Proceedings of The 2019 IEEE International
Conference on Real-time Computing and Robotics
August 4-9, 2019, Irkutsk, Russia

Two Frequencies Sequential Coding for the ASSR-based Brain-Computer Interface Application*

Guozhi Cao, *Student Member*, *IEEE*, Jun Xie, *Member*, *IEEE*, Guanghua Xu, *Member*, *IEEE*, Peng Fang, *Senior Member*, *IEEE*, Guangjing Du, *Student Member*, *IEEE*, Min Li, *Member*, *IEEE*, Qing Zhang, *Member*, *IEEE*, and Guanglin Li, *Senior Member*, *IEEE*

Abstract— **Auditory steady state response (ASSR) based brain-computer interface (BCI) shows unique advantages compared with other BCIs such as steady-state visual evoked potentials (SSVEP) based BCI paradigm due to its visual-independent characteristic, and has attracted much attention from researchers. In this paper, a novel ASSR-based BCI using multiple frequencies sequential coding was proposed with the aim of improving the performance of conventional ASSR BCI. Each audio stimulus stream was modulated by two sequential frequencies, i.e., 4Hz and 13Hz for left ear stimulation and 5Hz and 9Hz for right ear stimulation. An experiment consisting of 200-trial offline task and 50-trial online task was performed for each subject and four subjects were recruited to participate in this study. Canonical correlation analysis (CCA) method was used in offline task to extract the ASSR response features, with which LDA model was trained. By applying the trained model to online task, we found that the accuracy of the proposed multiple frequencies sequential coding method is higher than that using the single frequency. And we also draw such a conclusion that the low frequency range can be used in ASSR-based BCIs. All of the results proved the feasibility of the proposed method to provide a new possible paradigm for practical ASSR BCI applications.**

I. Introduction

Brain-computer interface (BCI) is a technology that bypasses the normal peripheral nerve and muscle channels to provide an alternative way of communication. By measuring and translating brain activities into digital commands, BCIs can convey brain intent into messages or control commands. This technology was initially developed to help patients with severe motor impairment, such as suffering from serious injury or neurological diseases, e.g., amyotrophic lateral sclerosis (ALS) disease, to interact with the external world [1]. There are various types of human brain mapping techniques used for implementing BCI systems, such as functional magnetic resonance imaging (fMRI), near infrared spectroscopy (NIRS), and Electroencephalography (EEG). Among them, EEG-based BCIs are extensively studied because they do not require surgery and the acquisition devices used to obtain EEG signals are portable and cost effective [2]. Several paradigms have been developed in EEG-based BCIs, such as motor imagery (MI), P300, steady-state visual evoked potentials (SSVEP), auditory steady-state response (ASSR), and so on. The development of such BCIs brings hope for those disables especially for locked-in syndrome (LIS) patients.

LIS is a disease in which the patient is awake and conscious but most of the voluntary muscles are paralyzed [3]. LIS can be classified into three categories of total LIS, classic LIS and incomplete LIS [4]. Total LIS patients are completely immobility and inability to communicate, but with full consciousness. Classic LIS patients preserve vertical eye movement and blinking. And incomplete LIS patients partially recover muscle control and eye movements. For incomplete LIS patients, vision-based BCIs such as SSVEP-based BCIs can be applied. However, for total and classic LIS patients, using vision-based BCI can be difficult because they may lose the ability to gaze at visual targets or to maintain gaze fixation. To overcome the limitations of conventional vision-based BCIs, some researchers have turned to auditory stimuli as an alternative way instead of visual stimuli [5]. When an auditory stimulus with a specific frequency is applied, ASSR responses can be elicited in human EEGs. This kind of ASSR can be used to implement BCI applications and many researchers have concentrated their focus on ASSR-based BCIs.

Galambos et al. [6] found that ASSR is phase-locked to the stimulus and is strongest at around 40 Hz. Lopez et al. [7] conducted experiments to test the possibility of using ASSR as a new BCI paradigm. Their study provided the evidence that selective attention can modulate ASSR, which proved the feasibility of ASSR-based BCIs. Kim et al. [8] implemented an online ASSR-based BCI system which used beat frequencies (i.e., 37 and 43 Hz). Recently, Kaongoen et al. [9] proposed a novel hybrid BCI by combining ASSR with P300, which achieved better results than implementing ASSR or P300 individually in the binary classification. However, ASSR is relatively weak and cannot achieve comparable BCI performance as the SSVEP-based BCI, even in binary classification tasks. Although these studies, to some extent, have explored the applicability of ASSR-based BCI and have also made some contributions to the improvement of the ASSR-based BCIs, they have not involved the specificity

*Research supported by the National Key Research & Development Program of China (2017YFA0701103) , the Shenzhen Basic Research Program (JCYJ20170818163724754), the National Natural Science Foundation of China (61503298 and 61773364), the Natural Science Foundation of Shaanxi Province of China (2016JQ6002), and the Fundamental Research Funds for the Central Universities (xjj2016006).

J Xie (Corresponding author) and G Xu are with State Key Laboratory for Manufacturing Systems Engineering, and School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049 China (e-mail: xiejun@mail.xjtu.edu.cn; ghxu@mail.xjtu.edu.cn).

G Cao, G Du, M Li and Q Zhang are with School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049 China (e-mail: caodd250@stu.xjtu.edu.cn; duguangjing@stu.xjtu.edu; min.li@mail.xjtu.edu.cn; zhangq@mail.xjtu.edu.cn).

P Fang and G Li are with the CAS Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institutes of Advanced Technology & Shenzhen Engineering Laboratory of Neural Rehabilitation Technology, Shenzhen 518055 China (e-mail: peng.fang@siat.ac.cn; gl.li@siat.ac.cn).

978-1-7281-3726-1/19/$31.00 © 2018 IEEE

characteristic in ASSR signals, which would be helpful in improving the ASSR BCI performance and needs further study.

In the research field of SSVEP-based BCIs, some studies have been carried out in improving EEG signal specificity. Chang et al. [10] proposed an amplitude-modulated visual stimulation method, which used a lower frequency (i.e., 9-12Hz) to modulate a higher frequency (i.e., higher than 40Hz), for the formation of stimulation targets. Experimental results showed that the proposed method could alleviate low eye fatigue and enhance the BCI performance. Shyu et al. [11] used a combination of dual frequencies to modulate the stimulation target. Each target consisted of two LED flickering at different frequencies simultaneously. Despite the improvement of the signal specificity derived from the combination of dual frequencies, this method would also bring in complex harmonic components in the power spectrum, making it more difficult to draw a conclusion. Zhang et al. [12] presented a novel SSVEP-based BCI paradigm in which multiple frequencies were sequentially arranged to code the stimulation target and confirmed its efficiency. This design improves specificity while avoiding complex harmonic problems.

Inspired by this work, we introduced this method to ASSR with the aim of enhancing the specificity of ASSR signals and improving its performance. In the proposed system, sounds with different pitches and amplitude modulation (AM) frequencies are simultaneously presented to the subject through separate sound channel. Each channel was modulated by two frequencies combined sequentially, i.e., 4Hz, 13Hz for the left ear channel and 5Hz, 9Hz for the right ear channel. Classified by LDA model with the feature extraction via CCA algorithm, we confirmed that the performance of the proposed paradigm is better than that using a single frequency. Besides, we found that lower frequencies can also be used in ASSR-based BCI. The system will be described in detail in the following sections.

II. Methods

A. Subjects

Four subjects from Xi'an Jiaotong University participated in the experiments. Two of them are males and two are females (aged 23 ± 1 years old). Before the experiment, all subjects were given a detailed, written summary of the experimental procedures. And all of them signed a written consent before the experiment. None of the subjects reported neurological disorders or hearing problems that might affect the experiments, nor had they experienced any kind of auditory BCI experiments before. The study was approved by the local ethics committee of the Xi'an Jiaotong University.

B. Stimuli design

Two sound sources with different pitch and AM frequency were used as the auditory stimuli (see Fig. 1). Both sounds were simultaneously presented separately in the left and right ear channels of the subject. In traditional coding methods, the source is usually coded by only one frequency. Different from the conventional way, this experiment used two frequencies to code one target. The two frequencies were sequentially presented in two periods separately. For the left ear channel, a stream of 2.6 kHz sounds was modulated by 4 Hz in the first 3 s and 13 Hz in the second 3 s. For the right ear channel, a stream of 540 Hz sounds was modulated by 5 Hz in the first 3 s and 9 Hz in the second 3 s. Thus, one stimulus cycle was divided into two periods (see Fig. 2). In one trial of stimuli, either the left or right ear channel was chosen as the target channel. Before stimuli, a sound stimulus lasting one second was presented to the subject to indicate the target channel as an indicator cue. The indicator cue had the same carrier frequency as the target channel but without AM modulation.

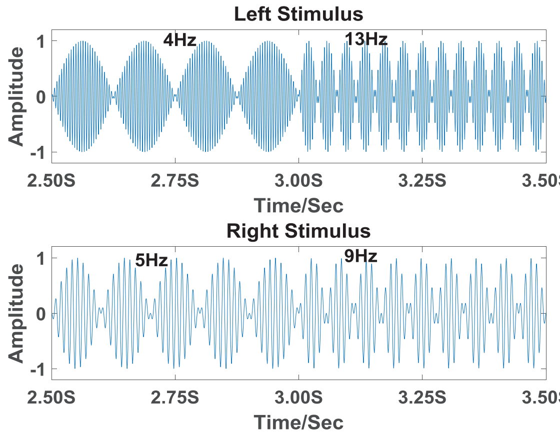

Figure 1. The two-frequencies sequential coding based stimulus design

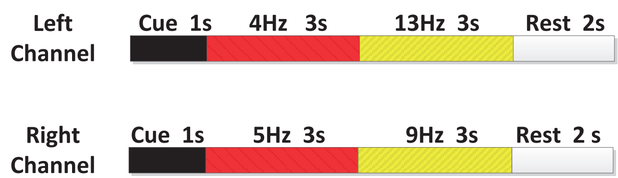

Figure 2. The stimulus cycle

C. Experimental setup

The experiment was conducted in a quiet room. Before the experiment, the experimental procedures were explained to the subject via both written and verbal instructions. Then the subject was asked to sit on a comfortable seat and put on an in-ear earphone (Sennheiser, IE 80S, German), and the auditory volume was adjusted to a comfortable level beforehand.

The entire experiments consisted of two tasks for each subject, i.e., an offline training task and an online experimental task. The aim of the offline task was to collect data from the subjects in order to train LDA model. It was composed of 20 runs lasted for approximately 40 min. Each run consisted of five trials. Each trial started with an indicator cue speaking 'left' or 'right', followed by stimuli lasting 6 s in both ear channels. Two trials were separated with a time interval of 2 s. And after 5 runs there will be a 2-min rest. The subject can also stop and rest whenever she or he felt tired.

After the offline task, an online task was implemented which consisted of 10 runs. The stimuli design of online experiments is similar to that in the offline experiments. The difference is that the LDA model trained in offline tasks would be utilized to classify the data in online tasks and hence real-time feedback would be provide in online tasks to indicate that whether the subject has identified the correct target or not. So, after auditory stimuli, an audio indicator of 'correct' or 'wrong' recorded from a live person would be played in the target channel and a letter of ‘L’ or ‘R’ would visually appear on the screen, according to the classification results.

D. Data Acquisition

Six EEG electrodes, including T7, T8, P7, P8, TP7 and TP8 according to the International 10-20 System, were selected as recording channels to acquire ASSR signals from the primary auditory cortical areas (see Fig. 3). The ground electrode was placed behind the left ear with the reference electrode on the opposite side. To amplify and acquire EEG signals, a 16-channel biosignal amplifier system of g.USBamp (g.tec Medical Engineering GmbH, Austria) was used. EEG data were sampled at a rate of 1200 Hz with a band-pass filter with the cut-off frequency between 2 Hz and 200 Hz and a notch filter with the cut-off frequency between 48 Hz and 52 Hz.

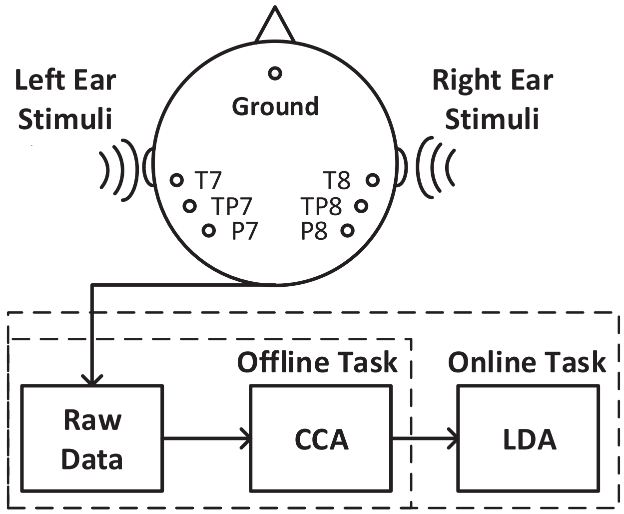

Figure 3. The experimental setup

E. Feature selection and classification using CCA and LDA

- **CCA based feature extraction**

Canonical correlation analysis (CCA) is one of the most commonly used algorithms for mining data association and was chosen as the feature extraction method in this study. In this case, we used CCA algorithm to compare the actual EEG signals with the reference signals to find out their correlation coefficients. The reference signals are defined as a set of Sin-Cos signals with the ASSR modulation frequency as follows:

$$Y_{\mathrm{i}} = \begin{Bmatrix} \cos(2\pi f_i t) \\ \sin(2\pi f_i t) \\ \cos(2\pi \cdot 2 f_i t) \\ \sin(2\pi \cdot 2 f_i t) \\ \cos(2\pi \cdot 0.5 f_i t) \\ \cos(2\pi \cdot 0.5 f_i t) \end{Bmatrix}, t = \frac{1}{F_s}, \frac{2}{F_s}, \cdots, \frac{S}{F_s} \quad (1)$$

Where f_i is the AM frequency of the ASSR stimuli, S is the number of data point, and F_s is the sampling rate. In this study, the reference signals used fundamental frequency, harmonic frequency, and sub-harmonic frequency signals.

For two variables X and Y, their correlation coefficients can be obtained by linear regression with N pairs of samples of these two variables. For groups of variables X and Y, they can be projected into one dimension by two weight vectors W_x and W_y, and then their correlation coefficients can be obtained by the same way as two variables. CCA seeks such weight vectors W_x and W_y to maximize their linear correlation $\rho_{x,y}$:

$$\begin{aligned} x &= X^T \cdot W_x \\ y_i &= Y_i^T \cdot W_y \\ \rho_{x,y_i} &= \frac{\mathrm{cov}(x, y_i)}{\sqrt{D(x)} \cdot \sqrt{D(y_i)}} \end{aligned} \quad (2)$$

- **LDA based ASSR classification**

In the online experiments, after CCA correlation coefficients were obtained, a linear discriminant analysis (LDA) method was employed for the classification. The two-class LDA method first projects the data onto a straight line. Then the within-class covariance S_B and the between-class covariance S_W can be expressed as follows:

$$\begin{aligned} S_B &= \sum\nolimits_L + \sum\nolimits_R \\ &= \sum_{x \in X_L} (x-\mu_L)(x-\mu_L)^T + \sum_{x \in X_R} (x-\mu_R)(x-\mu_R)^T \end{aligned} \quad (3)$$

$$S_W = (\mu_L - \mu_R)(\mu_L - \mu_R)^T \quad (4)$$

Where X_L and X_R represent the two classes, μ_L and μ_R represent their sample mean.

LDA seeks to minimize within-class covariance while maximize between-class covariance. Therefore, the optimization goal can be expressed as follow:

$$J_W = \max \frac{W^T S_B W}{W^T S_W W} \quad (5)$$

For each frequency (f_i, i=1, …, k), the correlation coefficient ρ_i is obtained through CCA. All CCA correlation coefficients form a multidimensional vector defined as:

$$\rho_{\mathrm{i}} = [\rho_1, \rho_2, ..., \rho_k] \quad (6)$$

By using this dataset labeled by frequency f_i, the projection matrix W can be obtained. To classify the input vector ρ, the distance between $W^T\rho$ and μ_L, μ_R can be calculated. In this way, ρ can be classified into the nearest class.

III. Results

Table I presents the averaged accuracy of each subject in the online experiments. The average accuracy of the four subjects is 70%. Subject S4 achieved the best accuracy of 74%. Besides, Mann-Whitney U test was used to detect the difference between the accuracies of the left ear target and the right ear target and no significant difference was found (p = 0.407).

TABLE I. THE ONLINE EXPERIMENTAL ACCURACIES

	S1	S2	S3	S4
Left Ear Target	70%	70%	75%	76%
Right Ear Target	75%	85%	70%	72%
Mean	72.5%	72%	72.5%	74%

Since the duration of each single trial can be divided into two parts as the first 3-s period for one frequency and the second 3-s period for another frequency, we analyzed the two parts separately. Fig. 4 shows the mean accuracy of the first 3-s period, the second 3-s period and the multi-frequency 6-s period for each subject, respectively.

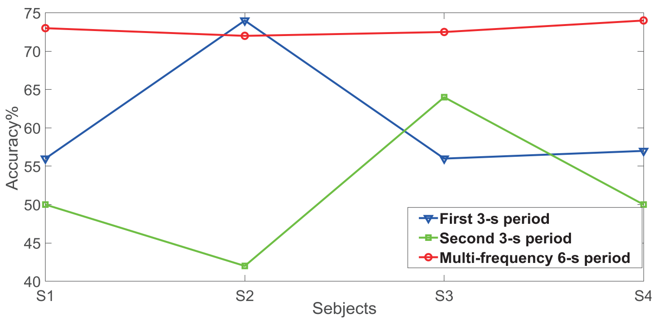

Figure 4. The mean accuracy of separate periods and the overall period for each subject

There is an obvious tendency that the proposed two frequencies sequential coding method of 6-s period can achieve higher accuracies than the single frequency stimulation of only 3 s in the first or second period within each trial. We used one-way analysis of variance test to compare the accuracies between the first 3-s period and the multi-frequency 6-s period, and the accuracies between the second 3-s period and the multi-frequency 6-s period. For both the former and latter cases, the accuracies in the condition of the multi-frequency 6-s period were significantly higher than the first 3-s period (p = 0.025) and the second 3-s period (p = 0.001) conditions, respectively. Because as we know, there is a positive correlation between the accuracy rate and the stimulation duration, the significant accuracy difference may be caused by different stimulation durations. But in the following study, the factor of duration length could be excluded because the accuracy difference between left and right ear target detection is very low in multiple frequencies coding way, while using a single frequency coding method may cause "accuracy imbalance" between left and right ear target detection (see Fig. 5), which means that one side ear target detection would always achieve high recognition accuracy while the other side ear target detection would achieve the opposite effect. By using one-way analysis of variance test, it is revealed that the "accuracy imbalance" from multi-frequency way to single frequency way is significantly decreased (p = 0.038). So if it is the stimulation duration that raised the accuracy, the "accuracy imbalance" will also be raised instead of being decreased. It is verified that the accuracy improvement was achieved by the proposed multi-frequency method instead of the stimulation duration effect.

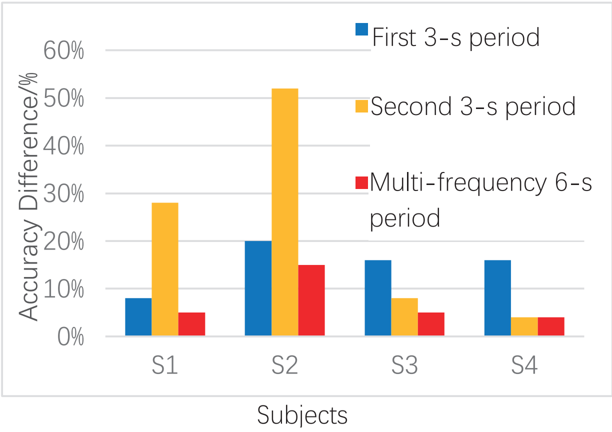

Figure 5. The accuracy difference of separate periods and the overall period for each subject

IV. Discussion

BCI is a very promising tool for those disables such as ALS patients. ASSR-based BCIs use audio stimulation to modulate information. However, ASSRs are relatively weak, which leads to low performance in ASSR-based BCIs. In this paper, a novel developed auditory stimuli sequential coding method for ASSR-based BCI was proposed to improve its performance. Results indicated that this method could make use of various frequencies and achieved higher accuracies.

For traditional ASSR-based BCI that used single frequency to code each target, one target only corresponds to one frequency. Such coding way limited the number of achievable targets, because the available frequency was limited. Here we used the sequential combination of two frequencies to code one target, thus has the potential to increase the number of available targets. Different from the method of using a combination of dual frequencies to modulate the target [11], the proposed method avoids the complex harmonic component problem. This method could be extended to other steady-state evoked potential related BCIs that are limited by the number of available frequencies.

And with the implementation of multiple frequencies for a single target, another question would arise: Will the first frequency affect the detection of the second frequency since they are sequentially presented in one target? From Fig. 4 we can see that the accuracy of the second 3-s period is lower than that of the first 3-s period. It may be because the stimulation frequency of the first 3-s period is lower than the second 3-s period, which would be prone to elicit stronger steady-state responses. It is also possible that the second 3-s period was affected by the first 3-s period since the two periods were presented continuously and there was no time interval between them. Further study is needed to explore whether using time interval between two periods could affect this phenomenon.

Many previous studies reported that ASSRs have the largest amplitudes on modulation frequencies of approximately 40 Hz [13]-[15]. Different from the previous studies, much lower frequencies such as 4 Hz, 5Hz, 9Hz and 13Hz were utilized in this paper. Results showed that the low-frequency range could also work well in the ASSR-based BCI. This is in accordance with Manuela Jaeger's study [16] that the ASSR could be elicited at 4 Hz and 7 Hz in human EEG.

However, the recognition accuracy of around 70% is still not high enough to be used in real-life BCI applications. And the trial duration of 6 s is still relative long, which may lead to a lower information transfer rate in BCI communication. Besides, subjects also reported that paying selective attention to a specific sound stream, which plays an important role in ASSR-based BCI recognition, was very tired. All those factors reduced user acceptance and limited its possible application. Besides, we used four subjects in the experiments, which made the conclusion not so sufficient. More subjects would be included in further study to sophisticatedly evaluate the extensibility of our present study. In future studies, we intend to use a more comfortable modulation carrier like nature sounds or music instead of flat tones and test it on more subjects. We expect this would make it easier and more comfortable for users.

V. Conclusion

In this study, we presented a novel BCI paradigm based on two frequencies sequential coded ASSR. Aimed to improve the performance of common ASSR-based BCIs and achieve higher accuracies, we introduced a multiple frequencies sequential coding method by combining two frequencies in one ear channel. In this way, more frequency information was used to modulate one target and the recognition accuracy was increased accordingly. Based on the experimental results from our four healthy subjects, our proposed system achieved considerable performance. For patients who lost their eyesight or have difficulties to use vision based BCI systems, our BCI design could provide an alternate way for them to communicate with external environments.

Acknowledgment

We want to thank the subjects for participating in these experiments and anonymous reviewers for their helpful comments.

References

[1] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, et al., "Brain-computer interface technology: a review of the first international meeting," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp: 164–173, June 2000.

[2] C. H. Chen, M. S. Ho, K. K. Shyu, K. C. Hsu, K. W. Wang, and P. L. Lee, "A noninvasive brain computer interface using visually-induced near-infrared spectroscopy responses," *Neurosci. Lett.*, vol. 580, pp: 22–26, Sept. 2014.

[3] G. Bauer, F. Gerstenbrand, and E. Rumpl, "Varieties of the locked-in syndrome," *J. Neurol.*, vol. 221, no. 2, pp: 77–91, Aug. 1979.

[4] E. Smith and M. Delargy, "Locked-in syndrome," *BMJ*, vol. 330, no. 7488, pp: 406–409, Feb. 2005.

[5] M. Schreuder, B. Blankertz, and M. Tangermann, "A new auditory multi-class brain-computer interface paradigm: spatial hearing as an informative cue," *PLoS ONE*, vol. 5, no. 4, pp: e9813, Apr. 2010.

[6] R. Galambos, S. Makeig, and P. J. Talmachoff, "A 40-Hz auditory potential recorded from the human scalp," *P. Natl. Acad. Sci. USA*, vol. 78, no. 4, pp: 2643–2647, Apr. 1981.

[7] M. A. Lopez, H. Pomares, F. Pelayo, J. Urquiza, and J. Perez, "Evidences of cognitive effects over auditory steady-state responses by means of artificial neural networks and its use in brain-computer interfaces," *Neurocomputing*, vol. 72, no. 16-18, pp: 3617–3623, Oct. 2009.

[8] D. W. Kim, H. J. Hwang, J. H. Lim, et al., "Classification of selective attention to auditory stimuli: toward vision-free brain-computer interfacing," *J. Neurosci. Meth.*, vol. 197, no. 1, pp: 180–185, Apr. 2011.

[9] N. Kaongoen and S. Jo, "A novel hybrid auditory BCI paradigm combining ASSR and P300," *J. Neurosci. Meth.*, vol. 279, pp: 44–51, Mar. 2017.

[10] M. H. Chang, H. J. Baek, S. M. Lee, and K. S. Park, "An amplitude-modulated visual stimulation for reducing eye fatigue in SSVEP-based brain-computer interfaces," *Clin. Neurophysiol.*, vol. 125, no. 7, pp: 1380–1391, July 2014.

[11] K. K. Shyu, P. L. Lee, Y. J. Liu, and J. J. Sie, "Dual-frequency steady-state visual evoked potential for brain computer interface," *Neurosci. Lett.*, vol. 483, no. 1, pp: 28–31, Oct. 2010.

[12] Y. Zhang, P. Xu, T. Liu, J. Hu, R. Zhang, and D. Yao, "Multiple frequencies sequential coding for SSVEP-based brain-computer interface," *PLoS ONE*, vol. 7, no. 3, pp: e29519, Mar. 2012.

[13] B. Ross, A. T. Herdman, and C. Pantev, "Right hemispheric laterality of human 40 Hz auditory steady-state responses," *Cereb. Cortex*, vol. 15, no. 12, pp: 2029–2039, Dec. 2005.

[14] C. Wienbruch, I. Paul, N. Weisz, T. Elbert, and L. E. Roberts, "Frequency organization of the 40-Hz auditory steady-state response in normal hearing and in tinnitus," *Neuroimage*, vol. 33, no. 1, pp: 180–194, Oct. 2006.

[15] J. Bohórquez and Ö. Özdamar, "Generation of the 40-Hz auditory steady-state response (ASSR) explained using convolution," *Clin. Neurophysiol.*, vol. 119, no. 11, pp: 2598–2607, Nov. 2008.

[16] Jaeger, M, et al. "Did You Listen to the Beat? Auditory Steady-State Responses in the Human Electroencephalogram at 4 and 7 Hz Modulation Rates Reflect Selective Attention." Brain Topography(2018):1-16.