



حذف نویز صوتی تناوبی با روش فعال و هوشمند

بهرروز رئیسی^{۱*}، شاپور گلپه‌ار حقیقی^۲ و سید علی اکبر صفوی^۳

^۱ دانشجوی دکتری مخابرات، دانشکده برق و کامپیوتر، دانشگاه شیراز، شیراز

^۲ استادیار، دانشکده برق و کامپیوتر، دانشگاه شیراز، شیراز

^۳ استاد، دانشکده برق و کامپیوتر، دانشگاه شیراز، شیراز

^۴ کارشناس ارشد کنترل، پژوهشکده مکانیک پژوهشگاه فضایی ایران، سازمان فضایی ایران، شیراز

چکیده

در این تحقیق، روشی جدید مبتنی بر تکنیک‌های یادگیری تقویتی با تکیه بر روش تفاوت زمانی صفر در مبحث یادگیری تقویتی، برای حذف نویز صوتی به صورت فعال در سیگنال‌های باند باریک ارائه گردیده است. برای پیاده‌سازی این روش مبتنی بر بازخورد، اطلاعاتی از دینامیک مسیر اولیه و ثانویه لازم نمی‌باشد و در نتیجه این روش نسبت به تغییر این پارامترها به صورت مقاوم عمل می‌نماید. حجم بسیار کم محاسبات ریاضی و حافظه بسیار اندک محاسبات، از دیگر مزایای روش پیشنهادی می‌باشد. بدین منظور در اولین گام با انجام تعاریف مناسب برای متغیرهای حالت، اعمال ممکن و سیگنال پاداش، مسئله به یک مسئله مناسب برای حل توسط یادگیری تقویتی تبدیل شده و با روش یادگیری کیو حل می‌گردد. در گام بعدی با استفاده از اطلاعات به دست آمده از مدل‌سازی در حوزه فرکانس، مسئله ساده گردیده و با روش تفاوت زمانی صفر مسئله حل خواهد گردید. آخرین مرحله پیشنهاد یک روش چند سطحی می‌باشد که در این روش اخیر بدون نیاز به بالاتر بردن حافظه مورد نیاز می‌توان، دقت را تا مقدار لزوم افزایش داد.

کلمات کلیدی: هوش مصنوعی؛ یادگیری تقویتی؛ حذف نویز به صورت فعال؛ سیگنال باند باریک؛ روش تفاوت زمانی صفر.

Periodic acoustic noise cancellation with intelligent and active method

B. Raeisy^{1,4*}, S. G. Haghghi² and A. A. Safavi³

¹ Ph.D. Student of Communications. Eng., Shiraz Univ., Shiraz, Iran

² Assistant Prof., Electrical Eng., Shiraz Univ., Shiraz, Iran

³ Prof., Process Systems Eng., Shiraz Univ., Shiraz, Iran

⁴ Control Eng., Institute of Mechnics, Iranian Space Research Center, Iranian Space Agency, Shiraz, Iran

Abstract

In this paper a new method based on temporal difference reinforcement learning algorithm is proposed to active noise control of periodic acoustic signal. This method does not need dynamic information and therefore it is fully robust with respect to the dynamic changes. Very low computational burden and low needed memory are some advantages of this method. In the first step, through well definition of the state variables, actions, and reward signal, a reinforcement learning problem is formed and then it is solved by Q-learning technique. In the next step, the problem is simplified with the aid of frequency domain modeling information and it is solved by temporal difference methods. Finally, a multi-level approach is presented which can increase the precision without increasing of the memory size. It is shown by simulation that this method works well and the improvement of the multi-level temporal difference method with respect to the Q-learning method is addressed.

Keywords: Artificial intelligence; Reinforcement learning; Active noise control; Narrow band signal; Temporal difference method.

۱- مقدمه

ساخت ماشینی که همانند بشر قادر باشد با تعامل با محیط اطراف خود و شرایط موجود مشکلاتش را از پیش رو برداشته و به سمت یک هدف مشخص پیش رود یکی از رؤیا پردازی-های دیرین بشر می‌باشد. در این رفتار تقلید گونه ماشین از انسان، سیستم باید بتواند با توجه به تجربیات قبلی در مواجهه با حالات جدید نیز عکس‌العمل‌های مناسبی از خود نشان دهد. آنچه انسان را در این‌گونه رفتارها از ماشین متمایز می‌سازد هوش اوست و برای تحقق هدف فوق بایستی بتوان هوش را برای ماشین شبیه‌سازی نمود. در این نمونه هوش که با عنوان هوش مصنوعی شناخته می‌شود سیستم روش آموختن انسان را تقلید می‌نماید و به جذب اطلاعات جدید جهت به‌کارگیری در مراحل بعدی می‌پردازد. راسل^۱ و نورینگ^۲ در تعریف سیستم هوشمند می‌گویند "یک سیستم هوشمند سیستمی است که با شناخت محیط اطراف خود شانس موفقیت و رسیدن به هدف خود را بالا می‌برد" [۱].

امروزه هوش مصنوعی در شاخه‌های مختلفی رشد نموده و ابزارهای بسیاری نظیر شبکه عصبی و سیستم‌های فازی را در اختیار دارد که در نتیجه این موضوع باعث گردیده که بتوان آن را در حوزه‌های وسیعی از علوم و مهندسی به کار گمارد. سیستم‌های مبتنی بر هوش مصنوعی معمولاً نیازی به دینامیک محیط ندارند، به صورت وقتی عمل می‌نمایند و در صورت تغییرات جزئی و یا اثرگذاری عوامل مزاحم خارجی به راحتی آن را پشت سر خواهد گذاشت.

در این تحقیق استفاده از روش یادگیری تقویتی در موضوع حذف نویز صوتی برای سیگنال‌های تناوبی با روش‌های فعال^۳ (ANC) مد نظر می‌باشد. روش یادگیری تقویتی یکی از روش‌های موجود در مبحث یادگیری ماشین می‌باشد [۲ و ۳]. در حل مسائل با ابزار یادگیری تقویتی همانند سایر روش‌های مبتنی بر هوش مصنوعی، سیستم خود را با تغییرات وفق داده و چندان نیازی به مدل‌سازی سیستم و محیط وجود ندارد. از طرف دیگر پیاده‌سازی آن نسبتاً ساده و با محاسبات بسیار کم و بر اساس روش‌های مبتنی بر تکرار^۴

می‌باشد به نحوی که در هر گام به هدف نهایی نزدیک و نزدیک‌تر می‌گردد.

روش‌های ANC عمدتاً بر روی تخمین دینامیک مسیر ثانویه^۵ استوار می‌باشند [۴] که اگر این تخمین در حین عملکرد سیستم به صورت بلادرنگ صورت نپذیرد دیگر سیستم نسبت به تغییر دینامیک مسیر ثانویه به صورت وقتی عمل نمی‌نماید و در صورت عملکرد بلادرنگ^۶ آن بار محاسباتی بالا می‌رود. با روش یادگیری تقویتی سیستم بدون داشتن مدل دینامیکی مسیر ثانویه و یا تخمین آن در حین عملکرد و تنها با نگاه کردن به سیگنال نویز موجود در محیط، بتواند یاد بگیرد که چگونه این سیگنال را به سمت صفر میل دهد. بار محاسباتی این روش بسیار کم بوده و به تدریج سعی در محو خطا خواهد نمود.

با توجه به نقش اساسی یادگیری تقویتی در این مقاله پس از مقدمه در دومین بخش مروری بر مبحث یادگیری تقویتی صورت خواهد پذیرفت. از طرف دیگر روش حذف نویز به صورت فعال به کمک یادگیری کیو [۵] به عنوان پایه‌ای برای روش فعلی می‌باشد که بدین دلیل در سومین بخش به صورت خلاصه این روش شرح داده خواهد شد. این روش برای تعداد محدود هارمونیک غالب^۷ در محیط پیاده‌سازی می‌گردد که با توجه به تعدد عمل‌ها و حالت‌های موجود نشان داده می‌شود که تعداد اعضاء جدول مقدار (حالت-عمل)^۸ بزرگ خواهد گردید. بنابراین هر چند که اگر تعداد محدود هارمونیک مد نظر باشد، بار بسیار کم محاسباتی روش می‌تواند توجیه‌گر نقیصه قبلی باشد ولی در صورت بالاتر رفتن هارمونیک‌های غالب در محیط مزیت نسبی کاهش می‌یابد. به همین منظور در بخش چهارم روش تفاوت زمانی صفر^۹ شرح داده شده و در بخش بعد روشی بر آن مبنا برای حذف نویز ارائه می‌گردد. در این روش به جای استفاده از جدول مقدار (حالت-عمل) از جدول مقدار حالت^{۱۰} استفاده می‌گردد که نتیجه آن کاهش ابعاد حافظه مورد لزوم

^۵ دینامیکی که رفتار محیط از خروجی کنترل کننده نویز تا منطقه‌ای که در آن نقطه بایستی نویز در آن حذف شود را نشان می‌دهد (Secondary path).

^۶ online

^۷ dominant

^۸ $Q(s,a)$

^۹ TD(0)

^{۱۰} $V(s)$

^۱ Russel

^۲ Noring

^۳ active noise control

^۴ iterative

صورت درازمدت است. در حالت کلی در یک سیستم یادگیری تقویتی شش قسمت اصلی وجود دارد:

الف- **عامل^۲ یادگیرنده**: که در واقع عامل تصمیم گیرنده است. این عامل اعمال را انتخاب می‌کند، محیط را اکتشاف می‌کند و یک راه مناسب و بهینه به منظور بیشینه کردن پاداشی که از محیط دریافت کرده، انتخاب می‌کند.

ب- **محیط^۳ یادگیری**: محیط منبع بازخوردهایی^۴ است که به عامل بر می‌گردد. این بازخورد اصولاً به صورت سیگنال پاداش مطابق با عملی که عامل انجام می‌دهد، به عامل بازمی‌گردد.

پ- **سیاست^۵**: خط مشی یادگیری عامل را مشخص می‌کند. در واقع سیاست یک نگاشت میان حالت‌های موجود در محیط و اعمالی که در موقع رخ دادن آن حالت اجرا می‌شوند، می‌باشد.

ت- **تابع پاداش^۶**: که به نوعی هدف مسئله یادگیری را تعریف می‌کند. این تابع یک نگاشت بین هر جفت حالت-عمل و یک مقدار عددی مشخص به نام پاداش می‌باشد که در مقابل انجام آن عمل توسط عامل در آن حالت، به آن داده می‌شود. وظیفه عامل بیشینه کردن مقدار پاداش کل البته در درازمدت است. اگر پاداشی که عامل از محیط در مرحله t ام می‌گیرد به صورت r_t نشان داده شود آنگاه پاداش کل^۷ به صورت رابطه (۱) تعریف می‌گردد:

$$R_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \quad 0 < \gamma < 1 \quad (1)$$

γ به نام نرخ تنزیل^۸ شناخته شده که برای مسائلی که حالت پایانی ندارند، از بی‌نهایت شدن پاداش جلوگیری می‌کند. این رابطه نشان می‌دهد که در تصمیم‌گیری نهایی عامل، پاداشی که k مرحله دیرتر دریافت شود ارزشی γ^{k-1} برابر کمتر از حالت ابتدایی خود دارد.

ث- **تابع مقدار (ارزش)^۹**: در حالت کلی این تابع مقدار خوب بودن اعمال انجام شده توسط عامل را در درازمدت

در نگهداری مقادیر داده‌های مورد نیاز می‌باشد. در همان بخش پیشنهاد یک روش بهبودیافته به نام الگوریتم تفاوت زمانی چند سطحی نیز ارائه می‌گردد که در آن بدون نیاز به بالاتر بردن حافظه مورد نیاز، می‌توان دقت را تا مقدار لزوم افزایش داد. این متد علاوه بر کاهش بعد فراوان در سیستم، باعث کاهش زمان همگرایی نیز خواهد گردید. در بخش شش نیز با شبیه‌سازی، صحت عملکرد رویه‌های پیشنهادی نشان داده شده و نهایتاً جمع‌بندی و نتیجه‌گیری در انتها صورت خواهد پذیرفت.

۲- مروری بر یادگیری تقویتی [۲]

تعامل^۱ با محیط اصلی‌ترین منبع به دست آوردن اطلاعات در زندگی محسوب می‌شود و این موضوع تقریباً می‌تواند ایده‌آسی برای همه نظریه‌های هوش مصنوعی و یادگیری محسوب شود. وقتی انسان در کاری مهارت پیدا می‌کند در واقع می‌آموزد که چگونه محیط اطراف نسبت به کنش‌های وی واکنش نشان می‌دهد و چگونه می‌تواند با کنش‌های مناسب این واکنش‌ها را تحت کنترل درآورد. یادگیری تقویتی به فرایندی گفته می‌شود که در طی آن یک عامل می‌آموزد چه اعمالی باید انجام دهد تا پاداش بیشتری از محیط دریافت کند. در این فرایند به عامل گفته نمی‌شود که چه عمل‌هایی را باید انجام دهد بلکه خود عامل است که باید طی این فرایند به دست آورد که کدام اعمال پاداش بیشتری حاصل می‌کند. نکته مهمی که باید به آن توجه شود این است که باید پاداش را به گونه‌ای اختصاص داد که عامل با بیشینه کردن آن هدف را تأمین سازد و نباید به او یاد داد که چگونه هدف را برآورده سازد. در واقع سیگنال پاداش کانال ارتباطی شما با عامل است که به وسیله آن به او می‌گویید به چه هدفی برسد نه اینکه چگونه به هدف برسد. نکته دیگری که در این یادگیری قابل توجه است آن است که انجام یک عمل توسط عامل نه تنها شامل یک پاداش آنی می‌شود بلکه بر موقعیت‌های بعدی و همچنین مجموع پاداش‌های به دست آمده طی فرایند نیز تأثیرگذار است. خصوصیت منحصر به فرد این نوع یادگیری نسبت به سایر روش‌های موجود جستجوی سعی و خطاوار عامل و همچنین در نظر گرفتن پاداش به

² agent

³ environment

⁴ feedback

⁵ policy

⁶ reward

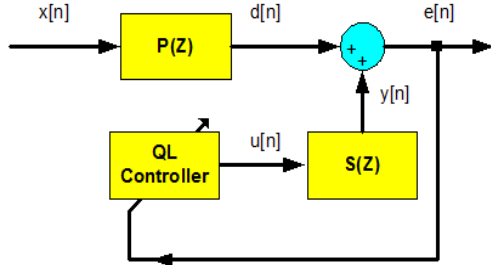
⁷ return

⁸ discount factor

⁹ value function

¹ interaction

نویز به صورت فعال با روش یادگیری کیو (QL-ANC) را نشان می‌دهد.



شکل ۱- بلوک دیاگرام حذف نویز با روش یادگیری تقویتی [۵]

در این شکل $x[n]$ نویز تناوبی تولید شده است که تغییر یافته آن پس از گذر از محیط، به نقطه‌ای می‌رسد که بایستی در آن مکان محیطی ساکت ایجاد شود. دینامیک محیط در این مسیر با $P(Z)$ نشان داده شده و به نام مسیر اولیه شناخته می‌گردد. کنترل کننده یادگیری تقویتی با استفاده از سیگنال پس خور بایستی سیگنالی تولید کند که پس از گذر دینامیک مسیر ثانویه^۲ محیط ساکت را ایجاد نماید. این سیستم در حالت کامل شده برای حذف نویز یک سیگنال تناوبی با M هارمونیک پیشنهاد گردیده است. کنترل کننده پیشنهادی برای این کار، دارای $2M$ متغیر و به صورت زیر تعریف می‌گردد:

$$u(t) = \sum_{m=1}^M a_{2m-1} \cos m\omega_0 t + a_{2m} \sin m\omega_0 t \quad (2)$$

هدف در این مسئله یافتن مقادیر a_1 تا a_{2M} است به نحوی که سیگنالهای تولید شده در رابطه (۲) بتوانند پس از گذر از محیط، نویز موجود در محیط در ناحیه ساکت را خنثی و توان سیگنال صوتی در آن ناحیه را به سمت صفر ببرد. با توجه به اینکه روش یادگیری تقویتی برای این منظور انتخاب شده تعریف متغیر حالت، اعمال قابل قبول و مشخص نمودن سیگنال تقویتی یا پاداش از مراحل کار می‌باشد. موقعیت فعلی پارامترهای متغیر مشخص کننده وضعیت سیستم به صورت کامل می‌باشد لذا برای تعریف متغیر حالت می‌توان از مقدار عددی شده پارامترهای a_1 تا a_{2M}

مشخص می‌کند. آن چیزی که باید توجه شود آن است که باید به دنبال انتخاب عمل‌هایی بود که بالاترین مقادیر را تولید می‌کند نه بالاترین پاداش‌ها و این راه‌برد باعث خواهد شد که در درازمدت این عمل‌ها پاداش بیشتری را برای عامل حاصل کنند. توابع ارزش هم برای حالت و هم برای حالت-عمل تعریف می‌گردند. بیان مفهومی تابع ارزش حالت، امید ریاضی پاداش کلی است که عامل با شروع از حالت $s_t = s$ و با در پیش گرفتن سیاست π به دست می‌آورد. این مفهوم به صورت ریاضی نیز به صورت $E_{\pi}\{R_t | s_t = s\} = V^{\pi}(s)$ قابل بیان می‌باشد که نشان‌دهنده امید ریاضی^۱ عامل به پاداش است هنگامی که وی در مرحله t در حالت s قرار دارد و از سیاست π پیروی می‌نماید.

تابع حالت-عمل نیز به علامت $Q^{\pi}(s, a)$ نشان داده شده و بیانگر امید ریاضی پاداش کلی است که عامل با شروع از حالت $s_t = s$ و انجام عمل $a_t = a$ و سپس در پیش گرفتن سیاست π به دست می‌آورد و به صورت $Q^{\pi}(s, a) = E_{\pi}\{R_t | s_t = s, a_t = a\}$ تعریف می‌گردد.

ج- مدل محیط: وجود آن در سیستم یادگیری تقویتی الزامی نمی‌باشد ولی از آن می‌توان برای شبیه‌سازی رفتار محیط و همچنین طرح‌ریزی بدون انجام عمل واقعی استفاده نمود.

۳- مروری بر روش حذف نویز با روش یادگیری کیو

استفاده از یادگیری تقویتی با روش QL در حذف نویز به صورت فعال، ابتدا در مرجع [۵] توسط نویسنده پیشنهاد گردید که علیرغم نوآوری و مزایایی که در پی داشت در مسائلی که تعداد هارمونیک افزایش می‌یابد با چالش‌هایی روبرو می‌گردید. با عنایت به این موضوع که تحقیق جاری باهدف تکمیل کار و تبدیل آن به روشی با کاربری گسترده‌تر برای سیگنال‌های با تعداد هارمونیک بیشتر صورت می‌گیرد لذا در این بخش مروری بر سیستم پیشنهادی قبلی صورت خواهد پذیرفت. شکل ۱ بلوک دیاگرام پیشنهادی برای حذف

^۲ S(Z)

^۱ expectation

که در آن α اندازه گام^۵ بوده و بر نرخ یادگیری اثر می‌گذارد. لازم به ذکر است روش $\varepsilon - greedy$ تصحیح‌شده روش حریمانه می‌باشد. این روش ابتدا توسط Watkins در ۱۹۸۹ در یک تز دکتری ارائه گردید [۸]. در روش حریمانه تنها عملی انتخاب می‌شود که بهترین تشخیص داده شود در صورتی که در $\varepsilon - greedy$ شانس به سایر اعمال نیز داده می‌شود. بنابراین در این روش با احتمال $1 - \varepsilon$ بهترین عمل و سایر اعمال با احتمال $\varepsilon / (n - 1)$ انتخاب می‌گردد. معمولاً ε کوچک و مثبت نظیر ۰/۱ اختیار می‌گردد.

الگوریتم QL سعی در بیشینه کردن تابع پاداش خود در طولانی مدت را دارد بنابراین با توجه به هدف سیستم که کاهش توان نویز در ناحیه ساکت می‌باشد، منفی توان سیگنال در محیط به عنوان پاداش در نظر گرفته می‌شود. با این تعریف بیشینه کردن پاداش، معادل صفر نمودن توان نویز در خروجی و یا ساکت بودن مطلق ناحیه مورد نظر خواهد گردید. اگر دوره تناوب زمانی اعمال فرمان با T_a نشان داده شود آنگاه در بازه زمانی بین $(k-1)T_a$ و kT_a سیگنال پاداش با $R[k]$ نشان داده شده و به صورت زیر تعریف می‌گردد:

$$R[k] = -\overline{e^2(kT_a)} = -\frac{1}{T_a} \int_{t=(k-1)T_a}^{kT_a} e^2(t) dt \quad (6)$$

نکته ظریف دیگری که در این میان باقی می‌ماند نیاز به علامتی همزمان با فرکانس اصلی و یا به عبارت دیگر تأمین سیگنال‌های $\sin m\omega_0 t$ و $\cos m\omega_0 t$ در رابطه (۲) می‌باشد. در برخی از موارد برای این منظور پیشنهاد استفاده از روش‌هایی در حوزه فرکانس نظیر تبدیل فوریه سریع (FFT^۶) از سیگنال خروجی [۹] ارائه شده است اما، این روش بار محاسباتی مضاعفی را بر سیستم تحمیل می‌نماید. در هنگامی که سیگنال نویز از یک منبع تناوبی حاصل می‌گردد معمولاً این امکان وجود دارد که سیگنالی هم‌فاز بدون نمونه‌برداری صوتی ایجاد نمود. به عنوان مثال در شکل ۲ منبع نویز حاصل از دوران یک جسم گردنده می‌باشد و با تعبیه یک اندازه گیر زاویه بر روی محور دوران سیگنالی با

استفاده نمود. تابع S_{fM} برای نسبت دادن یک عدد صحیح به حالت سیستم به صورت رابطه (۳) تعریف می‌گیرد. به عبارت دیگر این تابع عدد صحیح منحصر به فردی را به هر ناحیه کوچک از فضای $2M$ بعدی از پارامترهای $a_1 \times a_2 \times \dots \times a_{2M}$ اختصاص می‌دهد که این عدد نمایانگری^۱ یگانه برای آن محدوده می‌باشد.

$$S_{fM}(a_1, \dots, a_{2M}) = \sum_{m=1}^{2M} N^{m-1} \text{Digi}(a_m) \quad (3)$$

در این تعریف تابع $\text{Digi}(a_m)$ به a_m از مقدار کمینه تا بیشینه خود یک عدد صحیح 0 تا $N-1$ را نسبت می‌دهد که به این فرایند رقومی سازی^۲ گفته می‌شود. به سادگی این موضوع قابل بررسی است که تابع S_{fM} مقداری صحیح بین صفر تا $N^{2M} - 1$ را به حالت سیستم نسبت خواهد داد. انتخاب N دلخواه می‌باشد. هر چند که با انتخاب عدد بزرگتری برای N با دقت بیشتری یک ناحیه به خصوص آدرس داده می‌شود اما تعداد نواحی به صورت نمایی افزایش می‌یابد و همان گونه که ذکر خواهد گردید، این افزایش مقدار حافظه در برخی از موارد ایجاد اشکال می‌نماید.

برای تعریف مجموعه $A(s)$ که نشان‌دهنده اعمال مجاز در حالت دلخواه s هست، عملگرهای inc و dec برای افزایش و کاهش پارامتر خود به اندازه دلخواه $step$ (به شرطی که محدوده مجاز خود خارج نشود) و noAction برای بدون تغییر ماندن همه پارامترها معرفی و مجموعه $A(s)$ با $4M + 1$ عضو به صورت زیر تعریف می‌گردد:

$$A(s) \triangleq \{\text{dec}(a_1), \text{inc}(a_1), \dots, \text{dec}(a_{2M}), \text{inc}(a_{2M}), \text{noAction}\} \quad (4)$$

احتمال انتخاب عمل بین اعمال فوق در حالت s بسته به مقدار متناظر آن برای هر عمل در جدول $Q(s, a)$ و بر طبق مدت ε -حریمانه^۳ می‌باشد و نحوه به‌روز آوری جدول Q مطابق فرمول مشهور یادگیری کیو^۴ (QL) به صورت زیر می‌باشد:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)] \quad (5)$$

¹ represent

² digitization

³ ε -greedy

⁴ Q-Learning

⁵ step-size

⁶ Fast Fourier Transform

کردن مقدار دلخواه $step$ سرعت همگرایی بیشتر شده و الگوریتم با گام‌های بلندتری به سمت هدف حرکت می‌نماید اما، در حوالی هدف رسیدن به مقدار مطلوب به دقت مناسب با احتمال کمتری تضمین می‌گردد. بنابراین باید به توافقی بین دو کمیت دقت و سرعت تن در داد و مقداری میانی را بین این دو برگزید. در روش پیشنهادی این تحقیق راه‌حلی برای برطرف کردن هر دو معضل ارائه می‌گردد.

۴- روش تفاوت زمانی در یادگیری تقویتی

این روش یک روش مابین روش مونت کارلو و برنامه‌ریزی پویا می‌باشد و جزو محبوب‌ترین و پرکاربردترین روش‌های موجود است. این روش از یک طرف مانند مونت کارلو نیازی به مدل محیط ندارد و از طرف دیگر مانند برنامه‌ریزی پویا تخمین‌های خود برای تابع مقدار را بر اساس سایر تخمین‌های یاد گرفته‌شده انجام می‌دهد.

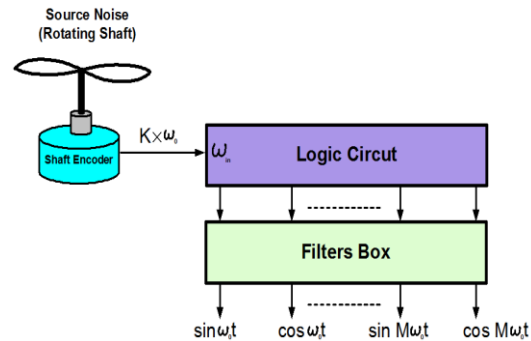
در این روش عامل سعی می‌کند به تقریب تابع مقدار بر اساس پاداش لحظه‌ای و مقدار تخمین زده‌شده برای حالت بعدی که در آن قرار خواهد گرفت بپردازد و در این راستا از روش تکرار مقدار نیز کمک می‌گیرد.

یکی از پرکاربردترین الگوریتم‌های موجود در دسته‌بندی فوق روش یادگیری کیو است که در بخش ۳ از این تکنیک استفاده گردید. اما اساس این تحقیق بر ساده‌ترین روش از این گروه از نظر به‌روز آوری یعنی TD(0) می‌باشد. در این روش به جای جدول مقدار حالت-عمل یا همان Q، جدول مقدار حالت با فرمول زیر به‌روز می‌شود:

$$V(s_t) \leftarrow V(s_t) + \alpha(r_{t+1} + \gamma V(s_{t+1}) - V(s_t)) \quad (8)$$

هر چند که استفاده از روش TD به صورت فوق ساده و بی‌عیب به نظر می‌رسد اما اشکال در عدم تعریف مشخصی از سیاست است. در اکثر راه‌حل‌های یادگیری تقویتی هدف یافتن سیاست بهینه با نسخه‌ای از روش عمومی تکرار سیاست (GPI) می‌باشد. در این روش بایستی پس از هر مرحله از ارزیابی سیاست قدمی نیز در بهبود سیاست برداشته شود تا تابع مقدار به سمت مقدار بهینه و سیاست نیز به سمت سیاست بهینه نزدیک شود [۲، ۳] که این موضوع در شکل ۳ نشان داده شده است.

فرکانس چند برابر فرکانس اصلی (K) تولید گردیده است. حال اگر مقصود تولید سیگنالی تا هارمونیک M ام باشد ($K > M$) به راحتی می‌توان با یک مدار منطقی فرکانس‌های مربعی هم‌فاز با سیگنال‌های خواسته‌شده ایجاد نمود و با گذر از یک مجموعه فیلتر سیگنال‌های مذکور را تولید نمود.



شکل ۲- تأمین سیگنال همزمان با نویز صوتی با نمونه‌برداری غیر صوتی

هرچند با تعریف کنترل‌کننده QL، متغیر حالت و مجموعه اعمال و سیگنال پاداش تعریف مسئله کامل گردید اما در اجرا روش مشکلاتی را خواهد داشت که ذیلاً بدان پرداخته می‌گردد.

روش QL بر اساس تخمین جدول $Q(s, a)$ و سپس تصمیم‌گیری انجام عمل بر اساس آن می‌باشد. تعداد اعضاء این جدول با ضرب تعداد حالت‌های ممکن در تعداد عمل برای یک حالت به دست خواهد آمد. تعداد حالت‌ها از رابطه (۳) و مساوی N^{2M} و تعداد اعمال تعریف‌شده در رابطه (۴) مساوی $4M + 1$ می‌باشد. بنابراین تعداد اعضاء این جدول در روش پیشنهادی به صورت زیر محاسبه خواهد گردید.

$$\mathbb{N}(Q_M(s, a)) = \mathbb{N}(S_M) \times \mathbb{N}(A_M(s)) = N^{2 \times M} \times (4 \times M + 1) \quad (9)$$

که M تعداد هارمونیک و $\mathbb{N}(\cdot)$ عملگری است که تعداد آرگومان خود را نشان می‌دهد. همان‌گونه که مشاهده می‌شود تعداد اعضاء این جدول با بالا رفتن تعداد هارمونیک‌ها به صورت نمایی افزایش می‌یابد. با توجه به اینکه برای رسیدن به دقت مناسب بایستی عددی نظیر ۵۰ برای N در نظر گرفت، در حالت کلی تهیه کردن حافظه در بیش از ۳ هارمونیک به سادگی ممکن نخواهد بود. نکته دیگر رابطه دقت روش و سرعت آن می‌باشد. هرچند که با بزرگ اختیار

¹ generalized policy iteration

با توجه به فرض خطی بودن کل سیستم، سامانه حذف نویز نیز بایستی سیگنالی پریودیک تولید نماید که در صورتی که Γ_m نیز مقدار تنظیم‌شده توسط کنترل‌کننده در هارمونیک m ام باشد آنگاه خروجی کنترل‌کننده را می‌توان به صورت $\sum_{m=1}^M \Gamma_m \delta(\omega - m\omega_0)$ نشان داد.

با توجه به شکل ۴ در حوزه فرکانس می‌توان روابط زیر را نوشت:

$$\begin{aligned}
 E(e^{j\omega}) &= Y(e^{j\omega}) + D(e^{j\omega}) \\
 Y(e^{j\omega}) &= S(e^{j\omega}) \sum_{m=1}^M \Gamma_m \delta(\omega - m\omega_0) \\
 &= \sum_{m=1}^M S(e^{jm\omega_0}) \Gamma_m \delta(\omega - m\omega_0) \quad (10) \\
 D(e^{j\omega}) &= P(e^{j\omega}) \sum_{m=1}^M \chi_m \delta(\omega - m\omega_0) \\
 &= \sum_{m=1}^M P(e^{jm\omega_0}) \chi_m \delta(\omega - m\omega_0)
 \end{aligned}$$

$$E(e^{j\omega}) = \sum_{m=1}^M [\chi_m P(e^{jm\omega_0}) + \Gamma_m S(e^{jm\omega_0})] \delta(\omega - m\omega_0) \quad (11)$$

با تعریف ξ_m به صورت $\xi_m \triangleq \chi_m P(e^{jm\omega_0}) + \Gamma_m S(e^{jm\omega_0})$ خطا به صورت رابطه (۱۲) قابل بیان خواهد بود

$$E(e^{j\omega}) = \sum_{m=1}^M \xi_m \delta(\omega - m\omega_0) \quad (12)$$

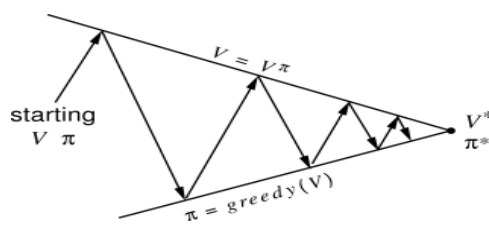
و در این صورت توان نیز مطابق رابطه زیر به دست خواهد آمد.

$$\overline{e^2} = \int_{-\pi}^{\pi} E^2(e^{j\omega}) d\omega = \sum_{m=1}^M |\xi_m|^2 \quad (13)$$

با تعریف α_m و β_m رابطه (۱۳) به صورت (۱۴) قابل بیان شدن است.

$$\left. \begin{aligned}
 \alpha_m &\triangleq \chi_m \frac{P(e^{jm\omega_0})}{S(e^{jm\omega_0})} \\
 \beta_m &\triangleq \frac{1}{S(e^{jm\omega_0})}
 \end{aligned} \right\} \rightarrow \overline{e^2} = \sum_{m=1}^M \frac{|\Gamma_m - \alpha_m|^2}{|\beta_m|^2} \quad (14)$$

هدف در این مسئله یافتن Γ_m با داشتن $\overline{e^2}$ می‌باشد. α_m و β_m نیز ناشناخته و مرتبط با مشخصات کلی سیستم و سیگنال ورودی می‌باشند. با تعریف اعداد حقیقی



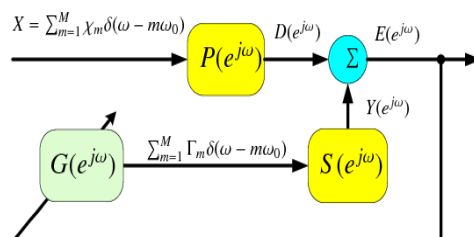
شکل ۳- روش عمومی تکرار سیاست [۲]

در روش QL در هر گام پس از به‌روز آوری جدول Q مطابق فرمول (۵) که در واقع همان یک مرحله ارزیابی سیاست است، روش انتخاب عمل هم که بر اساس جدول $Q(s, a)$ است بهبود می‌یابد و در نتیجه سیاست نیز بهینه می‌گردد و اصول GPI رعایت می‌گردد. اما در حالت کلی به‌روزرسانی و استفاده از $V(s)$ به تنهایی پیشنهادی در ارتباط با سیاست به عامل نمی‌دهد. برای حل این مشکل در ادامه بخش آتی و در این مسئله خاص، روشی بر اساس دینامیک پیشنهاد خواهد گردید.

۵- شرح الگوریتم پیشنهادی

در بخش قبل از روش یادگیری کیو بدون توجه به دینامیک مسئله استفاده گردید. در این بخش با کمک مدل‌سازی سیستم در حوزه فرکانس و توجه به مدل در حل مسئله، ساده‌سازی‌هایی صورت خواهد پذیرفت. بلوک دیاگرام بیان‌شده در شکل ۱ را می‌توان در فضای فرکانس به صورت شکل ۴ نشان داد. با توجه به تناوبی بودن نویز صوتی در صورتی که χ_m عددی موهومی و نمایانگر مؤلفه نویز صوتی ورودی در هارمونیک m ام باشد، آنگاه سیگنال صوتی را X نامیده و می‌توان آن را به صورت رابطه زیر نشان داد:

$$X = \sum_{m=1}^M \chi_m \delta(\omega - m\omega_0) \quad (9)$$



شکل ۱- نمایش سیستم حذف نویز در فضای فرکانس

S_{f_M} برای نسبت دادن یک عدد به حالت سیستم به صورت زیر تعریف می‌گیرد:

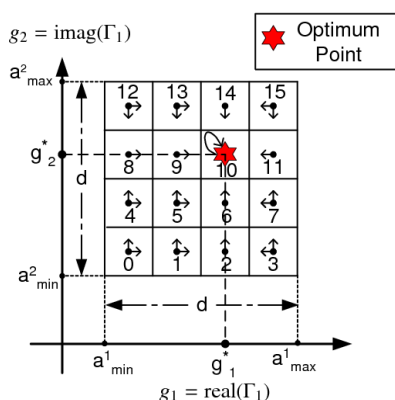
$$S_{f_M}(g_1, \dots, g_{2M}) = \sum_{m=1}^{2M} N^{m-1} \text{Digi}(g_m) \quad (17)$$

اعمال مجاز نیز مشابه حالت قبل و به صورت زیر تعریف می‌گردد:

$$A(s) \triangleq \{\text{dec}(g_1), \text{inc}(g_1), \dots, \text{dec}(g_{2M}), \text{inc}(g_{2M}), \text{noAction}\} \quad (18)$$

میزان افزایش و یا کاهش با پارامتر $step$ نشان داده شده و به اندازه d/N تعریف می‌گردد.

شکل ۶ حالت‌ها را برای ساده‌ترین حالت ممکن یعنی یک سیگنال با تک هارمونیک ($M=1$) و انتخاب $N=4$ نشان می‌دهد. در این حالت نمونه فرض گردیده که جواب نهایی بروی مقدار $\Gamma_1^* = g_1^* + jg_2^*$ قرار دارد.



شکل ۶ - نمایش حالت و سیاست بهینه برای یک مسئله نمونه

در روش پیشنهادی به هر گامی که منجر به کاهش توان خروجی شود پاداش +1 و هر گام که اثری معکوس در پی داشته باشد -1 اختصاص می‌یابد. در صوت عدم تغییر توان نیز عدد صفر به پاداش نسبت داده می‌شود که این موضوع در رابطه (۱۹) نشان داده شده است:

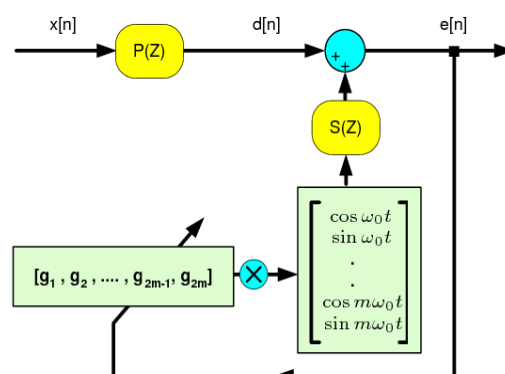
$$R[k+1] = \begin{cases} +1 & \overline{e^2}((k+1)T_a) < \overline{e^2}(kT_a) \\ 0 & \overline{e^2}((k+1)T_a) = \overline{e^2}(kT_a) \\ -1 & \overline{e^2}((k+1)T_a) > \overline{e^2}(kT_a) \end{cases} \quad (19)$$

یک مسئله یادگیری تقویتی با داشتن دینامیک مسئله به کمک روابط بلمن قابل حل می‌باشد. با توجه به مشخص

نهایتاً توان مشاهده‌شده مطابق رابطه زیر خواهد گردید:

$$\left. \begin{aligned} \Gamma_m &= g_{2m-1} + jg_{2m} \\ b_{2m-1} &= b_{2m} = |\beta_m|^2 \\ \alpha_m &= a_{2m-1} + ja_{2m} \end{aligned} \right\} \rightarrow \overline{e^2} = \sum_{m=1}^{2M} \frac{(g_m - a_m)^2}{b_m^2} \quad (15)$$

چون حل این مسئله بر اساس یادگیری تقویتی است بایستی متغیر حالت، مجموعه اعمال و پاداش تعریف گردد. هدف از طرح مسئله نیز یافتن بردار $\mathbf{g} = [g_1, \dots, g_{2M}]^T$ به نحوی که $\overline{e^2}$ را به سمت صفر ببرد، می‌باشد. Error! Reference source not found. پارامترهای $g_i (i = 1, \dots, 2M)$ در هنگام پیاده‌سازی عملی سیستم را نشان می‌دهد.



شکل ۵ - نقش پارامتر قابل تنظیم \mathbf{g} در سیستم حذف نویز

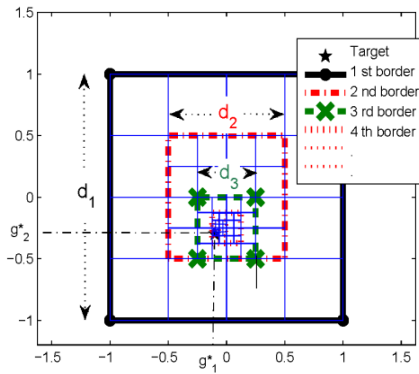
بنابراین با توجه به نقش کلیدی بردار \mathbf{g} ، مقدار رقومی شده آن بردار می‌تواند به عنوان متغیر حالت تعریف شود. برای این منظور، محدوده‌ای برای تغییر آرایه‌های \mathbf{g} به نحوی تعریف می‌گردد که جواب در آن ناحیه قرار گیرد. این موضوع در رابطه (۱۶) نشان داده شده است. در این رابطه پارامتر d بیانگر حداکثر بازه ممکن برای تغییر پارامترها می‌باشد که با اطلاعات اولیه‌ای که از مسئله مد نظر نزد طراح وجود دارد قابل تخمین زدن می‌باشد.

$$\forall m: 1 \leq m \leq M : a_{\min}^m \leq g_m \leq a_{\max}^m, d = a_{\max}^m - a_{\min}^m \quad (16)$$

همانند حالت قبل تابع $\text{Digi}(\cdot)$ برای رقومی نمودن تک پارامتر خود از مقدار حداقل و حداکثر آن به 0 تا $N-1$ و تابع

هر چند کاهش N به کمتر شدن حافظه مورد نیاز می‌انجامد اما باعث افزایش مقدار پارمتر دیگری یعنی $step = d / N$ خواهد. همان‌گونه که قبلاً هم بدان اشاره شد با افزایش مقدار $step$ دقت حذف نویز کاهش می‌یابد و بنابراین در این روش راهی برای رسیدن به هر دو مشخصه به صورت همزمان وجود ندارد. در ادامه روشی پیشنهاد می‌گردد که نقیصه فوق را مرتفع خواهد ساخت.

برای رفع نقیصه فوق روش تفاوت زمانی چند سطحی به شرح زیر و با نام $ML-TD-ANC^1$ پیشنهاد می‌گردد. در این روش ابتدا با انتخاب $d = d_1$ و $M = 1$: M محدودده‌ای که جواب در آن است، حدس زده می‌شود. این ناحیه با دقت کمی نظیر $N=4$ رقومی شده و با الگوریتم بیان‌شده در بخش قبل مسئله یادگیری با روش $TD(0)$ بر روی آن تعریف می‌گردد. با این انتخاب، با قدم‌های بزرگ با سرعت مناسب محدوده کوچک‌تری برای جواب یافته خواهد شد. بعد از اینکه با اطمینان زیادی زیر ناحیه‌ای از محدوده قبلی برای جواب تخمین زده شد، مسئله‌ای با بازه‌ای کوچک‌تر از حالت قبل تعریف شده ($d = d_2 : d_2 < d_1$) در این حالت گام پیشروی (d/N) کاهش یافته و در نتیجه دقت عمل افزایش خواهد یافت. همین رویه به هر تعداد که لازم باشد تکرار می‌گردد. این موضوع در شکل ۸ نشان داده شده است.

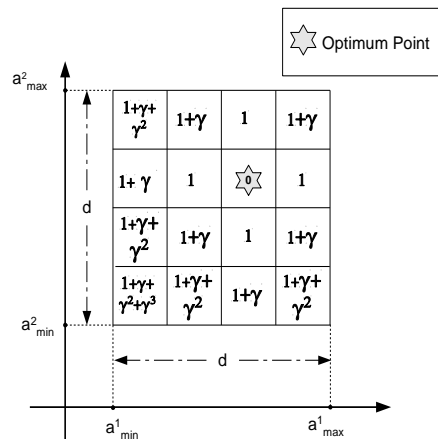


شکل ۸- روش تفاوت زمانی چند سطحی

استفاده از رویه فوق مزایای زیر را در پی خواهد داشت:

- ✓ برای رسیدن به دقت بالا نیازی به افزایش مقدار N نیست و به همین دلیل همزمان با امکان رسیدن

بودن کلیه مشخصات مسئله برای مسئله حالت قبل، قابل حل است و با فرض نرخ تنزیل γ در رابطه (۸) می‌توان $V(s)$ را برای حالت‌های مختلف را محاسبه نمود. شکل ۷ نمایش تصویری برای حالت نمونه فرض شده در شکل ۶ می‌باشد. این شکل نشان می‌دهد که با تعریف صورت گرفته حرکت صحیح در خلاف جهت مقدار $V^*(s)$ می‌باشد. بنابراین اگر سیاست، حرکت بر خلاف جهت مقدار $V(s)$ با الگوریتم $greedy - \epsilon$ تعریف گردد، با هر مرحله ارزیابی سیاست پس از به‌روز آوری تابع مقدار توسط رابطه (۸)، شروط GPI رعایت گردیده و در نتیجه سیاست نیز بهینه خواهید گردید و سیستم به سمت جواب صحیح میل خواهد نمود. البته باید توجه داشت که این نوع عملکرد مخصوص همین مسئله خاص و متأثر از دینامیک بحث شده در بخش قبل و نوع تعریف مسئله می‌باشد.



شکل ۷- محاسبه جدول $V(s)$ برای مسئله نمونه شکل قبل

از محسّنات این روش کاهش حافظه مورد لزوم می‌باشد. با تغییر از روش QL جدول $Q(s, a)$ به $V(s)$ تغییر یافته و در نتیجه نیاز به حافظه برای نگهداری آن به مرتبه تعداد اعمال مورد نیاز، یعنی $4M+1$ مرتبه، کاهش خواهد یافت. علیرغم کاهش ذکر شده هنوز N^{2M} حافظه برای نگهداری برای جدول $V(s)$ لازم است. این رابطه، رابطه‌ای نمایی است و در صورت بزرگ بودن N و مقدار بالای M می‌تواند مقادیر بزرگی اختیار کند. M تعداد هارمونیک موجود در محیط و توسط کاربر انتخاب نمی‌شود اما تعداد سطوح کوانتزه قابل تغییر است.

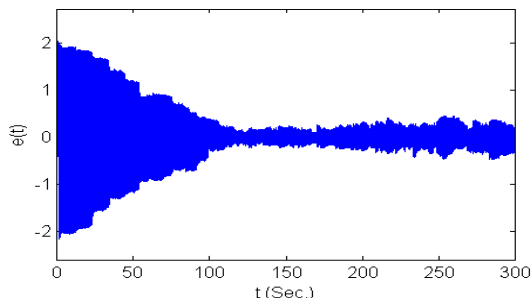
¹ multi-level temporal difference active noise control

$$y(n+1) = 0.9u(n-3) + 0.6u(n-3) + 0.1u(n-4) - 0.4u(n-5) - 0.1u(n-6) + 0.2u(n-7) + 0.1u(n-8) + 0.1u(n-9) + 0.001u(n-10)$$

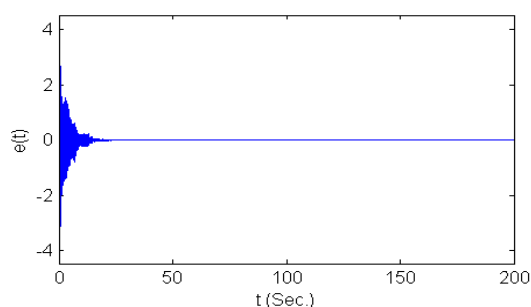
در مرحله اول برای مقایسه هر دو الگوریتم QL-AND و ML-TD-ANC شبیه‌سازی گردیده است. در روش اول $N=10$ ، $step = \frac{1}{2^4}$ و در روش دوم $N=4$ در نظر گرفته شد. مقدار بیش‌ترین و کمترین پارامترها مساوی ± 1 و در نتیجه $d = 2$ خواهد گردید. برای این کار $x(t)$ به صورت رابطه (۲۰) تعریف می‌گردد.

$$x(t) = \begin{bmatrix} 0.7 \\ 0.6 \\ 0.5 \end{bmatrix}^T \begin{bmatrix} \cos \omega_0 t \\ \cos 2\omega_0 t \\ \cos 3\omega_0 t \end{bmatrix} + \begin{bmatrix} 0.8 \\ -0.5 \\ 0.3 \end{bmatrix}^T \begin{bmatrix} \sin \omega_0 t \\ \sin 2\omega_0 t \\ \sin 3\omega_0 t \end{bmatrix} \quad (20)$$

که سیگنال $x[n]$ در شکل ۱ به صورت $x[n] = x(nT_s)$ و $\omega_0 = 2\pi 50$ در نظر گرفته شده است و عامل‌ها فرمان‌ها را با فرکانس 25Hz اعمال می‌دارند. در شکل ۹ مقدار نویز آکوستیک باقی‌مانده در محیط در دو روش برای این حالت مقایسه گردیده است.



الف) سیگنال خطا در QL-ANC



ب) سیگنال خطا در ML-TD-ANC

شکل ۹- مقایسه مقدار نویز آکوستیک باقی‌مانده در محیط در دو روش برای سیگنال با سه هارمونیک

به دقت بالا، حافظه مورد لزوم برای نگهداری جدول $V(s)$ در حد معقولی باقی می‌ماند. جدول ۱ مقایسه‌ای بین تعداد حافظه مورد نیاز در روش QL-ANC و ML-TD-ANC را نشان می‌دهد که در آن M تعداد هارمونیک N_1 و N_2 تعداد سطوح رقومی سازی در دو روش می‌باشد. با توجه به اینکه $N_2 \ll N_1$ در نظر گرفته می‌شود، در مقدار حافظه صرفه‌جویی شده مقدار قابل‌ملاحظه‌ای می‌گردد.

✓ قدم‌ها در ابتدا بزرگ و سپس با نزدیک‌تر شدن به محدوده هدف کوچک و کوچک‌تر می‌گردد. این گام برداشتن وقتی، همزمان با حفظ دقت، زمان همگرایی را نیز کاهش می‌دهد. در این تحقیق در هر مرحله گام نصف مرحله قبل می‌گردد.

جدول ۱- محاسبات حافظه مورد نیاز برای تابع $V(s)$ در

روش قبلی و روش پیشنهادی

نام روش	مقدار حافظه مورد نیاز
QL-ANC	$\ddagger N_1^{2 \times M} \times (4 \times M + 1)$
ML-TD-ANC	$\ddagger N_2^{2 \times M}$
	$\ddagger N_2 \ll N_1$

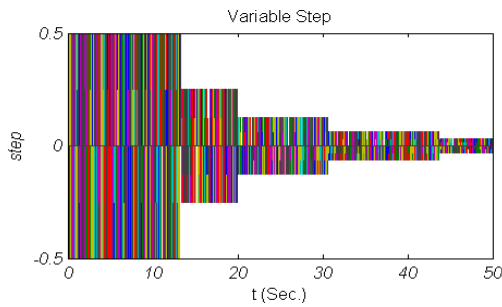
۶- نتایج شبیه‌سازی

با توجه به اینکه در این قسمت هدف مقایسه بهبود عملکرد نسبت به روش پیشنهادشده در [۵] می‌باشد لذا دینامیک محیط و شرایط دقیقاً به همان صورت اختیار می‌شود. البته لازم به ذکر است که این دینامیک یک محیط آزمایشگاهی می‌باشد که برخی دیگر از مقالات نیز از آن استفاده شده است [۶-۹، ۷].

در صورتی که فرکانس نمونه‌برداری مساوی $f_s = \frac{1}{T_s} = 1 \text{ KHz}$ در نظر گرفته شود، آنگاه تابع تبدیل مسیری پیشرو از منبع نویز تا ناحیه ساکت به صورت زیر می‌باشد:

$$d(n+1) = 0.8x(n-6) + 0.6x(n-7) - 0.2x(n-6) - 0.5x(n-9) - 0.1x(n-10) + 0.4x(n-11) - 0.05x(n-12)$$

و تابع تبدیل مسیری ثانویه، که از عملگر سیستم تا ناحیه ساکت است، به صورت زیر خواهد گردید:



شکل ۱۱- گام پیشروی متغیر در روش ML-TD-ANC

۷- نتیجه‌گیری

این تحقیق در تکمیل پژوهش قبل، که ایده‌ای مبتنی بر یادگیری تقویتی با روش کیو برای کاهش نویز صوتی به صورت فعال برای سیگنال‌های باند باریک بود (QL-ANC)، شکل گرفت. از مزایای هر دو روش که از این نوع یادگیری استفاده می‌نماید، حجم بسیار کم محاسبات مورد لزوم آن‌ها، عدم نیاز به دینامیک سیستم و در نتیجه مقاوم بودن آن‌ها نسبت به تغییرات می‌باشد.

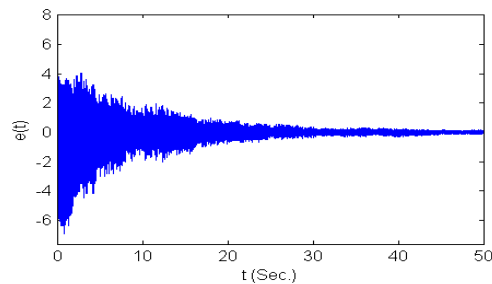
مشکل روش QL-ANC حافظه بسیار زیاد مورد لزوم و زمان زیاد همگرایی آن در صورت بالا رفتن تعداد هارمونیک‌های موجود در صوت مزاحم است. در روش جدید ML-TD-ANC در دو مرحله این مشکل حل گردید. در ابتدا با استفاده از بررسی این مسئله حذف نویز در حوزه فرکانس ثابت گردید که می‌توان به جای جدول $Q(s, a)$ از جدول $V(s)$ استفاده نمود. با این تغییر حافظه مورد نظر با مرتبه تعداد عمل‌ها کاهش یافت. نکته دیگر پیشنهاد روش چند سطحی است. در این راه حل ابتدا با دقت کم و دانه‌بندی^۱ درشت مسئله یادگیری حل می‌گردد. پس از اطمینان از رسیدن به محدوده جواب در مرحله بعد، همین مسئله اما این بار در بازه محدودتر اما دانه‌بندی ریزتری حل می‌گردد که این کار دقت را افزایش می‌دهد. تکرار این فرایند تا رسیدن به دقت مد نظر ادامه می‌یابد. به وسیله شبیه‌سازی توانمندی این روش در حل مسئله حذف نویز حتی تا ۷ هارمونیک مورد تحقیق قرار گرفت. علاوه بر کم شدن حافظه مورد نیاز، کم شدن زمان همگرایی و بهتر شدن میزان حذف نویز از محسّنات این روش می‌باشد.

همان‌گونه که در شکل دیده می‌شود الگوریتم ML-TD-ANC تفاوت بسیار معنی‌داری هم از نظر کم کردن توان نویز و هم از نظر زمان همگرایی پیدا نموده است. البته اگر به این مورد کاهش بسیار زیاد حافظه مورد نیاز در حالت قبلی به حالت فعلی نیز به محسّنات روش اضافه گردد آنگاه ارجح بودن روش جدید به صورت وضوح روشن می‌گردد.

در شبیه‌سازی بعدی در علامت ورودی تعداد هارمونیک‌ها افزایش یافته است که در رابطه (۲۱) نشان داده شده است. تأمین حافظه مورد لزوم برای این شبیه‌سازی با الگوریتم QL-ANC امکان‌پذیر نیست بنابراین تنها با روش ML-TD-ANC این کار انجام خواهد گردید.

$$x(t) = \begin{bmatrix} 0.8 \\ 0.5 \\ 0.6 \\ -0.3 \\ 0.9 \\ 0.8 \\ 0.4 \end{bmatrix}^T \begin{bmatrix} \cos \omega_0 t \\ \cos 2\omega_0 t \\ \cos 3\omega_0 t \\ \cos 4\omega_0 t \\ \cos 5\omega_0 t \\ \cos 6\omega_0 t \\ \cos 7\omega_0 t \end{bmatrix} + \begin{bmatrix} 0.9 \\ -0.5 \\ 0.3 \\ 0.6 \\ 0.9 \\ -0.3 \\ 0.6 \end{bmatrix}^T \begin{bmatrix} \sin \omega_0 t \\ \sin 2\omega_0 t \\ \sin 3\omega_0 t \\ \sin 4\omega_0 t \\ \sin 5\omega_0 t \\ \sin 6\omega_0 t \\ \sin 7\omega_0 t \end{bmatrix} \quad (21)$$

شکل ۱۰ مقدار نویز آکوستیک باقی‌مانده در محیط را برای این حالت نمایش می‌دهد که مقدار معقولی می‌باشد.



شکل ۱۰- مقدار نویز آکوستیک باقی‌مانده در محیط برای ۷ هارمونیک

در شکل ۱۱ نیز $step = d_i / N$ نشان داده می‌شود. در این شکل به وضوح نشان داده می‌شود که چگونه با نزدیک شدن به مقدار مطلوب اندازه گام کاهش می‌یابد.

^۱ grading

- مراجع
- [6] Zhou YL, Zhang QZ, Li XD, Gan WS (2005) Analysis and DSP implementation of an ANC system using a filtered-error neural network. *Journal of Sound and Vibration* 285: 1–25.
- [7] Qizhi Z, Yongle J (2002) Active noise hybrid feedforward/feedback control using neural network compensation. *Journal of vibration and acoustics* 124: 100-104.
- [8] Watkins CJCH (1989) Learning from delayed rewards. PhD Thesis.
- [9] Zhou YL, Zhang QZ, Li XD, Gan WS (2008) On the use of an SPSA-based model free feedback controller in active noise control for periodic disturbances in a duct. *Journal of Sound and Vibration* 317: 456–472.
- [1] Russell S, Norvig P (2009) *Artificial Intelligence: A Modern Approach*: Prentice Hall.
- [2] Sutton R, Barto A (1998) *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.
- [3] Kaelbling LP, Littman LM, Moore AW (1996) Reinforcement learning: A survey. *Journal of Artificial Intelligence Research* 4: 237–285.
- [4] Kuo SM, Morgan DR (1999) Active noise control: a tutorial review. in *IEEE Proceedings* 87: 943-973.
- [5] Raeisy B, Golbahar Haghghi S (2012) Active Noise Controller with Reinforcement Learning. presented at the The 16th CSI International Symposiums on Artificial Intelligence & Signal Processing (AISP 2012): 74-79, Shiraz, IRAN.