مجله کنترل
ISSN 2008-8345
جلد ۷، شماره ۳، پاییز ۱۳۹۲، صفحه ۳۰-۱۷

راهکارکنترل مقاوم مبتنی بر یادگیری تقویتی به منظور توانبخشی حرکتی بازوی دست

زهرا حسن‌زاده بنابیدی[۱]، حمیدرضا کبروی[۲]، سعید طوسی‌زاده[۳]، رضا بوستانی[۴]

[۱]دانشجو کارشناسی ارشد مهندسی پزشکی، گروه مهندسی پزشکی، دانشگاه آزاد اسلامی، واحد مشهد، hasanzadeh_67511@yahoo.com

[۲]استادیار، دانشکده فنی مهندسی، گروه مهندسی پزشکی، دانشگاه آزاد اسلامی، واحد مشهد، hkobravi@mshdiau.ac.ir

[۳]استادیار، دانشکده فنی مهندسی، گروه مهندسی برق، دانشگاه آزاد اسلامی، واحد مشهد، saeedtoosizadeh@yahoo.com

[۴]استادیار، گروه مغز و اعصاب، دانشگاه علوم پزشکی مشهد، BoostaniR@mums.ac.ir

(تاریخ دریافت مقاله ۱۳۹۲/۵/۲۳، تاریخ پذیرش مقاله ۱۳۹۲/۷/۲۸)

چکیده: توانبخشی حرکتی از مباحث مورد توجه محققان است. در این تحقیق، یک راهکار کنترلی به‌منظور کنترل حرکت مدلی از بازوی دست با سه مفصل ارائه شده است. در مدل مورد استفاده، اثر فعالسازی همزمان عضلات آگونیست و آنتاگونیست مفصل مچ دست با استفاده از مکانیزم فریزسازی لحاظ شده است. با الهام از عملکرد سیستم اعصاب مرکزی در کسب مهارتهای حرکتی، راهکار کنترلی ارئه شده مبتنی بر یکی از الگوریتم‌های یادگیری تقویتی توسعه داده شده است. ابتدا کارایی کنترل‌کننده کلاسیک مبتنی بر یادگیری تقویتی به منظور کنترل حرکت دست به سمت یک هدف مشخص، مورد ارزیابی قرار گرفت. نتایج نشان دادند که چنین کنترل‌کننده‌ای نمی‌تواند حتی بازای پس چندین اپیزود پایداری سیستم را حفظ نماید. در مرحله بعد به منظور بهبود وضعیت پایداری و همچنین کارایی کنترل‌کننده، ترکیبی از کنترل‌کننده تناسبی-انتگرالی-مشتقی (PID) و کنترل‌کننده مبتنی بر یادگیری تقویتی ارائه و مورد ارزیابی قرار گرفت. کنترل‌کننده PID در کنار یادگیری تقویتی کارایی خوبی داشته است، اما به لحاظ تئوری تضمینی برای اثبات پایداری وجود ندارد. لذا به منظور بهبود شرایط پایداری سیستم کنترل، از ترکیب کنترل‌کننده HTC و کنترل‌کننده PID همراه با کنترل‌کننده مبتنی بر راهکار یادگیری تقویتی استفاده شده است. به این ترتیب پایداری سیستم کنترل حلقه بسته مبتنی تئوری لیاپانوف تضمین شده است. نتایج نشان دادند که کنترل‌کننده ترکیبی با استفاده از HTC در برابر اغتشاش خارجی و تغییرات تصادفی پارامترهای سیستم، کارایی بسیار خوبی داشته است. پس از شبیه‌سازی-های کامپیوتری، به منظور ارزیابی کاملتر راهکار کنترلی پیشنهادی، آزمایش‌های انسانی طراحی و انجام شدند. در آزمایش‌های انسانی الگوی فعالیت عضلات مفاصل دست، طی عملکرد حرکت دست با استفاده از سیگنال‌های الکترومایوگرام عضلات، استخراج شدند. نتایج نشان می‌دهند که همزمانی قابل ملاحظه‌ای بین بازه زمانی مربوط به فعالیت و عدم فعالیت تغییرات سیگنال خروجی کنترل‌کننده در مطالعات شبیه‌سازی با الگوی فعالیت عضلات دست فرد سالم وجود دارد.

کلید واژه: کنترل‌کننده تناسبی-انتگرالی -مشتقی، کنترل‌کننده مقاوم (HTC)، مدل بازوی سه مفصل، مکانیزم فریزسازی، یادگیری تقویتی

A Robust Control Strategy Based on Reinforcement Learning Approach to Rehabilitat the Arm Movement

Z. Hasanzadeh Binabidi, Hamid Reza Kobravi, Saeed Toosizadeh, Reza Boostani

In this research, a control strategy has been presented to movement control of a three link model of human's arm. The freezing mechanism has been used to consider the role of antagonistic coactivation of wrist muscles in the used three link model. Inspired by motor learning process of central nervous system, the presented control strategy has been designed based on the reinforcement learning algorithm. At first, the performance of a control methodology based on reinforcement learning was evaluated. The results show the instability of control system even after numbers of leaning episode. Then, a combination of a proportional derivative integral (PID) controller and a reinforcement learning based controller were utilized to improve the stability conditions and performance of controller. Despite the good performance, there is no guarantee for stability of control system. So, to satisfy the stability conditions, a robust controller called HTC was added to the combination of a PID controller and a reinforcement learning based controller. According to the simulation results, the combinational controller accompany by HTC had good performance even in presence of external disturbance, measurement noise and random changes of model parameters. For more assessments, the muscle activation profile of involved muscles during the arm movement of an intact subject was compared with control signals obtained through the simulation studies. The results show an interesting timing synchronization between the activation and deactivation timing of control signals and muscle activation profiles.

[Downloaded from joc.kntu.ac.ir on 2026-11-08]
[DOR: 20.1001.1.20088345.1392.7.3.3.6]

Keyword: proportional derivative integral controller, robust controller (HTC), three link model of human's arm, freezing mechanism, reinforcement learning

۱- مقدمه

آسیب نخاعی[1]سطح بالا (سطح گردنیC1-C4)، می‌تواند موجب از دست دادن اکثر یا همه عملکردهای عضلانی ارادی در زیرگردن می‌باشد. تحریک الکتریکی عملکردی[2]می‌تواند به منظور ایجاد عملکردهای حرکتی در این افراد موثر باشد. در این راهکار با تحریک الکتریکی اعصاب حرکتی عضلات فلج و تنظیم شدت انقباض آنها، گشتاورهای مکانیکی حول مفاصل، کنترل و به طبع آن حرکت ایجاد می‌شود. بدلیل دینامیک متغیر، اغتشاشات، رفتار متغیر با زمان و عدم قطعیت‌های موجود در سیستمهای اسکلتی- عضلانی، توسعه سیستمهای مبتنی بر FES نیازمند استفاده از روش‌های کنترلی هستند. یکی از زمینه‌های کاربردی تحریک الکتریکی عملکردی در ایجاد و کنترل حرکت بازو دست است. حرکت بازو یکی از رایج‌ترین فعالیت‌ها در زندگی روزمره ما است و چگونگی حرکات هماهنگ موضوع تحقیق‌های گسترده‌ای بوده است [۶]-[۱]. از رویکردهای کنترل حرکت بازوی انسان می‌توان به استفاده از مدل‌های توصیفی، مدلهای کامل (شامل مدل‌های دینامیکی و مدل‌های تصادفی)، مدلهای اجرایی حرکت و مدل‌های دینامیک مفاصل انسان اشاره کرد [۱]. هدف مدل-های توصیفی، توصیف رفتار ظاهری حرکت انسان است. از جمله روابط تجربی این مدلها که در شناسایی حرکات بازوی انسان مورد استفاده قرار می‌گیرد می-توان به قانون فیت، پروفایل سرعت زنگوله‌وار در حرکات مستقیم و قانون ۲/۳ توان اشاده کرد. به رغم سادگی مدل‌های توصیفی، این مدل‌ها به علت در نظر نگرفتن جنبه‌های دینامیکی حرکت مورد انتقاد قرار می‌گیرند. از سوی دیگر مدل‌های دینامیکی قادر به پیش‌بینی برخی آزمایشات نبوده و در نتیجه رضایت-بخش نیستند. نشان داده شده است هنگامیکه فردی دست خود را بین موقعیت اولیه و نهایی حرکت می‌دهد، ممکن است مسیر دست خود را به منظور اصلاح خطسیر، با استفاده از فیدبکهای بصری تغییر دهد [۱]. مدلهای کینماتیک این تطبیق‌پذیری[3] را پیش‌بینی می‌کنند در حالی که مدلهای دینامیکی قادر به انجام چنین کاری نیستند زیرا محدوده کاری مدلهای دینامیکی در خطسیر تنها به موقعیت ابتدا و انتهای حرکت بستگی دارد و موقعیت میانی درنظر گرفته نمی-شود [۱]. هاریس و ولپرت مفهوم متفاوتی از طراحی حرکت را بیان می‌کنند که ادغامی از مفاهیم دینامیک و کینماتیک می‌باشد [۱]. سجنوسکی و همکاران، [۱]، به برخی از جنبه‌های مهم نظریه هاریس و ولپرت تحت عنوان حداکثر دقت اشاره می‌کنند و بیان می‌دارند که مدلهای مبتنی بر حداکثر بهره‌وری و حرکت بدون نوسان، نمی‌توانند توصیف خوبی از خود تطبیق‌پذیری سیستم عصبی دربرداشته باشند. لی و همکارانش یک روش کلی برای کنترل پیش‌خوردی سلسله‌مراتبی ارائه دادند که می‌تواند برای مشکلات کنترل حرکت دست مورد استفاده قرار گیرد[۷]. مدل دینامیکی ارائه شده توسط لی و همکارانش، یک مدل دو سطحی، شامل مدل دینامیکی سطح بالا و مدل دینامیکی سطح پایین هستند که با یکدیگر هماهنگی دارند. در تحقیق دیگری جیانگ‌یو و همکارانش، تئوری کنترل ثابت بهینه پیوسته در زمان را برای مدل‌سازی سیستم-های کنترل حرکتی انسان در حضور نویز وابسته به سیگنال، ارائه داده‌اند [۲]. مدل کنترلی آنها با برخی از مشاهدات مهم مانند عدم تقارن پروفایل سرعت سازگار است. جاجودنیک و همکاران سه کنترلر پس‌خوردی تناسبی-مشتقی(PD) برای تحریک شش عضله بازو، با استفاده از دو سنسور زاویه مفصل ارائه داده‌اند [۸]. استفاده از راهکارهای کنترل ترکیبی مانندترکیب کنترل پس-خوری و جلوسوی، یادگیری تقویتی و شبکه‌های عصبی مصنوعی نیز مورد ارزیابی قرار گرفته‌اند[۹]. اما با توجه به اینکه در فرآیند حرکت دست به سمت هدف، اولا خطسیر مطلوبی برای تغییرات موقعیت دست قابل طراحی نیست و ثانیا موقعیت نقاط ابتدا و انتها نیز ثابت نیستند، لذا استفاده از روش‌های کنترل مبتنی بر یادگیری تقویتی می‌تواند موثرتر باشد[۱۷]-[۱۰]. یادگیری تقویتی به معنای تصحیح عملکرد یک سیستم بر اساس پاداش و تنبیه‌هایی است که با توجه به شناخت از محیط و نتایج تعامل با محیط، طراحی شده‌اند. پژوهشگران دریافته اند که نرخ ارسال الکتریکی نورون های پخش کننده دوپامین در ناحیه تگمنتوم شکمی را می توان به تابع خطای الگوریتم تفاوت زمانی نسبت داد ،[۱۹]، که در فرایند حرکت دست به سمت هدف، این تابع در حقیقت وابسته به فاصله بین موقعیت فعلی و نهایی است. این امر موجب افزایش فعالیت مغزی در ناحیه گانگلیون بازال[4] می گردد، که این بخش از ساختار مغز در کنترل حرکات عضله و یادگیری فرد دخالت دارد [۱۹]. با الهام از عملکرد سیستم اعصاب مرکزی در کسب مهارتهای حرکتی، راهکار کنترلی مبتنی بر یکی از الگوریتم‌های یادگیری تقویتی توسعه داده شده است. به دلیل وجود عدم قطعیت، دینامیک-های مدل نشده و عدم وجود مدل ریاضی مناسب در سیستمهای اسکلتی-عضلانی، بکارگیری روشهای یادگیری تقویتی در کنار کنترل‌کننده‌هایی مقاوم می‌تواند، راهکاری موثر باشد. بر این اساس هدف اصلی این تحقیق، ارائه راهکار کنترلی جدیدی بر مبنای ترکیب یک کنترلگر با الگوریتم‌های یادگیری تقویتی می‌باشد. به منظور بهبود وضعیت پایداری و همچنین کارایی کنترل-کننده، ترکیبی از یک کنترل‌کننده پس‌خوردی و کنترل‌کننده مبتنی بر یادگیری تقویتی (RLPID) مورد استفاده قرار گرفته است. سپس به منظور تضمین کاملتر پایداری، ساختار کنترلی شامل ترکیب کنترل‌کننده مقاوم و کنترل‌کننده پس-خوردی همراه با کنترل‌کننده مبتنی بر راهکار یادگیری تقویتی (RLPIDHTC) استفاده شده است. به این ترتیب پایداری سیستم کنترل حلقه بسته با استفاده از تئوری لیاپانوف تضمین می‌شود. نتایج مطالعات مبتنی بر شبیه‌سازی‌های کامپیوتری نشان می‌دهند که راهکار کنترلی RLPIDHTC، دارای بهترین کاریی بوده است. پس از ارزیابی بوسیله شبیه‌سازی کامپیوتری، به منظور ارزیابی کاملتر راهکار کنترلی پیشنهادی، آزمایش‌های انسانی طراحی و انجام شده‌اند. بطوریکه

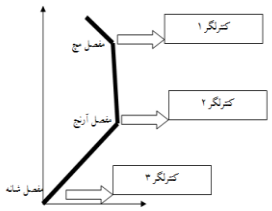
[1]Spinal Cord Injury(SCI)
[2]Functional Electrical Stimulation (FES)
[3]adaptation
[4]Basal ganglia

[Downloaded from joc.kntu.ac.ir on 2026-11-08]
[DOR: 20.1001.1.20088345.1392.7.3.3.6]

الگوی فعالیت عضلانی فرد سالم با سیگنال خروجی کنترل‌کننده توسعه یافته، مقایسه شد.

۲- ارائه راهکار کنترل

حرکت دست به سمت هدف، بوسیله ایجاد گشتاور حول سه مفصل مچ، آرنج و شانه ایجاد می‌شود. لذا برای کنترل چنین حرکتی دو راه‌حل وجود دارد. راه‌حل اول این است که، به از یک کنترل‌کننده چند ورودی-چند خروجی استفاده شود. اما طراحی این کنترل‌کننده مشکل بوده و به لحاظ پیاده‌سازی بار محاسباتی بالایی دارد، از طرفی تضمین پایداری چنین ساختاری بسیار مشکل است. راه‌حل دوم این است که یک استراتژی کنترل غیرمتمرکز توسعه داده شود بطوریکه برای کنترل حرکت هر مفصل از یک کنترل‌کننده مجزا استفاده شود، که ساختار کلی آن در شکل (۱) نشان داده شده است. در ساختار کنترلی غیر-متمرکز، کوپلینگ‌های بین زیر سیستم‌ها، به عنوان دینامیکهای مدل نشده یا اغتشاش برای هر زیر سیستم در نظر گرفته می‌شود.

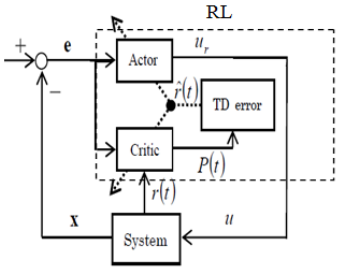


شکل ۱: سیستم کنترل غیرمتمرکز به منظور کنترل حرکت بازوی دست

ابتدا کنترل‌کننده مبتنی بر یادگیری تقویتی (RL) شامل دو قسمت عملگر[5] و ارزیاب[6] در جایگاه کنترل‌کننده قرار گرفت. پس از بررسی نتایج بدست آمده به منظور بهبود کارایی، کنترل‌کننده PID به ساختار کنترلی اضافه شد. سپس کنترل‌کننده HTC در جهت تضمین پایداری سیستم و رسیدن به نتیجه مطلوبتر به ساختار کنترلی اضافه گردید.

۲-۱- کنترل‌کننده عملگر-ارزیاب

ساختار کلاسیک کنترل‌کننده یادگیری تقویتی شامل دو بخش است: ۱) عملگر ؛ ۲) ارزیاب. عملگر سیگنال مطلوب را تولیدکرده و ارزیاب وظیفه ارزیابی عملکردشبکه عملگر را برعهده دارد، به عبارتی عملگر بر اساس خروجی‌های ارزیاب تطبیق می‌یابد[۱۸]. شکل (۲) ساختار کلی کنترل‌کننده مبتنی بر یادگیری تقویتی را نشان می‌دهد.



شکل ۲: ساختار کنترل‌کننده مبتنی بر یادگیری تقویتی

۲-۱-۱- عملگر

عملگر یک کنترل‌کننده با عملکرد ردیابی بالا است، که دارای ساختار شبکه عصبی-فازی می‌باشد[۱۱]. سیگنال یادگیری تقویتی بصورت معادله (۱) بیان می‌شود، که از روش greedy-ε به عنوان خط‌مشی عملگر به منظور تصمیم-گیری رفتار استفاده شده است .

$$u_r = \begin{cases} \sum_{j=1}^{R} w_j^{act} b_j^{act} \; probability \; 1-\varepsilon \\ \sum_{j=1}^{R} w_j^{act} b_j^{act} + n \; probability \varepsilon \end{cases} \qquad (۱)$$

R تعداد نرون‌های لایه میانی عملگر، w_j^{act} وزن از نرون لایه میانی jام به خروجی عملگر، n نویز برای یافتن سیگنال بهینه، b_j^{act} سازگاری نرون و ε نرخ جستجو که در معادله (۲) آمده است.

$$\varepsilon = \frac{\alpha_\varepsilon |e_t|}{1+\exp(t-10)} \qquad (۲)$$

که e_t خطای خروجی سیستم و α_ε ضریب ثابت است.

سازگاری گره از حاصلضرب تابع عضویت گوسین بدست می‌آید، که بصورت معادله (۳) بیان می‌شود.

$$b_j^{act} = \prod_{i=1}^{I} \exp\left\{-\frac{\left(e^{(i-1)}-c\right)^2}{(\sigma)^2+\bar{\omega}}\right\} \qquad (۳)$$

که c و σ مرکز و عرض تابع پایه لایه ورودی iام برای نرون jام لایه میانی، e^i خطای خروجی سیستم مربوط به لایه ورودی iام می‌باشند و $\bar{\omega}$ ثابت مثبتی است که از صفر شدن عرض تابع پایه جلوگیری می‌کند.

ساختار عملگر خودسازمانده است، به این صورت که تعداد نرون‌های لایه میانی شبکه عصبی بصورت خودکار اضافه و کم می‌شود تا ساختار ما بصورت موثر عمل نماید. مکانیزم خودسازمانده مشابه کار چنگ [۱۹] می‌باشد. افزودن نرون به این صورت انجام می‌گیرد که:

با فرض اینکه مجموعه سازگار از نرون b_j^{act} داریم، بزرگترین خروجی نرون‌ها برابر است با:

$$\Gamma_{max} = \max\left(b_j^{act}\right) \quad j = 1,2,\dots,R(t) \qquad (۴)$$

که R(t) تعداد نرون‌ها در زمان t است. شرایط اضافه شدن نرون جدید در لایه میانی به اینصورت است که:

$$\Gamma_{max}(t) \le \Gamma_{th} \qquad (۵)$$

$\Gamma_{th} \in (0,1)$ آستانه افزودن به منظور اضافه کردن نرون است. اگر معادله (۵) برآورده شود، نرون جدید به لایه میانی اضافه می‌شود. هر نرون دارای یک شاخص اهمیت I_j است [۱۹] که نرون با استفاده از آن حذف می‌شود. به-روزرسانی به اینصورت انجام می‌شود که:

$$I_j(t+1) = \begin{cases} I_j(t) & \text{if} \quad b_j^{act} \ge P_{th} \\ I_j(t).\exp(-\beta) & \text{if} \quad b_j^{act} < P_{th} \end{cases} \qquad (۶)$$

[5] Actor
[6] Critic

[Downloaded from joc.kntu.ac.ir on 2026-11-08]
[DOR: 20.1001.1.20088345.1392.7.3.3.6]

β شاخص طراحی و P_{th} آستانه تغییر است. اگر شاخص اهمیت نرون از آستانه تغییر کمتر باشد، در نتیجه این مقدار کاهش می‌یابد. مقدار اولیه شاخص اهمیت، I_j، برابر یک انتخاب شده است.

$$I_j \le I_{th} \quad (۷)$$

که I_{th} آستانه حذف است. اگر معادله (۷) برای نرون برآورده شود، یک نرون نامناسب تشخیص داده شده و حذف می‌شود.

۲-۱-۲- ارزیاب

شبکه عصبی مصنوعی قسمت ارزیاب را شکل داده است[۱۱]. ارزیاب پاداش p را محاسبه کرده تا خطای اختلاف زمانی[۷] را به حداقل برساند. پاداش مورد نظر برابر است با:

$$P(t) = \sum_{j=1}^{J} w_j^{cri} b_j^{cri} \quad (۸)$$

که در آن J تعداد نرون لایه میانی ارزیاب و w_j^{cri} وزن نرون jام لایه میانی به خروجی است و b_j^{cri} با استفاده از تابع هذلولی مماسی (tanh) تحقق می‌یابد.

ایده‌ی مرکزی و بدیع یادگیری تقویتی بدون شک یادگیری تفاضل زمانی خواهد بود[۲۰]. در اینجا یادگیری عملگر و ارزیاب با استفاده از خطای تفاضل زمانی انجام می‌گیرد، که بصورت معادله (۹) بیان می‌شود:

$$\hat{r}(t) = r(t) + \gamma \max P(t+1) - P(t) \quad (۹)$$

که r پاداش و γ نرخ نزول $(1 > \gamma > 0)$ می‌باشد.

$$r = \alpha_\theta e_t + \alpha_{\dot\theta} \dot{e}_t \quad (۱۰)$$

α_θ و $\alpha_{\dot\theta}$ مقادیر ثابت و e_t خطای خروجی سیستم هستند.

ویژگی خروجی برای دریافت بیشترین پاداش، گین عملگر است. به عبارتی یادگیری عملگر به منظور افزایش پاداش و یادگیری ارزیاب به منظور کاهش خطای TD انجام می‌گیرد. در نتیجه در این مقاله الگوریتم خطای پس انتشار خطا به منظور یادگیری شبکه عصبی استفاده می‌شود. در واقع یادگیری عملگر براساس معادله (۱۱) و یادگیری ارزیاب برطبق معادله (۱۲) خواهد بود.

$$\Delta w_j^{act} = \eta_w^{act} \frac{\partial u}{\partial w_j^{act}} \hat{r} \quad (۱۱)$$

$$\Delta w_j^{cri} = -\eta_w^{cri} \frac{\partial P}{\partial w_j^{cri}} \frac{\partial}{\partial P}\left(\frac{1}{2}\hat{r}^2\right) \quad (۱۲)$$

η_w^{act} و η_w^{cri} بترتیب نرخ یادگیری ارزیاب و عملگر و $\hat{r}$ سیگنال خطا می‌باشند.

از روش پروجکشن[۸] برای محدودکردن w_j^{act} استفاده می‌کنیم، به این معنا که یک قانون به‌روزرسانی برای w_j^{act} بصورت زیر تعریف می‌کنیم:

$$w_j^{act} = \begin{cases} w_j^{act} + \Delta w_j^{act} if \left|w_j^{act} + \Delta w_j^{act}\right| < D \\ w_j^{act} otherwise \end{cases} \quad (۱۳)$$

که D محدوده موثر وزن می‌باشد. در واقع اگر قدرمطلق وزن پس از یادگیری از D کمتر باشد، به‌روز رسانی w_j^{act} انجام می‌گیرد.

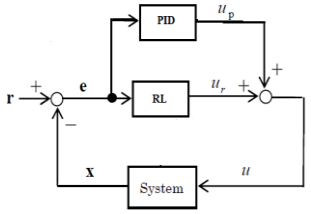
[7] Temporal Difference error (TD)
[8] projection

۲-۲- کنترل‌کننده مبتنی بر یادگیری تقویتی و PID

در ساختار نشان داده شده در شکل (۱)، از ترکیب کنترل‌کننده مبتنی بر یادگیری تقویتی و کنترل‌کننده پسخوردی به منظور کنترل حرکت دست و بهبود کارایی سیستم استفاده شده است. کنترل‌کننده تناسبی- انتگرالی- مشتقی کنترل‌کننده‌ای مبتنی بر فیدبک از خروجی می‌باشد. این کنترل‌کننده یکی از استراتژی‌های متداول در کنترل سیستم‌ها است. در ساختار پیشنهادی، از یک کنترل‌کننده PID، در کنار عملگر و ارزیاب استفاده شده است. با توجه به اینکه مدل بازو دارای سه مفصل بوده است، برای کنترل هر مفصل از یک کنترل‌کننده PID با ساختار تشریح شده در شکل (۳) استفاده می‌شود. تابع تبدیل کنترل-کننده‌PID عبارتست از:

$$G_c(s) = K_p + {K_I}/{s} + K_D s \quad (۱۴)$$

که در آن K_p گین تناسبی، K_I گین انتگرالی و K_D گین مشتقی است. تحقق فیزیکی این کنترل‌کننده توسط یک مدار الکتریکی امکانپذیر است. به منظور کنترل سیستم، از PID به عنوان جبرانساز در کنار راهکار کنترلی مبتنی بر روش یادگیری تقویتی (RLPID) استفاده می‌شود.



شکل ۳: کنترل‌کننده مبتنی بر یادگیری تقویتی و کنترل‌کننده PID

۲-۳- کنترل‌کننده مبتنی بر یادگیری تقویتی و PID و HTC

به منظور بهبود پایداری سیستم کنترل علاوه بر کنترل‌کننده PID یک کنترل‌کننده مقاوم به ترکیب افزوده شده است. این کنترل‌کننده مقاوم که از نوع H_∞ است، HTC[۹] (جبرانساز با دقت دنبال‌کنندگی بالا) نامیده می‌شود. ساختار این کنترل‌کننده در شکل (۴) آمده است. رابطه (۱۵) سیگنال کنترلی خروجی HTC را نشان می‌دهد. هدف عملگر توسط کنترل‌کننده مقاوم مشاهده و خروجی u_h مربوط به HTC بگونه‌ای طراحی می‌شود که عملکرد ردیابی H_∞ را تضمین نماید.

$$u_h = \frac{1}{8\tau^2} e^T P B, \quad B = [0\ 0\ \dots\ 0\ 1]^T \quad (۱۵)$$

τ ثابت است.

$$\delta = 2\tau \frac{g}{\sqrt{g}}$$

δ ثابت میرایی است.از آنجا که g ناشناخته است، فرض می‌کنیم:

$$g_{low} < g < g_{up}$$

در نتیجه داریم:

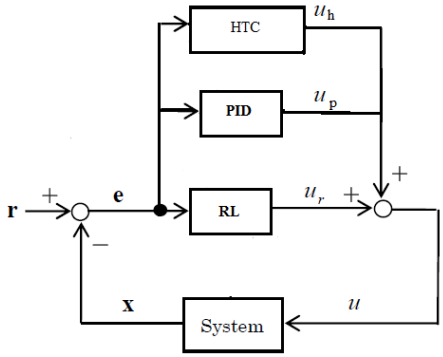
$$\delta = 2\tau \frac{g_{up}}{\sqrt{g_{low}}} \ge 2\tau \frac{g}{\sqrt{g}}$$

تا زمانیکه یادگیری به شکل مطلوبی اتفاق نیافتاده است، HTC نقش موثرتری در کنترل سیستم برعهده دارد، اما پس از اینکه یادگیری به سطح

[9] H_∞ Tracking Compensator

[Downloaded from joc.kntu.ac.ir on 2026-11-08]
[DOR: 20.1001.1.20088345.1392.7.3.3.6]

مطلوبی رسید، هر دو کنترل‌کننده مبتنی بر یادگیری تقویتی و HTC در کنترل عملکرد حرکتی موثر خواهند بود. خروجی H_∞ با استفاده از حل P در معادله لیاپانوف ایجاد می‌شود.



شکل ۴: کنترل‌کننده مبتنی بر یادگیری تقویتی و کنترل‌کننده PID و H_∞

در روش RLPID اثبات تئوری برای تضمین پایداری سیستم کنترل وجود ندارد. اما در حضور کنترل‌کننده HTC می‌توان با تعریف تابع لیاپانوف پایداری سیستم کنترل حلقه بسته را اثبات کرد. در زیر بخش بعد پایداری سیستم اثبات می‌شود.

۲-۳-۱- آنالیز پایداری سیستم کنترل

سیستم غیرخطی مرتبه اول را بصورت زیر در نظر می‌گیریم:

$$\dot{x} = A(x) + B(x)u \quad (۱۶)$$

که x بردار حالت سیستم، f تابع پیوسته محدود ناشناخته، g تابع پیوسته مثبت محدود ناشناخته و u ورودی کنترل به سیستم می‌باشد.

تابع لیاپانوف را با در نظر گرفتن ماتریس قطری ʌ که قطر اصلی آن شامل بهره پس‌خور k می‌باشد ($\Lambda = k^T e$، که e خطای دنبال‌کنندگی)، ماتریس مثبت معین P و ماتریس مثبت معین متعامد Q، که بصورت قراردادی تعیین شده، حل می‌نماییم با این شرط که پایداری هرویتز را برآورده نماید [۱۱].

$$\Lambda^T P + P\Lambda = -Q \quad (۱۷)$$

اگر A و B شناخته شده باشند، ورودی بهینه سیستم برابر است با:

$$u^* = \frac{1}{\mathrm{B}}(-\mathrm{A} + \dot{\mathrm{r}} + \mathrm{k^T e}) \quad (۱۸)$$

که ṙ مشتق اول سیگنال مرجع و e خطای ردیابی (e=r-x) است که باید صفر شود.

اگر سیستم (۱۶) را بصورت زیر در نظر می‌گیریم:

$$\dot{x} = A + Bu + Bu^* - Bu^* \quad (۱۹)$$

مشتق خطای ردیابی برابر است با:

$$\begin{aligned}\dot{\mathrm{e}} = \dot{\mathrm{r}} - \dot{\mathrm{x}} &= (\mathrm{A} + \mathrm{g}u^* - \mathrm{k^T e}) \\ &\quad - (A + Bu + Bu^* - Bu^*) \\ &= -\mathrm{k^T e} + \mathrm{B}(u^* - \mathrm{u})\end{aligned} \quad (۲۰)$$

معادله فضای حالت مشتق خطای ردیابی بصورت زیر می‌باشد:

$$\dot{e} = \Lambda e + B_l(u^* - \mathrm{u}) \quad (۲۱)$$

ورودی کنترلی سیستم مبتنی بر RLPIDHTC بصورت معادله (۲۲) بیان می‌شود:

$$\mathrm{u} = \mathrm{u_r} + \mathrm{u_h} + \mathrm{u_p} \quad (۲۲)$$

$\mathrm{u_h}$ و $\mathrm{u_p}$ و $\mathrm{u_r}$، بترتیب، سیگنال خروجی HTC و PID و RL هستند. با جایگزینی معادله (۲۲) در معادله (۲۱) داریم:

$$\dot{e} = \Lambda e + B_l(u^* - \mathrm{u_r} - \mathrm{u_h} - \mathrm{u_p}) \quad (۲۳)$$

خطای تقریب سیگنال یادگیری تقویتی، کنترل‌کننده PID و ورودی بهینه به صورت زیر می‌باشد:

$$\varepsilon_\mathrm{u} = \mathrm{u}^* - \mathrm{u_r} - \mathrm{u_p} \quad (۲۴)$$

با جایگذاری معادله (۲۴) در (۲۳) داریم:

$$\dot{e} = \Lambda e + B_l(\varepsilon_\mathrm{u} - \mathrm{u_h}) \quad (۲۵)$$

برطبق معادله لیاپانوف، تطبیق به این صورت است:

$$V = e^T P e \quad (۲۶)$$

با مشتق گیری از دو طرف معادله (۲۶) نسبت به زمان، داریم:

$$\dot{V} = \dot{e}^T P e + e^T P \dot{e} \quad (۲۷)$$

با استفاده از معادلات (۱۷) و (۲۳):

$$\dot{V} = -e^T Q e - 2e^T P B_l u_h + 2e^T P B_l \varepsilon_u \quad (۲۸)$$

با قراردادن معادله (۱۵) در (۲۸) داریم:

$$\begin{aligned}\dot{V} = -e^T Q e - \left\{\left(\frac{\sqrt{g}}{2\tau}\right)e^T P B - 2\tau\left(\frac{g}{\sqrt{g}}\right)\varepsilon_u\right\} \\ + (2\tau\frac{g}{\sqrt{g}})\varepsilon_u^T(2\tau\frac{g}{\sqrt{g}})\varepsilon_u{}^2\end{aligned} \quad (۲۹)$$

در نتیجه:

$$\dot{V} \leq = -e^T Q e + (2\tau\frac{g}{\sqrt{g}})\varepsilon_u^T(2\tau\frac{g}{\sqrt{g}})\varepsilon_u \quad (۳۰)$$

از دو طرف معادله (۳۰) انتگرال می‌گیریم:

$$v(t_f) - v(0) \leq -\int_0^{t_f} e^T Q e dt + (2\tau\frac{g}{\sqrt{g}})^2 \int_0^{t_f} \varepsilon_u^T \varepsilon_u dt \quad (۳۱)$$

از آنجا که $v(t_f) \geq 0$ و $v(0) = e^T(0)Pe(0)$، داریم:

$$\int_0^{t_f} e^T Q e dt \leq e^T(0)Pe(0) + (2\tau\frac{g}{\sqrt{g}})^2 \int_0^{t_f} \varepsilon_u^T \varepsilon_u dt \quad (۳۲)$$

که t_f زمان کنترل نهایی می‌باشد. همانطور که از معادله (۳۲) قابل مشاهده است، انتگرال خطای ردیابی e توسط مقادیر بدست آمده توسط ضریب خطای تقریب ε_u و ثابت میرایی δ محدود شده است. به عبارتی اگر خطای تقریب و ثابت میرایی هردو محدود باشد، خطای ردیابی نیز محدود است، در نتیجه پایداری سیستم کنترل اثبات می‌شود.

در نتیجه HTC بصورت معادله (۳۳) بیان می‌شود:

$$\int_0^{t_f} e^T Q e \leq e^T(0)Pe(0) + \delta^2 \int_0^{t_f} \varepsilon_t^T \varepsilon_t dt \quad (۳۳)$$

که t_f زمان کنترل نهایی، ε_t مجموع عدم قطعیت خطای تقریب و اغتشاشات در زمان t و δ ثابت مثبت غیر صفر می‌باشد [۱۱].

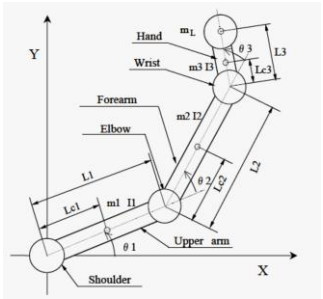
۳- مدل بازوی مجازی

مطالعات شبیه‌سازی مربوط به این تحقیق روی مدلی از بازوی دست انسان شامل سه مفصل انجام شده‌اند. شکل (۵) ساختار کلی این مدل را نشان می‌دهد [۶]. معادله دینامیکی بازوی سه مفصلی انسان که در صفحه افقی حرکت می‌کند، توسط معادله (۳۴) بیان می‌شود.

$$\begin{aligned}\begin{bmatrix}\tau_1\\ \tau_2\\ \tau_3\end{bmatrix} &= \begin{bmatrix}M_{11} & M_{12} & M_{13}\\ M_{21} & M_{22} & M_{23}\\ M_{31} & M_{32} & M_{33}\end{bmatrix}\begin{bmatrix}\ddot{\theta}_1\\ \ddot{\theta}_2\\ \ddot{\theta}_3\end{bmatrix} + \begin{bmatrix}h_1\\ h_2\\ h_3\end{bmatrix} + \begin{bmatrix}D_1 & 0 & 0\\ 0 & D_2 & 0\\ 0 & 0 & D_3\end{bmatrix}\begin{bmatrix}\dot{\theta}_1\\ \dot{\theta}_2\\ \dot{\theta}_3\end{bmatrix} \\ &= M\ddot{\theta} + h + D\dot{\theta}\end{aligned} \quad (۳۴)$$

[Downloaded from joc.kntu.ac.ir on 2026-11-08]
[DOR: 20.1001.1.20088345.1392.7.3.3.6]

که M ماتریس اینرسی، h بردار نیروی گریز از مرکز و جانب مرکز و D بردار نیروی وسیکوزیته می‌باشد.



شکل ۵: مدل بازوی ۳-مفصل انسان[۶]

لینک‌های ۱، ۲ و ۳ به ترتیب متناظر با بالای بازو، ساعد و مچ دست هستند و مفاصل ۱، ۲ و ۳ به ترتیب مربوط به مفاصل شانه، آرنج و ساعد می‌باشند. m_i و m_L به ترتیب جرم لینک i و جرم (1.2kg =) جسم گرفته شده[10]؛ L_i، L_{ci} و I_i به ترتیب طول، فاصله مرکز جرم[11] تا مفصل iام (i=1,2,3) است. مزیت این روش این است که مسیر بازتولید شده با حالت تجربی، بدون توجه به موقعیت هدف، همخوانی دارد [۶].

۳-۱- مکانیزم فریزسازی

در مدل بازوی مجازی به منظور لحاظ کردن اثر فعالیت عضلات آنتاگونیستی از مکانیزم فریزسازی استفاده شده است. سفتی مفاصل دست در اثر پدیده فریزسازی، که در آن عضلات آگونیست و آنتاگونیست همزمان منقبض می‌شوند، افزایش می‌یابد [۶]. در این مکانیزم یک مولفه گشتاور مطابق رابطه (۳۵) به گشتاور مچ دست افزوده شده است، چراکه در اثر گرفتن جسم، بیشترین گشتاور حاصل از فعالیت عضلات آنتاگونیستی به مچ اعمال می‌شود. گشتاور پسخوردی حاصل از مکانیزم فریزسازی τ_3^* در مفصل مچ دست بصورت زیر بیان می‌شود [۶]:

$$\tau_3^* = -k_a\ddot{\theta}_3 - k_v\dot{\theta}_3 - k_p\theta_3 \quad (۳۵)$$

k_a، k_v و k_p گین‌های فیدبک هستند. با لحاظ مکانیزم فریزسازی معادله دینامیکی بازو به صورت معادله (۳۶) تغییر می‌یابد.

$$\begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 + \tau_3^* \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \\ \ddot{\theta}_3 \end{bmatrix} + \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} + \begin{bmatrix} D_1 & 0 & 0 \\ 0 & D_2 & 0 \\ 0 & 0 & D_3 \end{bmatrix} \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{\theta}_3 \end{bmatrix} \quad (۳۶)$$

M ماترسی اینرسی، h بردار نیروی جانب مرکز و گریز از مرکز و D بردار نیروی ویسکوزیته می‌باشد. با این فرض که حرکت بازو در صفحه افقی و در ارتفاع شانه انجام می‌شود، معادلات حالت زیر بدست می‌آید:

$$\dot{x} = \begin{bmatrix} \dot{\theta} \\ -\bar{M}^{-1}\left(\bar{h} + \bar{D}\dot{\theta}\right) + \bar{M}^{-1}\tau \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ I \end{bmatrix} u = A(x) + Bu \quad (۳۷)$$

$$\theta = [\theta_1, \theta_2, \theta_3]^T \ , \tau = [\tau_1, \tau_2, \tau_3]^T \ , u = [\dot{\tau}_1, \dot{\tau}_2, \dot{\tau}_3]^T$$

$$x = [\theta_1, \theta_2, \theta_3, \dot{\theta}_1, \dot{\theta}_2, \dot{\theta}_3, \tau_1, \tau_2, \tau_3]^T$$

I بردار ۳×۳ واحد است. جدول (۱) خلاصه‌ای از پارامترهای فیزیکی بازو را نشان می‌دهد.

جدول ۱: پارامترهای دینامیکی بازو سه مفصل [۶]

پارامترها	شانه	آرنج	مچ
Li(m)	۰.۲۷۶	۰.۲۷۵	۰.۱۳۰
Lci(m)	۰.۱۴۱	۰.۰۸۶	۰.۰۴۳
mi(kg)	۱.۷۴۸	۰.۹۵۱	۰.۴۸۱
Ii($kg.m^2$)	۰.۰۱۲۷	۰.۰۰۴۱	۰.۰۰۰۵
Di(Nm.s/rad)	۰.۳۷۰	۰.۲۹۰	۰.۱۹۰

۴- مطالعات شبیه‌سازی

مطالعات شبیه‌سازی مربوطه در محیط سیمولینک متلب با زمان نمونه-برداری ۰/۰۴ ثانیه انجام شدند. در مطالعات شبیه‌سازی، قابلیت راهکارهای کنترلی مورد مطالعه، در رساندن موقعیت مفاصل بازوی دست از موقعیتهای مختلف به موقعیت تعادلی [0,0,0] ، مورد ارزیابی قرار گرفته است. همچنین کنترل‌کننده ها در حضور اغتشاش، تغییر پارامترهای سیستم و نویز اندازه‌گیری مورد ارزیابی قرار گرفته‌اند. در ادامه نتایج مطالعات شبیه‌سازی به تفصیل ارائه خواهند شد.

۴-۱- کنترل‌کننده RL

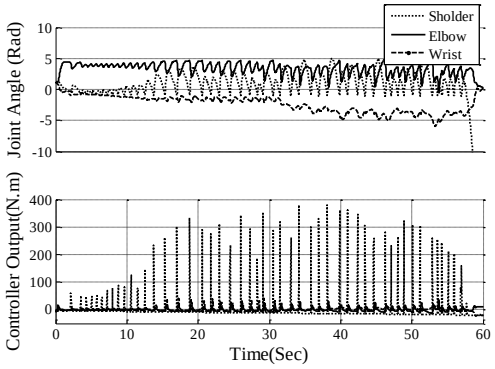
همانطور که پیشتر عنوان شد، به منظور کنترل حرکت بازو از سه کنترل-کننده مجزا برای هر مفصل استفاده می‌شود. در این قسمت سه کنترل‌کننده مبتنی بر یادگیری تقویتی (RL) به منظور کنترل حرکت مفاصل مورد استفاده قرار گرفته است. پارامترهای استفاده شده به منظور شبیه‌سازی کنترل‌کننده RL در جدول (۲) آمده است. کنترل‌کننده RL در چند اپیزود مختلف بر سیستم اعمال شد. اما نتایج بدست آمده در تمام اپیزودها مشابه بودند. لذا در شکل‌های (۶) تا (۸) تغییرات زوایای مفاصل و سیگنال خروجی کنترلی هر سه مفصل به ترتیب برای شرایط اولیه [1,1,1]، [5,5,5] و [-5,-5,-5] در اپیزود اول نشان داده شده‌اند.

جدول ۲: پارامترهای استفاده شده در شبیه‌سازی سیستم کنترل‌کننده RL

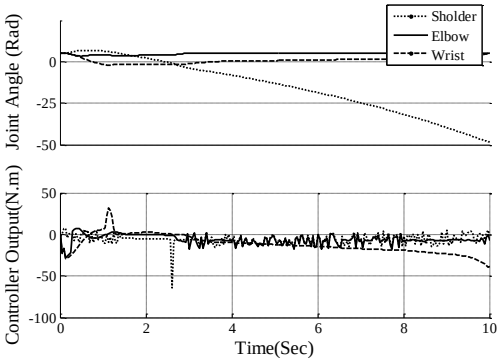
k	{1,5}	γ	0.95
Q	diag{10,10}	α_θ	5
δ	0.3	$\alpha_{\dot{\theta}}$	1
g_{low}	0.6	α_ε	0.1
g_{up}	1.5	η_w^{cri}	0.01
I	2	η_w^{act}	1
J	5	D	500
Γ_{th}	0.3	n	Random(-10:10)
σ_c	0.2	$\bar{\omega}$	0.001
I_{th}	0.01		

[10] handle
[11] Center of mass

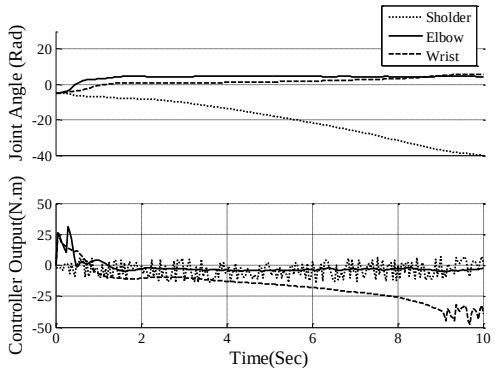
[Downloaded from joc.kntu.ac.ir on 2026-11-08]
[DOR: 20.1001.1.20088345.1392.7.3.3.6]



شکل ۶: زوایای مفاصل و خروجی کنترل‌کننده RL برای سه مفصل شانه، آرنج و دست با شرایط اولیه $\theta_0 = [1,1,1]$.



شکل ۷: زوایای مفاصل و خروجی کنترل‌کننده RL برای سه مفصل شانه، آرنج و دست با شرایط اولیه $\theta_0 = [5,5,5]$.



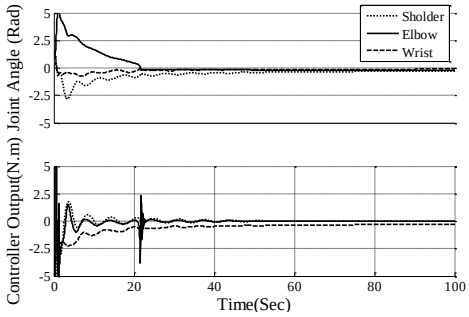
شکل ۸: زوایای مفاصل و خروجی کنترل‌کننده RL برای سه مفصل شانه، آرنج و دست با شرایط اولیه $\theta_0 = [-5,-5,-5]$.

٤-۲- کنترل‌کننده RLPID

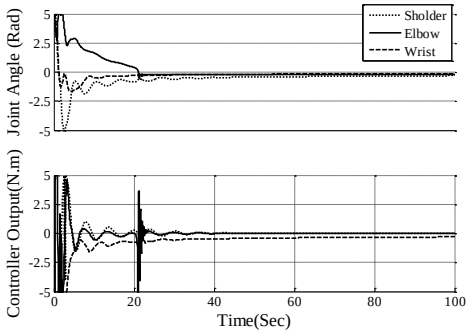
همانطور که از نتایج کنترل‌کننده RL قابل مشاهده است، این کنترل‌کننده ناپایدار است. به منظور بهبود کارایی و پایداری سیستم از کنترل‌کننده PID در کنار کنترل‌کننده RL استفاده خواهد شد. به منظور بررسی عملکرد کنترل‌کننده یادگیری تقویتی در کنار کنترل کنترل‌کننده‌PID، از ساختار شکل (۳) استفاده می‌شود. نتایج حاصل از اعمال همزمان کنترل‌کننده‌های یادگیری تقویتی و PID به مفاصل دست و همچنین خروجی‌های کنترلی با شرایط اولیه[1,1,1]، [5,5,5] و [-5,-5,-5] به ترتیب در شکل‌های (۹) تا (۱۱) آمده است. پارامترهای کنترل-کننده‌های PID بصورت زیر تنظیم شده‌اند:

$$\begin{cases} K_{p1} = 2, K_{i1} = 0.5, K_{d1} = 1 \\ K_{p2} = 5, K_{i2} = 0.5, K_{d2} = 1 \\ K_{p3} = 2, K_{i3} = 1, K_{d3} = 1 \end{cases} \quad (۳۸)$$

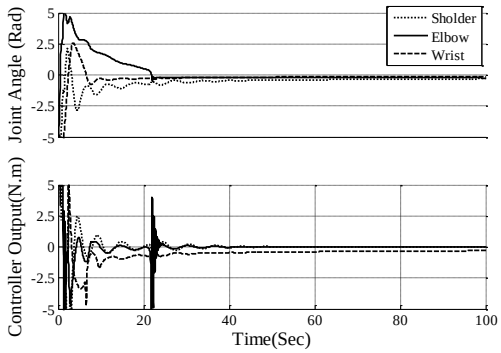
k_{p1}، k_{i1} و k_{d1} گین‌های مربوط به کنترل‌کننده PID مفصل شانه، k_{p2}، k_{i2} و k_{d2} گین‌های مربوط به کنترل‌کننده PID مفصل آرنج و k_{p3}، k_{i3} و k_{d3} گین‌های مربوط به کنترل‌کننده PID مفصل مچ دست می‌باشند، که ساختار کلی آن در شکل (۱) نشان داده شده است. میانگین مربعات خطا و همچنین خطای حالت ماندگار زوایای مفاصل دست با شرایط اولیه متفاوت ، به ترتیب، در جدول‌های (۳) و (۴) آمده است. در جدول (۵) حداکثر فراجهش برای شرایط اولیه [-5,-5,-5] و حداقل فروجهش برای شرایط اولیه [1,1,1] و [5,5,5] آمده است. زمان صعود برای شرایط اولیه [-5,-5,-5] و زمان نزول برای شرایط اولیه [1,1,1] و [5,5,5] در جدول (۶) محاسبه شده است.



شکل ۹: زوایای مفاصل و خروجی کنترل‌کنندهRLPID برای سه مفصل شانه، آرنج و دست با شرایط اولیه $\theta_0 = [1,1,1]$.



شکل ۱۰: زوایای مفاصل و خروجی کنترل‌کننده RLPID برای سه مفصل شانه، آرنج و دست با شرایط اولیه $\theta_0 = [5,5,5]$.



شکل ۱۱: زوایای مفاصل و خروجی کنترل‌کننده RLPID برای سه مفصل شانه، آرنج و دست با شرایط اولیه $\theta_0 = [-5,-5,-5]$.

[Downloaded from joc.kntu.ac.ir on 2026-11-08]
[DOR: 20.1001.1.20088345.1392.7.3.3.6]

جدول ۳: میانگین مربعات خطا زوایای مفاصل دست مربوط به کنترل‌کننده RLPID

	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۰.۴۸۹۶°	۱.۰۴۴۶°	۰.۶۷۴۸°
آرنج	۱.۰۵۲۷°	۱.۱۲۲۱°	۱.۱۸۵۶°
مچ	۰.۰۶۲۵°	۰.۳۰۰۲°	۰.۶۵۹۳°

جدول ۴: خطای حالت ماندگار زوایای مفاصل دست مربوط به کنترل‌کننده RLPID

	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۰.۴۲۵۵°	۰.۸۵۸۰°	۰.۶۰۵۲°
آرنج	۰.۹۹۱۵°	۱.۰۲۵۹°	۱.۰۵۶۴°
مچ	۰.۱۵۶۱°	۰.۵۱۲۰°	۰.۷۸۸۳°

جدول ۵: حداکثر فراجهش (فروجهش) زوایای مفاصل دست مربوط به کنترل‌کننده RLPID

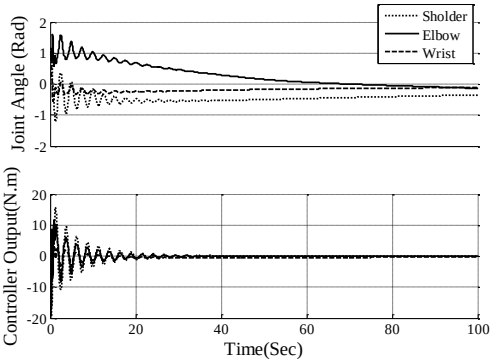
	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	−۲.۷۸۶۴°	−۵.۱۱۹۸°	۲.۱۱۱۱°
آرنج	−۰.۲۴۶۷°	−۰.۳۶۹۲°	۶.۰۵۹۶°
مچ	−۰.۷۴۸۴°	−۱.۶۷۹۱°	۲.۵۸۶۴°

جدول ۶: زمان صعود (نزول) زوایای مفاصل دست مربوط به کنترل‌کننده RLPID

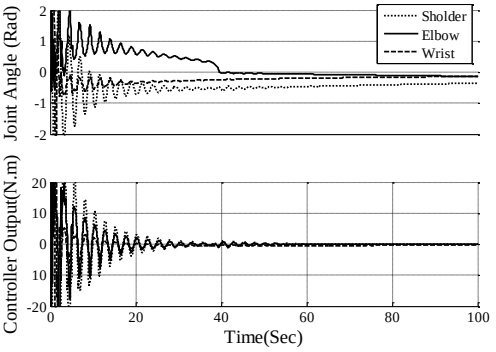
	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۵۲.۵۶	۱۴.۶۴	۱۵.۴
آرنج	۲۱.۸۴	۱۳.۶	۱۴.۵۲
مچ	۳۶.۵۲	۷.۱۲	۵.۶

٤-۳- کنترل‌کننده RLPIDHTC

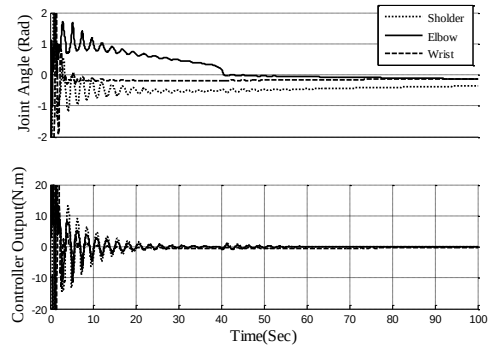
همانطور که پیشتر مشاهده شد کارایی کنترل‌کننده RL مطلوب نبود. در ادامه از کنترل‌کننده RLPID به منظور کنترل سیستم استفاده شد. با وجود کارایی قابل قبول، به منظور بهبود شرایط پایداری سیستم کنترل، یک کنترل کننده مقاوم HTC به سیستم کنترل مبتنی بر RLPID افزوده شد که با عنوان کنترل‌کننده RLPIDHTC معرفی شد. نتایج حاصل از اعمال همزمان کنترل‌کننده‌های یادگیری تقویتی، HTC و PID به منظور رساندن موقعیت مفاصل دست به موقعیت تعادلی [0,0,0] ، از شرایط اولیه [1,1,1]، [5,5,5] و [-5,-5,-5] برای هر سه مفصل در شکل‌های (۱۲) تا (۱۴) آمده است. در جدول‌های (۷) و (۸) نتایج مربوط به میانگین مربعات خطا و خطای حالت ماندگار حاصل از اعمال ساختار کنترلی RLPIDHTC آورده شده‌اند. در جدول (۹) حداکثر فراجهش برای شرایط اولیه [-5,-5,-5] و حداقل فروجهش برای شرایط اولیه [1,1,1] و [5,5,5] آمده است. زمان صعود برای شرایط اولیه [-5,-5,-5] و زمان نزول برای شرایط اولیه [1,1,1] و [5,5,5] در جدول (۱۰) محاسبه شده است.



شکل ۱۲: زوایای مفاصل و خروجی کنترل‌کننده RLPIDHTC برای سه مفصل شانه، آرنج و دست با شرایط اولیه $\theta_0 = [1,1,1]$.



شکل ۱۳: زوایای مفاصل و خروجی کنترل‌کننده RLPIDHTC برای سه مفصل شانه، آرنج و دست با شرایط اولیه $\theta_0 = [5,5,5]$.



شکل ۱۴: زوایای مفاصل و خروجی کنترل‌کننده RLPIDHTC برای سه مفصل شانه، آرنج و دست با شرایط اولیه $\theta_0 = [-5,-5,-5]$.

جدول ۷: میانگین مربعات خطا زوایای مفاصل دست مربوط به کنترل‌کننده RLPIDHTC

	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۰.۲۴۰۹°	۰.۵۱۹۵°	۰.۳۷۶۲°
آرنج	۰.۲۲۶۷°	۰.۲۸۳۰°	۰.۳۱۶۱°
مچ	۰.۰۴۰۲°	۰.۲۶۳۴°	۰.۲۵۵۴°

[Downloaded from joc.kntu.ac.ir on 2026-11-08]
[DOR: 20.1001.1.20088345.1392.7.3.3.6]

جدول ۸: خطای حالت ماندگار زوایای مفاصل دست مربوط به کنترل‌کننده RLPIDHTC

	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۰.۱۸۴۹°	۰.۵۵۹۴°	۰.۴۱۱۶°
آرنج	۰.۳۷۵۶°	۰.۴۴۵۴°	۰.۴۸۱۲°
مچ	۰.۰۹۱۰°	۰.۴۸۱۳°	۰.۴۷۲۸°

جدول ۹: حداکثر فراجهش و فروجهش زوایای مفاصل دست مربوط به کنترل‌کننده RLPIDHTC

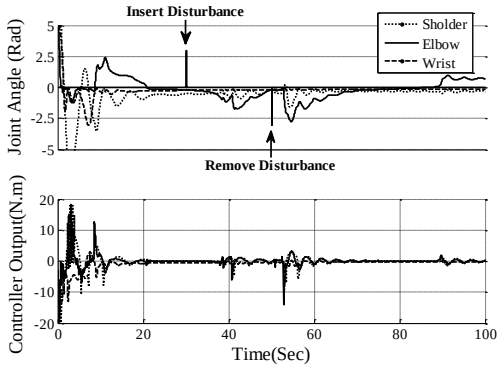
	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	−۱.۲۲۵۷°	−۴.۶۷۸۹°	۴.۱۶۸۲°
آرنج	−۰.۱۳۷۵°	−۰.۶۰۰۴°	۳.۴۰۱۵°
مچ	−۰.۵۵۹۵°	−۲.۵۴۱۳°	۲.۹۱۱۶°

جدول ۱۰: زمان صعود (نزول) زوایای مفاصل دست مربوط به کنترل‌کننده RLPIDHTC

	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۸۱.۲۴	۱۰.۷۲	۴.۴
آرنج	۴۶.۴۴	۱۴.۲۸	۱۲.۲۸
مچ	۳۸.۰۸	۱.۸۴	۲.۱۶

٤-٤- ارزیابی درحضور اغتشاش خارجی

پس از ارزیابی‌های اولیه ملاحظه شد که کنترل کننده‌های غیرمتمرکز مبتنی بر RLPID و RLPIDHTC در کنترل موقعیت دست، کارایی قابل قبولی داشته‌اند. در ادامه به‌منظور مقایسه دقیق‌تر کاریی دو کنترل کننده، قابلیت‌آنهادر مقابله با اعمال اغتشاش مکانیکی خارجی مورد بررسی قرارگرفت. به این منظور، در ثانیه ۳۰ پالسی با دامنه ۳ Nm و در ثانیه ۵۰ پالسی با دامنه ۳ Nm− در مدت ۱ ثانیه به مفاصل دست اعمال شدند. عملکرد کنترل‌کننده RLPID بازای در نظر گرفتن موقعیت اولیه [5,5,5] برای مفاصل، در شکل (۱۵) نشان داده شده است. میانگین مربعات خطا و خطای حالت ماندگار مربوط به زوایای مفاصل با شرایط اولیه مختلف در جدول‌های (۱۱) و (۱۲) آمده است. عملکرد کنترل‌کننده RLPIDHTC نیز بازای در نظر گرفتن موقعیت اولیه [5,5,5] برای مفاصل، در شکل (۱۶) نشان داده شده است. میانگین مربعات خطا و خطای حالت ماندگار مربوط به زوایای مفاصل با شرایط اولیه مختلف در جدول‌های (۱۳) و (۱۴) آمده است.



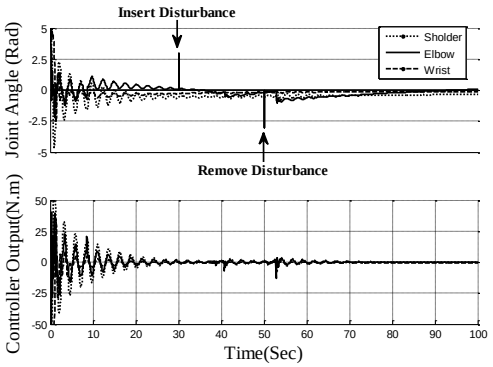
شکل ۱۵: زوایای مفاصل و خروجی کنترل‌کننده RLPID برای سه مفصل شانه، آرنج و دست با شرایط اولیه $\theta_0 = [5,5,5]$ در حضور اغتشاش (td=30s,50s).

جدول ۱۱: میانگین مربعات خطا زوایای مفاصل دست مربوط به کنترل‌کننده RLPID در حضور اغتشاش خارجی

	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۰.۸۲۴۶°	۱.۵۸۰۳°	۰.۷۶۷۵°
آرنج	۰.۶۷۹۸°	۰.۷۵۹۴°	۰.۸۰۵۲°
مچ	۰.۰۹۲۵°	۰.۴۳۴۴°	۰.۳۰۸۰ °

جدول ۱۲: خطای حالت ماندگار زوایای مفاصل دست مربوط به کنترل‌کننده RLPID در حضور اغتشاش خارجی

	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۰.۷۱۷۲°	۱.۱۲۱۷°	۰.۶۷۲۵°
آرنج	۰.۸۱۱۱°	۰.۸۵۶۴°	۰.۸۸۳۲°
مچ	۰.۲۳۵۰°	۰.۶۲۱۲°	۰.۵۱۲۲°



شکل ۱۶: زوایای مفاصل و خروجی کنترل‌کننده RLPIDHTC برای سه مفصل شانه، آرنج و دست با شرایط اولیه $\theta_0 = [5,5,5]$ در حضور اغتشاش(td=30s,50s).

[Downloaded from joc.kntu.ac.ir on 2026-11-08]
[DOR: 20.1001.1.20088345.1392.7.3.3.6]

جدول ۱۳: میانگین مربعات خطا زوایای مفاصل دست مربوط به کنترل-کننده RLPIDHTC در حضور اغتشاش خارجی

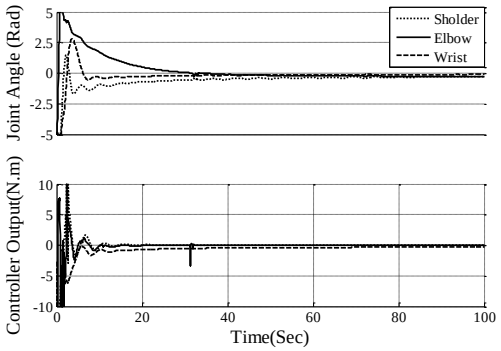
	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۰.۲۶۷۵°	۰.۵۳۰۴ °	۰.۴۸۳۸°
آرنج	۰.۱۳۵۰°	۰.۲۱۳۹ °	۰.۱۸۳۹°
مچ	۰.۰۳۹۵°	۰.۲۷۴۵°	۰.۲۱۹۱°

جدول ۱۴: خطای حالت ماندگار زوایای مفاصل دست مربوط به کنترل-کننده RLPIDHTC در حضور اغتشاش خارجی

	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۰.۲۴۷۱°	۰.۵۳۹۴°	۰.۵۲۳۵°
آرنج	۰.۳۴۵۵ °	۰.۴۱۵۳°	۰.۴۰۰۶°
مچ	۰.۰۹۴۰°	۰.۴۹۱۸°	۰.۴۲۴۲°

٤-٥- ارزیابی در حضور تغییر پارامترهای سیستم

در این مرحله، کارایی کنترل‌کننده‌های RLPID و RLPIDHTC در شرایطی که پارامترهای سیستم در طول زمان تغییر می‌کنند منظور بررسی قرار می‌گیرد. برای این منظور طول مفاصل، جرم لینک‌ها، طول و فاصله مرکز جرم تا مفصل را به اندازه ٪۵۰± مقدار اولیه آن تغییر می‌دهیم. نتایج مربوط به زاویه مفاصل و خروجی کنترلی کنترل‌کننده RLPID با در نظر گرفتن اثر تغییر پارامترها و شرایط اولیه [5,5,5] در شکل (۱۷) نشان داده شده است. میانگین مربعات خطا و خطای حالت ماندگار زوایای مفاصل در شرایط اولیه مختلف نیز محاسبه شده و در جدول‌های (۱۵) و (۱۶) آمده است. عملکرد کنترل‌کننده RLPIDHTC نیز بازای در نظر گرفتن موقعیت اولیه [5,5,5] برای مفاصل و اثر تغییر پارامتر، در شکل (۱۸) نشان داده شده است. میانگین مربعات خطا و خطای حالت ماندگار مربوط به زوایای مفاصل با شرایط اولیه مختلف در جدول‌های (۱۷) و (۱۸) آمده است.



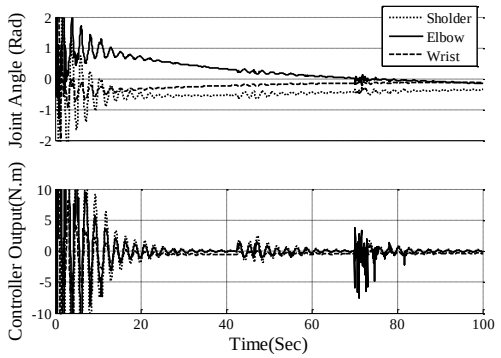
شکل ۱۷: زوایای مفاصل و خروجی کنترل‌کننده RLPID برای سه مفصل شانه، آرنج و دست با شرایط اولیه $\theta_0 = [5,5,5]$ در اثر تغییر پارامترهای سیستم.

جدول ۱۵: میانگین مربعات خطا زوایای مفاصل دست مربوط به کنترل-کننده RLPID در اثر تغییر پارامتر

	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۰.۴۷۴۱°	۰.۹۱۴۰°	۰.۶۶۵۳°
آرنج	۱.۰۵۰۲°	۱.۰۹۱۷°	۱.۲۱۹۶°
مچ	۰.۰۶۲۳°	۰.۳۰۶۸°	۰.۵۸۸۱°

جدول ۱۶: خطای حالت ماندگار زوایای مفاصل دست مربوط به کنترل-کنندهRLPID در اثر تغییر پارامتر

	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۰.۴۰۶۷°	۰.۷۷۸۱°	۰.۵۹۷۱°
آرنج	۰.۹۹۰۳°	۱.۰۱۱۰°	۱.۰۷۲۴°
مچ	۰.۱۵۵۵°	۰.۵۱۸۴°	۰.۷۴۱۷°



شکل ۱۸: زوایای مفاصل و خروجی کنترل‌کننده RLPIDHTC برای سه مفصل شانه، آرنج و دست با شرایط اولیه $\theta_0 = [5,5,5]$ در اثر تغییر پارامترهای سیستم.

جدول ۱۷: میانگین مربعات خطا زوایای مفاصل دست مربوط به کنترل‌کننده RLPIDHTC در اثر تغییر پارامتر

	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۰.۲۴۸۹°	۰.۴۸۶۹°	۰.۴۲۱۶°
آرنج	۰.۲۳۵۶°	۰.۲۸۷۷°	۰.۲۹۸۸°
مچ	۰.۰۴۱۰°	۰.۲۵۴۲°	۰.۲۲۶۴°

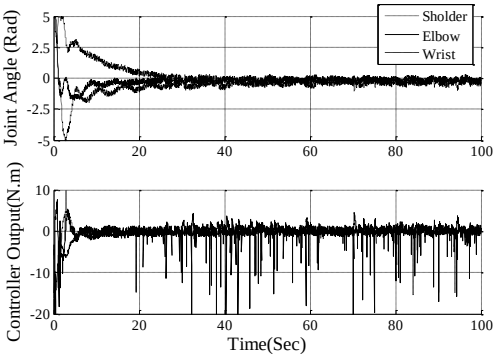
جدول ۱۸: خطای حالت ماندگار زوایای مفاصل دست مربوط به کنترل‌کننده RLPIDHTC در اثر تغییر پارامتر

	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۰.۲۰۷۳°	۰.۵۲۹۷°	۰.۴۶۱۳°
آرنج	۰.۳۸۶۷°	۰.۴۵۰۲°	۰.۴۶۴۰°
مچ	۰.۰۹۵۷°	۰.۴۷۱۶°	۰.۴۴۰۹°

[Downloaded from joc.kntu.ac.ir on 2026-11-08]
[DOR: 20.1001.1.20088345.1392.7.3.3.6]

٤-٦- ارزیابی درحضور نویز اندازه‌گیری

در این قسمت کارایی کنترل‌کننده‌های RLPID و RLPIDHTC در حضور نویز اندازه‌گیری به منظور بررسی می‌شود. به منظور لحاظ کردن عامل نویز در شبیه‌سازی‌ها، نویزی با توزیع یکنواخت با فاصله پیک تا پیک ۰.۵ به سیستم اعمال شده است. نتایج مربوط به زاویه مفاصل و خروجی کنترلی کنترل‌کننده RLPID با در نظر گرفتن اثر نویز اندازه‌گیری و شرایط اولیه [5,5,5] در شکل (۱۹) آمده است. نتایج محاسبه میانگین مربعات خطا مفاصل و خطای حالت ماندگار زوایای مفاصل با شرایط اولیه مختلف در جدول‌های (۱۹) و (۲۰) آمده است. نتایج مربوط به زاویه مفاصل و خروجی کنترلی کنترل‌کننده RLPIDHTC نیز بازای شرایط اولیه [5,5,5] و در نظر گرفتن اثر نویز اندازه‌گیری در شکل (۲۰) نشان داده شده است. میانگین مربعات خطا و خطای حالت ماندگار زوایای مفاصل نیز محاسبه شده و به ترتیب در جداول (۲۱) و (۲۲) آمده است.



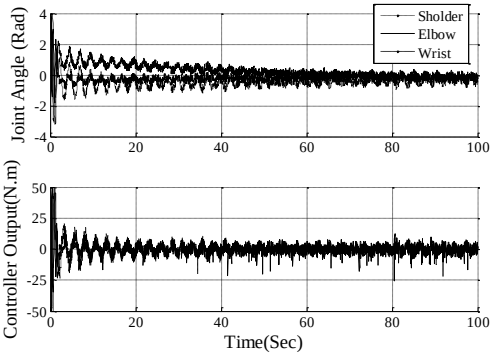
شکل ۱۹: زوایای مفاصل و خروجی کنترل‌کننده RLPID برای سه مفصل شانه، آرنج و دست با شرایط اولیه $\theta_0 = [5,5,5]$ در حضور نویز اندازه‌گیری.

جدول ۱۹: میانگین مربعات خطا زوایای مفاصل دست مربوط به کنترل‌کننده RLPIDدر حضور نویز اندازه‌گیری

	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۰.۵۱۸۸°	۱.۰۳۳۶°	۰.۷۹۹۹°
آرنج	۱.۰۶۶۷°	۱.۱۵۵۲°	۱.۲۹۸۸°
مچ	۰.۰۹۵۰°	۰.۳۳۴۰°	۰.۴۸۸۲°

جدول ۲۰: خطای حالت ماندگار زوایای مفاصل دست مربوط به کنترل‌کننده RLPID در حضور نویز اندازه‌گیری

	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۰.۴۶۰۰°	۰.۸۴۹۶°	۰.۶۹۹۱°
آرنج	۰.۹۹۹۲°	۱.۰۴۳۲°	۱.۱۰۹۸°
مچ	۰.۲۳۷۰°	۰.۵۴۲۹°	۰.۶۷۰۵°



شکل ۲۰: زوایای مفاصل و خروجی کنترل‌کننده RLPIDHTC برای سه مفصل شانه، آرنج و دست با شرایط اولیه $\theta_0 = [5,5,5]$ در حضور نویز اندازه‌گیری.

جدول ۲۱: میانگین مربعات خطا زوایای مفاصل دست مربوط به کنترل‌کننده RLPIDHTC در حضور نویز اندازه‌گیری

	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۰.۳۱۰۶°	۰.۴۳۱۵°	۰.۳۹۰۵°
آرنج	۰.۲۶۰۶°	۰.۳۳۴۸°	۰.۳۳۸۹°
مچ	۰.۰۷۲۴°	۰.۲۶۸۶°	۰.۱۸۱۰°

جدول ۲۲: خطای حالت ماندگار زوایای مفاصل دست مربوط به کنترل‌کننده RLPIDHTC در حضور نویز اندازه‌گیری

	[1,1,1]	[5,5,5]	[-5,-5,-5]
شانه	۰.۳۲۲۶°	۰.۴۷۴۷°	۰.۴۲۸۴°
آرنج	۰.۴۱۸۹°	۰.۵۰۰۸°	۰.۵۰۵۵°
مچ	۰.۲۰۰۹°	۰.۴۸۵۶°	۰.۳۸۵۴°

۵- ارزیابی مبتنی بر داده‌های انسانی

پس از شبیه‌سازی‌های کامپیوتری ملاحظه شد که کارایی کنترل‌کننده RLPIDHTC نسبت به RLPID بهتر است. لذا به منظور نوعی ارزیابی کاملتر و واقع‌گرایانه‌تر کنترل‌کننده RLPIDHTC، آزمایش‌های انسانی طراحی و انجام شدند. در آزمایش‌های انسانی الگوی فعالیت عضلات مفاصل دست، طی عملکرد حرکت دست با استفاده از سیگنال‌های الکترومایوگرام عضلات، استخراج می‌شوند. چهار داوطلب سالم (۱۹-۲۶ ساله) راست دست در آزمایش‌ها شرکت داشتند، هیچ یک از افراد شرکت‌کننده دارای مشکلات حرکتی نبودند. به منظور ثبت سیگنال‌های عضلانی از یک دستگاه پاورلب ۴ کاناله و دو دستگاه بایوامپ استفاده شد. نرم افزار استفاده شده به منظور ثبت لب‌چارت می‌باشد که قابلیت استفاده در محیط متلب را دارد. به منظورثبت سیگنال‌های الکترومایوگرام سطحی، تنظیمات زیر در سیستم قبل از شروع ثبت انجام شد. برای پوشش محدوده فرکانسی سیگنال EMG و نیز حذف انحراف از خط پایه و آرتیفکت‌های حرکتی و تداخلات مغناطیسی احتمالی، از یک فیلتر بالاگذر و یک فیلتر پایین‌گذر استفاده شد، همچنین برای حذف نویز برق شهر، فیلتر Notch اعمال گردید (فیلتر سیستم ثبت داده پاورلب). فرکانس نمونه‌برداری با در نظر گرفتن حداکثر فرکانس موجود در محتوای فرکانسی سیگنالEMG و

[Downloaded from joc.kntu.ac.ir on 2026-11-08]
[DOR: 20.1001.1.20088345.1392.7.3.3.6]

اعمال نرخ نایکوئیست برای اطمینان از عدم تداخل برابر ۱ کیلوهرتز و محدوده اندازه‌گیری دامنه سیگنالEMG، ۲میلی‌ولت در نظر گرفته شد. در پایان نیز به منظور انجام پردازش پوش سیگنال گرفته شده است.

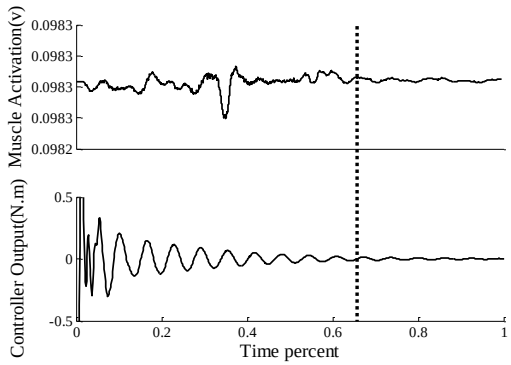
در این آزمایش الکترودهای ثبت بر روی عضلات دلتوئید[12] شانه، ترایسپس براچی[13] ساعد و اکستنسور کارپی ردیالیس برویس[14] قرار داده شده است. شکل (۲۱) سیستم آزمایش را نشان می‌دهد.از فرد خواسته شد در حالیکه نشسته است حرکات مسطح را برطبق الگو رسم شده انجام دهد. در این آزمایشات سه حالت بررسی می‌شود، که بر طبق آن هر سه مفصل دست در زوایای ۶۰ درجه، ۴۰ درجه و ۲۰ درجه قرار می‌گیرند. به منظور هماهنگی ثبت-ها از سوژه می‌خواهیم ابتدا دست خود را در موقعیت موردنظر قرار دهد و منتظر علامت اول باشد، با شنیدن علامت اول سوژه دست خود را به سمت هدف، هر سه مفصل به سمت زاویه صفر درجه، حرکت می‌دهد. در پایان از سوژه درخواست شد تا زمانیکه علامت دوم را بشنود، دست خود را در موقعیت هدف نگه دارد.

انجام حرکات به سه در جه آزادی محدود شده و مفصل شانه مبدا مختصات می‌باشد. هر آزمایش ۶ بار برای هر فرد تکرار و در کل برای هر فرد در مجموع سه حالت آزمایش، ۱۸ ثبت انجام گرفت. در بین هر حرکت به فرد فرصت کوتاهی برای استراحت داده می‌شود تا از ایجاد خستگی عضلانی جلوگیری شود.

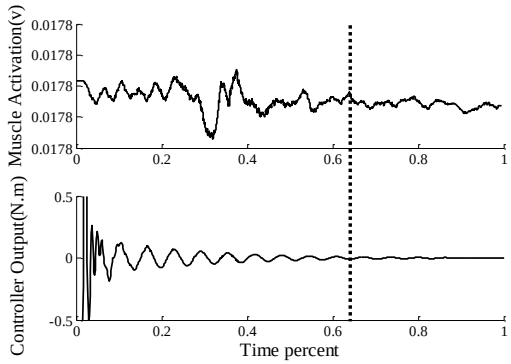


شکل ۲۱: نحوه انجام آزمایش‌های انسانی .

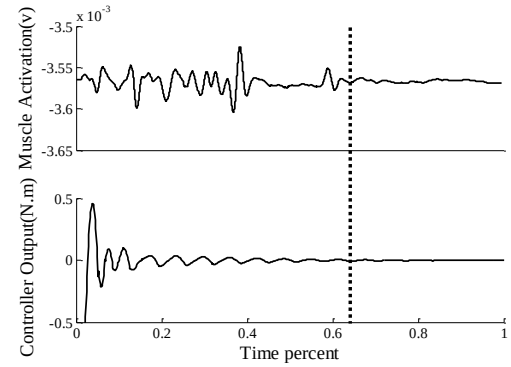
پس از ثبت داده‌های انسانی، زمانبندی تغییرات خروجی کنترل‌کننده در مطالعات شبیه‌سازی که همان گشتاور ایجاد شده در مفاصل هستند، با زمانبندی الگوی فعالیت عضلات که از داده‌های انسانی ثبت شده استخراج شده‌اند، مقایسه شده است. شایان ذکر است در مقایسه انجام شده تنها زمانبندی الگوی فعالیت عضلات و خروجی کنترل‌کننده مدنظر بوده است و هیچگونه مقایسه‌ای به لحاظ دامنه و فاز فعالیت عضلات و خروجی کنترل‌کننده انجام نشده است. در حقیقت تنها نوعی ارزیابی ابتدایی بر اساس داده‌های انسانی انجام گرفته است. شکل (۲۲) خروجی کنترل‌کننده پیشنهادی و الگوی فعالسازی عضله دلتوئید شانه را نشان می‌دهد. خروجی کنترل‌کننده پیشنهادی و الگوی فعالسازی عضله ترایسپس براچی در شکل (۲۳) نشان داده شده است. در شکل (۲۴) خروجی کنترل‌کننده پیشنهادی و الگوی فعالسازی عضله بازکننده مچ دست (کارپی ردیالیس برویس) آورده شده است. در واقع بوسیله نتایج ارائه شده در شکل‌های (۲۲) تا (۲۴)، نحوه زمانبندی فعالیت عضلات با زمانبندی فعالیت سیگنالهای کنترلی قابل مقایسه خواهند بود.



شکل ۲۲: خروجی کنترل‌کننده و الگوی فعالسازی عضله دلتوئید



شکل ۲۳: خروجی کنترل‌کننده و الگوی فعالسازی عضله ترایسپس براچی



شکل ۲۴: خروجی کنترل‌کننده و الگوی فعالسازی عضله اکستنسور کارپی ردیالیس برویس

٦- جمع‌بندی

در این پژوهش، یک راهکار کنترلی به‌منظور کنترل حرکت مدلی از بازوی دست با سه مفصل ارائه شد. در مدل مورد استفاده، اثر فعالسازی همزمان عضلات آگونیست و آنتاگونیست مفصل مچ دست با استفاده از مکانیزم فریزسازی لحاظ گردید. با الهام از عملکرد سیستم اعصاب مرکزی در کسب مهارتهای حرکتی، راهکار کنترلی ارئه شده مبتنی بر یکی از الگوریتم-هاییادگیری تقویتی توسعه یافت. مطابق نتایج به دست آمده، سیستم کنترل حرکت بازوی دست مبتنی بر کنترل‌کننده RL ناپایدار است. به منظور بهبود عملکرد سیستم از یک کنترل‌کننده پس‌خوردی، که این کنترل‌کننده PID بوده است، در کنار ساختار کنترلی RL استفاده شد. با توجه به نتایج مربوط به راهکار کنترلی RLPID، روشن می‌شود که کارایی کنترل‌کننده PID در کنار RL

[12]Deltoideus
[13]Triceps brachii
[14]Extensor carpi radialisbrevis

[Downloaded from joc.kntu.ac.ir on 2026-11-08]
[DOR: 20.1001.1.20088345.1392.7.3.3.6]

مطلوب است. میانگین مربعات خطا و خطای حالت ماندگار روش RLPID در مقایسه با محدوده تغییرات حرکت زوایای مفاصل کمتر است. نتایج نشان می‌دهد که کنترل‌کننده RLPID به ازای سه شرط اولیه مختلف برای مفاصل دست توانسته همگرایی با دقت قابل قبولی داشته باشد، البته نوساناتی دیده می‌شود که به دلیل دینامیک متغیر سیستم می‌باشد. با این حال به لحاظ تئوری اثباتی برای تضمین پایداری نداریم. به منظور تقویت این روش و تضمین پایداری سیستم از یک کنترل‌کننده HTC در کنار RLPID استفاده شد. از مقایسه روش‌های ارائه شده در قسمت نتایج روشن می‌شود که بکارگیری هر دو کنترل‌کننده PID و HTC در کنار کنترل‌کننده RL بمنظور کنترل مدل سه مفصل دست با در نظر گرفتن مکانیزم فریزسازی قابل قبول است است. همانطور که از جداول (۳) تا (۲۲) قابل مشاهده است، میزان میانگین مربعات خطا و خطای حالت ماندگار روش RLPIDHTC در مقایسه با محدوده تغییرات حرکت دست و هچنین روش RLPID به میزان قابل توجهی کمتر است. در مجموع کنترل‌کننده RLPIDHTC عملکرد بهتری را به نسبت کنترل‌کننده RLPID به لحاظ حداکثر فراجهش و طمان صعود داشته است. از طرفی این کنترل‌کننده در سه تست شبیه‌سازی طراحی شده به منظور ارزیابی ساختار کنترلی شامل اثر اغتشاش خارجی، تغییر پارامترهای سیستم و اثر نویز اندازه‌گیری نیز موفق عمل کرده است. میزان خطای محاسبه شده در سه تست شبیه‌سازی برای روش RLPIDHTC از روش RLPID به میزان قابل ملاحظه‌ای کمتر است. لذا ساختار کنترلی مبتنی بر RLPIDHTC روش پیشنهادی ما است.

در قسمت ارزیابی انسانی، الگوی فعالیت عضلات مفاصل دست، طی عملکرد حرکت دست با استفاده از سیگنال‌های الکترومایوگرام عضلات، استخراج شدند. همانطور که در شکل‌های (۲۲) تا (۲۴) نیز قابل مشاهده است شباهت قابل ملاحظه‌ای، به لحاظ زمانبندی تغییرات، بین خروجی کنترل‌کننده روش پیشنهادی و میزان فعالیت عضلات وجود دارد دارد. گشتاور ایجاد شده باعث تغییر فعالیت عضلانی می‌شود که کنترل‌کننده پیشنهادی به خوبی توانسته این تطابق با حالت طبیعی را ایجاد نماید. مجددا تاکید می‌شود که در مقایسه الگوی فعالیت عضلات و خروجی کنترل‌کننده، تنها زمانبندی فعالیت و عدم فعالیت مدنظر بوده است و هیچگونه مقایسه‌ای به لحاظ دامنه و فاز فعالیت عضلات و خروجی کنترل‌کننده انجام نشده است. در واقع پس از اینکه دست به هدف رسیده و میزان تغییر فعالیت عضلات کاهش می‌یابد، خروجی کنترل‌کننده نیز نوسانات کمی دارد، به عبارتی گشتاور تولید شده کاهش یافته است. نتایج نشان می‌دهند که همزمانی قابل ملاحظه‌ای بین زمان تغییرات سیگنال خروجی کنترل‌کننده با تغییرات الگو فعالیت عضلات دست فرد سالم وجود دارد. به عبارت دیگر نتایج نشان داده شده که به لحاظ بازه زمانی مربوط به فعالیت و عدم فعالیت عضلات و سیگنال خروجی کنترل‌کننده هم‌خوانی وجود دارد. البته به لحاظ دامنه تفاوت وجود دارد که دلیل آن عدم وجود عضله در مدل بیمار مجازی است و از طرفی مشخصات مدل بازو ارائه شده با حالت طبیعی یکسان نیستند.

یکی از موفق‌ترین راهکارهای موجود به منظور توانبخشی حرکتی، استفاده از پروتزهای عصبی است که از تحریک الکتریکی عملکردی به منظور فعال‌سازی عضلات از کار افتاده استفاده می‌کنند. در واقع از پروتزهای سیستم عصبی می‌توان به منظور انجام فعالیت‌هایروزانه زندگی برای افراد دچار فلج در بالا تنه و پایین تنه استفاده کرد [۲۳] - [۲۱]. افراد دچار آسیب‌های نخاعی اغلب برای حرکت دادن اندام‌های‌شان ناتوانند، اگرچه بیشتر اعصاب و عضلات آنها سالم هستند. تحریک الکتریکی عملکردی، این عضلات را برای انجام مجدد حرکات فعال می‌کند. کنترل حلقه بسته با استفاده از تحریک الکتریکی عملکردی، عملکرد عضله را بهبود می‌بخشد. در واقع آنچه در این مقاله ارائه شد، معرفی یک راهکار کنترلی حلقه بسته خودسازمانده مبتنی بر یادگیری تقویتی است که ضمن تضمین پایدای سیستم حلقه بسته و مقاوم بودن در برابر تغییر پارامترهای سیستم، اغتشاشهای خارجی و نویز اندازه‌گیری، قادر بوده است عملکرد کنترل حرکت مدلی از بازوی دست را با کارایی قابل قبولی انجام دهد. چنین راهکار کنترلی می‌تواند در راستای توسعه پروتزهای عصبی مربوط به کنترل حرکت دست، مثمر ثمر واقع شود.

مراجع

[1] F.M.M.O. Camposand J.M.F. Calado, "Approaches to human arm movement control—A review," *Annual Reviews in Control*, vol. 33, Issue. 1, pp. 69-77, April. 2009.
[2] JIANG Yu, JIANG Zhong-Ping and QIAN Ning, "Optimal Control Mechanisms in Human Arm Reaching Movements," *In Proc. 30th Chinese Control.* Conf. July 22-24. 2011.
[3] Magnus J. E. Richardson and Tamar Flash, "Comparing Smooth Arm Movements with the Two-Thirds Power Law and the Related Segmented-Control Hypothesis*," J.Neurosci*, vol. 22, no. 18, pp. 8201–8211, 2002.
[4] Elizabeth B. Torresand David Zipser, "Simultaneous control of hand displacements and rotations in orientation-matching experiments," *J. Appl. Physiol.*, vol. 96, pp.1978-1987, 2004.
[5] Yasuhiro Wada, Yuichi Kaneko, Eri Nakano, Rieko OsuandMitsuoKawato, "Quantitative examinations for multi joint arm trajectory planningÐusing a robust calculation algorithm of the minimum commanded torque change trajectory," *J Neural Networks*, vol. 14, pp. 381-393, 2001.
[6] Toshikazu Matsui, KoukiTakeshita and Takahisa Shibusawa "Effectiveness of Human Three-Joint Arm's Optimal Control Model Characterized by Hand- Joint's Freezing Mechanism in Reproducing Constrained Reaching Movement Characteristics," *In Proc. Int. Joint. Conf*, pp. 1206-1211, 209.
[7] Weiwei Li, Emanuel Todorov and Xiuchuan Pan, "Hierarchical Feedback and Learning for Multi-joint Arm Movement Control," *In Proc. 27thAnnu. Conf. EMBS*, 2005.
[8] Kathleen M.Jagodnik and Antonie J. vandenBogert "Optimization and evaluation of a proportional derivative controller for planar arm movement," *J. Biomech*., vol. 43, pp. 1086–1091, 2010.
[9] Qinmin Yang and SarangapaniJagannathan, "Reinforcement Learning Controller Design for Affine Nonlinear Discrete-Time Systems using Online Approximators," *IEEE Trans. Syst. Man, Cybern*. vol. 42, no. 2, pp.377-390, April. 2012.
[10] H.Kambara, J.Kim, M.Sato and Y.Koike, "learning arm's posture control using reinforcement learning and feedback-error-learning," *In Proc. 26thAnnu. Int. Conf. EMBS* , vol. 1, pp. 486 - 489, 2004.
[11] Shogo Uchiyama, Masanao Obayashi, Takashi Kuremoto and Kunikazu Kobayashi, "Robust Reinforcement Learning Control System with H∞ Tracking Performance Compensator," *In Proc. 11th Cont. Auto. Syst. Conf*, 2011.
[12] Magnus J. E. Richardson and Tamar Flash, "Comparing Smooth Arm Movements with the Two-Thirds Power Law and the Related Segmented-Control Hypothesis*," J.Neurosci*, vol. 22, no. 18, pp. 8201–8211, 2002.
[13] Daniele Caligiore, Eugenio Guglielmelli, Anna M. Borghi, DomenicoParisi and GianlucaBaldassarre "A Reinforcement Learning Model of Reaching Integrating Kinematic and Dynamic

[Downloaded from joc.kntu.ac.ir on 2026-11-08]
[DOR: 20.1001.1.20088345.1392.7.3.3.6]

Control in a Simulated Arm Robot," *In Proc. 9th IEEE Int. Conf. development. learning.*, pp. 211-218, 2010.

[14] Philip Thomas, Michael Branicky,Antonie van den Bogert and Kathleen Jagodnik, "Application of the Actor-Critic Architecture to Functional Electrical Stimulation Control of a Human Arm," *In Proc. Innov. Appl. Artif. Intell. Conf*., pp. 165–172, 2009.

[15] Qinmin Yang, Jonathan Blake VanceandS.Jagannathan, "Control of Nonaffine Nonlinear Discrete-Time Systems Using Reinforcement-Learning-Based Linearly Parameterized Neural Networks," *IEEE Trans. Syst. Man, Cybern*. vol. 38, no. 4, pp.994-1001, August. 2008.

[16] Lena AbbasiBrujeni, Jong Min LeeandSirish L. Shah, "Dynamic Tuning of PI-Controllers based on Model-free Reinforcement Learning Methods," *In Proc. Cont. Auto. Syst. Conf*, pp. 453-458, 2010.

[17] F.L. LewisandKyriakos G. Vamvoudakis, "Optimal Adaptive Control for Unknown Systems Using Output Feedback by Reinforcement Learning Methods," *In Proc. 8th Cont. Auto. Conf*, 2010.

[18] MinijaTamosiunaite, TamimAsfour and FlorentinWörgötter, "Learning to reach by reinforcement learning using a receptive field based function approximation approach with continuous actions," *Biol. Cybern*. vol. 100, no. 1, pp. 249–260, 2009.

[19] Kuo-Hsiang Cheng, "Auto-structuring fuzzy neuralsystem for intelligent control," *Journal of The FranklinInstitute*, vol.346, pp.267-288, 2009.

[20] Sutton, R. S. and Barto, A. G., "Reinforcement learning: An introduction,"

[21] C.L.LYNCH and M.R.POPOVIC, "Functional Electrical Stimulation," *IEEE CONTROL SYSTEMS MAGAZINE*, pp. 40-50, april 2008.

[22] P.S.Thomas, M.S.Branicky, A.V.D. Bogert, and K. Jagodnik, "FES Control of a Human Arm Using Reinforcement Learning" , 2007.

[23] E.B.Marsolais and R.Kobetic, "Development of a practical electrical stimulation system for restoring gait in the paralyzed patient," *Clin. Ortho. & Rel. Res.*, vol. 233, pp. 64–74, 1988.

[Downloaded from joc.kntu.ac.ir on 2026-11-08]

[DOR: 20.1001.1.20088345.1392.7.3.3.6]