Archive of SID

Stress Detection and Control According to the Skin Signal of Electrical Resistance and Heart Rate Using Reinforcement Learning

Mandana Sadat Ghafourian[1*], Amin Noori[2], Minoochehr Bashizade Fakhar[2]

[1]Department of Biomedical Engineering, Khaje Nasir Toosi University of Technology, Tehran, Iran
[2]Faculty of Electrical Engineering and Biomedical Engineering, Sadjad University of Technology, Mashhad, Iran

Article Info:
Received: **24 Jan 2017** ***Accepted:*** **6 May 2017**

Downloaded from shefayekhatam.ir at 15:58 +0330 on Wednesday January 31st 2018

ABSTRACT

Introduction: The anxiety disorders are the most common mental health condition. Through modulation of stress responses, the anxiety can be controlled. Stress clearly affects the autonomic nervous system and leads to excessive sweating and tachycardia. New methods to identify and manage stress are still needed. ***Materials and Methods:*** In this paper, anxiety is controlled by reinforcement learning (RL). For stress level measurement, galvanic skin resistance (GSR) and heart rate were measured. Skin electrical resistance and heart rates are measured in 15 subjects (8 females and 7 males), ages between 22 to 30 years, during exposure to melodies and pictures. By connecting bioelectric sensors to the hand, GSR signal can be tested. Anxiety level was measured by biological signals (biofeedback) through playing melodies and showing pictures. First, GSR and heart rate signals were measured (states of RL). Then, the application plays a melody (agent) and shows a picture (actions). After a period of time, GSR and heart rate were measured again. Based on the differences between initial and secondary values, the reward, indicating stress variations were calculated. Based on the received reward, the agent learns to control the stress level. ***Results:*** The intelligent agent of RL, learns which picture to show and what melody to play in order to manage the stress. For each person, different melodies and pictures reduce the stress. This learning procedure should be repeated for each subject. ***Conclusion:*** A novel method is proposed and investigated based on RL for anxiety control. The accuracy of diagnosis is improved in a simple, rapid, and low-cost way.

Key words:

1. Anxiety
2. Skin
3. Heart Rate

Corresponding Author: Mandana Sadat Ghafourian
E-mail: *m.ghafourian@email.kntu.ac.ir*

www.SID.ir

Archive of SID

تشخیص و کنترل استرس برمبنای سیگنال‌های مقاومت الکتریکی پوست و ضربان قلب با استفاده از یادگیری تقویتی

ماندانا سادات غفوریان[1*]، امین نوری[2]، مینوچهر باشی زاده فخار[2]

[1]گروه مهندسی پزشکی، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران

[2]دانشکده مهندسی برق و مهندسی پزشکی، دانشگاه صنعتی سجاد، مشهد، ایران

اطلاعات مقاله:

تاریخ دریافت: ۵ بهمن ۱۳۹۵

تاریخ پذیرش: ۱۶ اردیبهشت ۱۳۹۶

چـــــکیده

مقدمه: اختلالات اضطرابی شایع‌ترین وضعیت سلامت روانی است. از طریق تنظیم پاسخ به استرس می‌توان اضطراب را کنترل کرد. استرس به طور واضح بر سیستم اعصاب خودمختار اثر می‌گذارد و منجر به تعریق شدید و افزایش ضربان قلب می‌شود. روش‌های نوینی جهت شناسایی و کنترل استرس هنوز مورد نیاز است. ***مواد و روش‌ها:*** در این مقاله اضطراب از طریق یادگیری تقویتی کنترل شده است. برای اندازه‌گیری سطح استرس، مقاومت الکتریکی پوست و ضربان قلب اندازه‌گیری شد. مقاومت الکتریکی پوست و ضربان قلب در ۱۵ نفر (۸ زن و ۷ مرد)، رنج سنی بین ۲۲ تا ۳۰ سال، در طی پخش آهنگ و عکس اندازه‌گیری شد. با وصل‌کردن سنسورهای الکتریکی به دست، سیگنال‌های مقاومت الکتریکی پوست را می‌توان آزمایش کرد. سطح اضطراب از طریق سیگنال‌های زیستی (بیوفیدبک) از طریق پخش آهنگ و نشان دادن تصاویر اندازه‌گیری شد. اول، مقاومت الکتریکی پوست و سیگنال‌های ضربان قلب اندازه‌گیری شد (حالت‌های یادگیری تقویتی). سپس پخش آهنگ (عامل) و نمایش عکس (اقدامات) استفاده شد. پس از یک دورهٔ زمانی، مقاومت الکتریکی پوست و ضربان قلب دوباره اندازه‌گیری شد. بر اساس تفاوت‌های میان مقادیر اولیه و ثانویه پاداش، تغییرات استرس محاسبه شد. بر اساس دریافت پاداش، عامل یاد گرفت سطح استرس را کنترل کند. ***یافته‌ها:*** عامل باهوش یادگیری تقویتی یاد گرفت چه عکسی برای نمایش و چه آهنگی برای پخش کردن به‌منظور مدیریت استرس مناسب است. برای هر شخصی آهنگ‌ها و عکس‌های متفاوت استرس را کاهش می‌دهد. این روش یادگیری باید برای هر موضوع دیگری تکرار شود. ***نتیجه‌گیری:*** یک روش جدید بر اساس یادگیری تقویتی برای کنترل اضطراب مطرح شد و مورد بررسی قرار گرفت. دقت تشخیص در یک روش ساده، سریع و کم هزینه بهبود یافت.

کلید واژه‌ها:

۱. اضطراب
۲. پوست
۳. ضربان قلب

* **نویسنده مسئول:** ماندانا سادات غفوریان

آدرس الکترونیکی: *m.ghafourian@email.kntu.ac.ir*

Downloaded from shefayekhatam.ir at 15:58 +0330 on Wednesday January 31st 2018

www.SID.ir

Archive of SID

Downloaded from shefayekhatam.ir at 15:58 +0330 on Wednesday January 31st 2018

مقدمه

شمار زیادی از مراکز درمانی کشورهای جهان، برای برطرف کردن اختلالات استرس به سوی بیوفیدبک روی آوردند و این مراکز همواره به دنبال روش‌های کارآمد برای تشخیص استرس و کنترل آن به شیوه‌ای بهینه هستند. از دیرباز برای تشخیص استرس و برطرف کردن آن شیوه‌های گوناگونی را در نظر می‌گرفتند، از جملهٔ این شیوه‌ها این است که با ثبت گرفتن از سیگنال‌های مغزی و استخراج ویژگی‌ها و تحلیل این سیگنال‌ها به میزان استرس فرد پی برده می‌شود (۱). همچنین در زمینهٔ پردازش تصویر با ثبت گرفتن از قطر مردمک چشم[1] و ارتباط قطر مردمک چشم با میزان استرس (۲) و همچنین حرکات پرشی چشم[2] (۳) می‌توان به استرس فرد پی برد. تصویربرداری حرارتی[3] یکی دیگر از روش‌های پی بردن به استرس است که در این روش از چهرهٔ فرد طی مراحل مختلفی با دوربین‌های مادون قرمز حساس، تصویر گرفته می‌شود و اگر به سوژه در طی تصویربرداری استرسی وارد شود، به علت افزایش میزان جریان خون اطراف چشم به صورت نقاط گرم ظاهر می‌شود، در صورتی که در افراد آرام این نقاط به صورت سرد نمایان می‌شوند (۴).

تمامی مواردی که گفته شد نیازمند تخصص برای تشخیص استرس است، به‌عنوان مثال در ثبت سیگنال‌های مغزی سوژه، باید بدانیم هر حالت از امواج مغزی مربوط به چه وضعیتی است، مثلاً فرد در خواب، در حالت استرس و یا ... است. اولین بار در سال ۱۸۹۵ توسط لومبروسو از فشارخون و نرخ ضربان قلب برای تشخیص استرس استفاده شده و بعدها در سال ۱۹۲۱ جان لارسون، نرخ تنفس را به روند اضافه کرد و در سال ۱۹۳۹ هدایت الکتریکی پوست (GSR)[4] (۵، ۶) نیز توسط لئوناردکیلر اضافه شد و امروزه سیگنال‌های نرخ ضربان قلب، هدایت الکتریکی پوست، حجم خون و نرخ تنفس قابل اندازه‌گیری هستند (۷). در تمامی این روش‌ها، ثبت آسان و هزینهٔ کم حائز اهمیت است. بر این اساس ثبت سیگنال هدایت الکتریکی پوست و نرخ ضربان قلب بهترین گزینه در هنگام استرس است، زیرا در هنگام استرس ضربان قلب بالا رفته (۸) و فرد شروع به تعریق می‌کند (۹). در این تحقیق بعد از ثبت سیگنال هدایت الکتریکی پوست به وسیلهٔ ساخت دستگاه GSR و پی بردن به میزان سطح استرس، آرامش یا حالت فعال فرد، هدفمان این است که به کمک یادگیری تقویتی[5] و بیوفیدبک سطح استرس فرد را کنترل کنیم در این راستا در محیط نرم‌افزار متلب[6]، بر اساس راهبردی که توسط یادگیری تقویتی مشخص می‌شود عکس و آهنگ پخش می‌کنیم و این کار به کمک یادگیری تقویتی بارها و بارها انجام می‌شود تا فرد به حالت آرامش و حالت مطلوب برسد. در طی این عملیات فرد نیز باید تلاش کند که با تمرکز بر روی عکس و آهنگ پخش شده، میزان استرس خود را کاهش دهد.

مواد و روش‌ها

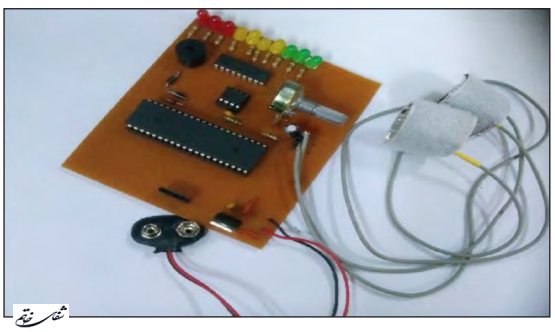
۱- ساخت دستگاه برای ثبت سیگنال فیزیولوژیکی

همانطور که گفته شد در این تحقیق برای تشخیص میزان استرس از دو پارامتر هدایت الکتریکی پوست و تعداد ضربان قلب جهت کاهش خطا در ثبت استرس استفاده کردیم که ثبت هر کدام از این سیگنال‌ها نیازمند مدار خاصی می‌باشد. بنابراین برای هر کدام از پارامترهای ذکر شده، مدار مربوط به آن را ساختیم. این مدارات بسیار ساده و کم هزینه می‌باشند که این نکته حائز اهمیت است. توجه داشته باشیم که دمای محیط نیز بر روی ثبت، اثرگذار بوده و ثبت باید در دمای اتاق (۲۵ درجهٔ سانتی‌گراد) صورت گیرد و بررسی این تحقیقات بر روی ۱۵ نفر، شامل ۸ زن و ۷ مرد در رنج سنی ۲۲ تا ۳۰ سال است که به صورت اتفاقی انجام شده است. در این تحقیق افرادی که نسبت به بقیه عصبی‌تر و یا دارای استرس بیشتری بودند حذف نشدند و افراد به صورت داوطلبانه از میان دانشجویان دانشگاه سجاد انتخاب شدند ولی این افراد دچار عارضه یا بیماری روانی نبودند.

۱-۱ مدار تشخیص استرس

مدار تصویر ۱ سطح استرس را با توجه به میزان هدایت الکتریکی پوست (انگشتان دست (۹)) تشخیص می‌دهد و داده‌ها را بر حسب ولتاژ توسط ارتباط سریال به کامپیوتر ارسال می‌کند، علاوه بر این می‌توانند این سطح استرس را توسط دیودهای نوری که بر روی آن تعبیه شده‌اند، نمایش دهد.

۱-۲ مدار ضربان سنج



تصویر ۱- تصویر گرفته شده از مدار تشخیص استرس در طی پژوهش.

[1] Pupilometry
[2] Saccadic eye movement
[3] Thermal imaging
[4] Galvanic skin resistant
[5] Reinforcement learning
[6] Matlab software

www.SID.ir

Archive of SID

Downloaded from shefayekhatam.ir at 15:58 +0330 on Wednesday January 31st 2018

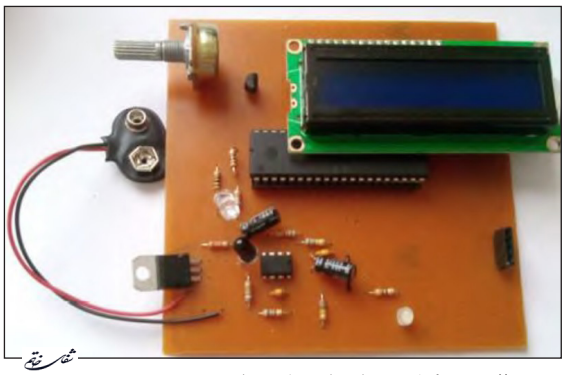
مدار ضربان سنج تصویر ۲، داری دو سنسور گیرنده[7] و فرستنده[8] مادون قرمز[9] می‌باشد که با قرار گرفتن انگشت بین آن‌ها، به ازای هر ضربان دیودهای نوری تعبیه شده بر روی آن چشمک می‌زند.

خروجی‌های دو مدار فوق شامل: ۱- سطح استرس فرد، ۲- میزان ضربان قلب از طریق ارتباط سریال[10] به نرم‌افزار متلب فرستاده می‌شود و به‌عنوان حالت‌ها در مسئلهٔ یادگیری تقویتی مورد استفاده قرار می‌گیرد.

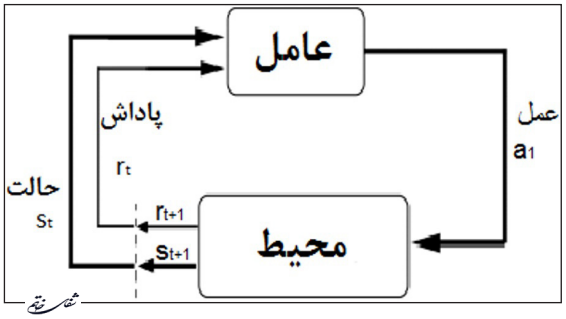
۲- یادگیری تقویتی

یادگیری تقویتی مسئله‌ای است که در آن، عامل می‌بایست رفتار خود را از طریق سعی و خطا در محیط پویای اطرافش بیاموزد. در واقع یادگیری تقویتی، یادگیری آنچه باید انجام بگیرد است، اینکه چگونه موقعیت‌ها را به عمل‌ها[11] نگاشت کنیم تا بتوانیم پاداش[12] عددی را بیشینه کنیم (۱۱، ۱۰). برای رسیدن به این مقصود روش‌های سعی و خطا[13] همراه با الگوریتم‌های یادگیری متنوعی وجود دارد که یکی از این الگوریتم‌ها، الگوریتم Q-Learning است که در سال ۱۹۸۹ توسط واتکینز ارائه گردیده است. مسئلهٔ یادگیری تقویتی در ساده‌ترین شکل به صورت تصویر ۳ بیان می‌شود (۱۳، ۱۲، ۱۰).

فرض کنید حالت‌ها و عمل‌ها به صورت زیر تعریف شوند:



تصویر ۲- تصویر گرفته شده از مدار ضربان در طی پژوهش.



تصویر ۳- ساختار کلی مسئلهٔ یادگیری تقویتی.

[13] Try and error
[14] State
[15] Voltage level
[16] Excitation
[17] Activity

$$S = \{s1 , s2 , , , , , sn\} \qquad (1)$$

$$A = \{a1 , a2 , , , , , am\} \qquad (2)$$

در هر تکرار با اعمال عمل a_t در حالت S_t حالت محیط تغییر می‌کند و به حالت S_t+1 می‌رود این حالت پاداش لحظه‌ای r_t را دریافت می‌نماید که مقدار این پاداش عددی بسته به هدف تعیین شده در مسئله متفاوت می‌باشد. مسئلهٔ مهم این می‌باشد که به دنبال ماکزیمم کردن این پاداش‌های دریافتی در درازمدت R_t می‌باشیم.

در یادگیری تقویتی پنج پارامتر مهم وجود دارد که در این مقاله به صورت زیر تعریف می‌شوند.

۲-۱ حالت یا وضعیت

در اینجا حالت یا وضعیت[14] همان سطح ولتاژ[15] خروجی مدار استرس‌سنج است. این سطح ولتاژ متناسب با شرایط افراد، متفاوت است. همچنین این ولتاژ با سطح استرس و هدایت الکتریکی پوست فرد رابطهٔ عکس دارد. در واقع با افزایش استرس ولتاژ خروجی کاهش می‌یابد، زیرا رطوبت دست فرد زیاد شده و مقاومت الکتریکی پوست کم می‌شود و در نتیجه در حضور جریان الکتریکی بسیار ناچیز، ولتاژ کمی را داراست و با کاهش استرس، ولتاژ خروجی افزایش می‌یابد (۷). در مسئلهٔ یادگیری تقویتی، محیط می‌بایست به صورت گسسته باشد بدین منظور حالت‌ها را به صورت گسسته درآورده و آن را بر اساس وضعیت استرس فرد به سه سطح تقسیم کردیم که به توضیح آن می‌پردازیم. سطح ولتاژ اول مربوط به زمانی است که فرد کاملاً آرام است و دیودهای نوری سبز روشن می‌شوند، در این حالت ولتاژ مدار مقادیر دو ولت و بیشتر از آن را شامل می‌شود، به‌منظور گسسته‌سازی، ولتاژهای دو و بالاتر از دو را به‌عنوان یک حالت در مسئلهٔ یادگیری تقویتی در نظر می‌گیریم. سطح ولتاژ دوم زمانی است که فرد کمی هیجان[16] دارد و در حالت فعالیت[17] است، مثلاً ذهنی درگیر دارد که بین حالت استرس‌دار و بدون استرس قرار دارد. در این حالت دیودهای نوری زرد روشن می‌شوند. سطح ولتاژ در این حالت بین یک تا دو ولت متغییر است که این سطوح ولتاژ را به‌عنوان حالتی دیگر در مسئلهٔ یادگیری تقویتی در نظر می‌گیریم. سطح ولتاژ آخر مربوط به زمانی است که فرد استرس و یا هیجان زیادی دارد. در این حالت دیودهای نوری قرمز روشن می‌شوند. سطح ولتاژ در این حالت مقادیر یک ولت و کمتر از یک ولت را شامل می‌شود که

[7] Receiver
[8] Transmitter
[9] Infrared
[10] Serial communication
[11] Action
[12] Reward

www.SID.ir

Archive of SID

Downloaded from shefayekhatam.ir at 15:58 +0330 on Wednesday January 31st 2018

این سطوح ولتاژ را نیز به‌عنوان حالتی دیگر در مسئلهٔ یادگیری تقویتی در نظر می‌گیریم. پس در کل سه حالت در اجرای این پروژه داریم که به صورت خلاصه در جدول ۱ آورده شده است.

جدول ۱- تعریف حالت‌ها بر اساس شدت ولتاژ دریافتی از مدار استرس سنج.

سطح ولتاژ	$2 \leq v$	$1 < v < 2$	$v \leq 1$
حالت	حالت ۱	حالت ۲	حالت ۳

۲-۲ عمل

در مقابل این حالت‌هایی که در یادگیری تقویتی داریم، بحث عمل مطرح میشود. به ازای هر حالتی که داریم، عامل یادگیر در محیط پویا[18] حرکت می‌کند و با سعی و خطا کردن در محیط به دنبال بهترین عمل برای حالت مورد نظر می‌گردد تا بیشترین پاداش را کسب کند (۱۱). در این پروژه عمل‌ها یک فضای دوبعدی از تصویر و آهنگ می‌باشد (۱۴، ۱۲). به این صورت که سه تصویر، شامل دو تصویر آرام‌بخش از دریا و منظره و یک تصویر ترسناک و استرس‌زا، همچنین سه آهنگ بی‌کلام، شامل آهنگ آرام‌بخش، معمولی و یک آهنگ ترسناک انتخاب شده است. سپس از همهٔ تصاویر با همهٔ آهنگ‌ها، ۹ عمل به صورت فضایی دوبعدی از یک تصویر و یک آهنگ، ایجاد کردیم. در این حالت می‌توان ترکیبی از آهنگ ترسناک و تصویر ترسناک، آهنگ آرام‌بخش و تصویر آرام‌بخش و همچنین آهنگ ترسناک و تصویر آرام‌بخش و بالعکس داشت. پس ما در کل ۹ عمل در اجرای این پروژه داریم. عمل اول مربوط به تصویر منظره به همراه آهنگ ترسناک است، عمل دوم مربوط به تصویر دریا و آهنگ آرام‌بخش است، عمل سوم تصویر ترسناک و آهنگ معمولی، عمل چهارم تصویر دریا و آهنگ ترسناک، عمل پنجم تصویر ترسناک و آهنگ ترسناک، عمل ششم تصویر منظره و آهنگ آرام‌بخش، عمل هفتم تصویر ترسناک و آهنگ آرام، عمل هشتم تصویر منظره و آهنگ معمولی و آخرین عمل مربوط به تصویر دریا و آهنگ معمولی است.

۲-۳ پاداش

همانطور که در بخش سوم گفته شده مهم‌ترین موضوع در یادگیری تقویتی محاسبهٔ پاداش است و در انتها مقایسه و انتخاب بهترین پاداش برای رسیدن به بهترین عملکرد است. برای محاسبهٔ پاداش در این پروژه ابتدا حالت فرد توسط مدار تشخیص استرس اندازه‌گیری می‌شود و متناسب با حالت فرد، عملی برای آن حالت در نظر گرفته می‌شود و دوباره حالت فرد بررسی می‌شود. تفاضل حالت دوم که بعد از اجرای عمل است با حالت قبل از اجرای عمل، پاداش ما را می‌سازد.

[20] Bellman
[21] Learning rate

۲-۴ سیاست و سیاست بهینه

منظور از سیاست[19] این می باشد که در هر حالت چه عملی انتخاب می‌شود. مهم‌ترین مسئله در یادگیری تقویتی محاسبهٔ سیاست بهینه می‌باشد بدین معنی که در هر حالت می بایست چه عملی را انجام دهیم تا بیشترین پاداش ممکن را در درازمدت دریافت کنیم. فرض کنید π یک سیاست می‌باشد و احتمال انتخاب عمل a در حالت s در تکرار t ام باشد آنگاه $v^{\pi}(s)$ ارزش حالت s تحت سیاست π می‌باشد که به صورت زیر تعریف می‌شود (۱۳، ۱۰):

$$V^{\pi}(s) = E_{\pi}\{R_t | s_t = s\} = E_{\pi}\{\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t = s\} \quad (3)$$

که در فرمول ۳، نرخ فراموشی می‌باشد و ما بین صفر و یک قرار دارد. به صورت مشابه نشان‌دهندهٔ ارزش یک زوج حالت -عمل تحت سیاست π می باشد که به صورت زیر تعریف می‌شود (۱۳، ۱۰):

$$Q^{\pi}(s, a) = E_{\pi}\{R_t | s_t = s, a_t = a\} = E_{\pi}\{\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t = s, a_t = a\} \quad (4)$$

سیاست بهینه نیز بر اساس فرمول بلمن[20] (۱۳، ۱۰) به صورت زیر برای حالت و زوج حالت -عمل انتخاب می‌شود:

$$V^*(s) = \max_{\pi} V^{\pi}(s) = \max_{a \in A(s)} \sum_{s'} P^a_{ss'} [R^a_{ss'} + \gamma V^*(s')] \quad (5)$$

$$Q^*(s, a) = \max_{\pi} Q^{\pi}(s, a) = \max_{a \in A(s)} \sum_{s'} P^a_{ss'} \left[R^a_{ss'} + \gamma \max_{a'} Q^*(s', a')\right] \quad (6)$$

که در فرمول‌های فوق نشان‌دهندهٔ احتمال انتقال از حالت s به حالت s' با انتخاب عمل a می‌باشد.

برای حل معادلهٔ بلمن روش‌های متفاوتی داریم که یکی از این روش‌ها یادگیری Q می‌باشد که از زیر شاخه‌های یادگیری تفاوت گذرا می‌باشد و به صورت زیر تعریف می‌شود (۱۲، ۱۰):

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha \left[r_{t+1} + \gamma Q \max_{a}(s_{t+1}, a) - Q(s_t, a_t)\right] \quad (7)$$

که در این فرمول فوق α نرخ یادگیری[21] می‌باشد و مقداری در بازهٔ (۰ و ۱) دارد.

[18] Dynamic environment
[19] Policy

www.SID.ir

Archive of SID

Downloaded from shefayekhatam.ir at 15:58 +0330 on Wednesday January 31st 2018

۵-۲ جدول ارزش Q

این جدول توسط پاسخ فرد در هر حالت، به عمل (تصویر و آهنگ پخش شده در آن حالت) ایجاد می‌شود. به این صورت که سطرهای این جدول، حالت‌ها و ستون‌های جدول، عمل‌ها می‌باشد. اما نکتهٔ مهم در اجرای این پروژه این است که هر فرد، جدول ارزش خاص خود را دارد، زیرا افراد مختلف در برابر عکس و آهنگ‌های یکسان واکنش‌های مختلفی از خودشان نشان می‌دهند. جدول ۲ نشان‌دهندهٔ جدول Q برای مسئلهٔ مورد نظر می‌باشد.

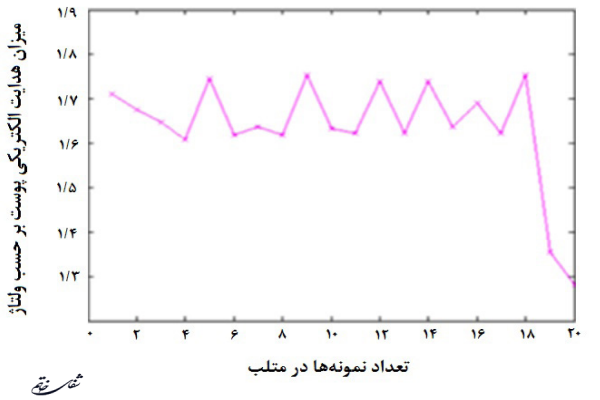
جدول ۲- جدول Q.

Q (s, a)	عمل ۱	عمل ۲	عمل ۹
آرامش کامل	$Q(s_1 , a_1)$	$Q(s_1 , a_2)$	$Q(s_1 , a_9)$
طبیعی	$Q(s_2 , a_1)$	$Q(s_2 , a_2)$	$Q(s_2 , a_9)$
استرس	$Q(s_3 , a_1)$	$Q(s_3 , a_2)$	$Q(s_3 , a_9)$

یافته‌ها

۳- الگوریتم Q-Learning

در ابتدای برنامه قبل از هر چیز، اول ماتریس ارزش Q را به صورت یک ماتریس صفر تعریف کردیم که این درایه‌های[22] صفر در طی اجرای برنامه آپدیت و غیر صفر می‌شوند. همچنین دستور بازکردن پورت سریال[23] جهت دریافت داده را در همین ابتدا قرار می‌دهیم، در غیر این صورت ممکن است در حین اجرای برنامه به مشکل عدم دریافت داده برخورد کنیم. سپس تصاویر و آهنگ‌های مورد نظر را در متلب تعریف می‌کنیم. برای تعریف هر عمل و قرار دادن تصویر و آهنگ مورد نظر در کنار هم باید تغییراتی بر روی ماتریس‌های آن‌ها انجام دهیم و آن‌ها را به صورت‌برداری در آوریم. باید توجه داشت که چون سایز این بردارها یکسان نیست، نمی‌توان آن‌ها را در یک ماتریس قرار داد، بنابراین برای قرار دادن آن‌ها در کنار هم باید از سلول[24] یا آرایه استفاده کرد. پس از تعریف عمل‌های مورد نظر، به قسمت اصلی برنامه که یادگیری می‌باشد، می‌رسیم. در این قسمت ابتدا داده که همان سطح ولتاژ استرس می‌باشد، توسط ارتباط سریال دریافت می‌شود و پس از مطابقت با بازه‌های تعریف شده، وضعیت فرد (state 1)، طبق حالت‌های گسسته که تعریف کردیم، مشخص می‌شود. جهت نمایش سطح ولتاژ متغییر فرد متناسب با وضعیت استرس او بر روی نمودار، می‌توان حلقه‌ای در برنامه قرار داد که پس از دریافت چند داده به صورت متوالی، این نمودار را نمایش دهد. به طور مثال همانطور که در نمودار ۱ مشاهده می‌شود در این ثبت با گذشت زمان ولتاژ کاهش یافته است که نشان‌دهندهٔ افزایش استرس می‌باشد. در این نمودار محور عمودی ولتاژ و محور افقی تعداد ثبت گرفته شده می‌باشد.



نمودار ۱- نمونه‌ای از ثبت میزان هدایت الکتریکی پوست بر حسب ولتاژ به تعداد نمونه‌ها در متلب.

حال پس از دریافت وضعیت استرس فرد و تعیین حالت مربوطه با توجه به مجموعه حالت‌های تعریف شده، نوبت به انتخاب عمل می‌رسد. یکی از مهم‌ترین موارد در مسئلهٔ یادگیری تقویتی ایجاد تعامل میان بهره‌برداری از اطلاعات یاد گرفته شده و جستجو در محیط می‌باشد. روش‌های متنوعی برای ایجاد این تعامل مطرح گردیده است (۱۲، ۱۱) که از جمله مهم‌ترین آن‌ها می‌توان به روش‌های ε-حریصانه[25] و سافت‌مکس[26] اشاره نمود. در این مقاله جهت ایجاد تعادل ما بین بهره‌برداری از اطلاعات یاد گرفته شده و جستجو در محیط از روش ε-حریصانه استفاده شده است، که ε عددی در بازهٔ (۰ و ۱) می‌باشد. در این روش در هر حالت، بهترین عمل با احتمال ε -1 و بقیهٔ اعمال با احتمال ε انتخاب می‌شوند. در اینجا منظور از بهترین عمل، در واقع عملی است که در سطر مربوط به حالت فعلی در جدول ارزش، بیشترین مقدار و بیشترین ارزش را دارد. جدول ۳ سیاست ε-حریصانه را برای حالت i ام نمایش می‌دهد. همانطور که در جدول ۳ نشان داده شده است در این حالت، عامل امکان انتخاب اعمال ۱ تا n را دارد ولی بهترین عمل که ارزش آن با Q* نشان داده شده است را با احتمال ε -1 انتخاب می‌کند و هر کدام از عمل‌های دیگر با احتمال ε/(n-1) انتخاب می‌شود که n نشان‌دهندهٔ تعداد اعمال ممکن قابل انجام برای حالت i ام می‌باشد.

جدول ۳- تشریح سیاست ε-حریصانه با توجه به ارزش‌ها.

	عمل ۱	عمل ۲	عمل j	عمل n
حالت i ام	$Q(s_i , a_1)$	$Q(s_i , a_2)$	$Q^*(s_i , a_j)$	$Q(s_i , a_n)$
احتمال انتخاب عمل	$\frac{\varepsilon}{n-1}$	$\frac{\varepsilon}{n-1}$	$1-\varepsilon$	$\frac{\varepsilon}{n-1}$

[22] Elements
[23] Serial port
[24] Cell
[25] ε-greedy
[26] Softmax

www.SID.ir

Archive of SID

Downloaded from shefayekhatam.ir at 15:58 +0330 on Wednesday January 31st 2018

پس از انتخاب شدن عمل، تصویر و آهنگ مربوط به آن عمل برای فرد پخش می‌شود و پس از اتمام پخش، مجدداً سطح ولتاژ استرس فرد توسط ارتباط سریال دریافت و وضعیت فرد (state 2) مشخص می‌گردد. حال همانطور که گفته شد از تفاضل وضعیت فعلی و وضعیت قبلی فرد، پاداش به دست می‌آید (فرمول ۸).

$$reward = v(2) - v(1) \qquad (8)$$

در فرمول فوق v(2) v(1) به ترتیب ولتاژ به دست آمده از مدار استرس‌سنج در حالت کنونی و ولتاژ به دست آمده پس از اعمال عمل مورد نظر می‌باشد. حال پس از به دست آمدن پاداش، جدول ارزش طبق فرمول (۷) در هر تکرار به صورت آنلاین، آپدیت می‌شود. تا همگرایی[۲۷] جدول Q، تکرارها انجام می‌شود. برای بهبود یادگیری، ε را در تکرارهای اول بزرگ انتخاب می‌کنیم و سپس به تدریج ε را کوچک می‌کنیم با این کار در گام‌های ابتدایی یادگیری، بیشتر بر روی جستجو در محیط تمرکز می‌کنیم و سپس از میزان جستجو در محیط کاسته و بر روی افزایش بهره‌برداری از اطلاعات یادگیری شده تمرکز می‌کنیم. به بیان دیگر در گام‌های ابتدایی احتمال انتخاب تمامی اعمال برای هر حالت یکسان می‌باشد تا برای هر شخص با توجه به حالتی که دارد (میزان استرس فرد –سطح ولتاژ و ضربان قلب) ابتدا امکان تست تمامی عمل‌ها وجود داشته باشد و بتوان برای شخص مورد نظر به تدریج عمل مناسب‌تر (تصویر و موسیقی که باعث آرامش بیشتر می‌شود) را تعیین کرد سپس در گام‌های بعدی برای ایجاد ثبات در آرامش فرد، بیشتر بر روی عملی که ارزش بیشتری دارد (آرامش بیشتری در فرد مورد نظر ایجاد می‌کند) تمرکز می‌کنیم. در واقع آرام‌بخش‌ترین عمل برای هر فرد انتخاب می‌شود و تمرکز برنامه بر روی پخش بیشتر این عمل جهت کاهش استرس فرد افزایش می‌یابد.

الگوریتم فوق توسط عامل برای یادگیری از طریق تجربیات یا آموزش استفاده می‌شود، هر تکرار معادل با یک دورهٔ آموزش استفاده می‌شود. در هر دورهٔ آموزش، عامل، محیط را که توسط ماتریس پاداش R نمایش داده می‌شود، کاوش می‌کند و تا زمانی که به حالت هدف نرسیده است، امتیازهای لازم داده می‌شود. هدف از آموزش، ساخت مغز عامل است که توسط ماتریس Q نمایش داده می‌شود. آموزش بیشتر منجر به ایجاد ماتریس Q بهتری خواهد شد، که می‌تواند توسط عامل برای حرکت در مسیر بهینه استفاده شود. بدین ترتیب با داشتن ماتریس Q، عامل می‌تواند در عوض کاوش و جلو و عقب‌های متعدد، با رجوع به ماتریس حالات و انتخاب گزینهٔ ماکزیمم، بهترین حالت را انتخاب کند. در جدول ۴ شبه کد مربوط به یادگیری تقویتی در این مسئله نشان داده شده است.

جدول ۴– شبه کد مربوط به یادگیری Q در کنترل استرس فرد.

ردیف	شبه کد	توضیحات
۱	تنظیم شرایط اولیهٔ یادگیری	➢ مقداردهی اولیهٔ جدول Q به صورت اختیاری و یا با توجه به نظرات افراد خبره ➢ $\alpha=0.1, \gamma=0.5, \varepsilon=0.5$
۲	تا زمانی که: تکرار < ماکزیمم تکرار	➢ مشاهده و تعیین حالت کنونی s_t (سطح ولتاژ دریافتی و یا به بیان دیگر میزان استرس فرد) ➢ انتخاب عمل a_t با توجه به سیاست ε-حریصانه ➢ با اعمال عمل a_t به فرد، حالت بعدی s_{t+1} را مشاهده می‌کنیم. ➢ پاداش آنی r_{t+1} را دریافت می‌کنیم. ➢ به‌روزرسانی جدول Q با استفاده از $Q(s_t, a_t) = Q(s_t, a_t) + \alpha\left[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)\right]$ ➢ حالت s_{t+1} را برابر حالت s_t قرار می‌دهیم.
۳	انجام عمل حریصانه	بهترین عمل جهت کنترل s با توجه به حالت استرس فرد انتخاب می‌شود.

جدول ۵– نمونه جدول Q به دست آمده از ارزش و عمل در این تحقیق.

	عمل ۱	عمل ۲	عمل ۳	عمل ۴	عمل ۵	عمل ۶	عمل ۷	عمل ۸	عمل ۹
آرامش کامل	-۰/۶۷۱۱	۰/۸۰۳۹	-۰/۳۴۵۰	-۰/۰۱۱۲	-۰/۹۹۸۹	۱/۰۰۳۷	-۰/۰۴۰۵	۰/۳۱۲۳	۰/۳۲۹۶
طبیعی	-۰/۰۵۳۴	۱/۲۰۷۸	-۰/۴۹۳۸	۰/۰۷۸۲	-۱/۳۱۰۶	۱۰/۰۱۹۰	۰/۰۷۷۶	۰/۱۹۸۵	۰/۱۱۲۳
استرس	۰/۰۰۴۸	۰/۸۷۵۰	-۰/۰۷۸۳	۰/۰۰۱۷	-۰/۸۳۷	۱/۵۶۲۳	۰/۵۵۹۸	۰/۵۳۲۴	۰/۶۹۷۷

[27] Convergence

www.SID.ir

Archive of SID

Downloaded from shefayekhatam.ir at 15:58 +0330 on Wednesday January 31st 2018

۴- تحلیل جدول ارزش

با توجه به ثبتی که گرفته شد جدول ۵ پس از ۱۰۰۰ بار تکرار و به‌روز رسانی به دست آمد. یعنی ۱۰۰۰ بار عکس و آهنگ پخش شد و این جدول ایجاد شد. در این مسئله در هر اپیزود ۱۰ تکرار داشتیم بدین معنا که پس از ۱۰ بار پخش موسیقی و تصاویر فرایند را به دلیل خسته‌کننده بودن برای فرد مورد نظر متوقف می‌کردیم و در زمان دیگر اپیزود بعدی را برای همان فرد تکرار می‌کردیم. حال می‌خواهیم از روی این جدول تشخیص بدهیم که برای فرد مورد نظر در هر حالت چه عملی باعث ایجاد آرامش و یا اضطراب می‌شود. زمانی که فرد در حالت سوم یعنی درحال استرس بوده است عمل ششم بیشترین آرامش را برای شخص ایجاد کرده است و سپس اعمال ۲ و ۹ می‌باشد. در حالتی که فرد در حالت طبیعی قرار داشته است به ترتیب اعمال ۲، ۶ و ۸ بیشترین آرامش را در فرد ایجاد کرده‌اند. در حالتی که فرد در حالت آرامش کامل قرار داشته است به ترتیب اعمال ۶، ۲ و ۹ بیشترین آرامش را در فرد ایجاد کرده‌اند. از روی جدول ۳ همچنین می‌توان فهمید که در هر حالت چه اعمالی باعث ایجاد استرس در فرد می‌شود. با توجه به جدول و اعمال تعریف شده مشخص می‌باشد که تأثیر موسیقی در مجموع بیش از تصویر در ایجاد آرامش و یا استرس در فرد مورد آزمایش می‌باشد. در حالت کلی می‌توان به این نتیجه رسید که تصویر و آهنگ خشن، سطح استرس فرد را بالا می‌برد، یعنی ضربان قلب فرد افزایش یافته و در نتیجه میزان تعریق پوست افزایش و مقاومت و سطح ولتاژ کاهش می‌یابد و با همین روال اگر تصویر و آهنگ آرامش‌بخش برای فرد گذاشته شود سطح استرس او کاهش می‌یابد. میزان تغییر استرس را علاوه بر ولتاژ می‌توان توسط دیودهای نوری که روی دستگاه تعبیه شده، مشاهده کرد.

بحث و نتیجه‌گیری

در این تحقیق با استفاده از یادگیری تقویتی، جدول Q ای شامل ارزش‌های حالت -عمل را آموزش دادیم که می‌توان از آن برای ایجاد آرامش در فرد استفاده کرد. با توجه به جدول Q که پس از آزمایش به دست می‌آید می‌توان برای فرد مورد آزمایش در هر حالتی که قرار دارد (آرامش، طبیعی و یا استرس) بهترین عمل را جهت از بین بردن استرس و یا ایجاد آرامش بیشتر انتخاب کرد. بدین ترتیب برای فرد مورد آزمایش می‌توان استرس فرد را تا حد مورد قبولی کنترل کرد. در مسئلهٔ یادگیری تقویتی مورد بحث، منظور از محیط در واقع همان شخص مورد آزمایش می‌باشد و جدول Q به دست آمده برای هر فرد متفاوت خواهد بود، چون محیط متفاوت می‌باشد. بنابراین برای هر فرد می‌بایست یک جدول Q مجزا که توسط یادگیری تقویتی به دست آمده، داشت که به وسیلهٔ آن شدت استرس افراد را کنترل نمود. منتهی نکته‌ای که حائز اهمیت می‌باشد این است که هر چند جدول Q از فردی به فرد دیگر متفاوت می‌باشد ولی برای هر فرد جدید می‌توان از میانگین جداول Q که تاکنون به دست آمده، آموزش فرد را شروع کرد و سپس با انجام آزمایشات بهترین اعمال را انتخاب کنیم با انجام این کار زمان سعی و خطا به مراتب کمتر خواهد شد. روال انتخاب بهترین اعمال یا به بیان دیگر بهره‌برداری از جداول Q به دست آمده می‌تواند به وسیلهٔ کامپیوتر نیز انجام شود و نیاز به حضور فردی جهت اعمال عمل‌های مطلوب نمی‌باشد. مسئلهٔ مهم دیگر این می‌باشد که در هنگام بهره‌برداری از جدول Q می‌توان همچنان یادگیری را به صورت آنلاین ادامه داد.

از مزایای این روش نسبت به سایر روش‌های کنترلی می‌توان به بی‌نیاز بودن به مدل ریاضی محیط، قابلیت تطبیق‌پذیری با تغییرات در کنترل اضطراب فرد و سریع بودن الگوریتم در فاز بهره‌برداری اشاره کرد. یکی از مسایل مهم در بحث یادگیری تقویتی خاصیت تطبیق‌پذیری با تغییرات در محیط می‌باشد که یادگیری تقویتی با آموزشی که در حین اجرا نیز می‌تواند داشته باشد، خود را با تغییرات تطبیق می‌دهد.

نکته‌ای که در مورد این تحقیق باید ذکر شود این است که می‌توان با استفاده از دستگاه‌های ساخته شده و استفاده از یادگیری تقویتی تا حد زیادی استرس افراد را کنترل کرد که از نظر هزینه بسیار پایین می‌باشد. نقش یادگیری تقویتی در حقیقت این است که بجای همان روانشناس در کلینیک‌ها، متناسب با شرایط درمانگر، استرس فرد را با پخش عکس و آهنگ کاهش دهد. بنابراین در این تحقیق روشی برای کنترل سطح استرس فرد به کمک یادگیری تقویتی برای بهبود صحت تشخیص، به شیوه‌ای آسان، با تشخیص سریع، هزینهٔ کم و سبک جدید، ارائه و ارزیابی شده است.

پیشنهاد می‌شود که این تحقیق با روش‌های دیگر یادگیری تقویتی همانند مونت کارلو[28] و سارسا[29] نیز انجام شود. با توجه به دستگاه استرس‌سنجی که ساخته شد که نمونهٔ آزمایشگاهی آن نیز وجود دارد، این روند یک کار آزمایشگاهی است که عیب کار آزمایشگاهی این است که سوژه ممکن است آزمون را جدی نگیرد و این کار باعث کاهش اعتبار و افزایش خطا شود به همین دلیل از سوژه‌ها خواسته می‌شود که تمرکز کافی را روی عکس و آهنگ داشته باشند. یعنی همان بیوفیدبک و اینکه تلاش کنند تا آرامش

[28] Monte-Carlo
[29] Sarsa

www.SID.ir

Archive of SID

بیشتری را به دست آورند. در سنجش استرس معمولاً دو نوع خطای FN[30] و FP[31] نامطلوب هستند. رخداد FN در واقع به این معناست که یک فرد دارای استرس زیاد،کم استرس تشخیص داده شود و رخداد FP به این معناست که فرد دارای استرس کم، با استرس زیاد نمایان شود. در مطالعات هدف کم کردن نرخ FN است که در این مقاله رعایت شدن این موضوع حائز اهمیت است.

منابع

1. Fabiani M, Gratton G, Coles M, Federmeier KD. Event-related brain potentials: methods, theory and applications. cacioppo j, tassinary lg. handbook of psychophysiology. 3rd ed. Cambridge University Press. 2007; p.85-111.

2. Pedrotti M, Mirzaei MA, Tedesco A, Chardonnet JR, Mérienne F, Benedetto S, et al. Automatic stress classification with pupil diameter analysis. J Human-Computer Interaction. 2014; 30(3): 220-36.

3. Cook AE, Hacker DJ, Webb AK, Osher D, Kristjansson SD, Woltz DJ, et al. Lyin'eyes: ocular-motor measures of reading reveal deception. J Exp Psychol Appl. 2012; 18(3): 301-13.

4. Pavlidis I, Levine J. Thermal image analysis for polygraph testing. IEEE Eng Med Biol. 2002; 21(6): 56-64.

5. Lykken DT, Vemables PH. Direct measurement of skin conductance: a proposal for standardization. J Psychophysiology. 1971; 8(5): 656-72.

6. Fredrickson M, Ohman A. Cardiovascular and electrodermal responses conditioned to fear- relevant stimuli. J Psycophysiology. 1979; 16(1): 1-7.

7. Abdullah AA, Hassan UH. Design and development of an ESI kit. IEEE Conference on Student. 2012; 253-7.

8. Choi J, Gutierrez-Osuna R. Using heart rate monitors to detect mental stress. Body Sensor Networks. 2009; p. 219-23

9. van Dooren M, Janssen JH. Emotional sweating across the body: comparing 16 different skin conductance measurement locations. Physiology & Behavior. 2012; 106(2): 298-304.

10. Richard S, Sutton RS, Barto AG. Reinforcement learning: an introduction. Cambridge: MIT Press. 1998.

11. Kaelbling LP, Littman ML, Moore AW. Reinforcement learning: a survey. JAIR. 1996; 4: 237-85.

12. Gosavi A. Reinforcement learning: a tutorial survey and recent advances. Informs J Computing. 2009; 21(2): 178-92.

13. Kakade SM. On the sample complexity of reinforcement learning. PhD thesis. University of London Gatsby Computational Neuroscience Unit. 2003.

14. Gomez P, Danuser B. Affective and physiological response to environmental noise and music. J Psychophysiology. 2004; 53(2): 91-103.

[30] False negative
[31] False positive

www.SID.ir

Downloaded from shefayekhatam.ir at 15:58 +0330 on Wednesday January 31st 2018