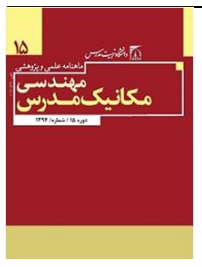
مجله مهندسی مکانیک مدرس، اردیبهشت 1395، دوره 16، شماره 2، ص‌ص 59-68

ماهنامه علمی پژوهشی

مهندسی مکانیک مدرس

mme.modares.ac.ir



کنترل‌کننده آموزش‌پذیر برگرفته از ساختار سیستم عصبی جهت تولید رفتارهای پیچیده در ربات انسان‌نما

ریحانه پرنده[1]، حامد شهبازی[2*]، کمال جمشیدی[3]، بهنام خدابنده جهرمی[4]

1- کارشناسی ارشد، مهندسی کامپیوتر، دانشگاه اصفهان، اصفهان
2- استادیار، مهندسی مکانیک، دانشگاه اصفهان، اصفهان
3- دانشیار، مهندسی کامپیوتر، دانشگاه اصفهان، اصفهان
4- کارشناسی ارشد، مهندسی مکانیک، دانشگاه اصفهان، اصفهان
*اصفهان، صندوق‌پستی 37935619 ، shahbazi@eng.ui.ac.ir

اطلاعات مقاله

مقاله پژوهشی کامل
دریافت: 21 مهر 1394
پذیرش: 19 دی 1394
ارائه در سایت: 07 بهمن 1394

کلید واژگان:
ربات انسان‌نما
شبکه عصبی
کنترل
یادگیری تقلیدی

چکیده

در این مقاله مسئله‌ی کنترل حرکت ربات‌های انسان‌نما مورد بررسی قرار می‌گیرد. مسئله‌ی یادگیری حرکات مشابه انسان به‌خصوص حرکات موزون و تکرارشونده به ربات‌های انسان‌نما چالشی بزرگ در زمینه علم رباتیک محسوب می‌شود. یادگیری تقلیدی که زیرمجموعه‌ای از یادگیری بانظارت است، یک فرم اصلی جهت یاد دادن کارهای پیچیده به ربات به شمار می‌آید و بر این اساس استوار است که یک سیستم مصنوعی می‌تواند حجم بالایی از اطلاعات را از طریق یادگیری تقلیدی از فرد آموزش‌دهنده یاد بگیرد. روش اصلی به کار گرفته شده در این تحقیق جهت یادگیری حرکات موزون و تکرارشونده به ربات انسان‌نمای نائو، یادگیری تقلیدی است. تکنیک اصلی به‌کار رفته در این پژوهش استفاده از ساختارهای عصبی موسوم به مولد مرکزی الگو می‌باشد که قادر است الگوی موردنیاز برای حرکت در یک ربات را براساس یک نوع آموزش نمایشی به دست آورد. نحوه‌ی طراحی سیستماتیک این ساختار عصبی مصنوعی که در علم کنترل کاربردهای فراوانی دارد، اصلی‌ترین چالش پیش روست که در این مقاله به ارائه‌ی روشی برای آن پرداخته شده است. روش سیستماتیک طراحی و آموزش مولدهای مرکزی الگو که از بازخوردهای حسی جهت تطابق خود با شرایط جدید استفاده می‌کند، در یک مدل دولایه‌ای مورد بحث و بررسی قرار داده شده است. مدل ارائه شده یک بستر مناسب جهت ارائه‌ی آموزش‌های نمایشی و مبتنی بر نمایش برای ربات‌های انسان‌نما محسوب می‌شود که نیاز به کنترل صریح و برنامه‌نویسی مستقیم را برطرف می‌سازد و امکان آموزش غیرمستقیم رفتارهای پیچیده روی انواع ربات‌ها را مهیا می‌سازد.

Design of a Trainable Controller Inspired from Neural System to Generate Complex Behaviors in Humanoid Robots

Reyhaneh Parandeh[1], Hamed Shahbazi[2*], Kamal Jamshidi[1], Behnam Khodabandeh Jahromi[2]

1- Department of Computer Engineering, University of Isfahan, Isfahan, Iran
2- Department of Mechanical Engineering University of Isfahan, Isfahan, Iran
* P.O.B. 37935619 Isfahan, Iran, shahbazi@eng.ui.ac.ir

ARTICLE INFORMATION

Original Research Paper
Received 13 October 2015
Accepted 09 January 2015
Available Online 27 January 2016

Keywords:
Humanoid Robot
Neural Network
Control
Imitation Learning

ABSTRACT

In this paper a new method for motion control in humanoid robots has been introduced. The problem of movement learning, especially dance and repetitive actions of human beings to humanoid robots is a major challenge in the field of robotics. Imitation learning, which is a subset of supervised learning, is one of the main forms used to teach complex tasks to the humanoid robot, and is, accordingly based on the concept that an artificial system can imitate a great deal of information through learning from a human trainer. The main technique uses Central Pattern Generators structures which are able to produce required motion trajectories based on imitation learning. Systematic design of these neural networks is the main problem that is solved in this paper. The proposed model is a basic paradigm for imitation learning in the humanoid robots which do not require direct design of controller and programming. The proposed model has many benefits including smooth walking patterns and modulation during imitation. Simulation results of this learning system in the robot simulator (WEBOTS) have been linked with MATLAB software and its implementation on a NAO robot demonstrates that the robot has learned the desired motion with high accuracy. This model can be extended and used in the Nao soccer player both for the standard platform and the 3D soccer simulation leagues of Robocup SPL competitions to train different types of motions.

1-مقدمه

حرکت کردن یکی از توانمندی‌های مهم در موجودات زنده است که نقش اساسی در رشد و بقای آن‌ها به‌جای می‌گذارد. توانمندی انجام انواع حرکات پیچیده در مهره‌داران و به‌خصوص انسان‌ها بخشی از هوشمندی آن‌ها محسوب می‌گردد که به‌طور ضمنی و در سلسله‌ی اعصاب و اندام‌های آن‌ها

برای ارجاع به این مقاله از عبارت ذیل استفاده نمایید:

Please cite this article using:

R. Parandeh, H. Shahbazi, K. Jamshidi, B. Khodabandeh Jahromi, Design of a Trainable Controller Inspired from Neural System to Generate Complex Behaviors in Humanoid Robots, *Modares Mechanical Engineering*, Vol. 16, No. 2, pp. 59-68, 2016 (in Persian)

[Downloaded from mme.modares.ac.ir on 2024-04-19]

[DOR: 20.1001.1.10275940.1395.16.2.27.2]

تعبیه‌شده است. هر موجود زنده که قادر به حرکت می‌باشد، فرم حرکت خاص خود را براساس شرایط محیطی و ساختار بدنی خود برگزیده است. علم مهندسی با الهام از انواع جانوران زنده، ربات‌ها و ماشین‌های مختلفی را ساخته است که نیازمند انجام حرکات مشابه با نمونه‌های زیستی متناظر با خود هستند. ازاین‌رو مسئله‌ی تولید حرکت در این‌گونه ربات‌ها یکی از چالش‌های اساسی در رباتیک محسوب می‌شود. تولید حرکت به معنای ایجاد مجموعه‌ای از خط سیرهای کنترلی است که در طول زمان به مفاصل یک ربات ارسال می‌گردند تا ربات یک حرکت یا مانور پیچیده را انجام دهد. خط سیرهای تولیدشده بایستی براساس تصمیم ربات جهت انجام یک کار و به شکل پویا و هوشمند تولید شوند و شرایط محیطی ربات و اطلاعات حسگرها را در نظر بگیرند. تولید پویای خط سیرهای کنترلی براساس یک برنامه‌ریزی[1] هوشمند، یک مسئله‌ی جستجوی فضای حالت در هوش مصنوعی به شمار می‌آید که راه‌حل‌های مختلفی برای آن وجود دارد. آن‌چه در حل این مسئله اهمیت دارد حل کارآمد و سریع آن است که به دلیل وجود توالی جستجو و اجرا در جستجوی برخط، بسیار مهم است.

برنامه‌ریزی و زمان‌بندی این خط سیرهای کنترلی برای ارضای یک هدف سطح بالا در یک ربات با درجات آزادی بالا، امری بسیار مشکل می‌باشد که نیازمند استفاده از تکنیک‌های یادگیری ماشین است. از آن‌جایی‌که برنامه‌نویسی مستقیم و صریح یک تولیدکننده‌ی حرکت بسیار پیچیده و زمان‌گیر است با کمک روش‌های یادگیری، روش‌هایی ارائه شده است که با استفاده از یک مربی و آموزش غیرمستقیم، یک رفتار پیچیده به ربات آموزش داده شود. در این روش‌های یادگیری رفتار مربی از طریق یک ابزار نمایشی به ربات ارسال می‌گردد و ربات سعی می‌کند که با تقلید از رفتار مربی این عمل را یاد بگیرد و دنباله‌های حرکتی موردنیاز جهت انجام آن را بازتولید نماید. ابزار نمایشی استفاده شده در این مقاله دوربین‌های دریافت تصویر مستقر در آزمایشگاه دانشکده توانبخشی دانشگاه علوم پزشکی اصفهان است. به این‌گونه روش‌های یادگیری، یادگیری نمایشی[2] یا یادگیری تقلیدی[3] اطلاق می‌شود. یادگیری نمایشی، زیرمجموعه‌ای از یادگیری با نظارت است که یک شکل اصلی از یاد دادن کارهای پیچیده به ربات به‌حساب می‌آید. همان‌گونه که انسان‌ها از طریق تقلید قادر به یادگیری هستند، یک سیستم مصنوعی نیز می‌تواند مقدار زیادی از اطلاعات مربوط به کار را به‌سادگی توسط تقلید از فرد دیگر، یا انسان یا ربات، به‌دست آورد. در یادگیری تقلیدی مفاهیم و ویژگی‌های اصلی رفتار آموزگار توسط ماشین کشف می‌شوند و ماشین قادر خواهد بود کلیات این رفتار را بازتولید کند. علاوه بر آن ربات قادر است رفتار یاد گرفته شده را با حالات و اشکال دیگری نیز به نمایش بگذارد که اصطلاحا آن‌را مدوله کردن[4] رفتار می‌نامیم. یعنی ربات قادر است کاری را که توسط انسان انجام شده تکرار کرده و پارامترهای مختلف این کار از جمله موقعیت، سرعت، دامنه و... را تغییر دهد تا رفتاری جدید تولید گردد. شکل 1 یادگیری تقلیدی حرکات دست را در یک ربات انسان‌نما نمایش می‌دهد.

آموزش رفتارهای پیچیده به ربات‌های انسان‌نما از طریق یادگیری نمایشی یک راه‌حل بسیار جالب و ساده است که به کمک آن می‌توان بسیاری از مشکلات موجود در تولید حرکت ربات انسان‌نما را حل نمود. در این مقاله به بررسی نحوه‌ی آموزش حرکت راه‌رفتن به ربات انسان‌نما از طریق یادگیری

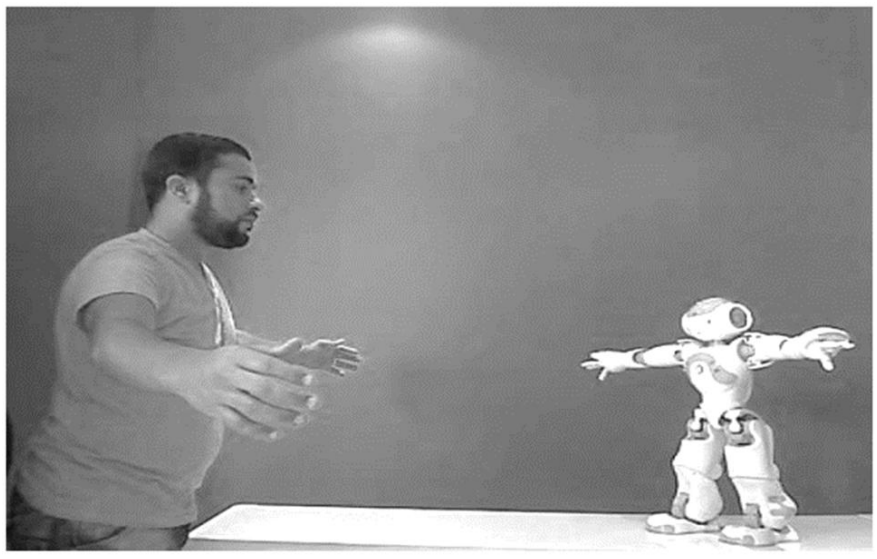


Fig. 1 Imitation learning of hand movement in humanoid robot [1]

شکل 1 یادگیری تقلیدی حرکات دست در ربات انسان‌نما [1]

نمایشی و به کمک یک لایه‌ی آموزش پذیر چندلایه می‌پردازیم. این مدل آموزش‌پذیر چندلایه‌ای الهام گرفته شده از ساختار سیستم عصبی انسان و مکانیسم‌های آموزشی آن می‌باشد که قابلیت یادگیری بسیاری از رفتارهای پیچیده را در ربات انسان‌نما دارا می‌باشد.

یک ربات انسان‌نما جهت راه‌رفتن روی یک محیط نامعین و انجام مانورهای مختلف بایستی همزمان نسبت‌های مختلفی را در مفاصل پاهای خود تعیین و گام‌های خود را طوری تنظیم کند تا هم روی مسیر در نظر گرفته شده با دقت حرکت کند وهم تعادل حرکتی خود را حفظ نماید. در دیدگاه کنترل بهینه ربات انسان‌نما ضمن راه رفتن و حفظ پایداری خود بایستی تعدادی پارامتر پاداش را نیز بیشینه‌سازی کند. این پارامترها می‌توانند زمان طی مسیر (که با بیشینه‌سازی سرعت به‌دست می‌آید) و خطای طی مسیر (که با بیشینه‌سازی دقت به‌دست می‌آید) و تعادل ربات باشند.

یکی از مشکلات بزرگ بر سر راه استفاده از ساختار سیستم عصبی عدم وجود یک روش سیستماتیک قوی برای تولید مدل و تطبیق ربات جهت استفاده از مدل‌ها است. در بخش دوم مقاله مولد مرکزی الگو و کاربردهای آن در رباتیک معرفی خواهد شد. این مدل‌ها روش‌های بسیار خوبی جهت مدل‌سازی حرکت در انواع ربات‌ها به شمار می‌آیند ولی بیشتر آن‌ها به‌صورت خاص و توسط طراحی با آزمون و خطا ساخته شده‌اند. در این میان روش‌هایی وجود دارند که یک مدل آموزش‌پذیر از مولد مرکزی الگو ارائه می‌کنند. مدل‌های آموزش‌پذیر، که عمدتا براساس مدل‌سازی نوسانگرهای غیرخطی جفت شده طراحی می‌شوند، قادر هستند خروجی موردنیاز یک ربات را براساس یک مرحله آموزش با سیگنال‌های آموزشی تولید کنند. این مدل‌ها ابزارهای مناسبی جهت به‌کارگیری یادگیری نمایشی روی انواع ربات‌ها به شمار می‌آیند.

در میان همه روش‌های گذشته برای آموزش و برنامه‌ریزی مولد مرکزی الگو مشکلاتی در حوزه‌ی طراحی وجود دارند که قابلیت به‌کارگیری آن‌ها را محدود می‌سازد. این مشکلات به شرح زیر می‌باشند:

در برنامه‌ریزی این موارد تعداد نوسان‌سازهای مورداستفاده همیشه ثابت است. از آن‌جا که الگوی آموزشی ورودی نامشخص است تعداد هماهنگ آن نیز معین نیست و هر نوسان‌ساز وفقی قادر است تا حداکثر با یکی از این هماهنگ‌ها تطابق یابد، مولد الگو همیشه قادر نیست تا به‌درستی الگوی ورودی را بازتولید کند. در صورتی‌که تعداد هماهنگ‌های الگو ورودی بیشتر از

1- Planning
2- Demonstration learning
3- Imitation learning
4- Modulation

[Downloaded from mme.modares.ac.ir on 2024-04-19]
[DOR: 20.1001.1.10275940.1395.16.2.27.2]

تعداد نوسان‌سازهای وفقی باشد بخشی از الگو آموزش داده نمی‌شود و از دست می‌رود. این امر موجب بالا رفتن خطای مرحله بازتولید الگو خواهد شد و کارایی آموزش را کاهش می‌دهد. از طرف دیگر بالا بودن تعداد نوسان‌سازهای وفقی در سیستم معادلات حالت منجر به افزایش زمان محاسبات هم در مرحله آموزش و هم در مرحله بازتولید بلادرنگ می‌گردد.

یکی دیگر از مشکلات این روش‌ها آن است که مقادیر اولیه‌ی نوسان‌سازهای وفقی نامشخص است و هیچ راهی برای تخمین آن‌ها وجود ندارد. این امر باعث عدم کشف برخی از هماهنگ‌ها می‌شود. به‌عبارت‌دیگر این روش‌ها قادر به کشف همه هماهنگ‌های الگو و بازتولید آن‌ها از هر نقطه شروعی نیستند و همگرایی آموزش آن‌ها به‌شدت وابسته به مقادیر اولیه معادلات حالت است که روش خاصی برای کشف آن پیشنهاد نشده است. معمولا این مقادیر به‌صورت تصادفی در یک بازه خاص به نوسان‌سازها انتساب داده می‌شوند. برای بسیاری از سیگنال‌ها الگوی ورودی در مرحله‌ی آموزش کشف نمی‌شود و خطای این مرحله به صفر میل نخواهد کرد. در بسیاری از موارد نوسان‌سازها به هماهنگ‌های کاملا مشابهی میل می‌کنند و این مسئله موجب کمبود نوسان‌ساز حتی در صورت معادل بودن تعداد هماهنگ‌ها با نوسان‌سازها می‌گردد. از دیگر ایرادات ذاتی روش‌های ارائه شده واگرایی سیستم معادلات حالت نوسان‌سازهای وفقی پس از قطع ورودی و ورود به مرحله تست است. این مسئله مربوط به اساس آموزش هبیان است که در همه این روش‌ها وجود دارد. این نوع آموزش مبتنی بر سنجش شباهت در هرلحظه و کاهش یا افزایش متغیرها براساس شباهت آنی است. این شیوه‌ی آموزشی پویا باوجود بسیاری از نقاط قوت، دارای ضعف پایداری است. این بدان معناست که آموزش به این شیوه همیشه به نقطه پایداری نمی‌رسد و ممکن است با صفر شدن ورودی، خطا نوسانی شده و مرتبا افزایش یابد. از دیگر مشکلات موجود در روش‌های برنامه‌ریزی مولد الگو براساس یادگیری هبیان وابسته بودن آن‌ها به نوع خاص نوسان‌سازهای هپ است. این روش‌ها صرفا بر پایه استفاده از این نوع نوسان‌ساز طراحی‌شده‌اند و قابلیت تعویض نوع نوسان‌ساز و استفاده از سایر انواع نوسان‌ساز برای طراحی لایه مولد الگو را ندارند. این محدودیت بزرگ مانع به‌کارگیری مزایای موجود در نوسان‌سازهای پیشرفته و مدل‌سازی شده کامل مفهوم مولد مرکزی الگو می‌باشد.

در این مقاله یک الگوریتم جدید برای برنامه‌ریزی و آموزش مولدهای الگو معرفی‌شده است که ایده‌های یادگیری هبیان و ایده یادگیری گرادیان خط‌مشی ماشین را با هم ترکیب کرده است تا راه‌کار بهتری ارائه نماید. این روش بسیاری از مشکلات موجود در طراحی خودکار مولد مرکزی الگو را حل نموده است و مزایای روش‌های گذشته را نیز داراست. از ویژگی‌های مهم آن اینست که نیاز به تعیین اولیه تعداد نرون‌های نوسان‌ساز ندارد و برای بسیاری از الگوهای پیچیده ورودی که در روش‌های قبلی قابلیت آموزش ندارند، آموزش‌پذیر است. این روش قادر است وظایف نوسان‌سازی نرون‌ها را به‌خوبی از هم تفکیک کرده و خطای مرحله آموزش و تست را تا حد دلخواه کاهش دهد.

در این مقاله بر روی تولید حرکات ریتمیک ورزشی برای ربات انسان نمای نائو به کمک روش‌های آموزشی مولد مرکزی الگو تمرکز شده است. در بخش دوم یادگیری تقلیدی توضیح داده می‌شود، سپس مفهوم مولد مرکزی الگو در بخش سوم تشریح می‌گردد. در بخش چهارم نحوه پیاده‌سازی عملی با استفاده از نرم‌افزارهای متلب و ویباتز بیان می‌شود و در بخش پنجم به ارزیابی نتایج و مقایسه روش‌های بیان‌شده پرداخته می‌شود.

2- یادگیری تقلیدی

مسئله یادگیری یک نگاشت بین حالت و عمل است که در قلب بسیاری از برنامه‌های کاربردی رباتیک نهفته است. این نگاشت که خط‌مشی ماشین[1] نیز نامیده می‌شود، ربات را قادر می‌سازد تا یک عمل براساس حالت کنونی محیطی که در آن است انتخاب کند. توسعه خط‌مشی ماشین به‌صورت دستی بسیار چالش‌برانگیز است و در نتیجه روش‌های یادگیری ماشین برای توسعه خط‌مشی ماشین اعمال می‌شوند. در یادگیری تقلیدی یک خط‌مشی ماشین از نمونه‌هایی که در طول نمایش توسط آموزگار انسانی به‌صورت مجموعه‌ای از جفت‌های حالت-عمل ضبط‌شده است، یاد گرفته می‌شود. یادگیری تقلیدی یک موضوع مرکزی رباتیک است که در تحقیقاتی چون تعامل انسان و ربات، یادگیری ماشین، بینایی ماشین و کنترل موتور به‌کار گرفته می‌شود. یادگیری تقلیدی یک مکانیزم قدرتمند برای کاهش پیچیدگی یادگیری در مقاله‌های تحقیقاتی به شمار می‌رود و یک ابزار قدرتمند برای بهبود و تسریع یادگیری هم در حیوانات و هم در مصنوعات است [2].

مسئله یادگیری تقلیدی شامل دو مرحله اصلی است. مرحله اول جمع‌آوری نمایش‌هایی است که آموزگار در کنار ربات انجام می‌دهد. مرحله دوم، استخراج پارامترها از این نمایش‌ها است، در واقع پس از ضبط خط سیر، مرحله بعدی نگاشت خط سیر به حرکت در ربات است.

اولین مرحله از یادگیری تقلیدی جمع‌آوری نمایش‌هایی است که آموزگار در کنار ربات انجام می‌دهد. در این مرحله باید دو تصمیم‌گیری انجام گیرد که این دو تصمیم‌گیری تحت تأثیر پیچیدگی ربات حاضر و کار موردنظر قرار می‌گیرد:

1- انتخاب نمایش‌دهنده: نمایش‌دهنده، یا ربات است، که در این صورت نمایش را یک ربات انجام می‌دهد و یک ربات دیگر از آن یاد گرفته، یا انسان است که در این صورت نمایش را یک انسان انجام می‌دهد و یک ربات از آن یاد می‌گیرد.

2- انتخاب روش نمایش: پس از آن‌که نمایش‌دهنده، کاری (مثلا نرمش کردن) را برای ربات انجام می‌دهد نوبت به این می‌رسد که خط سیر عمل انجام‌شده (نرمش کردن) ضبط شود تا از روی آن اندازه زاویه‌های هرکدام از درجه‌های آزادی ربات مشخص شود. انواع روش‌هایی که برای ضبط خط‌سیرها وجود دارد عبارت‌اند از: روش‌های بینایی کامپیوتر[2]، روش‌های فیزیکی (جنبشی) و استفاده از دستگاه ضبط حرکت[3].

در مرجع [3] از روش‌های بینایی کامپیوتر که هدف آن ضبط حرکات انسان بدون وجود مارکرهای ویژه روی بدن انسان، بلکه توسط وجود چشم‌های ربات یا وجود دوربینی که روی سر ربات است، استفاده کرده‌اند. این روش دارای خطای نسبتا زیادی است. در مرجع [5,4] از روش‌های فیزیکی که در آن یک معلم مفاصل ربات را مستقیما حرکت می‌دهد تا یک حرکت خاص را پیاده‌سازی کند، استفاده نموده‌اند. این روش احتیاج دارد که ربات بتواند تأثیرات نیروهای خارجی را جبران کند. مزیت این روش، این است که حرکات مستقیما روی ربات یادگیرنده ضبط می‌شود و احتیاج به انتقال بین سیستم‌ها از طریق دینامیک و سینماتیک مختلف ندارد اما با کمک این روش نمی‌توان هر حرکت پیچیده و موزونی را ضبط نمود.

برای ضبط انواع حرکات پیچیده، تکرارشونده و موزون می‌توان از دستگاهی به نام ضبط حرکت که براساس دستگاه‌های دنبال‌کننده نوری-

1- Policy
2- Computer vision technique
3- Motion capture

[Downloaded from mme.modares.ac.ir on 2024-04-19]
[DOR: 20.1001.1.10275940.1395.16.2.27.2]

مغناطیسی هستند، استفاده نمود. مزیت این روش نسبت به سایر روش‌ها در این است که سرعت و کیفیت این روش بسیار بیشتر می‌باشد و به‌راحتی از طریق این سیستم می‌توان خط سیر حرکات پیچیده چون نرمش کردن و سایر حرکات تکرارشونده را برای یاد دادن به ربات ضبط کرد بدون آن‌که ربات خود نیازی به انجام عملیاتی برای ضبط حرکت داشته باشد [6]. این سیستم، پروسه‌ی ضبط حرکات اشیا یا انسان را انجام می‌دهد و در کاربردهای مختلفی از جمله رباتیک، بینایی ماشین استفاده می‌شود. در جلسات ضبط حرکت، حرکات یک یا چند بازیگر چندین بار در هر ثانیه نمونه‌برداری می‌شود. این سیستم معمولا فقط حرکات بازیگر را، نه ظاهر بصری او را، ضبط می‌کند. این داده انیمیشن اغلب به یک مدل سه‌بعدی نگاشت می‌شود از همین رو، مدل، دقیقا عملیاتی مشابه با بازیگر را انجام می‌دهد. در این تحقیق، سعی بر آن است با استفاده از این سیستم که می‌تواند دنباله‌های حرکتی یک انسان یا یک شی را ضبط کند، دنباله‌های حرکتی دست یک آموزگار انسانی در غالب حرکات موزون را ضبط کرده و به ربات یاد داد. مزیت این روش نسبت به سایر روش‌ها آن است که روشی بسیار سریع است که می‌تواند حرکات پیچیده و تعاملات فیزیکی آموزگار را به‌طور بسیار دقیق ضبط کرده تا ربات بتواند رفتاری کاملا شبیه آن انجام دهد.

این سیستم شامل هفت دوربین با مدل پروریفلکس[1] است که دارای کیفیت بالا می‌باشد. این دوربین‌ها قابلیت ضبط و ثبت دقیق حرکات نشانگرهای[2] نوری نصب شده در موقعیت‌های مناسب بر روی بدن آموزگار انسانی را با اتصال به نرم‌افزار مرتبط با آن به نام کیو تی ام[3] دارند. این نرم‌افزار ردیابی، طراحی شده تا به‌صورت یکپارچه با هر مدل از دوربین، به‌صورت سریع و با دقت بالا برای جمع‌آوری داده، کار کند. سیستم اجازه می‌دهد تا کاربر داده‌ها را به‌صورت دو، سه و شش بعدی به‌صورت بلادرنگ با حداقل رکود ضبط نماید. وقتی دوربین‌ها راه‌اندازی می‌شوند، کیو تی ام به‌طور اتوماتیک تعداد دوربین‌ها را تشخیص می‌دهد، تنظیمات سخت‌افزار و نرم‌افزار آن را به کمک یک واسط کاربر گرافیکی بصری کنترل می‌کند و به‌راحتی اطلاعات ضبط شده را برای استفاده‌های بعدی ذخیره می‌کند.

دومین مرحله از یادگیری تقلیدی استخراج خط‌مشی ماشین از داده‌های به‌دست‌آمده از مرحله اول است، در واقع پس از ضبط خط سیر، مرحله بعدی نگاشت خط سیر به حرکت در ربات است. به‌بیان‌دیگر این خط سیرهای به‌دست آمده می‌بایستی کد شده و به فرمان موتور تبدیل شده تا حرکت موردنظر را در ربات ایجاد نماید که به آن خط‌مشی ماشین کنترلی[4] گویند و پس از آن باید به یادگیری پارامترهای خط‌مشی ماشین کنترلی به‌دست آمده پرداخته شود.

3- مولد مرکزی الگو

کنترل کردن ربات‌ها با چندین درجه آزادی برای انجام کارهای خودگردان یک مسئله‌ی مورد چالش است زیرا برنامه‌ریزی خط سیرهای پیچیده و وابسته به زمان، دشوار و یک پروسه هزینه‌بر است. به‌عبارت‌دیگر می‌توان گفت در این‌گونه ربات‌ها کنترلرهای سطح بالا اهداف موضعی ربات را تعیین می‌کنند و این اهداف در لایه‌های پایین باید به تعداد زیادی دنباله حرکتی در زمان نگاشته شوند. تولید و نظارت بر این دنباله‌های حرکتی توسط کنترلرهای سطح بالا، زمان‌بر است. برای حل این مشکل از یک میان لایه کنترلی به نام مولد مرکزی الگو که وظیفه‌ی دریافت تحریکات از لایه‌های بالاتر و تولید دنباله‌های حرکتی برای لایه‌های پایین‌تر دارد، استفاده می‌شود. مولدهای مرکزی الگو مدارهای عصبی هستند که عمدتا در ناحیه‌ی انتهایی مغز تا نخاع بیشتر مهره‌داران قرار گرفته‌اند و قادرند الگوهای هماهنگ و متناوبی را برای انجام حرکات گوناگون آن‌ها تولید کنند. ورودی این شبکه‌های عصبی، اطلاعات حسگری بسیار ساده و خروجی آن‌ها الگوهایی پیچیده، منظم و هماهنگ برای حرکت هستند و از این لحاظ موردتوجه فراوان در علوم رباتیک قرار گرفته‌اند [2-6]. مکانیزم تولید مرکزی الگو در نخاع انسان برای تولید حرکت در ربات انسان‌نما مدل‌سازی شده است. مولد مرکزی الگو، سازمانی از واحدهای پشت سر هم از عناصر با حداقل یک واحد برای هر درجه آزادی ربات (مفصل) می‌باشد که وظیفه‌ی تولید الگوهای تکرارشونده‌ی اولیه و پایه را دارد [7].

در اکثر مسائل رباتیک، مولد مرکزی الگوی مدل‌سازی شده با کمک نوسان‌سازهای غیرخطی، در یک لایه کنترلی بین کنترلرهای سطح بالا و سطح پایین ربات قرار داده می‌شود که به این لایه کنترلی میانه‌های دینامیک حرکت گفته می‌شود (شکل 2 (ب)). وظیفه‌ی این لایه‌ی کنترلی در ربات، همانند سطح نخاعی در بدن انسان است که دستورات سطح بالایی انسان یعنی مغز را به مفاصل بدن منتقل می‌کند. میانه‌های دینامیک حرکت براساس یک خط‌مشی ماشین داده شده در هر لحظه حالت فعلی سیستم را به یک عمل نگاشت می‌کند [8].

به‌طور کل می‌توان گفت هدف اصلی ایجاد یک سیستم آموزش‌پذیر مولد الگو است. این سیستم مدلی از یک سیستم عصبی مصنوعی است که قادر است با دریافت دنباله‌های ورودی آموزشی از یک مربی، پارامترها و اوزان مناسب را جهت تولید این دنباله‌ها، یادگیری کرده و به بازتولید دنباله‌ها بپردازد.

3-1- نوسان‌سازهای وفق‌پذیر

این مدل از نوسان‌سازها که در مرجع [9] برمبنای نوسان‌سازهای هپ طراحی شده‌اند، قادر هستند تا فرکانس ذاتی خود را با فرکانس یک سیگنال پیچیده تطبیق دهند. در مرجع [10] از روش یادگیری خاصی که معادلات دینامیکی نوسان‌ساز را توسعه می‌دهد تا قابلیت یادگیری فرکانس در سیستم به وجود آید، استفاده شده است. این روش یادگیری، برگرفته از مکانیزم یادگیری هبی در شبکه‌های عصبی مصنوعی می‌باشد. ویژگی اساسی این روش آن است که به هیچ اطلاعات آموزشی از سیگنال ورودی نیاز ندارد و آموزش را به‌صورت کاملا

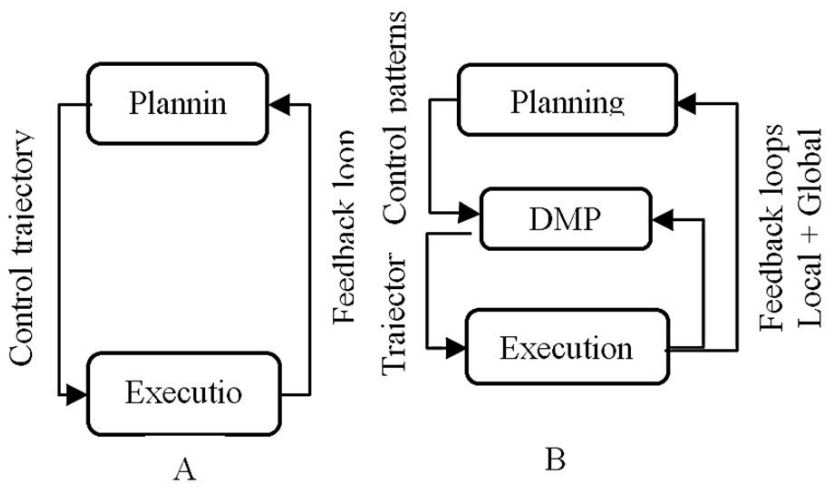


Fig. 2 DMPs for control. A) Old approach, B) DMP approach

شکل 2 میانه های دینامیک حرکت ها برای کنترل. الف) رهیافت قدیمی. ب) رهیافت میانه های دینامیک حرکت

1- ProReflex
2- Markers
3- QTM (Qualisys Track Manager)
4- Control policy

[Downloaded from mme.modares.ac.ir on 2024-04-19]
[DOR: 20.1001.1.10275940.1395.16.2.27.2]

برخط انجام می‌دهد. این بدان معناست که همه‌ی فرآیند یادگیری درون معادلات دینامیکی سیستم نهفته شده است. این روش یادگیری را علاوه بر نوسان‌سازهای هپ می‌توان به انواع دیگری از نوسان‌سازها نیز اعمال نمود.

معادلات کلی یک نوسان‌ساز [9] که توسط سیگنال ورودی متناوب F تحریک شده است به شکل (معادله 1) است:

$$\dot{x} = f(x, y, \omega) + \varepsilon F \quad (1)$$

$$\dot{y} = f(x, y, \omega) \quad (2)$$

که ω پارامتری است که بر روی فرکانس نوسان‌ساز اثرگذار است. در روش یادگیری هبی یک قانون یادگیری به شکل زیر به معادلات نوسان‌ساز افزوده می‌شود [9]:

$$\dot{\omega} = \pm\varepsilon \frac{y}{\sqrt{x^2 + y^2}} \quad (3)$$

علامت این رابطه به جهت چرخش سیکل حدی در صفحه (x, y) وابسته است. در این نوسان‌ساز ω به مقداری همگرا می‌شود که یک مؤلفه فرکانسی نوسان‌ساز با فرکانس یکی از مؤلفه‌های سیگنال ورودی F تطبیق یابد. در این معادلات ε یک ضریب کوچک است که به آن ثابت جفت شدن اطلاق می‌گردد.

در مرجع [11] یک روش دولایه‌ای برای یادگیری پارامترهای یک مولد الگوی برنامه‌پذیر بدون داشتن دانش قبلی از الگوی ورودی و فرکانس‌های آن ارائه شده است که قادر است دنباله‌های نوسانی یاد گرفته شده را در پاسخ به ورودی‌های خارجی مدوله‌سازی کند. این روش در این مقاله روش گمز نامیده می‌شود. این سیستم یادگیرنده‌ی مولد الگو جهت آموزش رفتارهای متناوب در یک ربات انسان نمای هآپ استفاده شده است. لایه‌ی اول این مولد الگو یک سیستم پویا کانونی است که از نوسان‌سازهای وفقی هپ مشابه روش قبلی ساخته شده است. از این لایه جهت کشف فرکانس بنیادین و فاز آن استفاده می‌شود. لایه‌ی دوم این سیستم یک سیستم پویای جدید برای تولید خروجی است که شکل موج خروجی را براساس ورودی‌های لایه‌ی اول یاد می‌گیرد. این دو سیستم پویا با یکدیگر ترکیب شده‌اند و به شکل همزمان دنباله‌های حرکتی ورودی را از مربی سیستم دریافت کرده و پارامترهای سیستم را کشف می‌کنند. شکل 3 نحوه عملکرد این سیستم دو لایه‌ای را نشان می‌دهد.

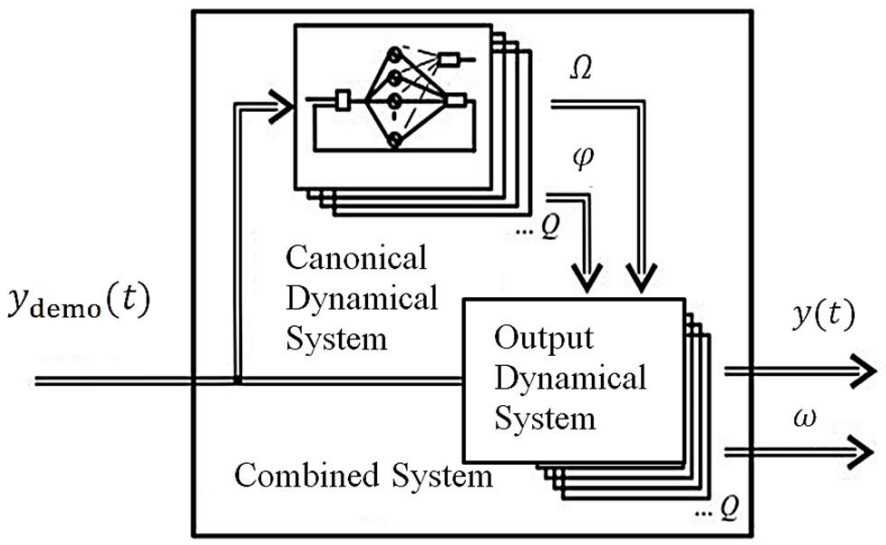


Fig. 3 Two-layer programmable pattern generator [10]

شکل 3 مولد الگوی برنامه‌پذیر دولایه‌ای [10]

3-2- یادگیری تقویتی در مولد مرکزی الگو

برای حل مسئله یادگیری پارامترهای شبکه عصبی در ابتدا رفتار سیستمی شبکه عصبی نوسانی را در قالب یک مسئله یادگیری تقویتی (عامل-نقاد) تعریف می‌شود. برای این کار لازم است تا مفاهیم تابع خط‌مشی ماشین، پارامترها، تابع پاداش- تنبیه، تابع ارزش و گرادیان خط‌مشی ماشین تعریف شود. در این‌جا کل یا بخشی از شبکه عصبی نوسانی را می‌توان در قالب یک تابع خط‌مشی ماشین نگاشت از حالت x و ادراک a به عمل خروجی u توصیف نمود. خط‌مشی ماشین توسط مجموعه پارامترهای θ پارامتربندی شده است که این پارامترها وزن‌ها و بایاس‌های شبکه عصبی نوسان‌کننده هستند. از آن‌جایی‌که خروجی شبکه عصبی همواره دارای مقداری خطا است رفتار این تابع خط‌مشی ماشین به شکل تصادفی تعریف می‌شود [10]:

$$\pi_\theta(x, a, u) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}(u - g_\theta(x, a))\right) \quad (4)$$

در این معادله خط‌مشی ماشین پارامتربندی شده ورودی‌های حالت x، ادراک a و عمل خروجی را به یک‌میزان احتمال نسبت می‌دهد که درون یک توزیع احتمال گوسی با مرکز $g_\theta(x, a)$ و واریانس σ قرار گرفته است. این شکل از تعریف تابع خط‌مشی ماشین احتمالی در بسیاری از کاربردها به کار گرفته شده‌اند که از جمله آن‌ها به روش‌های به‌کاررفته برای آموزش مولد مرکزی الگو در مراجع [12-14] می‌توان اشاره نمود. تابع $g_\theta(x, a)$ رفتار متوسط کل یا بخشی از شبکه عصبی موردنظر را با ورودی‌های x, a و پارامتر θ بیان می‌نماید که درون هر یک از الگوریتم‌های این بخش تعریف شده است. در تعریف این تابع حالت ورودی x اندیس زمانی تولید الگو و ادراک a مقدار ورودی حسگری سیستم است. در عمل به‌دلیل کاستن از پیچیدگی‌های طراحی این تابع، اثر ادراک a را در زمان آموزش نادیده می‌گیریم. این بدان معناست که ورودی‌های حسگری در آموزش لایه مولد مرکزی الگو دخیل نمی‌شود تا از پیچیدگی‌های مرحله آموزش کاسته شود.

تابع پاداش- تنبیه در چنین سیستمی خروجی واقعی شبکه عصبی را با خروجی مورد انتظار آن مقایسه می‌کند و برمبنای این مقایسه مقداری از پاداش یا تنبیه برای آن تولید می‌کند. یک شکل از تعریف آن با نادیده گرفتن ادراک ورودی در زمان آموزش به شکل زیر است [11]:

$$r(x, u) = (y_{\text{teach}}(x) - u)^2 \quad (5)$$

در این معادله پاداش داده شده به عمل u در حالت x برابر توان دوم اختلاف سیگنال آموزش موردنظر در حالت x $(y_{\text{teach}}(x))$ با مقدار عمل u تعریف می‌شود. به دلیل آن‌که سیگنال آموزش y_{teach} بدون در نظر گرفتن شرایط ورودی ادراک تولید شده است (ادراک ورودی در زمان تولید y_{teach} نادیده گرفته شده است) اثری از مقدار a در این معادله دیده نمی‌شود. تابع ارزش خط‌مشی ماشین به‌صورت جمع کاهش یافته مقادیر پاداش در طول یک دوره آزمایش تعریف می‌شود [12]:

$$V^\pi(x(t)) = E\langle \int_t^\infty e^{-\frac{s-t}{\tau}} r(x(s), u(s)) ds] \rangle \quad (6)$$

که در آن نماد E مقدار مورد انتظار و ضریب کاهش در حالت پیوسته به شکل ضرب عبارت $e^{-\frac{s-t}{\tau}}$ در معادله وارد شده است. در این معادله τ ثابت کاهش نامیده می‌شود. در عمل برای محاسبه این مقدار از جمع نمونه‌های به‌دست‌آمده از تابع پاداش $r(x(s), u(s))$ استفاده می‌شود.

روش ارائه شده یک نسخه تغییریافته از الگوریتم لونبرگ- مارکوآرت است که در مرجع [15] ارائه شده است. در این مرحله الگوریتم کل شبکه عصبی نوسان‌ساز را در قالب یک نگاشت‌کننده با تابع خط‌مشی ماشین تعریف شده در معادله (4) در نظر می‌گیریم که دارای مجموعه پارامتر زیر است:

$$\theta = [w_1, v_1^1, v_2^1, w_2, v_1^2, v_2^2, \dots, w_1, v_1^N, v_2^N] \quad (7)$$

[Downloaded from mme.modares.ac.ir on 2024-04-19]

[DOR: 20.1001.1.10275940.1395.16.2.27.2]

در این بردار پارامتر w_i وزن مابین نرون نوسان‌ساز i ام و نرون پست-سیناپتیک است و v_1^i, v_2^i دو بایاس مربوط به نرون i ام هستند که به عنوان مقادیر اولیه به نرون داده می‌شوند. تابع پارامتری $g_\theta(x,a)$ که در معادله (4) مورد استفاده قرار گرفت، به شکل زیر تعریف می گردد:

$$g_\theta(x,a) = \sum_{k=1}^{N} \theta_{3k}.O(1,\theta_{3k+1}.x + \theta_{3k+2}) \quad (8)$$

این تابع خروجی کلی نرون‌های نوسان‌ساز در حالت x را تولید می‌نماید. تابع $O(a,b)$ تابع کسینوس با دامنه a و زاویه b است که در آن مقدار سه پارامتر دامنه θ_{3k} فرکانس θ_{3k+1} و فاز θ_{3k+2} تعریف شده است. N تعداد نرون‌های موجود در شبکه عصبی است و تعداد پارامترهای بردار θ، برابر $3N$ می‌باشد. مقدار تابع پاداش $r(x,u)$ از معادله (5) بدست می‌آید که در آن u از معادله زیر بدست می‌آید:

$$u = \pi_\theta(x,a) = g_\theta(x,a) + \mathcal{N}(0,\sigma) \quad (9)$$

که در این‌جا $\mathcal{N}(0,\sigma)$ نویز گوسی به مرکز صفر و واریانس σ می‌باشد. این معادله به طور مستقیم از معادله (4) بدست می‌آید. تابع ارزش بردار θ در شکل ساده ($\gamma = 1$) بدون در نظر گرفتن نویز با معادله زیر تعریف می‌شود:

$$J(\theta) = \sum_{i=1}^{m} r(x,u) = \sum_{i=1}^{m} (y_{\text{teach}}(x) - g_\theta(x,a))^2 \quad (10)$$

پروسه یادگیری، یک روند تکراری است که بردار θ را با بهترین مقادیر اولیه بدست آمده در مرحله اول الگوریتم مقدار دهی اولیه می‌کند و هر بار بردار θ با بردار جدید $\theta + \Delta\theta$ جایگزین می‌شود تا به مقدار بهینه پارامترها دست یابیم. در این‌جا $\Delta\theta$ به کمک بردار گرادیان خط‌مشی ماشین برآورد می‌شود. برای برآورد $\Delta\theta$ در این روش از تخمین $\pi_{\theta+\Delta\theta}(x,a)$ با کمک خطی‌سازی استفاده می‌شود:

$$\pi_{\theta+\Delta\theta}(x,a) \approx \pi_\theta(x,a) + \nabla J.\Delta\theta \quad (11)$$

که در آن ∇J بردار گرادیان (مشتق) تابع ارزش نسبت به θ است. با مشتق گیری از J نسبت به θ در معادله (10) خواهیم داشت:

$$\nabla J_\theta = 2.\nabla(g_\theta(x,a)).\sum_{i=1}^{m}(y_{\text{teach}}(x) - g_\theta(x,a)) \quad (12)$$

که $\nabla(g_\theta(x,a))$ یا به طور نسبی $\frac{\partial}{\partial\theta_j}(g_\theta(x,a))$ از رابطه 13 محاسبه می‌شود:

$$\frac{\partial}{\partial\theta_j}(g_\theta(x,a)) = \begin{cases} O(1,\theta_{3k+1}.x + \theta_{3k+2}); j = 3k \\ x.\theta_{3k}.dO(1,\theta_{3k+1}.x + \theta_{3k+2}); j = 3k+1 \\ \theta_{3k}.dO(1,\theta_{3k+1}.x + \theta_{3k+2}); j = 3k+2 \end{cases} \quad (13)$$

که در آن از معادله (8) به طور نسبی مشتق‌گیری شده است. در این مشتق براساس جنبش مقدار های θ_j سه نوع مشتق بدست آمده است. مقدار $dO()$ مشتق تابع نوسان‌ساز $O()$ می‌باشد که رابطه آن با $O()$ مشابه رابطه کسینوس با سینوس است:

$$dO(1,c) = O(1,\frac{\pi}{2} - c) \quad (14)$$

با داشتن بردار گرادیان ∇J معادله گرادیان به شکل برداری در شکل زیر نوشته می‌شود:

$$(\overline{\nabla J}^{\mathrm{T}}\overline{\nabla J})\overline{\Delta\theta} = \overline{\nabla J}^{\mathrm{T}}[Y_{\text{teach}} - G_\theta] \quad (15)$$

که در آن Y_{teach} شکل برداری $y_{\text{teach}}(x)$ و G_θ شکل برداری $g_\theta(x,a)$ است که با مجتمع کردن مقادیر برای همه x ها بدست می‌آید. علاوه بر آن $\overline{\nabla J}$ به شکل ماتریس گرادیان‌ها (ماتریس ژاکوبین) در نظر گرفته شده است. در الگوریتم لونبرگ- مارکوآرت از یک نسخه میرا شده از این معادله استفاده می‌شود [15]:

$$(\overline{\nabla J}^{\mathrm{T}}\overline{\nabla J} + \lambda.\mathrm{diag}(\overline{\nabla J}^{\mathrm{T}}\overline{\nabla J}))\overline{\Delta\theta} = \overline{\nabla J}^{\mathrm{T}}[Y_{\text{teach}} - G_\theta] \quad (16)$$

در این روش از یک فاکتور میرایی λ در هر دوره از الگوریتم استفاده می‌شود. از مقادیر کوچک λ زمانی استفاده می‌شود که تغییرات مقدار تابع ارزش زیاد است که در این حالت الگوریتم به روش رگرسیون گوس- نیوتن نزدیک می‌شود و زمانی که مقدار تابع ارزش زیاد کاهش کافی پیدا کرد، λ افزایش می‌یابد تا الگوریتم را به روش گرادیان دیسنت نزدیک کند. در این‌جا ماتریس ژاکوبین به‌سادگی و به‌کمک روش مشتق گیری تحلیلی بدست می‌آید.

فلوچارت الگوریتم یادگیری اول برای هر دو مرحله در شکل 4 نمایش داده شده است. در سمت چپ صفحات موازی از الگوریتم‌های یکسان که روی هر مفصل از مولد مرکزی الگو اجرا می‌شود، دیده می‌شود. هر صفحه به یک مجموعه تک بعدی از نرون‌های جفت شده تعلق دارد. همه صفحات الگوریتم‌های مشابهی که در بالا شرح داده شد، را توصیف می‌کنند. فلوچارت از بالا سمت چپ شروع می‌شود. دو مرحله الگوریتم توسط بلوک پروسه تفکیک شده‌اند. در بالای پروسه اول یک پروسه اولیه برای محاسبه بایاس b_1 وجود دارد که الگوی ورودی آموزش را متقارن می‌سازد. پس از استخراج بایاس، پروسه مرحله اول آغاز می‌شود. در هر سیکل سیگنال خروجی از سیگنال ورودی کم شده و تفاضلی که سیگنال خطا را می‌سازد به معادلات سیستم کانونی وارد می‌شود.

این پروسه به دنبال یک مقادیر اولیه مناسب برای مرحله دوم می‌گردد و پس از طی شدن مدت زمان معین همگرا شده و مقادیر اولیه مناسب را تولید می‌کند. مقادیر اولیه بدست آمده در مرحله اول به پروسه دوم منتقل می‌شود و مرحله دوم الگوریتم از این مقادیر شروع به کار می‌کند. در هر سیکل از مرحله دوم، ماتریس گرادیان‌ها (ژاکوبین) به کمک معادله (12) به شکل تحلیلی محاسبه می‌شود و در گام بعد بردار تغییرات خط‌مشی ماشین $\overline{\Delta\theta}$ از معادله (16) محاسبه می‌گردد. با محاسبه بردار تغییرات خط‌مشی ماشین، بردار خط‌مشی ماشین θ در گام بعد بروزرسانی می‌شود. در گام چهارم خطای شبکه عصبی یا همان مقدار تابع ارزش محاسبه شده و در صورتی که خطا به اندازه کافی کوچک نباشد حلقه با بروزرسانی فاکتور میرایی λ ادامه می‌یابد. مقدار λ در ابتدا مقداری کوچک دارد و به تدریج افزایش می‌یابد. زمانی که همه پارامترها همگرا شوند و خطا به حد کافی کوچک شود، پارامترهای بایاس و وزن نرون‌ها در شبکه تنظیم می‌شود و مرحله نهایی الگوریتم شروع می‌شود.

در مرحله نهایی، مقادیر بایاس نرون‌ها به معادلات جفت‌سازی ارسال می‌شوند تا مقادیر جفت‌سازی بین نرون‌ها محاسبه گردد. این مقادیر بر در دو سطح؛ ما بین نرون‌های یک مجموعه و مابین نرون‌های اولیه دو مجموعه مجزا محاسبه می‌شوند. این مقادیر جفت‌سازی توسط معادلات زیر محاسبه می‌شوند:

$$d_{1i} = \frac{v_1^i}{v_1^1}v_2^1 - v_2^i \quad (17)$$

$$d_{ij} = \frac{v_1^{(i)}}{v_1^{(j)}}v_2^{(j)} - v_2^{(i)} \quad (18)$$

در این معادلات d_{1i} مقدار جفت‌سازی بین نرون اول و نرون i ام یک مجموعه است و d_{ij} مقدار جفت‌سازی بین نرون اول مجموعه i ام و نرون اول مجموعه j ام است که جفت‌سازی بین دو مجموعه را امکان‌پذیر می‌سازد. در نقشه جفت‌بندی استفاده شده در این تحقیق همه نرون‌های یک مجموعه تک بعدی با نرون اول آن مجموعه جفت می‌شوند و نرون‌های اول مجموعه های مجزا نیز به‌طور متوالی با هم جفت می‌گردند. برای جفت‌سازی و محاسبه مقدار P_{dc} در هر لحظه، نرون‌های اول باید فاز و فرکانس خود را به سایر نرون‌ها ارسال کنند. پس از محاسبه مقادیر جفت‌سازی بین نرون‌ها همه

[Downloaded from mme.modares.ac.ir on 2024-04-19]
[DOR: 20.1001.1.10275940.1395.16.2.27.2]

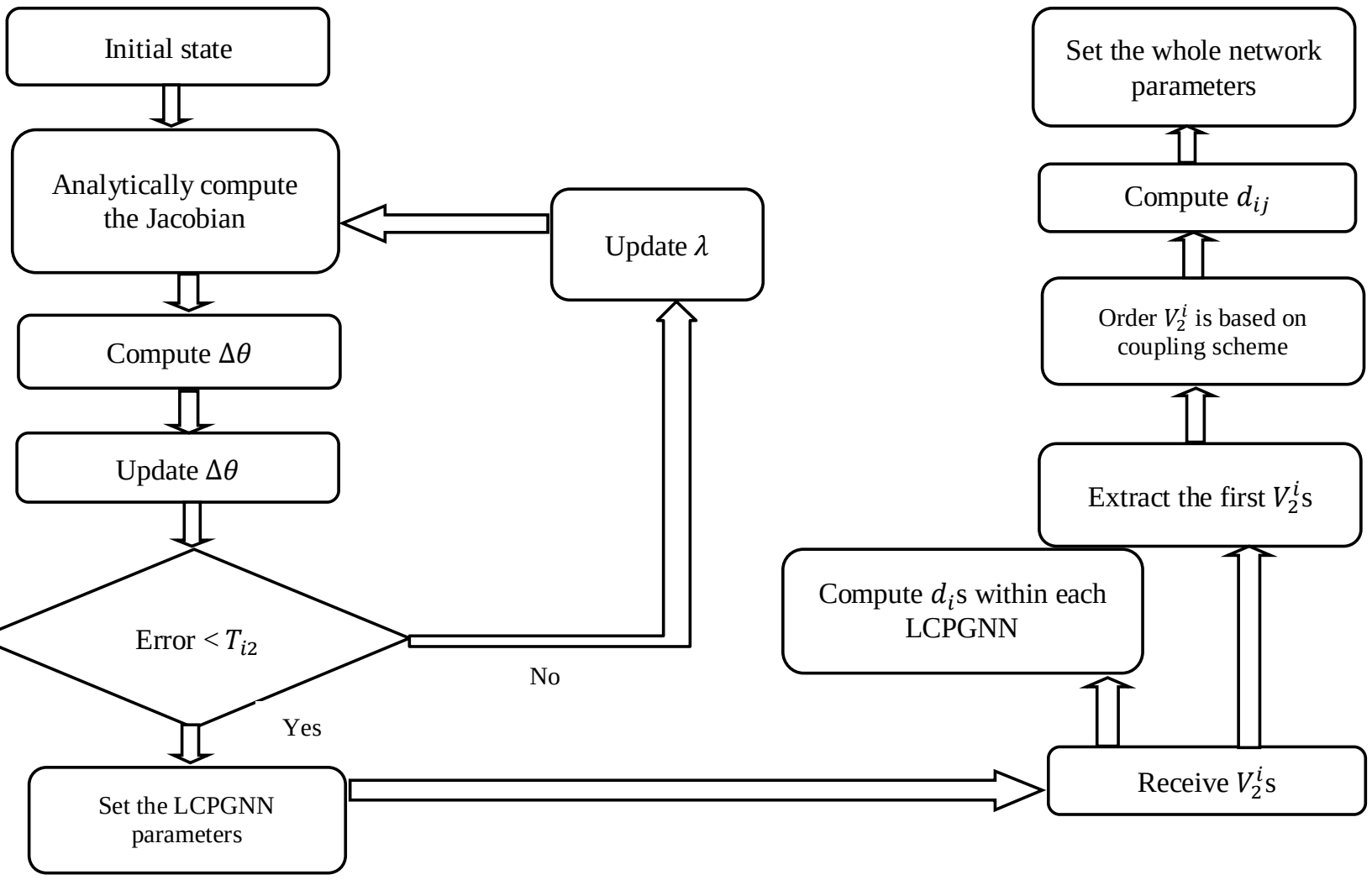


Fig. 4 Hebbian-gradient learning algorithm for planning the central pattern generator layer

شکل 4 الگوریتم یادگیری هبیان-گرادیان برای برنامه‌ریزی لایه مولد مرکزی الگو

پارامترهای مجهول شبکه عصبی نوسانی یاد گیرنده یافته شده‌اند. این پارامترها درون شبکه تنظیم شده و پس از آن شبکه عصبی در مرحله تست مورد استفاده قرار می‌گیرد.

4- پیاده سازی‌های عملی

در این بخش روش‌های پیشین یادگیری با کمک نوسان‌سازهای غیرخطی جفت شده (روش رقتی[1])، با روش پیشنهادی این مقاله مقایسه و مزایا و معایب آن‌ها بررسی می‌گردد. پس از آن به بررسی و تحلیل نتایج و ویژگی‌های آموزش دو نمونه حرکت، حرکات دست کاراته و نرمش پیش-بالا- طرفین، به ربات انسان‌نمای نائو در شبیه‌ساز ویباتز، به‌کمک روش پیشنهاد شده برای هر دو لایه‌ی سیستم یادگیری تقلیدی ارائه شده در این تحقیق، پرداخته خواهد شد.

برای شبیه‌سازی سیستم یادگیرنده‌ی تقلیدی لازم است در ابتدا الگوهای موردنیاز برای آموزش تقلیدی فراهم شود. در این تحقیق، تولید الگوها با استفاده از دستگاه ضبط حرکت انجام می‌گیرد. این دستگاه قادر است حرکات انجام شده توسط آموزگار انسانی، همچون حرکات دست کاراته، نرمش پیش-بالا- طرفین و دمبل زدن، را ضبط کرده و با اتصال به نرم‌افزار کیو تی ام، موقعیت‌های حرکتی هر مفصلی که به کمک نشانگرها روی بدن آموزگار انسانی نشان‌گذاری شده را به صورت خط سیر نسبت به زمان، به کاربر تحویل دهد. از آن‌جایی که نیاز به الگوهای حرکتی زوایای هر مفصل است، باید از طریق حل مسأله‌ی سینماتیک معکوس الگوی حرکتی زاویه‌ی هر مفصل به دست آید و پس از آن دنباله‌های حرکتی بدست آمده به نرم‌افزار ویباتز منتقل گردد تا به قالب قابل استفاده در آموزش (ماتریس‌های نرم‌افزار متلب) تبدیل شود. ویباتز با اتصال به نرم‌افزار متلب برنامه‌های آن را اجرا می‌کند و دنباله‌های آموزشی را به ربات انسان‌نمای نائو آموزش می‌دهد. این روش برای آموزش حرکات پیچیده‌ی تکرارشونده استفاده می‌شود. شکل 5 مراحل تولید الگوهای حرکتی مورد نظر را نمایش می‌دهد. این سیستم شامل هفت دوربین با قابلیت ضبط و ثبت دقیق مسیرهای حرکات نشانگرهای نصب شده در موقعیت‌های مناسب بر روی دو دست آموزگار انسانی است. دوربین‌ها موقعیت‌های نشانگرها را به صورت گسسته و در هر گام زمانی ثبت می‌کنند و به کمک تعریف و اختصاص یک دستگاه مختصات دکارتی متعامد سه بعدی اینرسی (با سه محور متعامد x,y,z) مختصات این موقعیت‌ها توسط نرم‌افزار *QTM* ذخیره می‌شود و قابل مشاهده و تحلیل است.

با توجه به داشتن مختصات نشانگرها در هر حرکت می‌توان به کمک حل مسئله سینماتیک معکوس به کمک الگوریتم ژاکوبین مختصات چرخشی این مختصات را که به اصطلاح مختصات فضای کار نامیده می‌شوند به مختصات

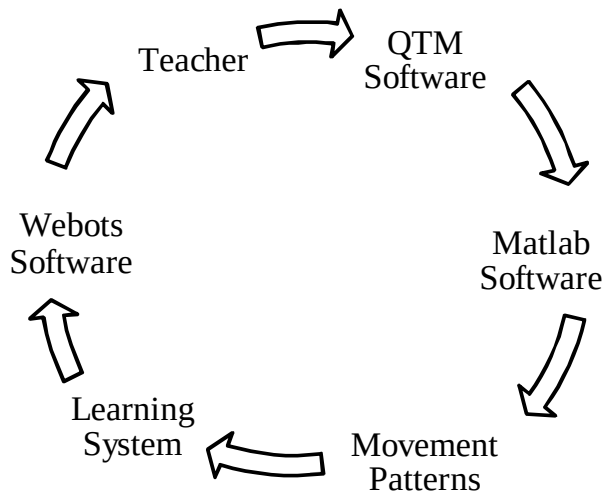


Fig. 5 Production processes of movement patterns for the first layer of system

شکل 5 مراحل تولید الگوهای حرکتی برای لایه اول سیستم

1- Righetti method

[Downloaded from mme.modares.ac.ir on 2024-04-19]
[DOR: 20.1001.1.10275940.1395.16.2.27.2]

نظیر آن‌ها در فضای مفاصل مربوط ساخت. به عبارت دیگر چگونگی تغییر و خط سیر زوایای مفاصل را به گونه‌ای تعیین نمود که نتیجه‌ی آن، انجام حرکات مورد نظر باشد. در این مطالعه، برای هر دست چهار مفصل و به عبارتی چهار درجه آزادی در نظر گرفته شده است.

با حل مسئله سینماتیک معکوس، زوایای هر حرکت به صورت یک خط سیر نسبت به زمان بدست می‌آید و برای آموزش به لایه بعدی داده می‌شود. طراحی داخلی سیستم یادگیری تقلیدی دو لایه براساس روش استفاده از نوسان‌سازهای وفق‌پذیر هپ [9] و نیز روش استفاده از نوسان‌سازهای دو لایه- (گمز) [10]، که در بخش سوم به طور کامل توضیح داده شد، صورت پذیرفته است. در ادامه نمایی از آموزش با هر کدام از این روش‌ها نشان داده خواهد شد.

نمایی از آموزش چهار مفصل از مفاصل به کار رفته در حرکت دست کاراته، با کمک روش رقتی اصلاح‌شده در شکل 6 نشان داده شده است. همان‌گونه که در تصاویر دید می‌شود از نقطه‌ای که خط تعیین شده در شکل‌ها نشان می‌دهد ورودی نوسان‌سازها ذره ذره کم می‌شود تا به صفر میل نماید. در واقع می‌توان گفت از این زمان به بعد مرحله‌ی آموزش به پایان رسیده و مرحله تست خط سیر یاد گرفته شده آغاز می‌شود. ورودی نوسان سازها در واقع سیگنال خطای تولید نوسان است که اختلاف سیگنال پریودیک آموزشی و خروجی غیرخطی پریودیک تولید شده است. این اختلاف به تدریج به سمت صفر میل می‌کند. خروجی‌های تولید شده سیستم در شکل پریودیک هستند.

به عنوان روش پیشنهادی، آن‌چنان که در بخش سوم شرح داده شد، هر خط سیر ورودی وارد شبکه‌ای از نوسان‌سازهای مبتنی بر هپ شده و

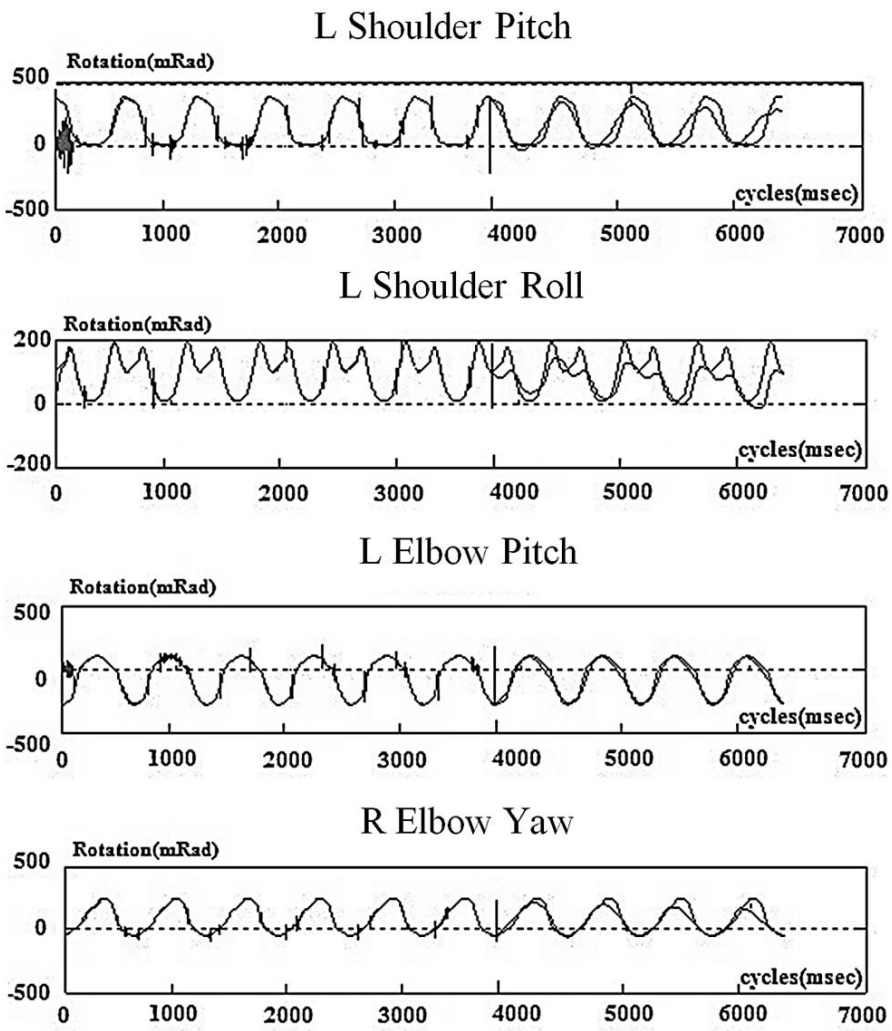


Fig. 6 A detailed view of four joints education of involved joints in hands movement of karate with Righetti approach

شکل 6 نمایی از آموزش چهار مفصل از مفاصل درگیر در حرکت دست کاراته با روش رقتی

مشخصه‌های هر کدام از این نوسان‌سازها، شامل دامنه، فاز و فرکانس به کمک الگوریتم ارائه شده استخراج می‌شود. نتایج این روش یادگیری در ادامه آمده است. همان‌گونه که مشاهده می‌شود این روش از دقت بالاتری برخوردار است و نرخ همگرایی مطلوب‌تری را فراهم می‌کند.

پس از اجرای فرآیند یادگیری در نرم‌افزار متلب و بدست آوردن پارامترهای نوسان‌سازها، می‌بایستی برای شبیه‌سازی سیستم موردنظر روی ربات انسان نمای نائو سراغ محیط شبیه‌ساز ویباتز رفت. به طور نمونه شکل 7 (الف) نمایی از ضبط حرکت کاراته‌ی مربی توسط دستگاه ضبط حرکت موجود در مرکز اسکلتی- عضلانی دانشکده توانبخشی دانشگاه اصفهان را نشان می‌دهد. شکل 7 (ب) نمایی از حرکت کاراته یاد گرفته شده توسط ربات نائو در شبیه‌ساز را نشان می‌دهد.

به منظور بررسی سیستم ارائه شده، سیستم موردنظر برمبنای روش پیشنهادی به ربات انسان‌نمای واقعی منتقل شده است. کدهای برنامه‌نویسی به کمک زبان پایتون به محیط برنامه‌سازی کروگراف منتقل و با اتصال کروگراف به ربات نائو واقعی بر روی آن پیاده‌سازی شده است. علی‌رغم تفاوت‌های بسیار مابین دینامیک ربات واقعی و ربات شبیه‌سازی شده، نائو واقعی قادر است حرکت کاراته که به عنوان نمونه از بین حرکات انجام شده در شبیه‌ساز، انتخاب شده را انجام دهد. شکل 8 پیاده‌سازی حرکت سوکی کاراته روی ربات نائوی واقعی را نشان می‌دهد. خط‌مشی ماشین اجرا شده تا حدود زیادی مشابه خط‌مشی ماشین اجرا شده در مدل شبیه‌سازی شده است .

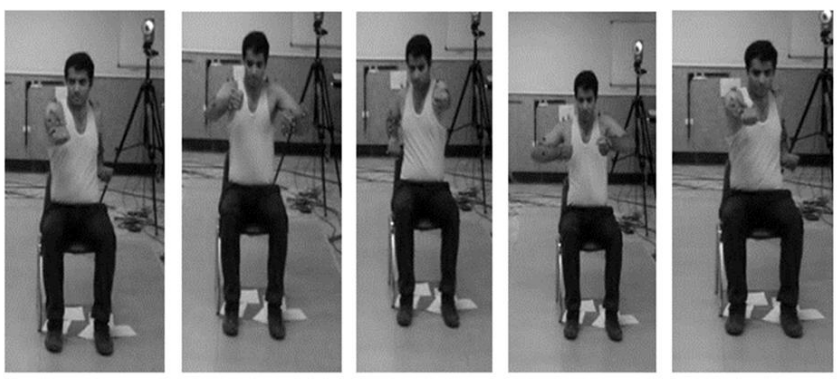


Fig. 7 Webots simulation system. (A) Recording of teacher karate movements by motion capture device, (B) karate movement samples that learned by Nao robot

شکل 7 شبیه‌سازی سیستم در ویباتز. (الف) ضبط نمونه حرکتی کاراته‌ی مربی توسط دستگاه ضبط حرکت. (ب) نمونه حرکت کاراته یاد گرفته شده توسط ربات نائو

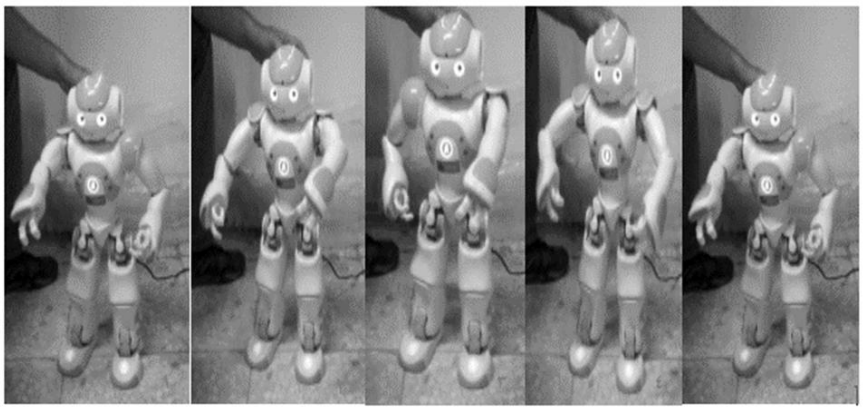


Fig. 8 The implementation of suki motion on real Nao robot

شکل 8 پیاده‌سازی حرکت سوکی کاراته روی ربات نائوی واقعی

[Downloaded from mme.modares.ac.ir on 2024-04-19]

[DOR: 20.1001.1.10275940.1395.16.2.27.2]

5- ارزیابی نتایج

در این بخش به مقایسه‌ی روش‌های گفته شده پرداخته می‌شود. جدول 1 مقایسه‌ی روش‌های رقتی، گمز و روش پیشنهادی را براساس معیارهای درصد میانگین نرخ همگرایی، متوسط درصد خطای تست و متوسط زمان آموزش برای تعدادی نمونه یکسان، نشان می‌دهد.

منظور از معیار نرخ همگرایی، درصد تعداد الگوهای حرکتی است که به کمک روش مورد نظر به صورت موفقیت آمیز و با دقت مطلوب آموزش می‌بینند و الگویی هماهنگ با الگوی حرکتی مربی دنبال می‌کنند. این معیار در جدول به صورت درصدی بیان شده است. منظور از معیار متوسط درصد خطای تست، متوسط توان دوم اختلاف الگوی حرکتی مربی و الگوی حرکتی یاد گرفته شده بر طول الگوهای حرکتی در میان چند الگوی حرکتی با روش مورد نظر می‌باشد. این معیار جهت سنجش دقت و کیفیت روش‌ها بسیار مهم می‌باشد. بالا بودن میزان این معیار نشان دهنده همگرا شدن بیشتر پارامترهای سیستم به پارامترهای درست موردنظر می‌باشد. منظور از معیار متوسط زمان آموزش، متوسط زمان لازم برای کاهش تغییرات پارامترهای سیستم و همگرا شدن پارامترها در روش موردنظر است که بر روی یک پردازنده واحد اندازه‌گیری می‌شود.

همان‌گونه که در جدول دیده می‌شود روش پیشنهادی دارای بیشترین نرخ همگرایی و کمترین خطای متوسط تست نسبت به دو روش دیگر می‌باشد یعنی این روش قادر به یادگیری الگوهای حرکتی بیشتر نسبت به دو روش دیگر است زیرا در این روش هیچ‌گونه وابستگی به مقادیر اولیه نوسان‌سازها ندارد و بدون وابستگی به مقادیر اولیه نوسان‌سازها تمام ریتم‌های الگوی حرکتی ورودی را یاد می‌گیرد. تفاوت بسیار زیاد خطای محاسبه شده در روش‌ها به دلیل عدم همگرایی روش‌های گذشته در تعدادی از حالت‌های آموزشی با وابستگی به مقادیر اولیه نوسان‌سازها است. درواقع روش‌های گذشته در تعدادی از حالت‌ها اصلا همگرا نشده‌اند و این امر سبب بالا رفتن متوسط خطای تست شده‌اند. در این روش ها با انتخاب بهتر چندین نقطه فضای حالت و انجام جستجو می‌توان به نتایج بهتری رسید. روش ارائه شده به‌طور همزمان از چندین نقطه فضای حالت جستجو را شروع می‌کند در حالی که در سایر روش‌ها یادگیری همواره از یک نقطه اولیه در فضای حالت آغاز می‌گردد. پس یادگیری روش‌های هبیانی وابستگی بسیاری به نقطه شروع و شیوه تنظیم مقادیر اولیه دارد و برای خیلی از نقاط اولیه پارامترها به سمت مقادیر نهایی همگرا نمی‌شوند درحالی‌که روش مبتنی بر الگوریتم‌های تکاملی مطرح شده در این مقاله با صرف زمان بیشتر قادر به یافتن نقاط بهینه مناسب برای همگرایی به سمت پارامترهای نهایی مطلوب می‌باشد. این روش مشکلات روش‌های قبل از جمله، قابل تغییر نبودن تعداد نوسان‌سازها، همگرا شدن یک نوسان‌ساز به هماهنگ‌های مشابه، کشف نشدن پارامترهای دقیق الگوی ورودی در مرحله آموزش و بنابراین میل نکردن خطای این مرحله به صفر را برطرف نموده است. در این روش یک نوع مصالحه بین زمان و کیفیت به وجود آمده است که در روش نهایی با افزایش معیار زمان آموزش، میزان معیار خطای تست کاهش و نرخ همگرایی افزایش یافته است.

جدول 1 مقایسه روش‌های گوناگون براساس معیارهای ارزیابی

Table 1 Compare different methods according to evaluation criteria

روش	درصد متوسط نرخ همگرایی	متوسط خطای تست	متوسط زمان آموزش (ثانیه)
رقتی (Righetti)	%71	% 26.62	477.466
گمز (Gams)	%76	% 2.31	489.1082
پیشنهادی	%94	% 0.03	231.7

6- نتیجه‌گیری

در این مقاله به آموزش تقلیدی یا نمایشی حرکات موزونی چون کاراته، نرمش پیش- بالا- طرفین، و دمبل زدن، به ربات انسان‌نمای نائو از طریق یک آموزگار انسانی پرداخته شد. برای آموزش تقلیدی، سیستمی دو لایه ارائه گردید که وظیفه لایه‌ی اول آن جمع‌آوری نمایش‌ها آموزگار انسانی و به دست آوردن خط سیرهای زوایای مفاصل درگیر در دو دست آن بود. به دلیل پیچیده بودن حرکات و نیاز داشتن به خط سیرهای دقیق متناظر با هر حرکت، در این پژوهش جهت جمع‌آوری و ضبط حرکات آموزگار انسانی از دستگاه ضبط حرکت موجود در مرکز تحقیقاتی اسکلتی- عضلانی دانشگاه علوم‌پزشکی اصفهان استفاده شد. در ادامه روشی مبتنی بر یادگیری گرادیان خط‌مشی ماشین روی مولدهای مرکزی الگو ارائه گردید که توسط آن آموزش حرکات ربات انجام گردیده است. الگوریتمی که در این مقاله ارائه شد، الگوریتم آموزش مولد مرکزی الگو است که براساس تکامل و بهینه‌سازی روش‌های موجود برای برنامه‌ریزی خودکار مولد مرکزی الگو پایه‌ریزی گشته‌اند. روش ارائه شده یکی از دغدغه‌های بزرگ در مسیر استفاده از مولدهای مرکزی الگو که شیوه برنامه‌ریزی و شکل‌دهی آن‌ها می‌باشد را تا حدود زیادی برطرف کرده است. در روش ارائه شده که به‌نوعی الگوریتم هبیان- گرادیان نامیده می‌شود، بسیاری از کاستی‌های موجود در روش‌های گذشته را جبران می‌شود. به کمک این روش یک سیستم برنامه‌پذیر مولد مرکزی الگو با نام شبکه عصبی نوسانی یاد گیرنده معرفی کردیم که قادر است با یادگیری یکی‌یکی نرون‌ها و اتصال آن‌ها به هم به تعداد لازم و کافی نرون را برای بازتولید الگوهای نوسانی فراهم سازد و خطای مرحله تست را تا حد دلخواه کاهش دهد. در پایان مزایای این روش نسبت به سایر روش‌ها تشریح گردید که از جمله آن‌ها نرخ همگرایی و آموزش‌پذیری بسیار بالاتر، عدم احتمال واگرایی و نرخ خطای پایین آن بود.

7- مراجع

[1] B. Sofiane, S. Anzalone, E. Tilmont, D. Cohen, M. Chetouani, Learning of social signatures through imitation game between a robot and a human partner, *Autonomous Mental Development*, IEEE Transactions on, Vol. 6, No. 3, pp. 213-225, 2014.

[2] A. Billard, S. Calinon, R. Dillmann, S. Schaal, *Robot Programming by Demonstration, Handbook of Robotics*, chapter 59, 2007.

[3] T. B. Moeslund, A. Hilton, V. Kr¨uger, A survey of advances in vision-based human motion capture and analysis, *Computer Vision and Image Understanding*, Vol. 104, No. 2, pp. 90–126, 2006.

[4] M. Hersch, F. Guenter, S. Calinon, A. Billard, Dynamical system modulation for robot learning via kinesthetic demonstrations, *Robotics*, IEEE Transactions, Vol. 24, No. 6, pp. 1463–1467, Dec. 2008.

[5] J. Peters, S. Schaal, Reinforcement learning of motor skills with policy gradients, *Neural Networks*, Vol. 21, No.4, pp. 682–697, 2008.

[6] S. Calinon, D. Halluin, L. Sauser, G. Caldwell, G. Billard, learning and reproduction of gestures by imitation, *Robotics & Automation Magazine*, IEEE , Vol. 17, No. 2, pp.44-54, 2010.

[7] S. Schaal, J. Peters, J. Nakanishi, A. Ijspeer, Learning movement primitives, *In Robotics Research*, *The Eleventh International Symposium*, pp. 561-572, Springer Berlin Heidelberg, 2005.

[8] T. L. Williams, The neural-mechanical link in lamprey locomotion, *Locomotor neural mechanisms in arthropods and vertebrates*, pp. 183-195, 1991.

[Downloaded from mme.modares.ac.ir on 2024-04-19]

[DOR: 20.1001.1.10275940.1395.16.2.27.2]

[9] L. Righetti, J. Buchli, A. Ijspeert, Dynamic hebbian learning in adaptive frequency oscillators, *Physica D: Nonlinear Phenomena*, Vol. 216, No.2, pp. 269-281, 2006.

[10] L. Righetti, A. J. Ijspeert, Programmable central pattern generators: an application to biped locomotion control, *In Robotics and Automation, Proceedings 2006 IEEE International Conference on 15 May*, IEEE, pp. 1585-1590, 2006.

[11] A. Gams, A. J. Ijspeert, S. Schaal, J. Lenarčič, On-line learning and modulation of periodic movements with nonlinear dynamical systems, *Autonomous robots*, Vol. 27, No.2, pp. 3-23, 2009.

[12] G. Endo, J. Morimoto, T. Matsubara, J. Nakanishi, G. Cheng, Learning CPG-based biped locomotion with a policy gradient method: Application to a humanoid robot, *The International Journal of Robotics Research*, Vol. 27, No.2, pp. 213-228, 2008.

[13] T. Matsubara, J. Morimoto, J. Nakanishi, M. A. Sato, K. Doya, Learning CPG-based biped locomotion with a policy gradient method, *Robotics and Autonomous Systems*, Vol. 54, No.11, pp. 911-920, 2006.

[14] Y. Nakamura, T. Mori, M. A. Sato, S. Ishii, Reinforcement learning for a biped robot based on a CPG-actor-critic method, *Neural Networks*, Vol. 20, No.6, pp. 723-735, 2007.

[15] J. J. Moré, The Levenberg-Marquardt algorithm: implementation and theory in Numerical analysis**,** *Numerical analysis*, pp. 105-116, Springer Berlin Heidelberg,1978.

[Downloaded from mme.modares.ac.ir on 2024-04-19]

[DOR: 20.1001.1.10275940.1395.16.2.27.2]