

Comparison of Spectral and Template Matching Features for SSVEP BCI Target Frequency Classification

Ibrahim Kaya*¹

Submitted: 28/02/2021

Accepted : 02/04/2021

Abstract: Brain-computer interfaces (BCI) provide new communication and control channels to restore and support these functions of the restricted users. Among these Visual Evoked Potential (VEP) based BCIs are the most promising in terms of ease of use and performance. The frequency following phenomenon of VEPs produce Steady State Visual Evoked Potentials (SSVEP) at the frequency of stimulation of the human visual system. In such interface systems, each target is encoded with a particular stimulation frequency and phase. In communication purpose speller interfaces each target flickers a letter or character with a particular stimulation frequency and phase. The detection of the focused target by the computer is required. In this process, classification methods and feature extraction method play critical roles. This study used a publicly available benchmark dataset of a 40 target SSVEP BCI. In the analysis, two feature vectors are obtained from power spectrum parameters and one from stimulus template matching correlation coefficients. The performance of the three classification methods, namely Fine Tree, Linear Discriminant Analysis (LDA) and K-Nearest Neighbors (KNN), are compared using these feature vectors. Spectral features performed better than the template matching features. Especially the feature vector of the target frequency signal ratio (TFSR) to the total stimulation band energy features provided better accuracy values. LDA and KNN performed better than decision tree in classification.

Keywords: Classification, Feature Extraction, SSVEP BCI

This is an open access article under the CC BY-SA 4.0 license. (<https://creativecommons.org/licenses/by-sa/4.0/>)

1. Introduction

Visual Evoked Potentials (VEPs) are electrophysiological signals from the electrical activity of the visual cortex and recorded with electrodes placed over the scalp. There exist three VEP types depending on the stimulus type, namely Flash VEP, Pattern-onset/offset VEP, and Pattern-reversal VEP [1]. In Flash VEP paradigm, a brief luminance increment or flash is presented to the subject, and it produces a characteristic response with the most stable peaks occurring at N2 (90ms) and P2 (120ms) [2]. Normally this response is difficult to observe due to background EEG activity. It requires many repetitions of the experiment and synchronous averaging to enhance the characteristic VEP signal. If the stimulus repetition interval is long enough, greater than 250ms (stimulation rate < 4Hz), such that the visual system can return to a stationary initial state before the beginning of each stimulus, a brief input of luminance change creates a characteristic response known as Transient VEP [3][4][5]. Transient VEP has a number of characteristic peaks time-locked to the stimulus. For this reason, Transient VEP has been widely studied in Event Related Potential (ERP) studies and monitoring diseases of visual pathway. On the other hand, rapid visual stimulation with fixed stimulus repetition interval less than 250ms (stimulation rate >4Hz) generates a Steady-State Visual Evoked Potential (SSVEP) that contains a constant fundamental frequency due to periodic overlapping of each evoked potential peaks [6][7][8]. SSVEP spectra contain a major frequency the

same as the stimulation as well as its harmonics and subharmonics.

Brain-Computer Interface (BCI) is a system developed to provide its users ability to interact with the environment by translating the specific brain signals into desired words or actions to improve the quality of life. A BCI records electrical brain activity and links it to the external environment actions or internal body parts in order to improve the natural brain outputs [9]. BCIs have gained tremendous popularity in the last decade. Especially SSVEP BCIs gained much interest during the last decade due to some particular properties. The advantages of SSVEP BCIs include easy system configuration without a need for synchronization [10]. Neither a training nor an initial recording to generate a VEP template is required. It provides a high Information Transfer Rate (ITR) (usually >50bpm). Although the LCD screen refresh rate limits the number of frequencies available for targets, phase encoding provides more targets assigned with one frequency [11]. Another advantage is that it is implementable by common LCD screens and stimuli can vary from color, pattern to faces possible. In a multiple target SSVEP BCI, target detection is challenging due to mixing of frequencies. There exist time and spectral features for EEG classification, there lacks an ideal feature set for multi-target SSVEP classification application. Since the optimal selection of features give rise to the best performance, the potential feature extraction methods for SSVEP need to be studied in detail. This research aims to investigate the effects of three feature extraction mechanisms and three classification methods on the accuracy of the classification. These feature extraction methods employ different domain features, either

¹ Department of Biomedical Eng., Izmir Katip Çelebi University, Izmir, 35620, Turkey, ORCID ID : 0000-0003-0802-4376

* Corresponding Author Email: ibrahimkaya21@yahoo.com

spectral components or time domain template matching based correlation coefficient components.

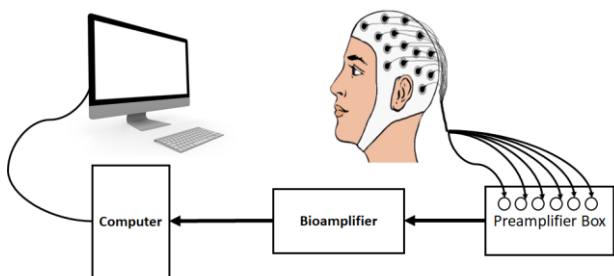


Fig. 1. A typical BCI recording system block diagram is shown.

In general, power spectral density (PSDA) or Signal to Noise Ratio (SNR) analysis of the target frequency is an easy and efficient solution. However, it can be further improved by incorporating the information in the other target frequencies in a feature vector as well. This study aims to illustrate the potential of this feature vector in SSVEP BCI classification.

2. Materials and Methods

2.1. Setup and Stimulation

A general SSVEP BCI recording system is shown in Figure 1. A stimulus matrix of 5 rows and 8 columns was created to present 40 characters to the subject using a 23.6 inch Acer GD245HQ LCD monitor [12]. The response time of the monitor was 2ms. Pixel resolution and refresh rate were 1920 x 1080 pixels and 60Hz. From 70cm away viewing distance, total speller area covered $34^\circ \times 24^\circ$ for horizontal and vertical axes respectively. Each target stimulus covered $3.2^\circ \times 3.2^\circ$ of visual field. The separation between stimulus targets were 1.14° for both horizontal and vertical axes. Joint frequency and phase modulation (JFPM) method was used to present the 40 visual target stimuli [13]. A sampled sinusoidal stimulation method was applied, the details can be found in [12][14][15]. Target frequencies to elicit SSVEP responses ranged from 8Hz to 15.8Hz with steps of 0.2Hz and target phases changed by multiples of $\pi/2$ as shown in Figure 2.

8Hz 0	9Hz 0.5π	10Hz π	11Hz 1.5π	12Hz 0	13Hz 0.5π	14Hz π	15Hz 1.5π
8.2Hz 0.5π	9.2Hz π	10.2Hz 1.5π	11.2Hz 0	12.2Hz 0.5π	13.2Hz π	14.2Hz 1.5π	15.2Hz 0
8.4Hz π	9.4Hz 1.5π	10.4Hz 0	11.4Hz 0.5π	12.4Hz π	13.4Hz 1.5π	14.4Hz 0	15.4Hz 0.5π
8.6Hz 1.5π	9.6Hz 0	10.6Hz 0.5π	11.6Hz π	12.6Hz 1.5π	13.6Hz 0	14.6Hz 0.5π	15.6Hz π
8.8Hz 0	9.8Hz 0.5π	10.8Hz π	11.8Hz 1.5π	12.8Hz 0	13.8Hz 0.5π	14.8Hz π	15.8Hz 1.5π

Fig. 2. 5x8 target stimulus matrix layout is shown. Each square represents one character and flickers with the frequency and phase given.

2.2. Dataset

The dataset is obtained from the University of Tsinghua Biomedical Engineering BCI Group webpage (<http://bci.med.tsinghua.edu.cn/download.html>). It contains data of 35 subjects tested on a BCI setup with a wide range of (40) target frequencies encoded with JFPM. For each subject, experimental paradigm consisted of gazing tasks at each targets.

One trial was recorded for each target. Each gazing and VEP recording task was repeated 6 times in one trial. Total number of epochs is 40 trials x 6 epochs/trial = 240 epochs. In each epoch, 5s simultaneous stimulation of all the targets on the screen was preceded by a 0.5s target cue. At the end of each epoch, there was a 0.5s rest interval. Total epoch duration was 6s. Each epoch was resulted from a 5s visual flicker stimulation of 60Hz LCD screen. Data was collected from 64 electrode channel EEG recording setup using international 10/20 EEG electrode system. Data was referenced to Cz. The sampling rate was 1kHz and a notch filter was applied to remove 50Hz power-line noise. To reduce data storage size, data was down-sampled to $f_s=250\text{Hz}$. Since the dataset includes both experienced and inexperienced subjects, the equal number of subjects (7subject) was selected for analysis for each group. The data for subject 5 was not available in the dataset and excluded from the analysis. The details of the subjects are shown in the Table-1.

Table 1. Subject properties

Subject ID	Gender	Age	Group
S1	Male	31	Experienced
S2	Male	29	Experienced
S3	Male	26	Experienced
S4	Female	25	Experienced
S6	Male	26	Experienced
S7	Female	28	Experienced
S8	Female	23	Experienced
S9	Female	31	Inexperienced
S10	Male	25	Inexperienced
S11	Female	18	Inexperienced
S12	Male	21	Inexperienced
S13	Female	17	Inexperienced
S14	Male	18	Inexperienced
S15	Female	19	Inexperienced

2.3. Data Pre-processing

In this study a laplacian spatial filtering of four channels is utilized such that the data is obtained from subtracting the average of O1 (61), O2 (63), and Pz (48) from the Oz (62) electrode recordings. Laplacian spatial filtering improves the BCI performance [16]. This spatial filtering reduced the common background EEG noise and enhanced the SSVEP response at the Oz channel. In addition to the spatial filtering a bandpass filter was applied. The filter was a 3rd order Butterworth band-pass filter with pass-band frequencies from 4Hz to 32Hz. The `filtfilt()` function in Matlab was used to prevent the phase delays in the signal due to filtering

2.4. Feature Extraction

Feature vectors based on three methods are compared in this study, namely time domain SSVEP template-matching cross-correlation coefficients (TMCC), PSDA (SNR) based feature vector, and target frequency power ratio (TFSR) in stimulation frequency band feature vector. Although there exists other time and spectral domain features in literature, their use in SSVEP detection have been found very limited in a preliminary analysis. Therefore, these basic features are excluded from further analysis. Each feature extraction method is described below.

2.4.1. Target Frequency SNRs as Feature Vector

PSDA based methods are the common target detection methods used in SSVEP BCIs [17][18][19]. In PSDA based target frequency detection, SNR in terms of energies of the each target frequency is computed using the following formula (1):

$$S(f_k) = \frac{n P(f_k)^2}{\sum_{m=1}^{n/2} P(f_k - m f_{res})^2 + P(f_k + m f_{res})^2} \quad (1)$$

where S is defined as the ratio of energy in k th frequency to the overall energy in the n neighbouring frequency bins. $P(f_k)$ is the amplitude FFT value at the k th frequency bin. f_k is the stimulus frequency, f_{res} is the frequency resolution of the FFT. $P=FFT(x)$ where $x(\)$ is the 1250-points temporal EEG data. Since the sampling rate is 250 Hz, f_{res} is $250/1250=0.2$ Hz. The number of neighboring frequencies is set to $n=5$ to include the -1Hz and 1Hz spectrum in the PSDA energy ratio $S(f_k)$ computation. Finally an output vector of 40 $S(f_k)$ values for each epoch is constructed as feature vector.

2.4.2. Target Frequency Signal Ratio (TFSR) in Stimulation Frequency Band Feature Vector

Target Frequency Signal Ratio (TFSR) is basically the ratio of the target frequency energy to the total stimulation frequency band energy. It incorporates the each target frequency energies which is not included in the SNR feature vector. The TFSR can be computed using (2) as follows:

$$TFSR(f_k) = \frac{N P(f_k)^2}{\sum_{m=0}^{N-1} P(f_k + m f_{res})^2} \quad (2)$$

where N is the number of total stimulus frequencies ($N=40$), k is the target frequency index for each target ($k=1:40$). A feature vector of all 40 TFSR is computed for each epoch for training and test data.

2.4.3. Template Matching Correlation Coefficient (TMCC) Feature Vector

Another method for target frequency detection is template matching based method. An SSVEP template is created for each target SSVEP using the stimulus properties. An example of an SSVEP recorded signal and template is shown in Figure 3.

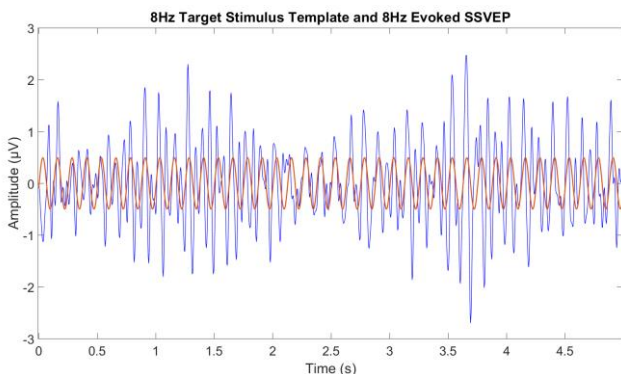


Fig. 3. SSVEP template from 8Hz target stimulus (red) and the corresponding recorded population average SSVEP (blue) are shown.

The SSVEP response is expected to be with the same stimulation frequency with a delay about 136ms since the flash VEP response first peak latency is about 136ms in [12]. Therefore, an SSVEP template $T_k(n)$ for each target frequency is created using (3).

$$T_k(n) = 0.5 \sin(2\pi f_k n + \theta_k - \Delta) \quad (3)$$

Here f_k is the target stimulus frequency given in Figure 2 and θ_k is the corresponding phase. The average latency delay (Δ) of the flash VEP is subtracted from the phase. TMCC feature vector is computed using correlation coefficient equation [19] as in (4).

$$TMCC(k) = \frac{\sum_{n=1}^N T_k(n) X(n)}{[\sum_{n=1}^N T_k(n)^2 \sum_{n=1}^N X(n)^2]^{1/2}} \quad (4)$$

2.5. Classification

In the classification stage, three classifiers are used namely; decision tree (fine tree), Discriminant Analysis (LDA) and KNN. To avoid possible errors due to the non-homogenous distributions in the data, a 5-fold cross validation is applied to the classification learning stage. The data is divided into 5 random partitions, among these 4 partitions are used in the training and 1 partition is used in the test process, this is repeated 5 times. The performance outputs at the end of classification with each data division process are gathered. The arithmetic mean of the performance outputs from each classification process is obtained as the final metric, particularly training accuracy. Two third of the data is used in training. For evaluation of the performance with unseen data, remaining one third of the data is used as test data. Performance assessment is done using accuracy metric which can be computed using (5).

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (5)$$

where TP, TN, FP and FN represent true positive, true negative, false positive and false negative classification counts respectively.

2.5.1. Decision Tree

Decision tree based classification simplifies a complex classification process into a collection of simpler decisions, [20]. Decision tree classifiers try to have the simplest structure, classify the most of the training data accurately as possible, achieve a generalization so as to perform good with unseen data, be updated easily with more training [20]. Fine tree method is selected for the advantages of fast implementation, high flexibility, and small memory usage. In fine tree many leafs are used in order to have a fine separation between classes .

2.5.2. LDA

LDA classifies the training feature vectors of different classes using hyperplanes [21][22]. For a two-class classification, the classes are defined using a linear discrimination function representing a hyperplane boundary in the feature space. The class to which the feature vector belongs depends on the side of the plane where the vector is found. To classify more than two classes, e.g. N-class problem ($N > 2$), several hyperplanes are used. For multiclass BCI classification one vs all separation is repeated for each class to find the correct class among all. Using the training data, the location and orientation of the hyperplane are set [24]. One decision plane can be defined as $g(x)$:

$$g(x) = w^T x + w_0 \quad (6)$$

where w is the weight vector, x is the input feature vector and w_0 is a threshold. The sign of $g(x)$ determines the input feature vector class [24]. This indicates which side or class the input belongs to.

LDA classifier is a simple low complexity classifier but efficient. It does not need high computation power yet achieves good accuracy for BCI classification purposes especially for online BCI systems [25][26]. Simplicity of this classifier makes it good at generalizing to unseen test data and yields good accuracy values in practice [26].

2.5.3. KNN

In K-Nearest Neighbours as the name suggests the unseen test sample class is set to the dominant class among its k nearest neighbours within the training set [22]. It relies on the principle that features related to each different class forms a separate cluster in feature space [23]. The power of KNN is that if a high k and enough training samples are used it can approximate any function, even nonlinear decision boundaries [26].

3. Results and Discussion

The accuracy for each subject, feature and classification method combination was calculated. Figure-4 illustrates accuracy plots for each feature using Fine Tree (A), LDA (B), and KNN (C) classifiers for all 14 subjects. Mean classifier training times are 0.66s, 0.48s, and 0.47s for Fine Tree, LDA and KNN respectively. An ANOVA with generalized linear model has been carried out to compare the effects of each feature, classifier, and the experience on the accuracy performance. This analysis revealed all three factors have significant effect on classification accuracy ($p < 0.01$). While all three classifiers mean accuracies for experienced subjects are 81.4%, 64.8%, and 86.3% for SNR, TMCC, and TFSR respectively, inexperienced subject accuracies

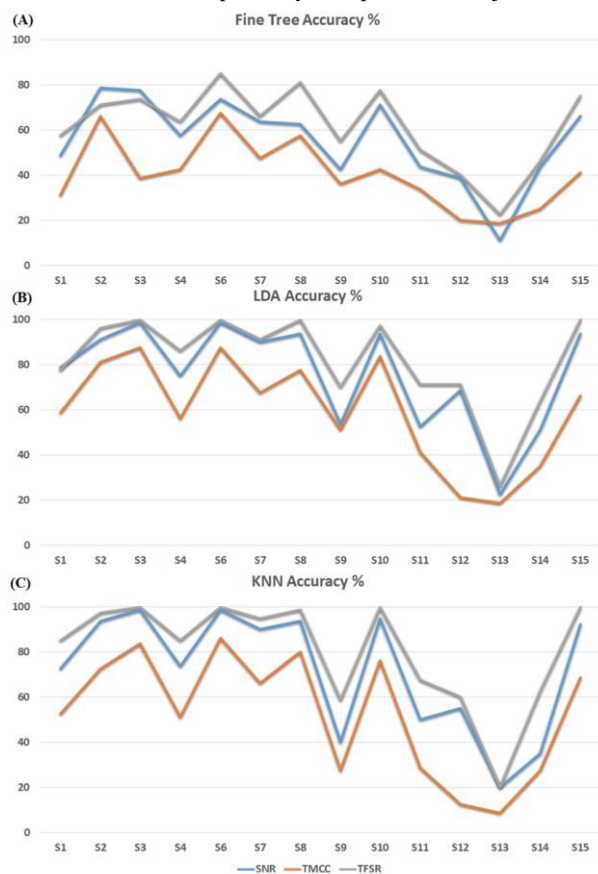


Fig. 4. Classification accuracies for each subject and each features of SNR, TMCC and TFSR using Fine Tree (A), LDA (B), and KNN (C) classifiers are depicted.

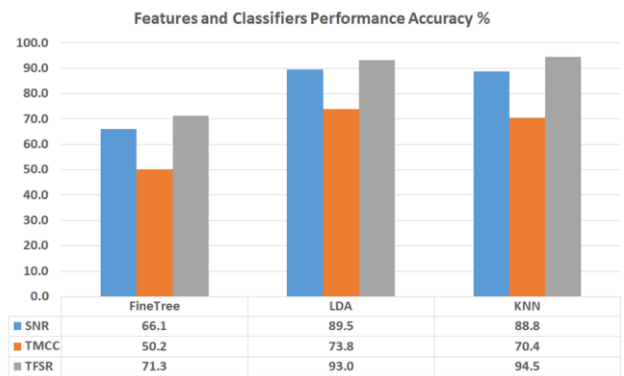


Fig. 5. Comparison chart of average performances for each classifier and feature for experienced subject group is illustrated.

are 54.3%, 37.4%, and 63.6%. Since there is a huge difference between inexperienced and experienced subject performances, I carried out further comparison analyses for each separate groups of experience independently. TFSR combined with KNN or LDA achieved the highest accuracy compared to other features and classifiers ($p < 0.05$), see Figure-5. TFSR significantly improved accuracies compared to SNR and TMCC by 6% and 33% respectively for experienced subjects. On the other hand, for inexperienced subjects 17% and 70% improvement is observed when TFSR is used instead of SNR or TMCC respectively. The correlation coefficient between averages performances of each set of features are computed. There is strong correlation between each feature performances (0.98 for TFSR-SNR, 0.93 for TFSR-TMCC, 0.93 for SNR-TMCC), however the TFSR has the highest performance in classification and can be used in SSVEP BCI classification tasks.

In conclusion, the inclusion of each target frequency signal ratio to the total stimulation frequency band energy is a promising feature extraction method to achieve higher accuracy. It is easy and efficient. It can replace the traditional PSDA methods. In addition, LDA and KNN classifiers provide acceptable accuracies, and high and medium prediction speeds respectively, see Figure-5. Moreover, these performances are obtained using only 4 electrodes which is feasible for a typical BCI realization for daily-life usage. For future studies, the performance of the proposed TFSR feature with other datasets and classification methods can be analyzed.

References

- [1] J. V. Odom et al., "ISCEV standard for clinical visual evoked potentials:(2016 update)," *Doc. Ophthalmol.*, vol. 133, no. 1, pp. 1-9, Aug. 2016.
- [2] F. B. Vialatte et al., "Steady-state visually evoked potentials: focus on essential paradigms and future perspectives," *Prog. in Neurobiol.*, vol. 90, no. 4, pp. 418-38, Apr. 1, 2010.
- [3] Ö. Özdamar, J. Bohórquez, "Signal-to-noise ratio and frequency analysis of continuous loop averaging deconvolution (CLAD) of overlapping evoked potentials," *JASA*, vol. 119, no. 1, pp. 429-38, Jan. 2006.
- [4] A. Capilla et al., "Steady-state visual evoked potentials can be explained by temporal superposition of transient event-related responses," *PloS one*, vol. 6, no. 1, e14543, Jan. 18 2011.
- [5] P. L. Nunez, "Electric and magnetic fields produced by the brain," in *Brain-Computer Interfaces: Principles and Practice*, pp. 171-212, Jan 24 2012.
- [6] G. G. Celesia, "Steady-state and transient visual evoked potentials in clinical practice," *Ann. N. Y. Acad. Sci.*, vol. 388, no. 1, pp 290-305, Jun. 1982.
- [7] D. Regan, "Human brain electrophysiology. Evoked potentials and

- evoked magnetic fields in science and medicine.” New York, Elsevier, 1989.
- [8] D. Regan, “Comparison of transient and steady-state methods,” *Ann. N. Y. Acad. Sci.*, vol. 388, no. 1, pp. 45-71, Jun. 1982.
- [9] J. R. Wolpaw, E. W. Wolpaw. “Brain-computer interfaces: something new under the sun,” in *Brain-computer interfaces: principles and practice*, 14, Jan. 24 2012.
- [10] G. Bin et al., “An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method,” *J. Neural Eng.*, vol. 6, no. 4, 046002, Jun. 3 2009.
- [11] Y. Wang et al., “Brain-computer interfaces based on visual evoked potentials,” *IEEE Eng. Med. Biol. Mag.*, vol. 27, no. 5, pp. 64-71, Sep. 16 2008.
- [12] Y. Wang et al., “A benchmark dataset for SSVEP-based brain-computer interfaces,” *IEEE Trans. Neural Syst.*, vol. 10, no. 25, pp. 1746-52, Nov. 10 2016.
- [13] Y. Zhang et al.. “Multivariate synchronization index for frequency recognition of SSVEP-based brain-computer interface,” *J. Neurosci. Methods*. Issue: 221, pp. 32-40, Jan. 15 2014.
- [14] X. Chen et al., “A high-itr ssvpe-based bci speller,” *Brain-Computer Interfaces*. vol. 1, issue 3-4, pp. 181-191, Oct. 2 2014.
- [15] N. V. Manyakov et al.. “Sampled sinusoidal stimulation profile and multichannel fuzzy logic classification for monitor-based phase-coded SSVEP brain-computer interfacing,” *J. Neural Eng.*, vol. 10, no. 3, 036011, Apr 18 2013.
- [16] D. J. McFarland “The advantages of the surface Laplacian in brain-computer interface research.” *Int. J. Psychophysiol.*, vol. 97, no. 3, pp. 271-276. Sep 1 2015.
- [17] J. Castillo-Garcia et al., “Comparison among feature extraction techniques based on power spectrum for a SSVEP-BCI,” In 2014 12th IEEE International Conference on Industrial Informatics (INDIN) , pp. 284-288. IEEE. July 2014.
- [18] Q. Wei, M. Xiao, Z. Lu, “A comparative study of canonical correlation analysis and power spectral density analysis for SSVEP detection,” In 2011 Third International Conference on Intelligent Human-Machine Systems and Cybernetics, vol. 2, pp. 7-10. IEEE. August 2011.
- [19] Richard M. G. Tello et al., "A comparison of techniques and technologies for SSVEP classification," 5th ISSNIP-IEEE Biosignals and Biorobotics Conference (2014): Biosignals and Robotics for Better and Safer Living (BRC). IEEE, 2014.
- [20] M. T. Puth, M. Neuhäuser, G.D. Ruxton, “Effective use of Pearson’s product-moment correlation coefficient,” *Animal behaviour*. vol. 93, pp. 183-189, Jul. 1 2014.
- [21] S. R. Safavian, D. Landgrebe, “A survey of decision tree classifier methodology,” *IEEE Trans. Syst., Man, Cybern.*, vol. 21, no. 3, pp. 660-674, May 1991.
- [22] R. O. Duda, P. E. Hart, D. G. Stork, *Pattern Classification*. John Willey and Sons, New York. 2001.
- [23] K. Fukunaga. *Statistical Pattern Recognition* 2nd edn (New York: Academic), 1990.
- [24] F. Lotte, “A tutorial on EEG signal-processing techniques for mental-state recognition in brain-computer interfaces,” In: Miranda E., Castet J. (eds) *Guide to Brain-Computer Music Interfacing*. Springer, London. pp. 133-161, 2014.
- [25] L. F. Nicolas-Alonso, J. Gomez-Gil. “Brain computer interfaces, a review,” *sensors*, vol. 12, no. 2, pp. 1211-1279, Feb. 2012.
- [26] F. Lotte , “A review of classification algorithms for EEG-based brain-computer interfaces,” *J. Neural Eng.* vol. 4, no. 2, R1. Jan. 31 2007.