Original Article ICNSJ
International Clinical Neuroscience Journal

SSVEP Extraction Applying Wavelet Transform and Decision Tree with Bays Classification

Hoda Heidari, Zahra Einalou

Department of Biomedical Engineering, North Tehran Branch, Islamic Azad University, Tehran, Iran

ABSTRACT

Background: SSVEP signals are usable in BCI systems (Brain-Computer interface) in order to make the paralysis movement more comfortable via his Wheelchair.

Methods: In this study, we extracted The SSVEP from EEG signals, next we attained the features from it then we ranked them to obtain the best features among all feature and at the end we applied the selected features to classify them. We want to show the degree of accuracy we applied in this work.

Results: In this study Bayes (applied for classifying of selected features) got the highest level of accuracy (83.32%) with t-test method, until the SVM took the next place of having the highest accuracy to itself with t-test method (79.62%). In the next place according to the feature selection method, decision tree took the next place with Bayes classification (79.13%) and then with SVM classification (78.70%).

Conclusion: Bays obtained the better results to itself rather than SVM with t-test.

Keywords: Brian Computer Interface; Steady State Visual Evoked Potentials; Bays classification

ICNSJ 2017; 4 (3) :91-97 www.journals.sbmu.ac.ir/neuroscience

Correspondence to: Zahra Einalou, PhD, Department of Biomedical Engineering, North Tehran Branch, Islamic Azad University, Tehran, Iran; Tel;+98 (21)-77009836; E-mail: z.einalou@srbiau.ac.ir

Received: Jun 6, 2017 **Accepted:** Jun 28, 2017

INTRODUCTION

People who suffer from server disabilities cannot connect to the world. In many cases, these people have the normal brain function and they can see, hear, understand and be able to talk but they do not have any way to react. In the field of rehabilitation engineering, this idea causes to invent the Brain–Computer Interface (BCI) system to recover the lost functions. One of the objectives of this technology is to help people with server movement disorders, which cause them not to do their daily activities and all of us need to use muscles to do them. In recent decades, many studies have been started in BCI system and have been built some cases. These systems have many differences in their inputs, their extraction methods and features' interpretation, outputs, and other characters such as accuracy and speed [1].

BCI is a connecting or controlling system which is not related to the output of muscles circumferential nerves [2]. It interprets the Brain signals (EEG) to the signals which control the external instruments [3]. Visual Evoked potential are signals which have been created by a visual simulation and also has been occurred in occipital region. Steady-State Visual Evoked Potential is a reaction to a visual simulation with a specified frequency. SSVEP is described increasing EEG activity at stimulation frequency [1]. SSVEP BCI systems use the reaction to a blink stimulation in visual cortex and this provides more efficiency with the lowest training time. The benefits of these systems are the high information transfer rate (ITR) and comfortable preparation of system [4]. The visual evoked potential (VEP) is a potential which has been simulated by a visual stimulant. VEPs which are recorded from visual cortex of skull, reflect the data processing of brightness [5]. There is one difference between impermanent and permanent VEP. Impermanent VEP is accrued when Electrical stimulations of sightedness system are. If the repetitions of stimulations are faster than 6 Hz, the reactions Start to combine and make SSVEP. The SSVEP signal is a repetition signal which its discrete frequency parameters can keep fixing their phase and amplitude in a long time period [6].

FEATURE EXTRACTION AND CLASSIFICATION METHOD

Data Acquisition

The 10-20 international system was applied in this study to determine the place of electrodes on the scalp as shown in the figure 1 [7].

For recording SSVEP in this study, three (3) experiments were done. In experiments 1 and 2, 128 active electrodes were with 2048 HZ sampling rate. In experiment three (3), the 256 HZ was applied which was a lower sampling rate. Inverted T-shape figure was used with 5 occipital electrodes. Due to ocular artifacts, FZ (an electrode located at the front) was used [8].

In this study, Discrete Wavelet Transform (DWT) has been applied as a method for feature extraction of SSVEP signals and it decomposes the signal in the 4 level of decomposing.

Wavelet Transform

Wavelets are a complex preliminary wave shapes which is decomposed an input signal. The signal analysis is prepared by wavelet coefficients. Wavelet transform is a time- frequency method which is usable for non-stationary signals such as EEG. Discrete Wavelet transform (DWT) decomposed the signal into "a coarse approximation and detail information" which can make us analyze the signal. The two collections of filters are decomposed by DWT which names are high pass ("wavelet functions") and low pass ("scaling functions"). High pass and low pass filters make the signal be decomposed into different frequency bands of time domain. These bands are called sub-bands. The approximation coefficients are the products of low-pass filter (Ca1 at the first decomposition level in Figure 2) and the detail coefficients are the products of high- pass filter (Cd1 at the first decomposition level in Figure 2). It is very important to recognize the number of levels of decomposition and suitable wavelet. Frequency components determine the number of decomposing and the parts of signals have been selected which are correlated with needed frequencies in the classification of signal [9,21,22].

At the first level of decomposing, samples have been passed from the low pass fitter with g response (convolution of x and g) [10,11].

$$y[n] = (x * g) = \sum_{k=-\infty}^{\infty} x[k]g[n-k] \tag{1}$$

Due to eliminating the half of signal frequencies, the half of samples can be failed by Naikouist rule. So, the outputs of filter are sampled with rate 2 as the following (g is related to high pass filter, h is related low pass filter) [10,11].

$$y_{low}[n] = \sum_{k=-\infty}^{\infty} x[k]h[2\,n-k] \tag{2}$$

$$y_{high}[n] = \sum_{k=-\infty}^{\infty} x[k]g[2\,n-k] \tag{3}$$

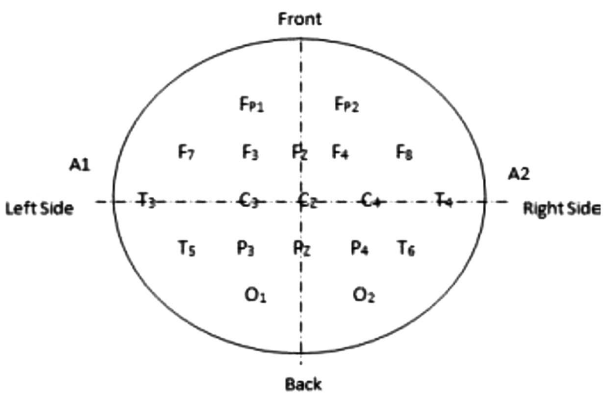

Figure 1. The 10-20 System of Electrode Placement [7].

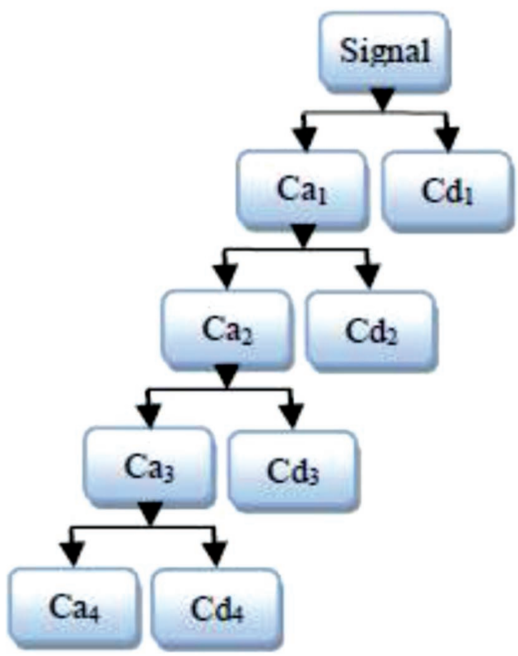

Figure 2. Decomposition of the signal in four levels [9].

Feature Extraction

Entropy Shannon

This Entropy is the famous Entropy among all kinds of Entropy. If p(x) is the probability x, the Shannon Entropy has been defined in for the continuous case [12].

$$H(x) = -\int_{-\infty}^{+\infty} P(x)\log_2(P(x))dx \quad (4)$$

The maximum of p(x) is occurred in continues case when the signal has the halberd distribution [12].

For the discrete case, p(x) has been considered as below [12]:

$$H(x) = -\sum_x (P(x)\log_2(P(x))) \quad (5)$$

The maximum of p(x) is happened in discrete case when the signal has the monotonous distribution. In practical, the repetition of values in the various slopes from considered signal can lead to achieving the spectrum of probability density for Entropy Shannon [12].

Ranking Features

t-test

T-test compares the mean and criterion deviation between two samples to be recognized that whether it is a meaningful difference between them or not. To salsify the previous subject, hypothesizes H0 and H1 should be considered then figure t should be calculated. H0 is a hypothesis which should determine the non-difference or similarity in the considered society. H1 is a hypothesis which is a reverse of H0 and shows the claim of researchers. t is calculated as following [13,23].

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{(s_1)^2}{n_1} + \frac{(s_2)^2}{n_2}}} \quad (6)$$

$\bar{x}_1$: The mean of the first sample
s_1: Criterion deviation of the first sample
n_1: The number of people in the first sample
$\bar{x}_2$: The mean of the second sample
s_2: Criterion deviation of the second sample
n_2: The number of people in the second sample

Then the calculated t compares with the distribution table of t. We need the freedom degree which is obtained as the following [13]:

$$df = (n_1 + n_2) - 2 \quad (7)$$

If the calculated t is lower than the figure t in the table, H0 will be accepted. And if the calculated t is higher than the figure t in the table, H1 will be accepted [13].

Decision tree

This is an interesting way of classification. The complex of decision nodes which are connected to each other by the branches, and branches go down from root nodes until they come to leaf nodes. The start point is from root which is at the top of decision tree according to the agreement standard. Indicators have been examined in the decision nodes and each result is placed in branches. Each branch leads to another final leaf node or another decision node. The figure 3 can depict the simple decision tree in the subject of risk and financial case [14].

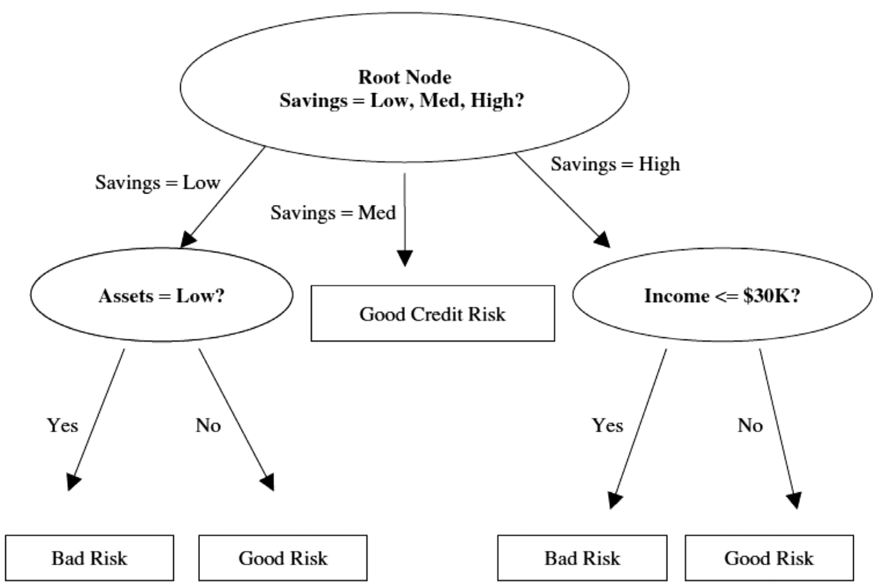

Figure 3. simple decision tree [14].

Bhattacharyya

Bhattacharyya is a feature selection or weighted class which is applied when we face a lack of features in the smallest set of classes for classifications from multiple class. The coefficient of Bhattacharyya is a novel way to sort and select those members iteratively which belongs to the minimum surface and the surface is defined as C (C- 1)/ 2 and pairs of C classes are a specified surface. Figure 4 illustrates three features' surfaces. Features x, y and z are defined in this figure with the means of 0.2, 0.5 and 0.6782 respectively. The best separation of the pair classes is done by the mean of surfaces. In fact, average or median is used as the mean of Bhattacharyya. There is no intelligent way to apply the best ordering process. Applying the minimum surface created by selecting features (for example in the first iteration x and z have been chose) helps strictly to optimize the separation of features ordering. Therefore, the separation of features is well done without needing to consider all combinations of features. X, z and y are the order of features which are selected by minimum surface respectively. The median values specify the minimum surface when all features put at the initial minimum surface [15].

Classification

Support Vector Machine (SVM)

Pattern recognition and classification cases are the two most important applications of support vector machine (SVM). It defines a hyperplane to divide all data belonged to one class from those of other class. The best hyperplane is the one to make the largest border between two classes. Border is defined as the largest width of the slab which parallels with the hyperplane and there is no data point in it. The support vectors are those which are near the hyperplane and they are on the boundary of slab. The addition of specification of SVM is to transfer data to the higher dimension space in order to optimize the best hyperplane. SVM applies the means of Kernel function to make a higher dimension space of features over this nonlinear mapping, then it builds the linear hyperplane which has been optimized separation between two classes. Gaussian, Radial Basis Functions and polynomials are the common way to apply Kernels [9,20].

SVM method has the less overfitting in contrasting to other ways. The SVM is depicted as figure 5 [16].

Bayes

The principle of this classification which is famous for the best classification is to compare the conditional happening probabilities in the classes. It considers that the data of x with its unknown class belongs to which possible classes and with what probability [13]. Each class results the highest probability recognized as the related class of data x. According to the mathematics' point of view, it defines as below [13]:

$$P(x|w1) > P(x|w2) \quad x \in w1 \tag{8}$$

$$P(x|w1) < P(x|w2) \quad x \in w2 \tag{9}$$

W_1 and W_2 are the possible classes and P shows the probability [13].

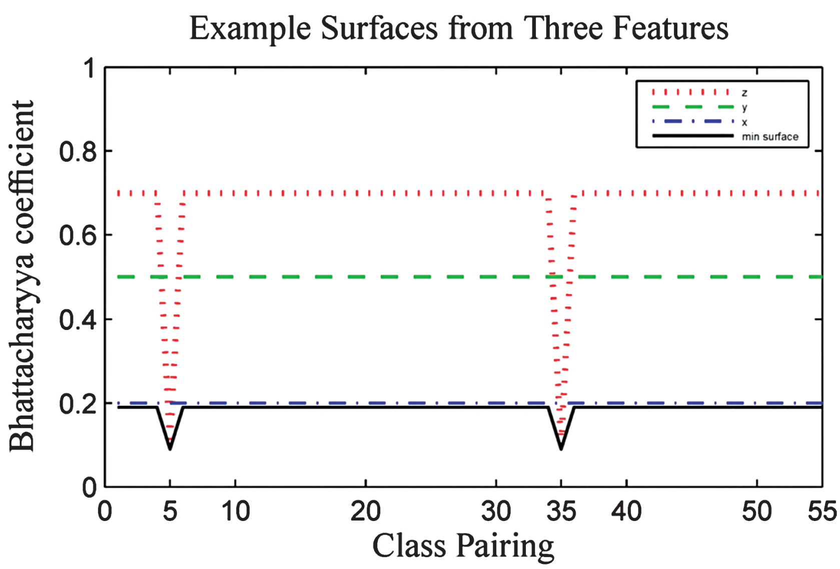

Figure 4. The mean performance for this example is 0.2, 0.5, and 0.6782 for *x, y,* and *z,* respectively. The current minimum surface consists mostly of feature *x* and two class pairs of feature *z* (5 and 35) [15].

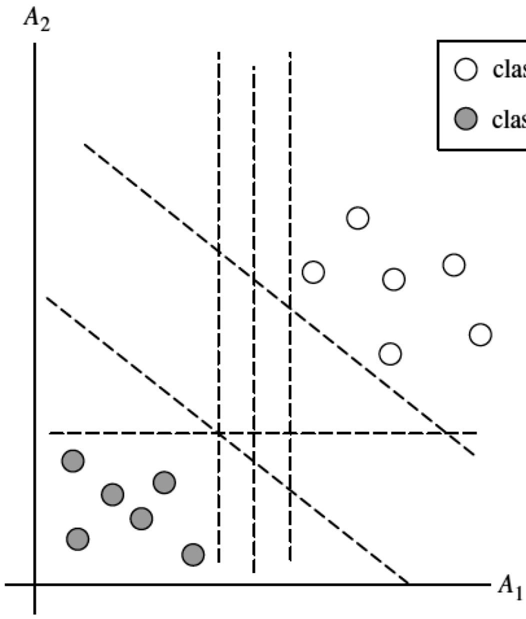

Figure 5. Support Vector Machine [16].

EXPERIMENTAL RESULT AND DISCUSSIONS

Dataset

Bioseme Inc. Has been used to record SSVEP signals which is built in nether lands. Data has been recorded in Brain Science Institute, Laboratory for Advanced Brain Signal Processing RIKEN. EEC signal has been registered with four (4) subjects from 128 active electrodes. The subjects were all aware of the purpose of the experiment. The high sensitive test was done for each subject toward epilepsy disease. The SSVEP stimulation was accomplished with black and white checkered sheets (6×6 screen). EEC signal was extracted and removed with central electrode (CZ), winking and artifacts of muscle to after referencing once more. The stimulation was done with a small checkered sheet for three (3) frequencies (8, 14, 28 HZ) in the second experiment [8].

The sampling frequency of signals was 256 HZ. The participants were seated at the 90 cm of monitor. The start of SSVEP is 5 seconds after the start of data and its finished point was 20 seconds from the start of data. The 15 seconds of SSVEP is available from four (4) subjects [8].

This data includes SSVEP and it doesn't contain EEG signals. Stimulation is done at three frequencies (8, 14, 28, HZ) and each frequency includes 5 experiments for each participants. The length of recorded data is about 24.8s [8]. For instance, the wave shape is depicted as figures 6 and 7 for the first subject in the first experiment at 28 HZ frequency. The vertical axis in terms of signal amplitude and the horizontal axis in terms of samples.

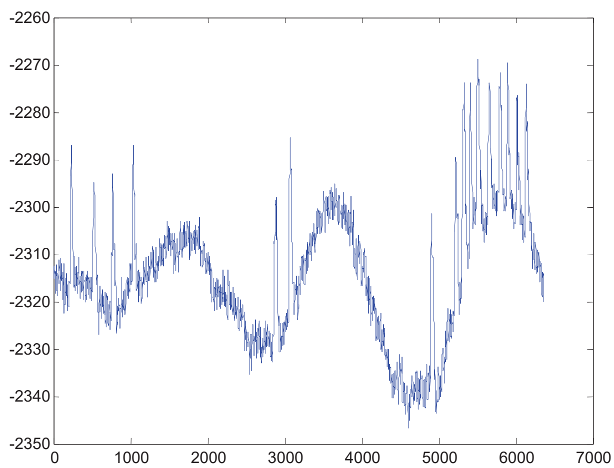

Figure 6. The first signal from the first participant at 28HZ.

Processing

The SSVEP signals are decomposed with Wavelet transform (DWT) in four (4) levels of decomposing. Features are ready to be extracted at this level including entropy Shannon. We normalize futures and then we apply decision tree, Bhattacharyya and t-test. The Bays and SVM will classify our data.

RESULTS

The selected features were rated by ranking methods applied before (including t-test and Bhattacharyya). The

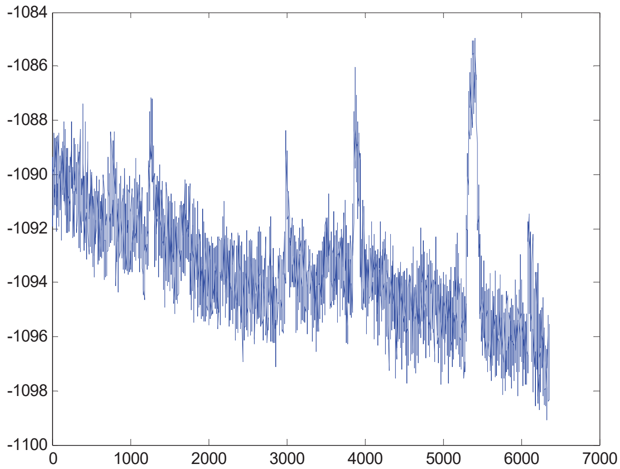

Figure 7. The first signal from the second participant at 8HZ.

three (3) tabled below show the rate of features (Table 1).

After rating futures, all data divided into two branches (test and training). The training group of date was given to SVM and Bayes classifications. The test group of data was kept for considering the classifications. The classification was done with four (4) classes (Table 2-5).

Table 1. The normal ranking scale in comparison with other frequencies.

	t-test		Bhattacharyya	
	Number of feature	Rate of *P*-value	Number of feature	Score
1	2086	3.44×10^{-13}	1025	Inf
2	2405	3.35×10^{-13}	1026	0.018852
3	2165	3.35×10^{-13}	1027	0.018852
4	2421	3.22×10^{-13}	1	7.04×10^{-7}
5	2169	3.22×10^{-13}	2	7.04×10^{-7}

Table 2. The rank scaling of 8HZ in comparison with other frequencies.

	t-test		Bhattacharyya	
	Number of feature	Rate of *P*-value	Number of feature	Score
1	2158	2.6×10^{-13}	1025	Inf
2	2161	2.29×10^{-13}	1027	Inf
3	2240	4.8433×10^{-14}	1026	0.004
4	2384	2.2095×10^{-13}	2	2.52×10^{-7}
5	2138	2.133×10^{-13}	3	2.52×10^{-7}

Table 3. The rank scaling of 14 HZ in comparison with other frequencies.

	t-test		Bhattacharyya	
	Number of feature	Rate of *P*-value	Number of feature	Score
1	2331	3.27×10^{-13}	1025	Inf
2	2058	2.49×10^{-13}	1026	Inf
3	2413	2.16×10^{-13}	1027	0.004087
4	2327	2.1×10^{-13}	1	2.52×10^{-7}
5	2137	2.304×10^{-13}	2	2.52×10^{-7}

Table 4. the rank scaling of 28HZ in comparison with other frequencies.

	t-test		Bhattacharyya	
	Number of feature	Rate of *P*-value	Number of feature	Score
1	2405	3.48×10^{-13}	1025	Inf
2	2141	2.9×10^{-13}	1026	Inf
3	2180	2.8×10^{-13}	1027	Inf
4	2421	2.6322×10^{-13}	1	2.52×10^{-7}
5	2169	2.47×10^{-13}	2	2.52×10^{-7}

Table 5. the results of all classifications with feature selections methods.

Feature selection method	Classification	Accuracy
t-test	SVM	79.62±3.25
t-test	Bayes	83.32±3.41
Bhattacharyya	SVM	65.11±3.24
Bhattacharyya	Bayes	73.01±3.23
Decision tree	SVM	78.70±0.18
Decision tree	Bayes	79.13±0.12

The three (3) of them belong to 8, 14, 28 HZ and the another one is related to the normal class. SVM was trained with linear and non-linear of Kernal (with degree three (3)).

The results show that SVM with non-linear Kernal was better than its linear one. The all results of classifications and ranking methods are in the table as following:

According to the above table, the undermentioned results have been received us below:

- Bayes has the better operation with accuracy of 83.32%.
- Bhattacharyya has a lower operation in contrast to other classifications.

DISCUSSION

The results of this study show that the taken accuracies by SVM and Bayes are higher than other resembled studies. Korn and Aunon reached to 53.98 in 2006 with applying SVM [18]. It is hard to compare our study to others because of the level of decomposing in wavelet, participates situations, EEC signals [17]. we used linear and non-linear classification and it brings a point to our study because we can understand the difference of these two ways. The non-linear classifications are hard to be applied because they are hare to lime consuming and over fitting caused [19].

CONCLUSION

Seeing the results in the tables one to five show that the highest accuracy was about 83.32% which is related to Bayes classification. Bays obtained the better results to itself rather than SVM with t-test and then decision tree took the next place to itself as the best ranking feature method after t-test.

ACKNOWLEDEMENT

The authors would like to thank God for his invaluable help.

REFERENCES

1. Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM. Brain-computer interfaces for communication and control. Clin Neurophysiol. 2002 Jun;113(6):767-91.
2. Wolpaw JR, Birbaumer N, Heetderks WJ, McFarland DJ, Peckham PH, Schalk Get al. Brain computer interface technology: A review of the first international meeting. IEEE Trans Rehab Eng. 2000;8:164–173.
3. Kübler A, Kotchoubey B, Kaiser J, Wolpaw JR, Birbaumer N. Brain-computer communication: unlocking the locked in. Psychol Bull. 2001 May;127(3):358-75.
4. Wang Y, Wang R, Gao X, Hong B, Gao S. A practical VEP-based brain-computer interface. IEEE Trans Neural Syst Rehabil Eng. 2006 Jun;14(2):234-9.
5. E. Donchin, W. Ritter, and W.C. McCallum. Event-related brain potentials in man, volume Cognitive psychophysiology: the endogenous components of the ERP. Academic Press; 1978.
6. D. Regan, Human Brain Electrophysiology: Evoked potentials and evoked magnetic fields in science and medicine. North Holland, The Netherlands: Elsevier; 1987.
7. Lin CJ, Hsieh MH. Classification of mental task from EEG data using neural networks based on particle swarm optimization. Neurocomputing. 2009;75:1121-30.
8. Bakardjian H, Tanaka T, Cichocki A, Optimization of SSVEP brain responses with application to eight-command Brain–Computer Interface, Neurosci Lett. 2010; 469(1):34-38.
9. Talwar TS, Matharu SS. Classification of SSVEP Based Brain Signals using Discrete Wavelet Transform. International Journal for Research in Applied Science & Engineering Technology (IJRASET). 2016; 4(5), 421-426.
10. Mallat SA wavelet tour of signal processing: the sparse way. Academic press; 2008.
11. Akansu A N, Haddad RA. Multiresolution signal decomposition: transforms, subbands, and wavelets. Academic Press; 2001.
12. Tong S, Bezerianos A, Malhotra A, Zhu Y, Thakor N. "Parameterized entropy analysis of EEG following hypoxic–ischemic brain injury," Physics Letters A. 2003;314, 354–361.
13. Luger, G.F. Artificial Intelligence. Addison Wesley; 2005.
14. Larose DT. Discovering Knowledge in Data: An Introduction to Data Mining. USA: Wiley & Sons; 2005.
15. Gonzalez JA, Mendenhall MJ, Merenyi E. Minimum Surface Bhattacharyya Feature Selection. In Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, 2009. WHISPERS'09. First Workshop on (pp. 1-4). IEEE.
16. Chen YF, Atal K, Xie SQ, Liu Q. A new multivariate empirical mode decomposition method for improving the performance of SSVEP-based brain-computer interface. J Neural Eng. 2017 Mar 30;14(4):046028.
17. Amin HU, Malik AS, Ahmad RF, Badruddin N, Kamel N, Hussain M, et al. Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques. Australas Phys Eng Sci Med. 2015 Mar;38(1):139-49.
18. Liang N-Y, Saratchandran P, Huang G-B, Sundararajan N. Classification of mental tasks from EEG signals using extreme learning machine. Int J Neural Syst. 2006; 16:29–38.
19. Witten IH, Frank E. Data mining: practical machine learning tools and techniques. Morgan Kaufmann, Burlington; 2005.