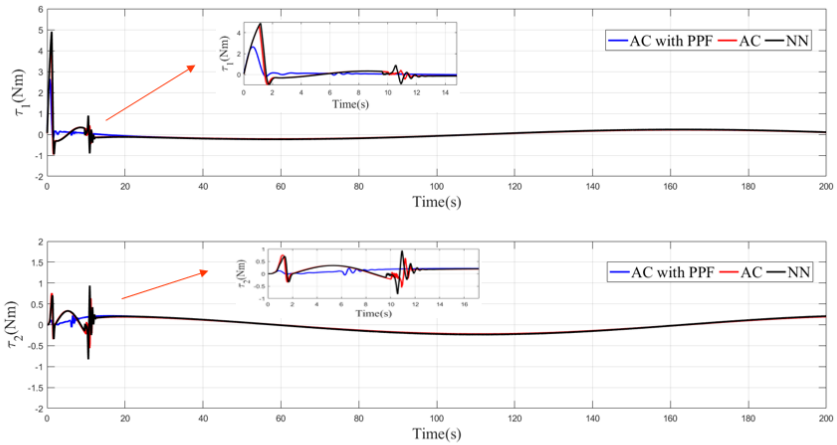
نتایج به دست آمده از شبیه سازی سه کنترل کننده (AC with PPF)، (AC) و (NN) حاکی از آن است که کنترل کننده‌های مبتنی بر یادگیری تقویتی (AC with PPF) و (AC) خط سیر زمانی مطلوب را نسبت به کنترل کننده (NN) که تنها دارای یک شبکه عصبی است ، به طور ایده آل تر ردیابی کرده است. هم‌چنین خطاهای ردیابی برای کنترل کننده‌های (AC with PPF) و (AC) به ناحیه کوچکتری اطراف مبدأ همگرا شده‌اند این در حالی است که برای کنترل کننده (NN) همگرایی خطاهای ردیابی به ناحیه بزرگتری اطراف مبدأ بوده است و در واقع باند نهایی خطاهای ردیابی برای کنترل کننده‌های مبتنی بر یادگیری تقویتی (AC with PPF) و (AC) کوچکتر از باند نهایی خطای ردیابی برای کنترل کننده (NN) است. این اختلاف به دلیل آن است که استفاده از شبکه عصبی منتقد در کنار شبکه عصبی بازیگر باعث تقویت تخمین عدم قطعیت‌ها و غیرخطی‌های نامعلوم شده است.



شکل ۱۲: مقایسه ردیابی خط سیر زمانی



شکل ۱۳: مقایسه سیگنال های خطای ردیابی



شکل ۱۴: مقایسه سیگنال کنترلی

[Downloaded from joc.kntu.ac.ir on 2026-11-10]
[DOR: 20.1001.1.20088345.1401.16.4.5.3]
[DOI: 10.61186/joc.16.4.57]

شاخص‌های عملکردی زیر برای اندازه گیری کمی دقیق تر کنترل کننده‌ها پیشنهاد می‌شود [۴۵]:

$$RMS\left(e_j(t)\right) = \sqrt{(1/T_f)\int_0^{T_f}\left|e_j(t)\right|^2 dt}\ , j = x, y \quad (۷۲)$$

که T_f نشان دهنده کل زمان اجرا و $e_j(t)$ بیان کننده خطای ردیابی برای j اُمین خروجی است. RMS بیان شده برای ارزیابی میانگین عملکرد ردیابی استفاده می‌شود.

$$RMS\left(u_j(t)\right) = \sqrt{(1/T_f)\int_0^{T_f}\left|u_j(t)\right|^2 dt}\ , j = v, w \quad (۷۳)$$

که RMS سیگنال‌های کنترل برای ارزیابی هزینه کنترل استفاده می‌شود.

$$e_{M,j} = max\{\left|e_j(t)\right|\}, j = x, y \quad (۷۴)$$

که ماکزیمم مقدار مطلق خطای ردیابی برای تعیین عملکرد گذرای کنترل کننده‌ها استفاده می‌شود.

$$e_{f,j} = max_{T_f - T_L} \le t \le T_f\{\left|e_j(t)\right|\}, j = x, y \quad (۷۵)$$

که ماکزیمم مقدار مطلق خطای ردیابی در طول T_L ثانیه آخر به عنوان شاخص اندازه گیری دقت ردیابی کنترل کننده‌ها استفاده می‌شود و در این شبیه سازی $T_L = 5s$ است.

جدول (۳): مقایسه عددی کنترل کننده یادگیری تقویتی با سایر کنترل کننده ها

شاخص های عملکرد	AC with PPF	AC	NN
$rms(e_x(t))(m)$	0.7003	1.5879	2.4144
$rms(e_y(t))(m)$	0.0653	1.4403	2.4088
$rms(u_1(t))(V)$	1.5498	2.1612	2.3587
$rms(u_2(t))(V)$	1.0427	1.0933	1.1565
$e_{M,1}$	2.05	2.05	2.05
$e_{M,2}$	0.1027	1.5024	1.5208
$e_{f,1}$	0.002	0.0012	0.3558
$e_{f,2}$	0.0043	0.0027	0.1708

جدول ۳ مقایسه عددی شاخص‌های عملکردی ذکرشده را نشان می‌دهدکه عملکرد کلی و سیگنال‌های کنترلی کنترل‌کننده‌ها را مقایسه می‌کند. همانطور که مشاهده می‌شود میانگین عملکرد ردیابی $rms(e_j(t))(m)$، دقت ردیابی نهایی $e_{f,j}$، سیگنال کنترل، عملکرد گذرای کنترل کننده پیشنهاد شده $e_{M,1}$ و ماکزیمم مقدار مطلق خطای ردیابی در T_L ثانیه آخر برای کنترل کننده‌های مبتنی بر یادگیری تقویتی (AC with PPF) و (AC) به دلیل استفاده از شبکه عصبی منتقد در آن‌ها از کنترل کننده شامل یک شبکه عصبی (NN) کوچکتر می‌باشد اما کنترل کننده پیشنهادی در این مقاله (AC with PPF) به دلیل استفاده از تابع عملکرد از پیش تعریف شده نتایج مطلوب تری نسبت به کنترل کننده (AC) تنها داشته است. نتیجه کلی به دست آمده از جدول ۳ آن است که عملکرد کنترل کننده یادگیری تقویتی با عملکرد از پیش تعریف شده بسیار موثرتر نسبت به دیگر کنترل‌کننده‌ها می‌باشد.

۷ - نتیجه گیری

در این مقاله پژوهشی، به منظور حل مسئله ردیابی خط سیرزمانی مرجع برای یک ربات شبه اتومبیل با استفاده از روش کنترل نگاه به جلو در فضای دو بعدی، از یادگیری تقویتی با استفاده از عملکرد از پیش تعریف شده با به کار گیری تکنیک کنترل سطح دینامیک پیشنهاد گردید. استفاده از یک فیلتر مرتبه اول در تکنیک کنترل سطح دینامیک نیاز به محاسبه مشتق سیگنال کنترل مجازی را در هر مرحله از طراحی کنترل کننده کاهش داده است و در نتیجه منجر به کاهش حجم محاسبات شده است. ساختار یادگیری تقویتی استفاده شده در این مقاله بر اساس شبکه‌های عصبی بازیگر- منتقد می‌باشد که شبکه عصبی بازیگر برای تخمین غیرخطی‌های نامعلوم و عدم قطعیت‌ها و تولید سیگنال کنترلی سیستم است و شبکه عصبی منتقد به ارزیابی عملکرد سیستم می‌پردازد. در نهایت از یک کنترل کننده تطبیقی مقاوم به منظور مقابله با پارامترهای نامعلوم و اغتشاشات خارجی استفاده شده است. به منظور اثبات کرانداری نهایی یکنواخت خطاهای ردیابی از روش لیاپانوف مستقیم برای سیستم حلقه بسته استفاده گردید. نتایج شبیه سازی، عملکرد کنترل کننده پیشنهادی را نشان داده است که مطابق با نتایج به دست آمده خطاهای ردیابی در نزدیکی مبدأ همگرا شده‌اند. نتایج به دست آمده حاکی از آن است که روش کنترل پیشنهادی یک انتخاب مناسب برای پیاده‌سازی بر روی یک ربات در دنیای واقعی محسوب می‌شود.

[Downloaded from joc.kntu.ac.ir on 2026-11-10]
[DOR: 20.1001.1.20088345.1401.16.4.5.3]
[DOI: 10.61186/joc.16.4.57]

۸- مراجع

[1] Dhall D, Kaur R, Juneja M. Machine learning: a review of the algorithms and its applications. Proceedings of ICRIC 2019. 2020:47-63.

[2] Nasteski V. An overview of the supervised machine learning methods. Horizons. b. 2017 Dec;4:51-62.

[3] Sen PC, Hajra M, Ghosh M. Supervised classification algorithms in machine learning: A survey and review. InEmerging technology in modelling and graphics 2020 (pp. 99-111). Springer, Singapore.

[4] Rajoub B. Supervised and unsupervised learning. InBiomedical Signal Processing and Artificial Intelligence in Healthcare 2020 Jan 1 (pp. 51-89). Academic Press.

[5] Mahesh B. Machine learning algorithms-a review. International Journal of Science and Research (IJSR).[Internet]. 2020 Oct;9:381-6.

[6] Alpaydin E. Introduction to machine learning. MIT press; 2020 Mar 24.

[7] Qiang W, Zhongli Z. Reinforcement learning model, algorithms and its application. In2011 International Conference on Mechatronic Science, Electric Engineering and Computer (MEC) 2011 Aug 19 (pp. 1143-1146). IEEE.

[8] Kurinov I, Orzechowski G, Hämäläinen P, Mikkola A. Automated Excavator Based on Reinforcement Learning and Multibody System Dynamics. IEEE Access. 2020 Nov 24;8:213998-4006.

[9] Sutton RS, Barto AG. Reinforcement learning: An introduction. MIT press; 2018 Nov 13.

[10] Hidalgo-Paniagua A, Vega-Rodríguez MA, Ferruz J. Applying the MOVNS (multi-objective variable neighborhood search) algorithm to solve the path planning problem in mobile robotics. Expert Systems with Applications. 2016 Oct 1;58:20-35

[11] Pourboghrat F, Karlsson MP. Adaptive control of dynamic mobile robots with nonholonomic constraints. Computers & Electrical Engineering. 2002 Jul 1;28(4):241-53.

[12] Tzafestas SG. Introduction to mobile robot control. Elsevier; 2013 Oct 3.

[13] Simpson J, Jacobsen CL, Jadud MC. Mobile robot control. Communicating Process Architectures. 2006 Sep 6:225.

[14] Lee K, Chung W. Calibration of kinematic parameters of a car-like mobile robot to improve odometry accuracy. In2008 IEEE International Conference on Robotics and Automation 2008 May 19 (pp. 2546-2551). IEEE.

[۱۵]کاظمی پور ملیحه، شجاعی خوشنام. کنترل آرایش بندی سطح دینامیکی زمان محدود فازی تطبیقی مبتنی بر رؤیتگر اغتشاش غیرخطی ربات های متحرک چرخ دار تراکتور- تریلر. مجله کنترل. ۱۴۰۰؛ ۱۵ (۱): ۱۱۲-۹۳

[16] Lee JH, Lin C, Lim H, Lee JM. Sliding mode control for trajectory tracking of mobile robot in the RFID sensor space. International Journal of Control, Automation and Systems. 2009 Jun;7(3):429-35.

[17] Künhe F, Gomes J, Fetter W. Mobile robot trajectory tracking using model predictive control. In IEEE latin-American Robotics Symposium 2005 Sep (Vol. 51).

[18] Koubaa Y, Boukattaya M, Damak T. Adaptive sliding mode control for trajectory tracking of nonholonomic mobile robot with uncertain kinematics and dynamics. Applied Artificial Intelligence. 2018 Nov 26;32(9-10):924-38.

[19] Mac Thi T, Copot C, De Keyser R, Tran TD, Vu T. MIMO fuzzy control for autonomous mobile robot. Journal of Automation and Control Engineering. 2016;4(1):65-70.

[20] Mirzaeinejad H. Optimization-based nonlinear control laws with increased robustness for trajectory tracking of non-holonomic wheeled mobile robots. Transportation Research Part C: Emerging Technologies. 2019 Apr 1;101:1-7.

[21] Sabiha AD, Kamel MA, Said E, Hussein WM. ROS-based trajectory tracking control for autonomous tracked vehicle using optimized backstepping and sliding mode control. Robotics and Autonomous Systems. 2022 Jun 1;152:104058.

[22] Shafei HR, Bahrami M. Trajectory tracking control of a wheeled mobile robot in the presence of matched uncertainties via a composite control approach. Asian Journal of Control. 2021 Nov;23(6):2805-23.

[23] Chen Z, Liu Y, He W, Qiao H, Ji H. Adaptive Neural Network-Based Trajectory Tracking Control for a Nonholonomic Wheeled Mobile Robot with Velocity Constraints. IEEE Transactions on Industrial Electronics. 2020 Apr 28.

[24] Fan QY, Yang GH, Ye D. Quantization-based adaptive actor-critic tracking control with tracking error constraints. IEEE transactions on neural networks and learning systems. 2017 Feb 1; 29 (4): 970-80.

[25] Pan G, Xiang Y, Wang X, Yu Z, Zhou X. Research on path planning algorithm of mobile robot based on reinforcement learning. Soft Computing. 2022 Jul 21:1-0.

[26] Kiumarsi B, Vamvoudakis KG, Modares H, Lewis FL. Optimal and autonomous control using

[Downloaded from joc.kntu.ac.ir on 2026-11-10]
[DOR: 20.1001.1.20088345.1401.16.4.5.3]
[DOI: 10.61186/joc.16.4.57]

reinforcement learning: A survey. IEEE transactions on neural networks and learning systems. 2017 Dec 7;29(6):2042-62.

[27] Yang Y, Gao W, Modares H, Xu CZ. Robust actor-critic learning for continuous-time nonlinear systems with unmodeled dynamics. IEEE Transactions on Fuzzy Systems. 2021 Apr 26.

[28] Zhang H, Wang H, Niu B, Zhang L, Ahmad AM. Sliding-mode surface-based adaptive actor-critic optimal control for switched nonlinear systems with average dwell time. Information Sciences. 2021 Nov 1;580:756-74.

[29] Yang X, Li B, Wen G. Adaptive Neural Network Optimized Control Using Reinforcement Learning of Critic-Actor Architecture for a Class of Non-Affine Nonlinear Systems. IEEE Access. 2021 Oct 15;9:141758-65.

[30] Wen G, Xu L, Li B. Optimized backstepping tracking control using reinforcement learning for a class of stochastic nonlinear strict-feedback systems. IEEE Transactions on Neural Networks and Learning Systems. 2021 Aug 26.

[31] Yan L, Liu Z, Chen CP, Zhang Y, Wu Z. Reinforcement learning based adaptive optimal control for constrained nonlinear system via a novel state-dependent transformation. ISA Transactions. 2022 Jul 12.

[32] Wei C, Luo J, Yin Z, Yuan J. Leader-following consensus of second-order multi-agent systems with arbitrarily appointed-time prescribed performance. IET Control Theory & Applications. 2018 Nov;12(16):2276-86.

[33] Yin Z, Luo J, Wei C. Robust prescribed performance control for Euler–Lagrange systems with practically finite-time stability. European Journal of Control. 2020 Mar 1;52:1-0.

[34] Li J, Du J, Lewis FL. Distributed three-dimension time-varying formation control with prescribed performance for multiple underactuated autonomous underwater vehicles. International Journal of Robust and Nonlinear Control. 2021 Sep 10;31(13):6272-87.

[35] Kiumarsi B, Lewis FL. Actor–critic-based optimal tracking for partially unknown nonlinear discrete-time systems. IEEE Transactions on Neural Networks and Learning Systems. 2014 Oct 8;26(1):140-51.

[36] Shojaei K. Neural adaptive PID formation control of car-like mobile robots without velocity measurements. Advanced Robotics. 2017 Sep 17;31(18):947-64.

[37] Sarkar N, Yun X, Kumar V. Control of mechanical systems with rolling constraints: Application to dynamic control of mobile robots. The International Journal of Robotics Research. 1994 Feb;13(1):55-69.

[۳۸]کیماسی خلجی علی، یزدانی ابوالفضل. کنترل تعقیب مسیر یک ربات چرخدار در حرکت‌های رو به عقب و رو به جلو. مجله کنترل. ۱۳۹۹؛ ۱۴ (۲) :۳۵-۴۵

[39] Elhaki O, Shojaei K. “Neural network-based target tracking control of underactuated autonomous underwater vehicles with a prescribed performance.” *Ocean Engineering*. Vol. 167, pp. 239-56, 2018.

[40] Yun X, Yamamoto Y. Stability analysis of the internal dynamics of a wheeled mobile robot. Journal of Robotic Systems. 1997 Oct;14(10):697-709.

[41] Slotine JJ, Li W. Applied nonlinear control. Englewood Cliffs, NJ: Prentice hall; 1991 Jan.

[42] Swaroop D, Hedrick JK, Yip PP, Gerdes JC. Dynamic surface control for a class of nonlinear systems. IEEE Transactions on Automatic Control. 2000 Oct;45(10):1893-9.

[43] Bechlioulis CP, Rovithakis GA. Robust adaptive control of feedback linearizable MIMO nonlinear systems with prescribed performance. IEEE Transactions on Automatic Control. 2008 Oct 7;53(9):2090-9.

[44] Elhaki O, Shojaei K. A novel model-free robust saturated reinforcement learning-based controller for quadrotors guaranteeing prescribed transient and steady state performance. Aerospace Science and Technology. 2021 Dec 1;119:107128.

[45] Xu L, Yao B. Output feedback adaptive robust precision motion control of linear motors. Automatica. 2001 Jul 1;37(7):1029-39.

[Downloaded from joc.kntu.ac.ir on 2026-11-10]
[DOR: 20.1001.1.20088345.1401.16.4.5.3]
[DOI: 10.61186/joc.16.4.57]