



# کنترل کننده آموزش پذیر برگرفته از ساختار سیستم عصبی جهت تولید رفتارهای پیچیده در ربات انسان نما

ریحانه پرنده<sup>1</sup>، حامد شهبازی<sup>2\*</sup>، کمال جمشیدی<sup>3</sup>، بهنام خدابنده جهرمی<sup>4</sup>

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، صندوق پستی 37935619، eng.ui.ac.ir، shahbazi@

## اطلاعات مقاله

مقاله پژوهشی کامل

دریافت: 21 مهر 1394

پذیرش: 19 دی 1394

ارائه در سایت: 07 بهمن 1394

کلید واژگان:

ربات انسان نما

شبکه عصبی

کنترل

یادگیری تقلیدی

## چکیده

در این مقاله مسئله کنترل حرکت ربات‌های انسان‌نما مورد بررسی قرار می‌گیرد. مسئله 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 ناو، یادگیری تقلیدی است. تکنیک اصلی به کار 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. نحوه طراحی سیستماتیک این ساختار عصبی مصنوعی که در علم کنترل کاربردهای فراوانی دارد، اصلی‌ترین چالش پیش روست که در این مقاله به ارائه 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 ارائه 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<sup>1</sup>, Hamed Shahbazi<sup>2\*</sup>, Kamal Jamshidi<sup>1</sup>, Behnam Khodabandeh Jahromi<sup>2</sup>

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

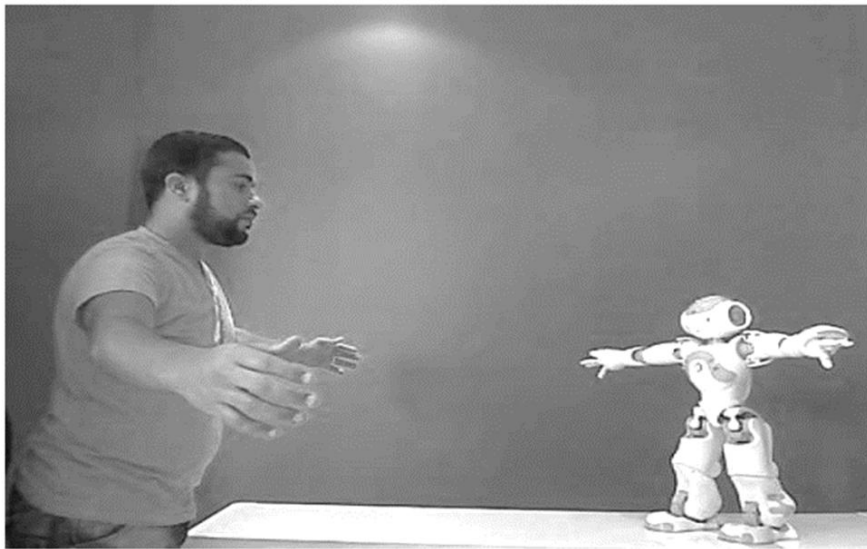
اساسی در رشد و بقای آن‌ها به‌جای می‌گذارد. توانمندی انجام انواع حرکات پیچیده در مهره‌داران و به‌خصوص انسان‌ها بخشی از هوشمندی آن‌ها محسوب می‌گردد که به‌طور ضمنی و در سلسله‌ای اعصاب و اندام‌های آن‌ها

حرکت کردن یکی از توانمندی‌های مهم در موجودات زنده است که نقش

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)

برای ارجاع به این مقاله از عبارت ذیل استفاده نمایید:



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

شکل 1 یادگیری تقلیدی حرکات دست در ربات انسان نما [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 انسان نما دارا می‌باشد.

یک ربات انسان نما جهت راه رفتن 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 و هم تعادل حرکتی خود را حفظ نماید. در دیدگاه کنترل بهینه ربات انسان نما 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 به دست می‌آید) و خطای طی مسیر (که با بیشینه‌سازی دقت به دست 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 به صورت 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 عمدتاً 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 به کارگیری 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 به کارگیری آن‌ها را محدود می‌سازد. این مشکلات به شرح زیر می‌باشند:

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

تعبیه شده 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. از این رو مسئله‌ی تولید حرکت در این گونه 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 تولید شده 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<sup>1</sup> 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 آنجایی که 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 این گونه روش‌های یادگیری، یادگیری نمایشی<sup>2</sup> یا یادگیری تقلیدی<sup>3</sup> اطلاق می‌شود. یادگیری نمایشی، زیرمجموعه‌ای از یادگیری با نظارت است که یک شکل اصلی از یاد دادن کارهای پیچیده به ربات به حساب می‌آید. همان گونه که انسان‌ها از طریق تقلید قادر به یادگیری هستند، یک سیستم مصنوعی نیز می‌تواند مقدار زیادی از اطلاعات مربوط به کار را به سادگی توسط تقلید از فرد دیگر، یا انسان یا ربات، به دست 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 اصطلاحاً آن را مدوله کردن<sup>4</sup> 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 انسان نما نمایش می‌دهد.

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

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

## 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<sup>1</sup> 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 (مثلاً 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<sup>2</sup>، روش‌های فیزیکی (جنبشی) و استفاده از دستگاه ضبط حرکت<sup>3</sup>.

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 نسبتاً زیادی است. در مرجع [4,5] از روش‌های فیزیکی که در آن یک معلم مفاصل ربات را مستقیماً 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 مستقیماً 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-

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. معمولاً 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 کاملاً 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 هر لحظه 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 مرتباً 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 صرفاً 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 معرفی شده 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 ویب‌تاز بیان می‌شود و در بخش پنجم به ارزیابی نتایج و مقایسه روش‌های بیان‌شده پرداخته می‌شود.

1- Policy

2- Computer vision technique

3- Motion capture

مولد مرکزی الگو که وظیفه‌ی دریافت تحرکات از لایه‌های بالاتر و تولید دنباله‌های حرکتی برای لایه‌های پایین‌تر دارد، استفاده می‌شود. مولدهای مرکزی الگو مدارهای عصبی هستند که عمدتاً 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 مورد توجه فراوان در علوم رباتیک قرار گرفته‌اند [2-6]. مکانیزم تولید مرکزی الگو در نخاع انسان برای تولید حرکت در ربات انسان نما مدل سازی 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 مدل سازی شده با کمک نوسان سازهای 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 لایه 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 آموزش پذیر 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- نوسان سازهای وفق پذیر

این مدل از نوسان سازها که در مرجع [9] بر مبنای نوسان سازهای 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 نوسان ساز 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 به صورت کاملاً

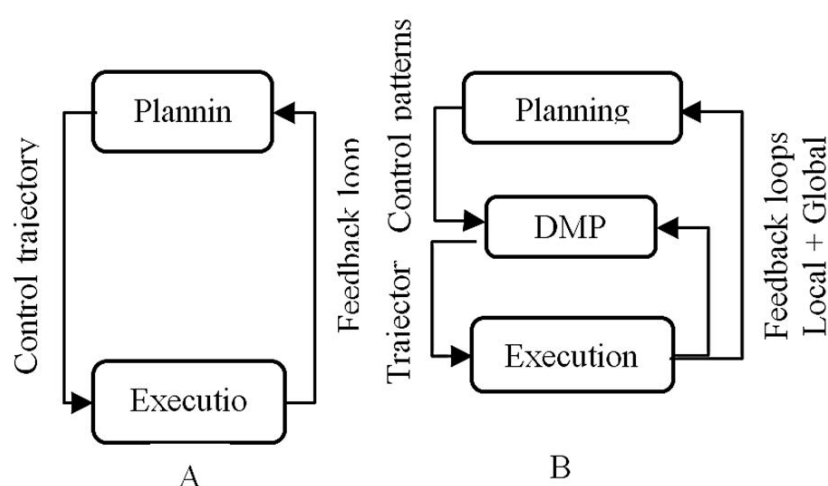
مغناطیسی هستند، استفاده نمود. مزیت این روش نسبت به سایر روش‌ها در این است که سرعت و کیفیت این روش بسیار بیشتر می‌باشد و به راحتی از طریق این سیستم می‌توان خط سیر حرکات پیچیده چون نرمش کردن و سایر حرکات تکرارشونده را برای یاد دادن به ربات ضبط کرد بدون آن‌که ربات خود نیازی به انجام عملیاتی برای ضبط حرکت 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 نمونه برداری می‌شود. این سیستم معمولاً فقط حرکات بازیگر را، نه ظاهر بصری او را، ضبط می‌کند. این داده انیمیشن اغلب به یک مدل سه بعدی نگاشت می‌شود از همین رو، مدل، دقیقاً 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 کاملاً شبیه آن انجام دهد.

این سیستم شامل هفت دوربین با مدل پروریفلکس<sup>1</sup> است که دارای کیفیت بالا می‌باشد. این دوربین‌ها قابلیت ضبط و ثبت دقیق حرکات نشانگرهای<sup>2</sup> نوری نصب شده در موقعیت‌های مناسب بر روی بدن آموزگار انسانی را با اتصال به نرم افزار مرتبط با آن به نام کیو تی ام<sup>3</sup> دارند. این نرم افزار ردیابی، طراحی شده تا به صورت یکپارچه با هر مدل از دوربین، به صورت سریع و با دقت بالا برای جمع آوری داده، کار کند. سیستم اجازه می‌دهد تا کاربر داده‌ها را به صورت دو، سه و شش بعدی به صورت بلادرنگ با حداقل رکود ضبط نماید. وقتی دوربین‌ها راه اندازی می‌شوند، کیو تی ام به طور اتوماتیک تعداد دوربین‌ها را تشخیص می‌دهد، تنظیمات سخت افزار و نرم افزار آن را به کمک یک واسط کاربر گرافیکی بصری کنترل می‌کند و به راحتی اطلاعات ضبط شده را برای استفاده‌های بعدی ذخیره می‌کند.

دومین مرحله از یادگیری تقلیدی استخراج خط‌مشی ماشین از داده‌های به دست آمده از مرحله اول است، در واقع پس از ضبط خط سیر، مرحله بعدی نگاشت خط سیر به حرکت در ربات است. به بیان دیگر این خط سیرهای به دست آمده می‌بایستی کد شده و به فرمان موتور تبدیل شده تا حرکت مورد نظر را در ربات ایجاد نماید که به آن خط‌مشی ماشین کنترلی<sup>4</sup> گویند و پس از آن باید به یادگیری پارامترهای خط‌مشی ماشین کنترلی به دست آمده پرداخته شود.

### 3- مولد مرکزی الگو

کنترل کردن ربات‌ها با چندین درجه آزادی برای انجام کارهای خودگردان یک مسئله‌ی مورد چالش است زیرا برنامه ریزی خط سیرهای پیچیده و وابسته به زمان، دشوار و یک پروسه هزینه بر است. به عبارت دیگر می‌توان گفت در این گونه 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، زمان بر است. برای حل این مشکل از یک میان لایه کنترلی به نام



**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

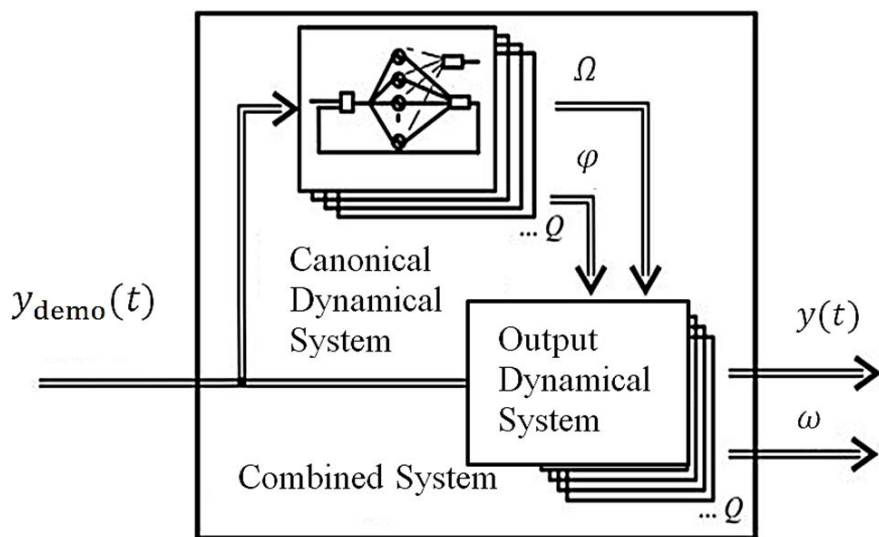


Fig. 3 Two-layer programmable pattern generator [10]

شکل 3 مولد الگوی برنامه پذیر دولایه ای [10]

احتمال گوسی با مرکز  $g_\theta(x, a)$  و واریانس  $\sigma$  قرار گرفته است. این شکل از تعریف تابع خط مشی ماشین احتمالی در بسیاری از کاربردها به کار گرفته شده اند که از جمله آن ها به روش های به کار رفته برای آموزش مولد مرکزی الگو در مراجع [12-14] می توان اشاره نمود. تابع  $g_\theta(x, a)$  رفتار متوسط کل یا بخشی از شبکه عصبی مورد نظر را با ورودی های  $x, a$  و پارامتر  $\theta$  بیان می نماید که درون هر یک از الگوریتم های 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 به دلیل کاستن از پیچیدگی های طراحی این تابع، اثر ادراک  $a$  را در زمان آموزش نادیده می گیریم. این بدان معناست که ورودی های حسگری در آموزش لایه مولد مرکزی الگو دخیل نمی شود تا از پیچیدگی های مرحله آموزش کاسته شود.

تابع پاداش - تنبیه در چنین سیستمی خروجی واقعی شبکه عصبی را با خروجی مورد انتظار آن مقایسه می کند و بر مبنای این مقایسه مقداری از پاداش یا تنبیه برای آن تولید می کند. 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 مورد نظر در حالت  $x$  ( $y_{\text{teach}}(x)$ ) با مقدار عمل  $u$  تعریف می شود. به دلیل آن که سیگنال آموزش  $y_{\text{teach}}$  بدون در نظر گرفتن شرایط ورودی ادراک تولید شده است (ادراک ورودی در زمان تولید  $y_{\text{teach}}$  نادیده گرفته شده است) اثری از مقدار  $a$  در این معادله دیده نمی شود. تابع ارزش خط مشی ماشین به صورت جمع کاهش یافته مقادیر پاداش در طول یک دوره آزمایش تعریف می شود [12]:

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

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  $\tau$  ثابت کاهش نامیده می شود. در عمل برای محاسبه این مقدار از جمع نمونه های به دست آمده از تابع پاداش  $r(x(s), u(s))$  استفاده می شود.

روش ارائه شده یک نسخه تغییر یافته از الگوریتم لونیگ-مارکوآرت است که در مرجع [15] ارائه شده است. در این مرحله الگوریتم کل شبکه عصبی نوسان ساز را در قالب یک نگاشت کننده با تابع خط مشی ماشین تعریف شده در معادله (4) در نظر می گیریم که دارای مجموعه پارامتر زیر است:

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

برخط انجام می دهد. این بدان معناست که تمامی فرآیند یادگیری درون معادلات دینامیکی سیستم نهفته شده است. این روش یادگیری را علاوه بر نوسان سازهای هپ می توان به انواع دیگری از نوسان سازها نیز اعمال نمود.

معادلات کلی یک نوسان ساز [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  $\omega$  پارامتری است که بر روی فرکانس نوسان ساز اثرگذار است. در روش یادگیری هپی یک قانون یادگیری به شکل زیر به معادلات نوسان ساز افزوده می شود [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 نوسان ساز  $\omega$  به مقداری همگرا می شود که یک مؤلفه فرکانسی نوسان ساز با فرکانس یکی از مؤلفه های سیگنال ورودی  $F$  تطبیق یابد. در این معادلات  $\varepsilon$  یک ضریب کوچک است که به آن ثابت جفت شدن اطلاق می گردد.

در مرجع [11] یک روش دولایه ای برای یادگیری پارامترهای یک مولد الگوی برنامه پذیر بدون داشتن دانش قبلی از الگوی ورودی و فرکانس های آن ارائه شده است که قادر است دنباله های نوسانی یاد گرفته شده را در پاسخ به ورودی های خارجی مدوله سازی کند. این روش در این مقاله روش گمز نامیده می شود. این سیستم یادگیرنده ی مولد الگو جهت آموزش رفتارهای متناوب در یک ربات انسان نمای هآپ استفاده شده است. لایه ی اول این مولد الگو یک سیستم پویا کانونی است که از نوسان سازهای وقتی هپ مشابه روش قبلی ساخته شده است. از این لایه جهت کشف فرکانس بنیادین و فاز آن استفاده می شود. لایه ی دوم این سیستم یک سیستم پویای جدید برای تولید خروجی است که شکل موج خروجی را براساس ورودی های لایه ی اول یاد می گیرد. این دو سیستم پویا با یکدیگر ترکیب شده اند و به شکل همزمان دنباله های حرکتی ورودی را از مربی سیستم دریافت کرده و پارامترهای سیستم را کشف می کنند. شکل 3 نحوه عملکرد این سیستم دو لایه ای را نشان می دهد.

### 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 می شود. برای این کار لازم است تا مفاهیم تابع خط مشی ماشین، پارامترها، تابع پاداش - تنبیه، تابع ارزش و گرادیان خط مشی ماشین تعریف شود. در این جا کل یا بخشی از شبکه عصبی نوسانی را می توان در قالب یک تابع خط مشی ماشین نگاشت از حالت  $x$  و ادراک  $a$  به عمل خروجی  $u$  توصیف نمود. خط مشی ماشین توسط مجموعه پارامترهای  $\theta$  پارامتر بندی شده است که این پارامترها وزن ها و بایاس های شبکه عصبی نوسان کننده هستند. از آنجایی که خروجی شبکه عصبی همواره دارای مقداری خطا است رفتار این تابع خط مشی ماشین به شکل تصادفی تعریف می شود [10]:

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

در این معادله خط مشی ماشین پارامتر بندی شده ورودی های حالت  $x$ ، ادراک  $a$  و عمل خروجی را به یک میزان احتمال نسبت می دهد که درون یک توزیع

در این روش از یک فاکتور میرایی  $\lambda$  در هر دوره از الگوریتم استفاده می شود. از مقادیر کوچک  $\lambda$  زمانی استفاده می شود که تغییرات مقدار تابع ارزش زیاد است که در این حالت الگوریتم به روش رگرسیون گوس- نیوتن نزدیک می شود و زمانی که مقدار تابع ارزش زیاد کاهش کافی پیدا کرد،  $\lambda$  افزایش می یابد تا الگوریتم را به روش گرادیان دیسنت نزدیک کند. در این جا ماتریس ژاکوبین به سادگی و به کمک روش مشتق گیری تحلیلی بدست می آید.

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

این پروسه به دنبال یک مقادیر اولیه مناسب برای مرحله دوم می گردد و پس از طی شدن مدت زمان معین همگرا شده و مقادیر اولیه مناسب را تولید می کند. مقادیر اولیه بدست آمده در مرحله اول به پروسه دوم منتقل می شود و مرحله دوم الگوریتم از این مقادیر شروع به کار می کند. در هر سیکل از مرحله دوم، ماتریس گرادیان ها (ژاکوبین) به کمک معادله (12) به شکل تحلیلی محاسبه می شود و در گام بعد بردار تغییرات خط مشی ماشین  $\Delta \theta$  از معادله (16) محاسبه می گردد. با محاسبه بردار تغییرات خط مشی ماشین، بردار خط مشی ماشین  $\theta$  در گام بعد بروزرسانی می شود. 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  $\lambda$  ادامه می یابد. مقدار  $\lambda$  در ابتدا مقداری کوچک دارد و به تدریج افزایش می یابد. زمانی که همه پارامترها همگرا شوند و خطا به حد کافی کوچک شود، پارامترهای بایاس و وزن نرون ها در شبکه تنظیم می شود و مرحله نهایی الگوریتم شروع می شود.

در مرحله نهایی، مقادیر بایاس نرون ها به معادلات جفت سازی ارسال می شوند تا مقادیر جفت سازی بین نرون ها محاسبه گردد. این مقادیر بر دو سطح؛ ما بین نرون های یک مجموعه و مابین نرون های اولیه دو مجموعه مجزا محاسبه می شوند. این مقادیر جفت سازی توسط معادلات زیر محاسبه می شوند:

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

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

در این معادلات  $d_{1i}$  مقدار جفت سازی بین نرون اول و نرون  $i$  ام یک مجموعه است و  $d_{ij}$  مقدار جفت سازی بین نرون اول مجموعه  $i$  ام و نرون اول مجموعه  $j$  ام است که جفت سازی بین دو مجموعه را امکان پذیر می سازد. در نقشه جفت بندی استفاده شده در این تحقیق همه نرون های یک مجموعه تک بعدی با نرون اول آن مجموعه جفت می شوند و نرون های اول مجموعه های مجزا نیز به طور متوالی با هم جفت می گردند. برای جفت سازی و محاسبه مقدار  $P_{dc}$  در هر لحظه، نرون های اول باید فاز و فرکانس خود را به سایر نرون ها ارسال کنند. پس از محاسبه مقادیر جفت سازی بین نرون ها همه

در این بردار پارامتر  $w_i$  وزن مابین نرون نوسان ساز  $i$  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 می شوند. تابع پارامتری  $g_\theta(x, a)$  که در معادله (4) مورد استفاده قرار گرفت، به شکل زیر تعریف می گردد:

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

این تابع خروجی کلی نرون های نوسان ساز در حالت  $x$  را تولید می نماید. تابع  $O(a, b)$  تابع کسینوس با دامنه  $a$  و زاویه  $b$  است که در آن مقدار سه پارامتر دامنه  $\theta_{3k}$  فرکانس  $\theta_{3k+1}$  و فاز  $\theta_{3k+2}$  تعریف شده است.  $N$  تعداد نرون های موجود در شبکه عصبی است و تعداد پارامترهای بردار  $\theta$ ، برابر  $3N$  می باشد. مقدار تابع پاداش  $r(x, u)$  از معادله (5) بدست می آید که در آن  $u$  از معادله زیر بدست می آید:

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

که در این جا  $\mathcal{N}(0, \sigma)$  نویز گوسی به مرکز صفر و واریانس  $\sigma$  می باشد. این معادله به طور مستقیم از معادله (4) بدست می آید. تابع ارزش بردار  $\theta$  در شکل ساده ( $\gamma = 1$ ) بدون در نظر گرفتن نویز با معادله زیر تعریف می شود:

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

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

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

که در آن  $\nabla J$  بردار گرادیان (مشتق) تابع ارزش نسبت به  $\theta$  است. با مشتق گیری از  $J$  نسبت به  $\theta$  در معادله (10) خواهیم داشت:

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

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

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

که در آن از معادله (8) به طور نسبی مشتق گیری شده است. در این مشتق براساس جنبش مقدار های  $\theta_j$  سه نوع مشتق بدست آمده است. مقدار  $dO()$  مشتق تابع نوسان ساز  $O()$  می باشد که رابطه آن با  $O()$  مشابه رابطه کسینوس با سینوس است:

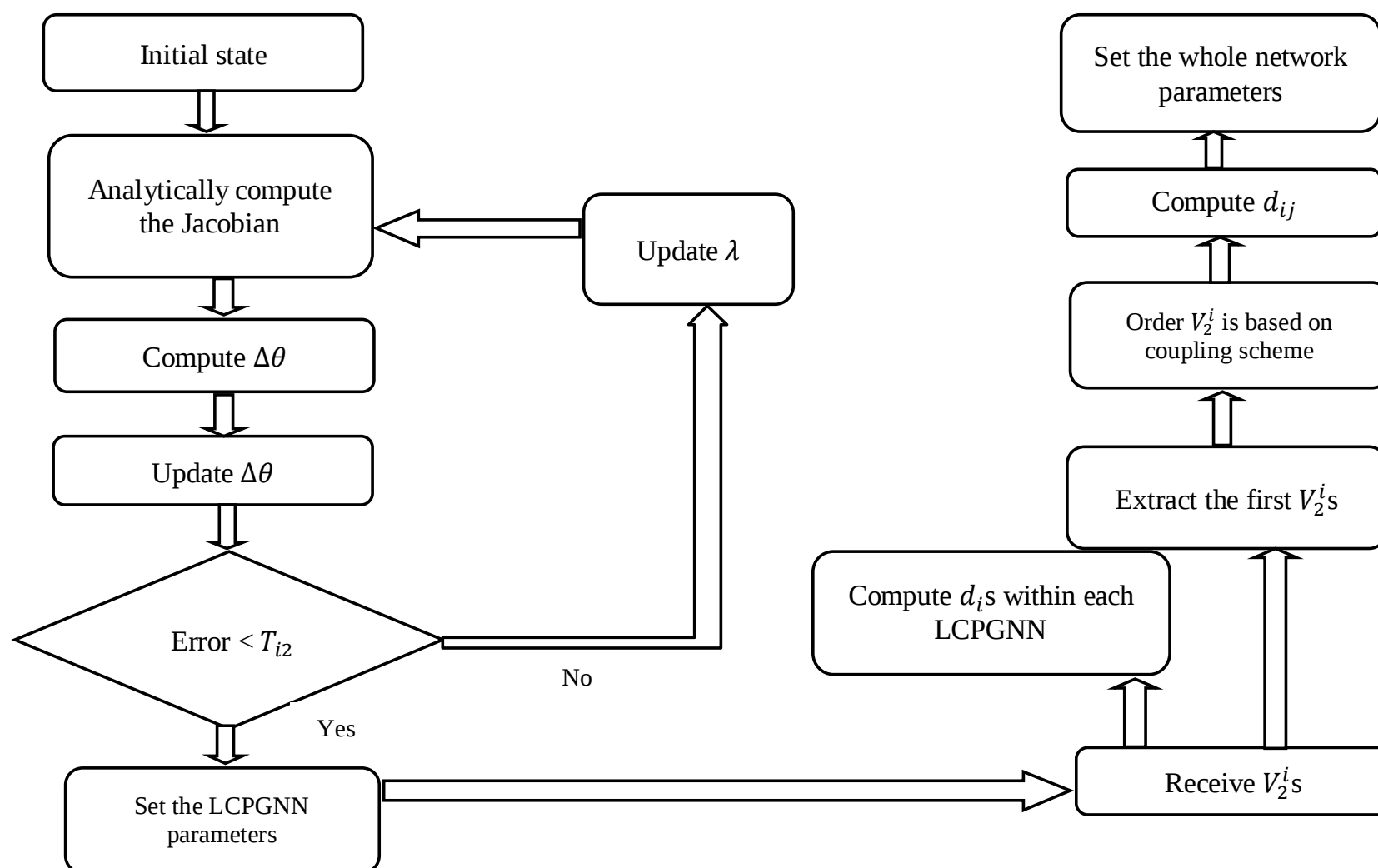
$$dO(1, c) = O(1, \frac{\pi}{2} - c) \quad (14)$$

با داشتن بردار گرادیان  $\nabla J$  معادله گرادیان به شکل برداری در شکل زیر نوشته می شود:

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

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

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



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

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

اجرا می کند و دنباله های آموزشی را به ربات انسان نمای نائو آموزش می دهد. این روش برای آموزش حرکات پیچیده تکرار شونده استفاده می شود. شکل 5 مراحل تولید الگوهای حرکتی مورد نظر را نمایش می دهد. این سیستم شامل هفت دوربین با قابلیت ضبط و ثبت دقیق مسیرهای حرکات نشانگرهای نصب شده در موقعیت های مناسب بر روی دو دست آموزگار انسانی است. دوربین ها موقعیت های نشانگرها را به صورت گسسته و در هر گام زمانی ثبت می کنند و به کمک تعریف و اختصاص یک دستگاه مختصات دکارتی متعامد سه بعدی اینرسی (با سه محور متعامد  $z, y, x$ ) مختصات این موقعیت ها توسط نرم افزار  $QTM$  ذخیره می شود و قابل مشاهده و تحلیل است.

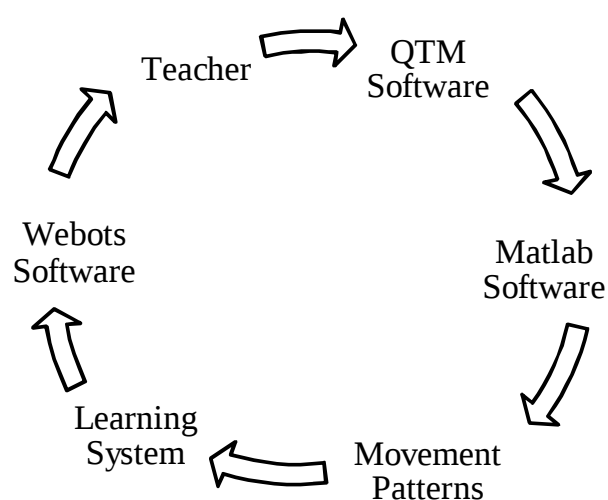
با توجه به داشتن مختصات نشانگرها در هر حرکت می توان به کمک حل مسئله سینماتیک معکوس به کمک الگوریتم ژاکوبین مختصات چرخشی این مختصات را که به اصطلاح مختصات فضای کار نامیده می شوند به مختصات

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

#### 4- پیاده سازی های عملی

در این بخش روش های پیشین یادگیری با کمک نوسان سازهای غیرخطی جفت شده (روش رقتی<sup>1</sup>)، با روش پیشنهادی این مقاله مقایسه و مزایا و معایب آن ها بررسی می گردد. پس از آن به بررسی و تحلیل نتایج و ویژگی های آموزش دو نمونه حرکت، حرکات دست کاراته و نرمش پیش-بالا-طرفین، به ربات انسان نمای نائو در شبیه ساز ویباز، به کمک روش پیشنهاد شده برای هر دو لایه ی سیستم یادگیری تقلیدی ارائه شده در این تحقیق، پرداخته خواهد شد.

برای شبیه سازی سیستم یادگیرنده ی تقلیدی لازم است در ابتدا الگوهای مورد نیاز برای آموزش تقلیدی فراهم شود. در این تحقیق، تولید الگوها با استفاده از دستگاه ضبط حرکت انجام می گیرد. 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 نرم افزار کیو تی ام، موقعیت های حرکتی هر مفصلی که به کمک نشانگرها روی بدن آموزگار انسانی نشان گذاری 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 مسأله ی سینماتیک معکوس الگوی حرکتی زاویه ی هر مفصل به دست آید و پس از آن دنباله های حرکتی بدست آمده به نرم افزار ویباز منتقل گردد تا به قالب قابل استفاده در آموزش (ماتریس های نرم افزار متلب) تبدیل شود. ویباز با اتصال به نرم افزار متلب برنامه های آن را



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

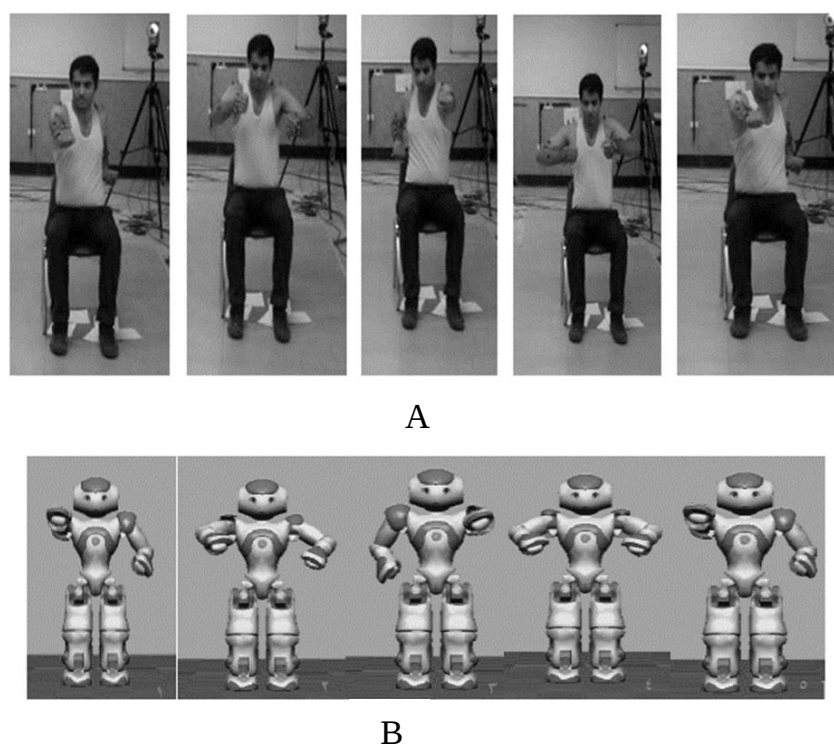
شکل 5 مراحل تولید الگوهای حرکتی برای لایه اول سیستم

1- Righetti method

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. همان گونه که مشاهده می‌شود این روش از دقت بالاتری برخوردار است و نرخ همگرایی مطلوب‌تری را فراهم می‌کند.

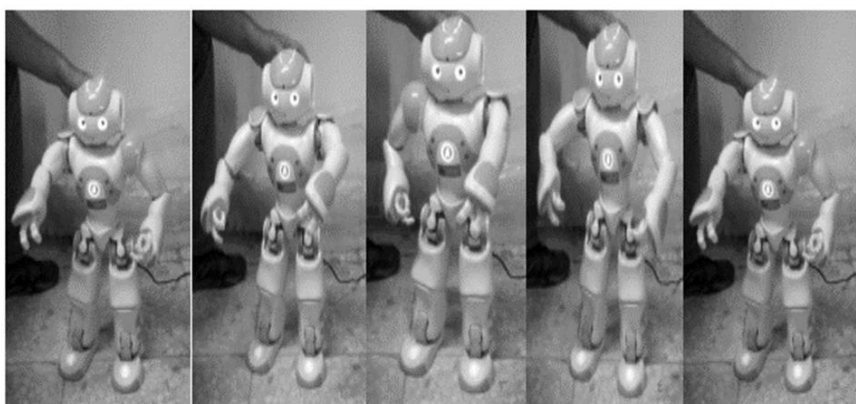
پس از اجرای فرآیند یادگیری در نرم‌افزار متلب و بدست آوردن پارامترهای نوسان‌سازها، می‌بایستی برای شبیه‌سازی سیستم موردنظر روی ربات انسان نمای نائو سراغ محیط شبیه‌ساز ویباتز رفت. به طور نمونه شکل 7 (الف) نمایی از ضبط حرکت کاراته‌ی مربی توسط دستگاه ضبط حرکت موجود در مرکز اسکلتی- عضلانی دانشکده توانبخشی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

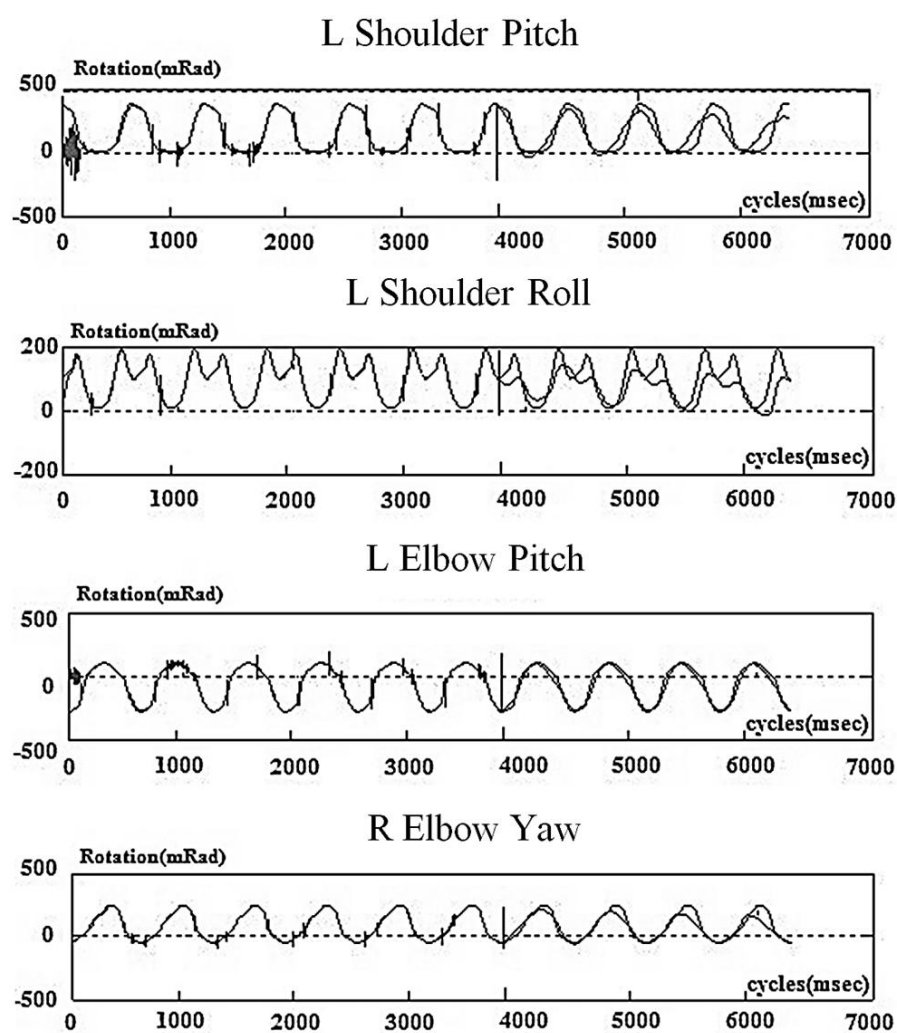
شکل 8 پیاده‌سازی حرکت سوکی کاراته روی ربات نائوی واقعی

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 وفق پذیر 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 اصلاح شده در شکل 6 نشان داده شده است. همان گونه 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

## 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.

همان گونه 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 هیچ گونه وابستگی به مقادیر اولیه نوسان سازها ندارد و بدون وابستگی به مقادیر اولیه نوسان سازها تمام ریتیم‌های 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 نوسان سازها است. درواقع روش‌های گذشته در تعدادی از حالت‌ها اصلاً همگرا نشده‌اند و این امر سبب بالا رفتن متوسط خطای تست شده‌اند. در این روش‌ها 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 نوسان سازها، همگرا شدن یک نوسان ساز 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

**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

وجود آمده است که در روش نهایی با افزایش معیار زمان آموزش، میزان معیار خطای تست کاهش و نرخ همگرایی افزایش یافته است.

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 علوم پزشکی 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 یکی-یکی 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. Krüger, 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. Ijspeert, 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.

- [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.
- [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.