IOP Conference Series: Materials Science and Engineering

PAPER • **OPEN ACCESS**

Enhancement of Motor Imagery Brain Computer Interface Performance Using Channel Reduction Method based on Statistical Parameters

To cite this article: Mohd Shuhanaz Zanar Azalan *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **557** 012016

View the article online for updates and enhancements.

This content was downloaded from IP address 121.15.254.156 on 17/10/2019 at 14:58

IOP Conf. Series: Materials Science and Engineering **557** (2019) 012016 doi:10.1088/1757-899X/557/1/012016

Enhancement of Motor Imagery Brain Computer Interface Performance Using Channel Reduction Method based on Statistical Parameters

Mohd Shuhanaz Zanar Azalan[1], M P Paulraj[2], Abdul Hamid Adom[1]

[1]School of Mechatronic Engineering, Universiti Malaysia Perlis, Arau, Perlis, Malaysia [2]Sri Ramakrishna Institute of Technology, Coimbatore, Tamilnadu, India

Abstract. In this paper, a novel method to reduce the number of EEG channels for a Motor Imagery-based Brain Computer Interfaced (BCI) system without compromising its performance is proposed. By reducing the number of EEG channels, the number of features can be reduced and this has to be achieved without sacrificing the classification accuracy and computational time of the BCI. EEG signals were recorded from 10 subjects using a 19-channel EEG amplifier. Higuchi Fractal features were extracted from the recorded signals and modelled using Neural Networks (NN). A simple statistical analysis based on standard deviation was then used for the channel reduction process. The classification accuracy of the NN model formulated with the 19 channels features were compared to that of the model with features selected using statistical method. From the results it was observed that using this approach, the number of EEG channels can be reduced up to 30% without sacrificing its classification performance.

1. Introduction

A BCI is a system that acts as link for the brain signal to communicate with computer system without going through a usual route of peripheral nerves and muscles [1]. Research about BCI have been explored for over two decades, and is considered as a new research area. The interest in BCI has been increasing rapidly and this is evidence in the number of publications on the subject matter. The group of researcher who had done a research about BCI also increase year to year [2].

However, there are still great opportunities in this research area. Many improvements in many aspects of BCI performance can be done for example by removing features that may act as noise to the system through channel selection process. New channel selection algorithm based on high order statistical parameters using motor imagery data is presented in this paper. The proposed channel selection algorithm was used to reduce the feature dimension in order to enhance the classification performance of the recorded motor imagery data features.

The motor imagery data are based on a novel asynchronous data collection protocol. Motor imagery can be defined as the process to carry out pretended movement of arm or other parts of human body. The concept of pretended movement of arm can be defined such as preparation for movement, passive observations of action, and mental operations of sensorimotor representations [3]. The selection of channel is used to reduce the size of data dimension and has been widely used to increase the performance of BCI systems. This refers to the process of choosing a subset of variables from the

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

original dataset such that the optimal features are retained. Some researchers claim that the selection of variables will improve classification accuracy or decreasing the size of features without significantly decreasing the accuracy of the classifiers [4].

In order to find suitable channel, S. M. Park et al.[5] study four channel selection methods namely binary particle swarm optimization (BPSO), BPSO with a channel impact factor, a genetic algorithm (GA), and a harmony search (HS) using a support vector machine (SVM).The performance for each method are compared in term of the classification accuracy and the number of selected electrodes. The results showed that BPSO with a channel impact factor selected least number of electrodes and improved accuracy by 20.4%.

Hossain et al.[6] proposed channel selection approach using wavelet-based Maximum Entropy on the Mean (wMEM). They used the concept of sparsity assumption. The channels located around the areas with high MI-related amplitude is selected. It is observed that the number of channel and the computation time can be reduced without reducing the classification accuracy of the system.

Qiu et al.[7] proposed a channel selection method by improving the Sequential floating forward selection (SFFS) method. The result indicates that the improved SFFS yielded better performance than using all channels and support vector machine recursive feature elimination method with an average improvement of 3.8%. The number of channel is reduced with the use of only 30.8 out of 118 channels.

Shan et al.[8] proposed channel selection based on Relief algorithm. Support vector machine was used as the classifier. Three different datasets were used to validate the method. The results indicate that the proposed method yielded average classification accuracies of 85.2%, 94.1%, and 83.2 %, respectively for Dataset 1, Dataset 2, and Dataset 3. The average accuracies were improved by 31.7%, 8.0% and 19.7 % for Dataset 1, Dataset 2 and Dataset 3 respectively.

Yang et al.[9] proposed channel selection method based on Fisher's discriminant analysis for evaluating the discriminative power features for each channel. The result indicates that the method can reduce the number of channels (from 118 channels to 9 in average) without a decrease in mean classification accuracy.

The channel selection plays a major role in reducing the setup time, to minimize the time for placing the cap on the scalp and to avoid the increase in computation time due to the presence of number of channels. In spite of numerous approaches published in the literature, data dimensional reduction through variable selection is still an on-going research and there is opportunity to improve it. It is impossible to say which technique amongst all the reported is the best for channel selection. It is highly subjective in nature and very application specific. Researchers are still looking for new techniques which are relatively simple and robust learning model to improve multiclass motor imagery classification accuracy and learning time.

In this paper a channel selection methods known as Channel Selection based on Standard Deviation of raw energy signal (CSSD) is proposed to enhance the classification performance by selecting significant channel and discard a channel that may give irrelevant information. The Higuchi fractal feature from four different bands namely alpha 2 (11-12 Hz), beta 1 (13-15 Hz) and beta 2 (16-18 Hz) and beta 3 (19-25 Hz) were extracted from all the 19 channels of the EEG raw data. The extracted features were then classified into four different types of hand rotational movements using a feed-forward neural network model. The extracted features were then minimized using CSSD and a network model was subsequently developed using the minimized features. The performances of the two network model were then compared.

2. Methodology

2.1. Data Acquisition

The EEG raw signal was recorded from ten healthy subjects. All the subjects were from University Malaysia Perlis. The experimental work was conducted in air-conditioned laboratory with normal level of noise. This condition makes it more natural since BCI systems would usually be implemented in situations with many distractions. The age range of the participated subjects was between 22 to 34 years. All the chosen subjects must fulfil the criteria such that they must be medication

free, free from illness, enough slept at least six hours in the night the day before the experiment was conducted and have no neurological or psychiatric disorders. All the ten subjects who have participated in the data recording process are right handed and not familiar with the EEG recording equipment.

A non-cue-based protocol using motor imagery as a predefined task is proposed in this experimental work. All the ten subjects performed four different imaginary tasks. They were asked to go through four different imagery task sessions. Each session represents a specific hand movement task. The four different imaginary tasks employed in the experimental procedure were relax, right arm movement, left arm movement and both arm movement. The subjects were given a video demonstration showing a person moving his right forearm up and down for 10 seconds. After the video presentation, the subjects were given a relaxation period of 20 seconds. Then the subjects were requested to imagine moving that they were making up and down right forearm movements. The subjects were asked to repeat the same imagination process for 10 times; simultaneously, the EEG signals corresponding to the 10 trials were recorded. Between the trails, the subjects were given a resting period of 15 seconds. For 'relax' task, the subjects will be in a state of rest and do not perform any task. Subjects were requested as much as possible not to think of anything.

The EEG raw signals were acquired using Mindset-24 Topographic Neuro Mapping Instrument of Nolan Computer System LLC. This EEG instrument is a portable acquisition system which consists of 19 EEG bipolar channels. All the 19 electrodes were positioned as per the international 10- 20 method of electrode placement.

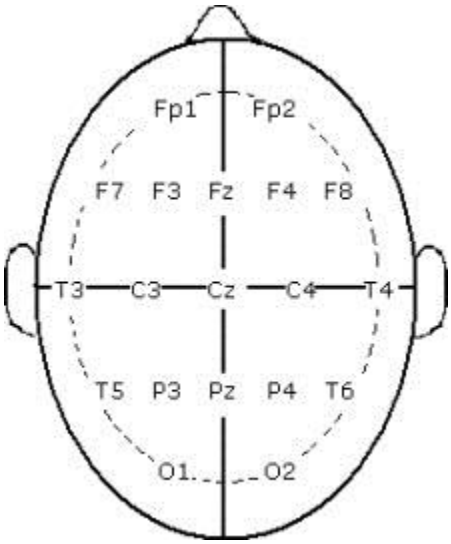


Figure 1. ''10–20" system of electrode placement. F = frontal, T = temporal, C = central, O = occipital, P = parietal. Odd numbers = left hemisphere, even numbers = right hemisphere. [10]

2.2. Pre-processing and Feature Extraction

The signals were recorded for 10 seconds. The EEG signals were sampled at the rate of 256 samples per seconds. The signals were then segmented in such a way that all the segmented frames have equal number of samples. Each frame has 128 sample data points (corresponding to 0.5 s) with 50% overlapping between two consecutive signal frames. A notch filter is used to remove the 50 Hz power line frequency noise from the raw EEG signal. This 50Hz power line frequency noise is due to electrical interference from the EEG recording equipment. Every signal frame is band pass filtered using a second order Chebyshev bandpass filter into four different sub-band frequencies namely alpha 2 (11-13 Hz), beta 1 (13-15 Hz), beta 2 (16-18 Hz) and beta 3 (19-30 Hz). Due to extracted frequency band lies between 8 – 30 Hz, so there's no additional step to remove ocular artefact and muscle artefact energy. It is because most of the ocular artefacts generally appear under 4 Hz frequency band and most of the muscle artefact energy is not contained in 8-30 Hz frequency band [11].

In this paper, Higuchi fractal dimension features were extracted from the recorded EEG. Fractal Dimension (FD) is a measurement process to quantify the signal self-similar characteristic based

on the illustrative presentation of the signal. A single non integer value (fractional) is obtained through the process. The fractal feature values corresponding to the EEG signals lies between 1 and 2. Higuchi proposed a new mathematical process to calculate the time series average. The proposed mathematical process differentiated between Higuchi method and Burlaga and Klein method. The proposed mathematical process is suitable for short time series signals in order to obtain a solid fractal dimension value. By using Higuchi's method, a solid fractal dimension features are brought out from the Motor Imagery EEG signals. The fractal dimension value F_d can be computed by using following Equation (1):

$$F_d = \frac{log(L_k)}{log(k)} \tag{1}$$

where L_k is the mean length value for all the curve length of the subset time series[12].

2.3. Channel Selection

In this paper, a channel selection based on statistical parameters using standard deviation of raw energy signals are proposed. The algorithm is discussed using following step:

Step 1: Segmentation of the recorded EEG signals into number of frame is lead oft in the process of reducing the channel. After the segmentation process, total of 39 frames per channel are gained. Each frame comprises of 128 number of sample with an overlapping of 50% between the successive frames. Thus for 10 trials we have F (390) frames.

Step 2: Compute and formulate raw energy signals matrix $\boldsymbol{R}^k$. $\boldsymbol{R}^k$ represents raw energy signals obtained from all the *F* frames and *C* channels for the k^{th} task. It can be formulate as follow:

$$\boldsymbol{R^k} = \begin{bmatrix} r^k_{1,1} & r^k_{1,2} & r^k_{1,3} & \cdots & r^k_{1,j} & \cdots & r^k_{1,C} \\ r^k_{2,1} & r^k_{2,2} & r^k_{2,3} & \cdots & r^k_{2,j} & \cdots & r^k_{2,C} \\ r^k_{3,1} & r^k_{3,2} & r^k_{3,3} & \cdots & r^k_{3,j} & \cdots & r^k_{3,C} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r^k_{i,1} & r^k_{i,2} & r^k_{i,3} & \cdots & r^k_{i,j} & \cdots & r^k_{i,C} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r^k_{F,1} & r^k_{F,2} & r^k_{F,3} & \cdots & r^k_{F,j} & \cdots & r^k_{F,C} \end{bmatrix} \tag{2}$$

where $r^k_{i,j} = \sum_{q=1}^{N} \left[x^k_{i,j}(q)\right]^2$ is the raw energy value of the signals in the i^{th} frame, j^{th} channel and for the k^{th} task. $x^k_{i,j}(q)$ is the q^{th} data sample of the i^{th} frame, j^{th} channel and k^{th} task where *N* is total number of sample in each frame.

Step 3: Compute standard deviation for column of matrix $\boldsymbol{R}^k$ for k^{th} task. Matrix *D* is then formulated and written as:

$$\boldsymbol{D}=\begin{bmatrix} sd_1^1 & sd_2^1 & sd_3^1 & \cdots & sd_j^1 & \cdots & sd_C^1 \\ sd_1^2 & sd_2^2 & sd_3^2 & \cdots & sd_j^2 & \cdots & sd_C^2 \\ sd_1^3 & sd_2^3 & sd_3^3 & \cdots & sd_j^3 & \cdots & sd_C^3 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ sd_1^k & sd_2^k & sd_3^k & \cdots & sd_j^k & \cdots & sd_C^k \\ \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ sd_1^T & sd_2^T & sd_3^T & \cdots & sd_j^T & \cdots & sd_C^T \end{bmatrix} \tag{3}$$

where sd_j^k is standard deviation of the j^{th} channel and corresponding to the k^{th} task.

Step 4: Using the standard deviation matrix *D*, sum all standard deviation values for each column. These values are then arranged in the descending order.

$$\boldsymbol{KC}=\left[kc_1,kc_2,kc_3,\cdots,kc_j,\cdots kc_c\right] \tag{4}$$

where $kc_i < kc_{i+1}, i=1,2,\cdots,C-1$ and the associated channels are represented as $\boldsymbol{K_S}$:

$$\boldsymbol{K_S}=\left\{ks_1,\ ks_2,\ ks_3,\ \cdots\ ks_j,\ \cdots\ ks_C\right\} \tag{5}$$

Step 5: Build up the neural network model by utilizing the feature set extracted from all the *C* channels. The recorded EEG signal from *'B'* selected frequency band which is derived from band frequency were considered in this process. Classification accuracy gain by developed neural network model is marked as a main classification performance (MCP).

Step 6: Set channel index value, set of channel S and channel performance tolerance (CPT) as follows: index=2

$$\boldsymbol{S}=\left\{ks_1,ks_2\right\}$$

channel performance tolerance (CPT) = 2%.

Step 7: Develop neural network model using the feature set extracted from channel sets. Obtain its classification accuracy and mark it as CP.

Step 8: If (MCP-CP > 2%)

Set index= index+1;

Reformulate the channel set as $S=S\cup ks_{index}$

Go to step 7.

Else

Terminate the process.

2.4. Classification

In this paper standard feed-forward neural network with one hidden layer was developed and implemented to classify the four imagery tasks for each subject. Feed-forward neural network was used because of its simplicity and generalization capabilities. One of the advantages of using artificial neural networks as a classification method is because of their robustness to choices of parameter values and their similarity to other nonlinear regression method.

Higuchi fractal features of four different bands from the selected channel were extracted. A neural network model is developed using the extracted features. The trained neural network model was tested and the classification performance with computation time was obtained and tabulated. The hidden

neurons and network learning rate for the network were chosen as 6 and 0.5 respectively. Through simulation, the number of hidden neurons and the network learning rate was chosen in such a way to give the highest classification accuracy. Both hidden and output neurons were activated using log sigmoid activation function. Training tolerance and testing tolerance were set to 0.01 and 0.1 respectively. The network iteration process was performed until the Mean Square Error (MSE) value reached below 0.0001 or a maximum epoch values of 500 has been reached. The network was trained using Levenberg-Marquardt (LM) algorithm. Binary normalization algorithm was used to normalize the training and the testing samples.

The performances of the channel selection algorithm are evaluated and its performance in each subject is observed. Set of feature from four different bands is obtained from each selected channel. This set of feature is used as an input feature and a neural network model is developed using the input feature. The trained neural network model is tested and the classification performance with its computation time is obtained and tabulated.

In the analysis, the classification performance of selected channel subset is evaluated by 10-fold cross validation. All datasets are split into ten subsets of approximately equal size. Randomly, one dataset is used for testing and the remainders are training. The same procedure is repeated ten times and the observation of performance for each network model is made. The performance is observed base on the average, maximum, minimum classification rate and their average computation time. The proposed algorithms are implemented in MATLAB 2015 on Intel(R) Core(TM) i5-2450 CPU @ 1.8 GHz; RAM: 4 GB.

3. Results and Discussions

The channel selection method based on the standard deviation of raw energy signal was applied to the Higuchi fractal features. The list of the best channels after the channel reduction process is obtained for each subject. Table 1 depicts the result obtained by each subject after reducing the channel.

Table 1. Classification Performance of Channel Selection using CSSD

Subject	All channel		CSSD		
	Mean Accuracy (%)	Mean Computation time (sec)	Number selected channels	Mean Accuracy (%)	Mean Computation time (sec)
Subject 1	95.0	4.1	15	94.5	3.1
Subject 2	96.9	2.8	13	97.1	2.0
Subject 3	95.5	2.8	16	94.5	2.0
Subject 4	99.2	2.5	13	**98.4**	2.2
Subject 5	95.4	3.7	17	95.1	3.3
Subject 6	97.5	2.9	16	97.7	2.1
Subject 7	98.2	2.6	14	98.3	2.1
Subject 8	93.8	2.4	16	**93.7**	2.3
Subject 9	96.9	3.7	18	96.8	2.8
Subject 10	97.2	2.9	14	97.3	1.8
Mean	96.5	3.0	15.2	96.3	2.3
SD	1.6	0.5	1.6	1.7	0.5

A channel selection method based on CSSD algorithm can reduced the number of channel between 1 to 6 channels. The number of reduce channel also depend on the subject. From Table 1, it is

observed that subject #4 has the highest mean classification accuracy of 98.4% and the subject #8 has the lowest mean classification accuracy of 93.7%.

It is clear from the Table 1 that a reduction in the number of channels not only maintains the same classification accuracy but also increases the classification accuracy for some subject. The classification performances of each the subjects were being remain in the range of -1% to +0.5% when compared to the network model classification rates of the same subject obtained using 19 channels. It is observed that classification accuracy for subject #2, #6, #7 and #10 is increase in the range of 0.1% to 0.2%.

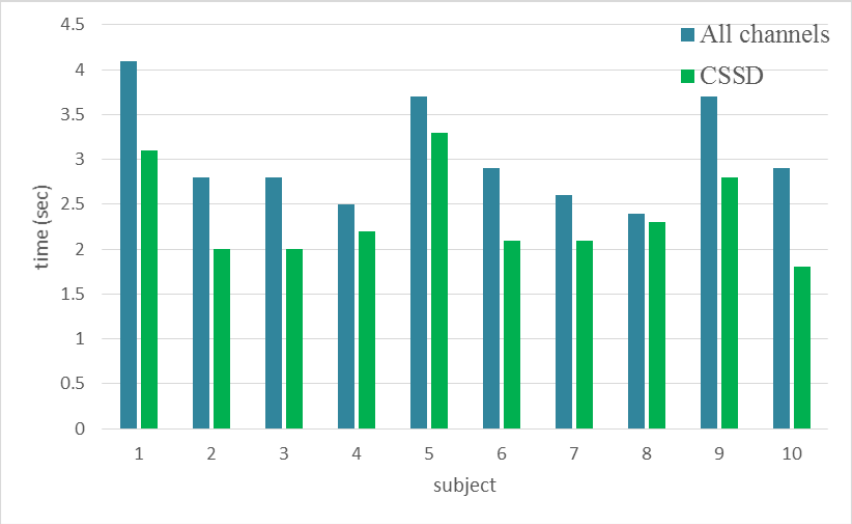


Figure 2. Computation time comparison between All Channels and CSSD

The comparison of computation time is shown in Figure 2. The computation time of the CSSD was significantly decrease in the range of 4% to 37% for each subject. On average, CSSD method reduced the computation time by 22%.

The percentage of channel selected by the CSSD are shown in Figure 3. Channel FP1, FP2, F7, F3, Fz, F4 and P3 were selected by all the subjects. It is observed that channel F8, Cz and P4 were selected by more than 90% of subjects. All those channels are distributed in frontal lobe and parietal lobe of the cerebral cortex. The less significant channel are located in occipital lobe and temporal lobe of the cerebral cortex. It can be seen in Figure 3 where channel in this area were selected by less than 80% of the subjects.

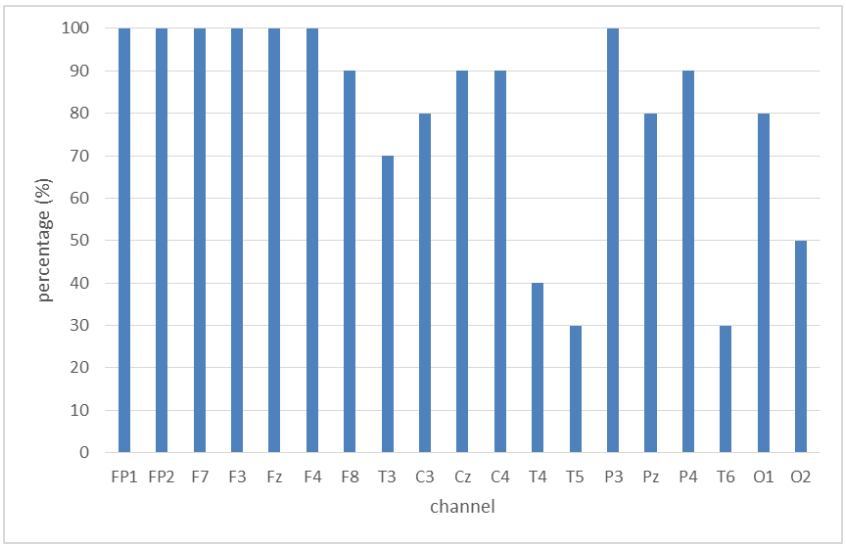


Figure 3. Percentage of channel selected using CSSD

Overall the number of channels is reduced in the range of 5% to 20%. The distributions of channels selected by the CSSD are shown in Figure 3. The distributed channels shows are referred to the number of times for each channels was selected by every subjects. Overall, it is observed that most channels that covers the frontal lobe are selected by 100% of subjects. The motor areas of cerebral cortex are covered in frontal lobe. Channels that related to motor imagery also located in this area. Due to this, most selected channels are cover in this area. Primary somatosensory cortex is part of the parietal lobe. The information related to the perception from external stimuli are processed in this area.

The proposed protocol involved the perception from external stimuli in order to induce the motor imagery signal. Due to this, most channel that distributed in parietal lobe of the cerebral cortex is selected by more than 90% of the subjects. Most channels that are located in the occipital and temporal lobes of the cerebral cortex were selected by less than 70% of the subjects. This is because the channels covered in this area are not related with the motor imagery and coordination movement.

The combination of selected data set channels in different area of the brain that related motor movement will give the best performance and accuracy to discriminate four different motor imagery task. The improved performances and the correct of selection channels in motor cortex area proof that the proposed channel selection method can determine the significant channel that relate to the motor imagery task.

4. Conclusions

The paper shows that the reduction in channels using the proposed algorithm do not result in significant loss in performance compared to that using the whole 19 channels. It also has the benefit of reduced training and testing time. It was also observed that the number of selected channel depends on the individual subjects. This may be due to the concentration levels of individuals vary.

From the above, it can be concluded that the proposed channel reduction method helps to find the reduced number of channels and indirectly can reduce the number of input features and computation time. In the future work, the research work can be extended with the fusion of different channel selection techniques to improve the performance of the classification accuracy.

References

[1] J. R. Wolpaw *et al.*, “Brain-computer interfaces for communication and control.,” *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–91, 2002.

[2] Shih et al, “Brain-computer interfaces in medicine,” *Mayo Clinic Proceedings*, vol. 87, no. 3. pp. 268–279, 2012.

[3] M. Jeannerod, “The representing Brain : Neural correlates of motor imagery and intention,” *Behav. Brain Sci.*, vol. 17, pp. 187–245, 1994.

[4] D. Koller and M. Sahami, “Toward {Optimal} {Feature} {Selection},” *Int. {Conference} {Machine} {Learning}*, pp. 284–292, 1996.

[5] S. M. Park, J. Y. Kim, and K. B. Sim, “EEG electrode selection method based on BPSO with channel impact factor for acquisition of significant brain signal,” *Optik (Stuttg).*, vol. 155, pp. 89–96, 2018.

[6] M. S. Hossain, S. Saha, M. A. Habib, A. Al Noman, T. Sharfuddin, and K. I. Ahmed, “Application of wavelet-based maximum entropy on the mean in channel optimization for BCI,” *1st Int. Conf. Med. Eng. Heal. Informatics Technol. MediTec 2016*, 2017.

[7] Z. Qiu, J. Jin, H. K. Lam, Y. Zhang, X. Wang, and A. Cichocki, “Improved SFFS method for channel selection in motor imagery based BCI,” *Neurocomputing*, vol. 207, pp. 519–527, 2016.

[8] H. Shan, H. Xu, S. Zhu, and B. He, “A novel channel selection method for optimal classification in different motor imagery BCI paradigms.,” *Biomed. Eng. Online*, vol. 14, no. 1, p. 93, 2015.

[9] Y. Yang, I. Bloch, S. Chevallier, and J. Wiart, “Subject-Specific Channel Selection Using Time Information for Motor Imagery Brain–Computer Interfaces,” *Cognit. Comput.*, vol. 8, no. 3, pp. 505–518, 2016.

[10] T. Thompson, T. Steffert, T. Ros, J. Leach, and J. Gruzelier, “EEG applications for sport and

performance," *Methods*, vol. 45, no. 4, pp. 279–288, 2008.
[11] E. Gysels and P. Celka, "'Phase synchronization for the recognition of mental tasks in a brain-computer interface.,' {IEEE} {Transactions} on {Neural} {Systems} and {Rehabilitation} {Engineering},," *Ieee Trans. Neural Syst. Rehabil. Eng.*, vol. 12, no. 4, pp. 406–415, 2004.
[12] T. Higuchi, "Approach to an irregular time series on the basis of the fractal theory," *Phys. D Nonlinear Phenom.*, vol. 31, no. 2, pp. 277–283, Jun. 1988.

Acknowledgments
The authors would like to thank Dato' Professor Dr. Zul Azhar Zahid Jamal, the vice chancellor, Universiti of Malaysia Perlis for his continuous support and encouragement.