A CNN-based compare network for classification of SSVEPs in human walking

Chenyao Wu[1,2], Shuang Qiu[1], Jiezhen Xing[1,2], and Huiguang He[1,2,3,*], Member, IEEE

Abstract— **Brain-computer interface (BCI) can provide a way for the disabled to interact with the outside world. Steady-state visual evoked potential (SSVEP), which evokes potential through visual stimulation is one of important BCI paradigms. In laboratory environment, the classification accuracy of SSVEPs is excellent. However, in motion state, the accuracy will be greatly affected and reduce quite a lot. In this paper, in order to improve the classification accuracy of the SSVEP signals in the motion state, we collected SSVEP data of five targets at three speeds of 0km/h, 2.5km/h and 5km/h. A compare network based on convolutional neural network (CNN) was proposed to learn the relationship between EEG signal and the template corresponding to each stimulus frequency and classify. Compared with traditional methods (i.e., CCA, FBCCA and SVM) and state-of-the-art method (CNN) on the collected SSVEP datasets of 20 subjects, the method we proposed always performed best at different speeds. Therefore, these results validated the effectiveness of the method. In addition, compared with the speed of 0 km / h, the accuracy of the compare network at a high walking rate (5km/h) did not decrease much, and it could still maintain a good performance.**

I. Introduction

A brain–computer interface (BCI) is a system that decodes brain activities to provide users with alternative ways to control various computer-based applications and assistive devices [1]. The non-invasive acquisition method based on electroencephalogram (EEG) is much more studied because its acquisition equipment is cheaper and easier to carry than large-scale equipment such as fMRI. Existing EEG-based BCI include slow cortical potential (SCP) [2], p300 [3], motor imagery (MI) [4] [5], steady-state visual evoked potential (SSVEP), steady-state auditory evoked potential (SSAEP), movement-related cortical potential, etc. The SSVEP-based BCI system has been widely used because of its simple experiment, high information transfer rate and easily extracted feature. SSVEPs are brain responses elicited by stimulating the retina of the eyeball at a fixed frequency.

At present, most of the studies on BCI are carried out in a quiet laboratory environment, and subjects are required to stay still, which is inconsistent with the real-life application scenario. Recent studies have found that when people walk or run, their brains work differently compared with when they are still [6]. Nowadays, many methods based on SSVEP paradigm can achieve high accuracy in the laboratory environment. However, it is difficult for them to maintain high accuracy while walking. With the development of portable EEG acquisition equipment, the research on SSVEP signal in motion has attracted much attention. In recent years, some articles have found that the classification accuracy of SSVEP were affected to some extent when people are moving [7], [8]. At the same time, movement of the human body will induced more noise into the collected signals [9], which added extra difficulty to classification. In many BCI applications, such as controlling exoskeleton to walk by EEG, it's hard to avoid the movement of the subjects. However, few methods can still maintain a high accuracy while walking. In addition, most of the current research on SSVEP doesn't include the classification of rest state, which is also a problem that needs to be solved in practical application.

Nowadays, Canonical Correlation Analysis (CCA), MwayCCA and other traditional methods are still widely used in the research of SSVEP [10]-[12]. In recent years, deep learning has made great progress in speech recognition, image recognition and other fields. With the development of deep learning, it has been applied in EEG signal processing, such as motor imagery, P300 and performs well. In the laboratory environment, CNN [13]-[15] has excellent performance in processing SSVEP signals. However, there are seldom studies using methods related to deep learning to classify SSVEP signals in human walking.

In this paper, we aim to use deep learning methods to improve classification performance of SSVEPs in the moving state. We collected EEG data at three walking speeds, including four visual stimuli with different frequencies and one without flickering. In order to maintain a high accuracy in the processing of SSVEP signals in the moving state, a compare network was designed to classify the SSVEP signals in the state of moving, hoping to get better performance.

Compare network we proposed is based on CNN, combining the advantages of deep learning and task-related component analysis (TRCA). The input of compare network is frequency domain signal, including signal input and template input. Network learns the relationship between signal and template rather than frequency feature. Several common methods used to classify SSVEPs were reproduced and their classification performance was compared with compare network.

*This work was supported in part by the National Natural Science Foundation of China under Grant 61976209, Grant 81701785, and Grant 61906188, and in part by the Strategic Priority Research Program of CAS under Grant XDB32040200, and in part by the China Postdoctoral Science Foundation 2019M650893.

1 Research Center for Brain-inspired Intelligence, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Science, Beijing, China

2 University of Chinese Academy of Sciences, Beijing, China

3 Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Science, Beijing, China

*Corresponding author huiguang.he@ia.ac.cn

II. EXPERIMENT

A. Subject

20 healthy adults (4 females; 19-30 years of age) with normal or corrected-to-normal vision participated in our experiment. Each participant signed an informed consent in advance. This study was proved by the Research Ethics Committee of the Institute of Automation, Chinese Academy of Science.

B. EEG recording

We recorded EEG signals with a LiveAmp wearable EEG system (gUSBamp and Ladybird electrodes, g.tec Guger Technologies, Austria). The data sampling rate was 500 Hz. The ground electrode was AFz and the reference electrode was Fcz. The impedances of all electrodes were kept below 10KΩ throughout the experiment. During the experiment, EEG signal of 31 electrodes (excluding EOG) in accordance with the international 10-20 system and 3-axis acceleration signal were collected.

C. Experiment setup

In this study, we designed an experiment at three speeds : 0km/h, 2.5km/h, 5km/h. The speed of 0 km/h is standing. The speed of 2.5 km/h is slowly walking. And the speed of 5 km/h is fast walking.

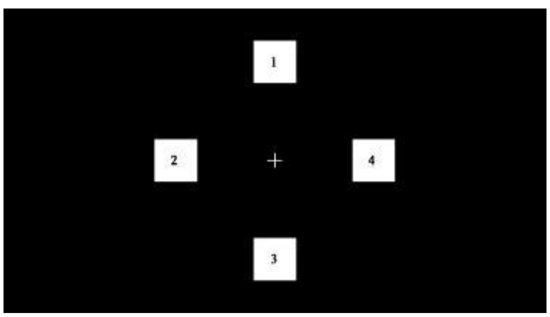

Figure 1. Stimulus design of the 5-target BCI system.

There were four squares flickering at 7 Hz, 8 Hz, 9 Hz and 10 Hz, and a non-flickering cross on the screen. Four squares were distributed at the top, bottom, left and right of the screen respectively, and the non-flickering cross was located in the center of the screen. There were totally 25 trials in each session. Every trial included briefing period and gazing period. During the briefing period which lasts 2 seconds, one of the four squares or the middle cross was marked red randomly. In each session, every target was marked red 5 times. During this period, all the squares were not flickering. Subjects were asked to shift their vision to the target marked red during this period. During the gazing period which lasts 5 seconds, four squares and the cross all stayed white. The four squares were flickering at their frequencies respectively, and the cross did not flicker. Subjects need to keep gazing during this period. Each session lasted about 3 minutes, and the whole experiment was about 24 minutes at each speed.

Participants need to maintain the speed through a treadmill and had to focus their attention on the stimulus throughout the whole course. During the experiment, the subjects were asked to hold the handle of the treadmill to keep the distance between the head and the screen unchanged. The voltage value of amplifier remained at working voltage throughout the whole experiment, and the data is reliable.

III. METHODS

A. Data preprocessing

EEG signals need to be preprocessed before classification and data analysis. Because the experiment used visual stimulation to evoke, the signals at P3, P4, Pz, O1 and O2 were processed and analyzed. Data filtering with a band-pass filter from 5-70Hz was used in order to remove the noise and artifacts carried by raw EEG data. In this paper, we used the Butterworth filter to filter data. Then, EEG signals were down-sampled to 250 Hz. A sliding window with a shift size of 40ms was also used to process EEG signals. After sliding window, each sample lasts for 2 seconds. After transforming data to frequency domain, we intercepted each sample from 5Hz to 68hz. After the whole preprocessing, each sample has 128 points and 5 channels.

B. Compare network

We proposed a compare network, combining the strength of CNN and TRCA, as they are two methods that perform well when classifying SSVEP signal. The method used in this paper compares EEG signals with template signals generated by stimulation frequency and TRCA. This is similar to Canonical correlation analysis (CCA) which calculates the correlation between EEG signals and reference signals. Based on CNN, compare network changes the input from EEG signals to signal input as well as template input. Signal input means task-related components extracted by TRCA from EEG signal. Template input means template generated by TRCA from part of train data. EEG signal used here has been preprocessed and then processed with FFT and standardization.

TRCA [16] is an important part of compare network, which uses a matrix to map data to a new space, in order to extract task-related components from EEG signals. EEG signal contains task-related component as well as task-unrelated component. Multiply mapping matrix with EEG signal x(t) to generate new data y(t). y(t) can be expressed as:

$$y(t)=\sum_j \omega_j x_j(t)=\sum_j(\omega_j a_{1,j} s(t)+ \omega_j a_{2,j} n(t)) \quad (1)$$

Add up the covariance of all possible combinations of trials. The sum obtained is described as:

$$\sum_{h_1h_2} C_{h_1h_2}=\sum_{h_1h_2,\, h_1\neq h_2}\sum_{j_1,j_2} \omega_{j_1}\omega_{j_2} \mathrm{Cov}(x_{j_1}^{(h_1)},x_{j_2}^{(h_2)})= \boldsymbol{\omega}^T\mathbf{S}\boldsymbol{\omega} \quad (2)$$

To obtain a finite solution, the variance of y(t) is constrained as:

$$\mathrm{Var}(y(t))=\sum_{j_1,j_2} \omega_{j_1}\omega_{j_2} \mathrm{Cov}(x_{j_1}(t),x_{j_2}(t))= \boldsymbol{\omega}^T\mathbf{Q}\boldsymbol{\omega}=1 \quad (3)$$

The optimum weight matrix can be obtained by calculating the eigenvalues and eigenvectors of $Q^{-1}S$. The weight matrix obtained is used to map the EEG signal to get the signal input. As the experiment designs five different stimulation targets, there are five different TRCA matrix in total as well as five signal inputs generated by EEG data. TRCA matrix is calculated by part of the train data. Template data is the average of the same part of train data. There are also five template inputs generated by template data corresponding to five stimulation targets.

The Compare network is based on CNN. The architecture of compare network includes input layer I, convolutional layer C1, convolutional layer C2, full connected layer F and output layer O. Because the size of TRCA matrix is 5*5, the size of input that has already been extracted by task-related component is 128*5. The size of the convolution kernel in the first layer is 1×5. Five feature maps are extracted from the first layer, and each map contains 128 units. In the second layer, the size of convolution kernel is 11×1, which also has 5 feature maps. The first layer of convolution convolutes the data between channels to extract the spatial features of SSVEP signal and the second convolution layer convolutes the data in the channel to extract the frequency features of SSVEP signal. The output of the second convolutional layer has 128 units and 5 maps. The input of the full connected layer has 6400 (5*128*10) units. The output of the full connected layer has 512 units.

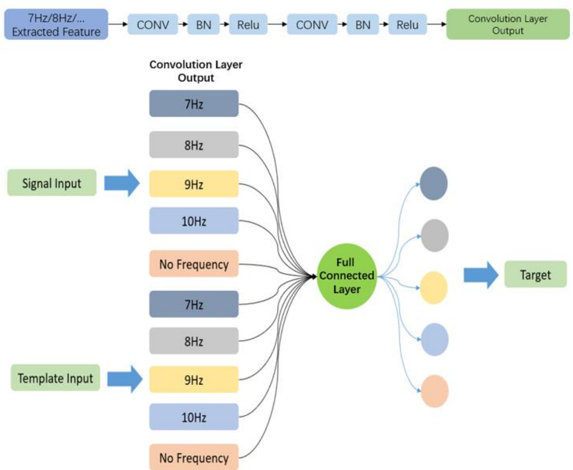

Figure 2. Sketch Map of compare network.

There are totally 10 inputs of the compare network. Five of them are signal inputs and five of them are template inputs. Ten inputs have to be calculated by two convolutional layers separately as well as to be processed by the full connected layer together. Each convolution layer includes convolution, batch normalization and linear rectification function (ReLU). After full connected layer, five results corresponding to five classes are obtained using softmax. The loss function is cross-entropy function and dropout rate is set to 0.5. Part of the train data is randomly selected for one training and the accuracy is calculated. For each kind of walking speed and each person, train data is randomly selected three times, and the final accuracy rate is the average of three accuracy.

C. Compared methods

To validate the performance of compare network, we reproduced several methods that are widely used in SSVEP. We used these methods to classify EEG signals obtained by our experiment, and compared the results with the accuracy calculated by compare network. These methods include CCA, FBCCA, SVM and CNN.

CCA is a multivariate statistical analysis method to grasp the overall correlation between two groups of indicators. Gao Xiaorong et al [17] calculated the correlation between EEG signal and reference signals like trigonometric function to find the most possible frequency to get the EEG signal. Filter bank canonical correlation analysis (FBCCA) is an extension method of CCA, proposed in 2015 [18]. FBCCA uses several filters to filter EEG signal and superposes the correlation of different filters to analyze. Support Vector Machine (SVM) extracts features from the original data as input, and it performs well in solving small sample problems and non-linear classification problems. SVM has been widely used in the fields of SSVEP [19][20]. CNN is also widely used in EEG signal processing which performs well. CNN is one of the earliest proposed deep learning algorithms and can effectively save training costs and improve operation efficiency because of adopting the strategy of "weight sharing".

IV. Result

A. Spectra analysis of EEG signals in walking state

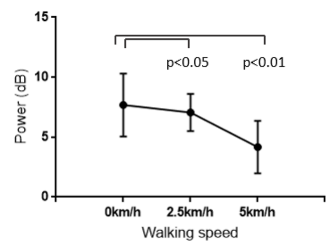

Figure 3. The averaged power of SSVEP signals at different walking speeds.

Under the visual stimulation of each frequency, the subtraction of the average power at this frequency between the spectrum of this frequency and the spectrum without frequency stimulation was calculated. Then the subtraction between these two spectrums at double frequency was also calculated. The power difference at this frequency was obtained by averaging the result of these two subtractions. Then, the power difference of four different frequencies was summed to get the power difference at a certain speed. Figure 3 shows the result of subtracting the average power of frequency from the average power of non-frequency visual stimulation at different speeds. The average power difference at three speeds is 0km/h-7.43(dB), 2.5km/h-5.40(dB) and 5km/h-4.26(dB). We used one-way ANOVA to compare different under each speed and find that speed has a significant impact on power. The decreasing range at 5km/h step speed is greater than that at 2.5km/h step speed. At the speed of 2.5km/h, the power is slightly lower ($P < 0.05$). At the speed of 5km/h, the power is significantly lower ($P < 0.01$), and the effective SSVEP signal decreased significantly. This may be due to the subject's fatigue during walking or the relative movement between EEG acquisition equipment and the head of subject. It's also a possible reason that the subjects are difficult to concentrate while moving.

B. Classification results of four-class

We used the compare network to classify the data under the stimulation of four different frequencies, and calculated the classification accuracy. The classification methods of CCA, FBCCA, SVM and CNN were also reproduced, and

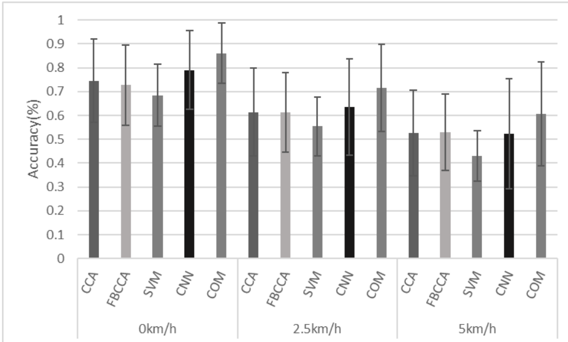

Figure 4. Classification results of four classification methods.

were used to classify the same data. Figure 4 shows the results under 4 classes.

Two-way repeated measures ANOVA analysis of variance (ANOVA) was used to analyze the classification accuracy of 20 subjects. The result shows that different walking speeds and different classification methods both have a significant impact on the accuracy, with $F = 46.61$, $P < 0.0001$, under the speed factor, and $F = 7.67$, $P < 0.0001$ under the classification method factor. Then, we used one-way ANOVA to compare the different methods. The result shows that the accuracy of different classification methods is significantly different at 0 km/h ($F = 3.62$, $P < 0.01$), and the accuracy of compare network classification method is the highest. The t-test among each other was also analyzed at three walking speeds, and the result is as follows: The performance of compare network is significantly better than the other four methods. The result of t-test shows that the classification accuracy of CCA, FBCCA and CNN is similar, and there is no significant difference. The accuracy of SVM is significantly lower than compare network and CNN. The accuracy and standard deviation of the compare network from the lowest walking speed to the fastest walking speed are 86.08/12.68, 71.53/18.25 and 60.63/21.69, respectively. Therefore, compare network improves the decoding of SSVEPs in walking state, and the increase of speed reduces the accuracy.

C. Classification results of five-class

Because CCA and FBCCA can't classify non-flickering stimulation with others, we compared the results of three methods for five-class data here. The compare network was used to classify the data under the stimulation of five different stimuluses, and the classification accuracy was calculated. Reproducing the SVM and CNN methods, the accuracy histogram of 5-class was obtained. The accuracy of compare network is the highest among the three methods. ANOVA and t-test were used to analyze the classification accuracy of 20 subjects. The result at the speed of 0km/h is as follows. At 5-class, compare network classification method performances significantly better than SVM ($F = 23.35$, $P < 0.0001$, t-test: 2.24E-5) and slightly better than CNN ($F = 2.70$, $P > 0.05$, t-test: 0.109). The results at other speeds are different in numerical value, but still can reach the same conclusion. The accuracy and standard deviation of the compare network from the lowest speed to the fastest speed are 77.26/14.40, 62.00/19.78, 51.22/21.52. Thus, compare network may

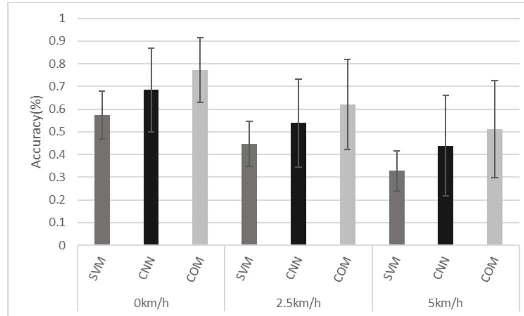

Figure 5. Classification results of five classification methods.

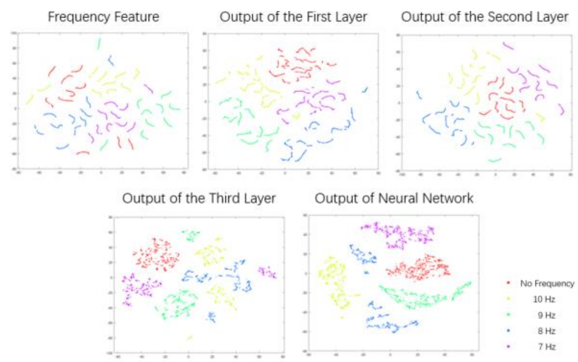

Figure 6. Reduced dimension data of the frequency feature and output of each layer from five-class compare network.

perform excellently while classification task including distinguishing non-flickering stimulation.

In order to see the effect of each layer of the compare network more clearly, we took a subject as an example and projected the output of each layer into two dimensions from the five-class compare network. The project method used in this paper is t-distributed stochastic neighbor embedding (t-SNE). The projected output is shown in Figure 6. Through Figure 6, we learn that frequency feature is more staggered and complex while the output of neural network is more gathered and easier to be distinguished.

V. Conclusion

In this paper, we designed the BCI experiment based on SSVEP at three walking speeds. By analyzing collected data, we found that the head would shake while participants were walking and effective EEG power would decrease during exercise. With the increase of walking speed, the reduction of effective EEG power became more and more significant, which makes it difficult to classify SSVEP signals in motion. We used the compare network to classify collected data. The classification results showed that compare network performed better than other comparable and could maintain a good classification performance at high walking speed. Comparing the classification performance of the compare network with the classification results of other methods, the compare network has the highest classification accuracy whether it is four or five classification problem.

References

[1] Rami Alazrai, Hisham Alwanni and Mohammad I. Daoud, “EEG-based BCI system for decoding finger movements within the same hand,” Neuroscience Letters, vol. 698, pp. 113-120, 2019.

[2] Han-Gue Jo, Jose Raul Naranjo, Thilo Hinterberger, Ulf Winter and Stefan Schmidt, “Phase synchrony in slow cortical potentials is decreased in both expert and trained novice meditators,” Neuroscience Letters, vol. 701, pp. 142-145, 2019.

[3] Yeom SK, Fazli S, Müller KR and S.-W Lee, “An Efficient ERP-Based Brain-Computer Interface Using Random Set Presentation and Face Familiarity,” PLOS ONE, vol. 9, no. 11, pp. e111157, 2014.

[4] Pfurtscheller G and Neuper C, “Motor imagery and direct brain-computer communication,” Proceedings of the IEEE, vol. 89, no. 7, pp. 1123-1134, 2002.

[5] Razi S, Karami Mollaei M R and Ghasemi J, “A novel method for classification of BCI multi-class motor imagery task based on Dempster–Shafer theory,” Information Sciences, vol. 484, pp. 14-26, 2019.

[6] Lin Y P, Wang Y and Jung T P, “A mobile SSVEP-based brain-computer interface for freely moving humans: the robustness of canonical correlation analysis to motion artifacts,” International Conference of the IEEE Engineering in Medicine & Biology Society. IEEE, 2013.

[7] Joseph T. Gwin, Klaus Gramann, Scott Makeig and Daniel P. Ferris, “Removal of movement artifact from high-density EEG recorded during walking and running,” Journal of Neurophysiology, vol. 103, no. 6, pp. 3526-3534, 2010.

[8] Kwak N S , KlausRobert Müller and Lee S W , “A convolutional neural network for steady state visual evoked potential classification under ambulatory environment,” Plos One, vol. 12, no. 2, pp. e0172578, 2017.

[9] Lin Y P, Wang Y and Jung T P, “Assessing the feasibility of online SSVEP decoding in human walking using a consumer EEG headset,” Journal of NeuroEngineering and Rehabilitation, vol. 11, no. 1, pp. 119, 2014.

[10] Zhao J, Mao X, Hu H, et al. “Behavior-Based SSVEP Hierarchical Architecture for Telepresence Control of Humanoid Robot to Achieve Full-Body Movement,” IEEE Transactions on Cognitive and Developmental Systems, vol. 9, no. 2, pp. 197-209, 2017.

[11] Zhonglin Lin, Changshui Zhang, Wei Wu, and Xiaorong Gao, “Frequency Recognition in SSVEP-based BCI using Multiset Canonical Correlation Analysis,” International Journal of Neural Systems, vol. 24, no. 4, pp. 1172-1176, 2013.

[12] Nakanishi M, Wang Y, Wang Y T, and Yasue Mitsukura, “A HIGH-SPEED BRAIN SPELLER USING STEADY-STATE VISUAL EVOKED POTENTIALS,” International Journal of Neural Systems, vol. 24, no. 6, pp. 1450019, 2014.

[13] Kwak N S, KlausRobert Müller and Lee S W, “A convolutional neural network for steady state visual evoked potential classification under ambulatory environment,” Plos One, vol. 12, no. 2, pp. e0172578, 2017.

[14] Nik Khadijah Nik Aznan, Stephen Bonner, Jason D. Connolly, Noura Al Moubayed and Toby P. Breckon, “On the Classification of SSVEP-Based Dry-EEG Signals via Convolutional Neural Networks,” arXiv preprint arXiv:1805.04157, 2018.

[15] Nik Khadijah Nik Aznan, Jason D. Connolly, Noura Al Moubayed and Toby P. Breckon, “Using Variable Natural Environment Brain-Computer Interface Stimuli for Real-time Humanoid Robot Navigation,” arXiv preprint arXiv: 1811.10280, 2019.

[16] Nakanishi M, Wang Y, Chen X, et al, “Enhancing Detection of SSVEPs for a High-Speed Brain Speller Using Task-Related Component Analysis,” IEEE Transactions on Biomedical Engineering, vol. 65, no. 1, pp. 104-112, 2017.

[17] Zhonglin Lin, Changshui Zhang, Wei Wu, and Xiaorong Gao, “Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs,” IEEE Transactions on Biomedical Engineering, vol. 54, no. 6, pp. 1172-1176, 2007.

[18] Chen X, Wang Y, Gao S, et al, “Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain-computer interface,” Journal of Neural Engineering, vol. 12, no. 4, pp. 046008, 2015.

[19] Jian H L and Tang K T, “Improving classification accuracy of SSVEP based BCI using RBF SVM with signal quality evaluation,” International Symposium on Intelligent Signal Processing & Communication Systems. IEEE, 2014.

[20] Anindya S F, Rachmat H H and Sutjiredjeki E, “A prototype of SSVEP-based BCI for home appliances control,” 2016 1st International Conference on Biomedical Engineering (IBIOMED). IEEE, 2016.