Developing an Optimized Single-Trial P300-Based Brain Computer Interface System

F.E. Motlagh[1] and F. Ibrahim[1]

[1]Department of Biomedical Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur, Malaysia

Abstract— **Brain event related potentials (ERP) have been used in developing brain computer interface (BCI) systems. P300 as a robust ERP has been utilized in BCI and clinical researches. A common P300-based BCI system consist of brain signal recording, pre-processing, P300 features extraction, and classification units. Achieving a high accuracy in detection of single-trial P300, using fast computational algorithms is the main challenge of designing these systems. However, there is trade-off between accuracy and computational time. In this study, various well-developed algorithms controlled by a rule-based platform to optimize the detection algorithm. P300 feature extraction algorithms has been developed by using wavelet transform techniques, while SVM with linear/Gaussian kernels and logistic regression applied as alternative supervised learning classifiers. Principle component analysis also was used for feature selection in order to speed up the classification procedure. This optimization system make decision on selecting the proper P300 detection method via selecting the group of channels, feature extraction algorithm, number of selected principle components, and type of classifier. Controller used cross validation data set to calculate the accuracy and ratio of computational time for each possible combination, and the optimized method was assessed using test data set. The results suggest that designing a P300-BCI system with the ability to select the proper method of detection can be utilized in different applications to benefit the user with a better performance.**

Keywords— **Brain computer interface, Event related potentials, Wavelet transform, Support vector machine, Supervised learning.**

I. INTRODUCTION

BCI systems are designed to transform brain electrophysiological signals into commands for computers. Brain computer interface (BCI) systems has been used excessively for research purposes and clinical diagnostics in the past decade [1].There have been various BCI systems based on different attributes of brain signals. However, event-related potential (ERP)-based BCI systems are known as effective systems in this field. P300 is a robust positive ERP which has been utilized in BCI systems and shows promising results in terms of accuracy and robustness [2]. P300 occurs as a response to rare task-relevant stimuli in a series of task irrelevant stimuli around 300 ms after stimuli [3]. Farwell and Donchin described P300-BCI system to communicate with computers without utilizing voluntary muscle activity [4], using oddball paradigm to evoke P300 component [5]. Oddball paradigm demonstrate a random sequence of desired and undesired events, which is supposed to probe P300 during the novel desired stimuli. The aim of BCI detection algorithm is to detect the target signals (carry P300) among non-targets.

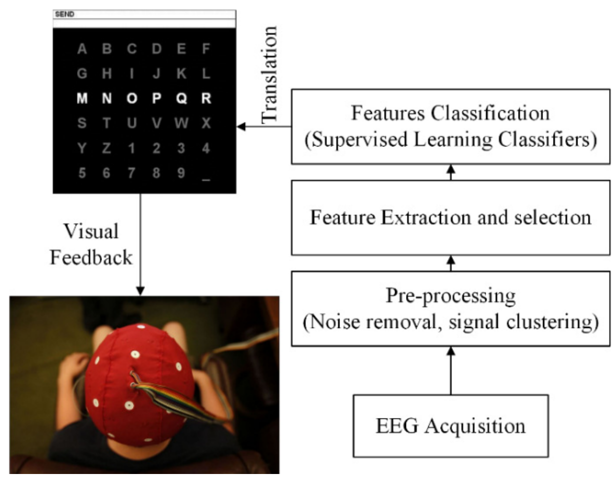

Table 1 Fig. 1 BCI system components, including data acquisition, signal processing and BCI application units. Brain signals are recorded using an EEG machine. Signal processing using is consist of Pre-prosessing, feature extraction, and classification blocks. BCI application includes translated command of brain and a paradigm for stimulating the brain.

Typical P300-BCI system consists of three major blocks: (1) signal acquisition and pre-processing, (2) P300 features extraction, and (3) classification (Figure 1). There are challenges in developing components of this system; each unit of the system should be developed in order to optimize the performance of whole system. Designing an algorithm for providing noise and artifact free signals, accurate P300 feature extraction and selection, and finally efficient classification of these features are the main criteria of a robust P300 detection method. However, when it comes to real time and online applications, processing time, number of channels, accuracy of single-trial detection, and complexity of the classifier are significance issues to be considered.

This paper aims to utilize a rule-based controller for optimizing the single-trial P300 detection accuracy and

F. Ibrahim et al. (eds.), *International Conference for Innovation in Biomedical Engineering and Life Sciences*,
IFMBE Proceedings 56,
DOI: 10.1007/978-981-10-0266-3_2

enhancement of processing time in real-time P300-BCI systems. This optimization platform evaluates the EEG recording channels and employs the proper algorithm for pre-processing, feature extraction, and classification units.

II. P300-BCI SYSTEM

A. Signal Acquisition

Although various types of current scalp electroencephalogram (EEG) equipment record efficient signals using convenient and user-friendly electrodes; there are two significant considerations about the recording signals and number of recording channels in a P300-BCI system. Developing a P300-BCI system using large number of electrodes cause user discomfort as well as longer processing time. Therefore, employing smaller subset of electrodes was suggested to reduce the processing time and increasing user comfort, while providing enough information for an accurate detection of P300. Some studies aimed to define an optimal subset of electrodes applicable in P300-BCI systems, e.g. Krusienski et al, suggested 8-channel electrodes set (Fz, Cz, P3, Pz, P4, PO7, PO8, Oz) [6], and Motlagh et al, suggested five channels (C1, Cz, Cpz, Pz, Fcz) [7].

In this study, dataset were obtained based on 10–20 system using 19 EEG Channels ($Fp_{1/2}$, $F_{7/8}$, F_z, $F_{3/4}$, $T_{3/4}$, $C_{3/4}$, C_z, $T_{5/6}$, $P_{3/4}$, P_z, $O_{1/2}$) with average of $A_{1/2}$ as reference during the performance of an oddball paradigm. Nicolet EEG diagnostics system (Care Fusion Corporation, 3750 Torrey View Court, San Diego, CA 92130) was used to capture the EEG activities within the frequency band of 0.5-70Hz (with a sampling rate of 256 Hz). Before data collection, the impedances of all the electrodes were monitored during the EEG recording, to verify its value to be under 5 kΩ, and the paradigm timing system and the EEG recorder were synchronized.

B. Pre-Processing

EEG is highly susceptible to various forms and sources of noises, which present significant difficulties and challenges in analysis and interpretation of EEG data. Preprocessing the data as the most essential step in development of a reliable BCI system should be accurate and efficient. In this study, an automated standard pre-processing steps was applied on the signals in the initial phase.

Each channel's signal band-pass filtered between 0.1-45 Hz using *slepian multitaper* spectrum (MATLAB "*pmtm*" function) by applying four orthogonal tapers, (a combination of a high pass and low pass filter), in order to remove power line, high frequency noises and DC biases.

Signal mean, standard deviation, skewness, kurtosis and median (five first cumulates of distribution) were calculated and stored. Signals data-points distribution from each channel shows an estimation of quality of EEG recording from that channel. Kolmogrov-smirnov test applied to estimate the distribution of the signal of each channel subsequently. The result of this test at a significant level of P≤0.05 shows whether the data distribution of signal is Gaussian or not; thus, each channel would be labeled based on the equation (1) criteria.

$$Channel\ Number \begin{cases} 1 & P\ value \leq 0.05 \\ 0 & P\ value > 0.05 \end{cases} \tag{1}$$

Channels with label "0" should be eliminated from the rest of the procedure. In order to have a uniform and standard procedure for detecting the EEG artifacts, all Gaussian signals were divided to epochs with duration of one-second period, and following steps were applied:

Removing linear trend: During acquisition of EEG, recording-induced current drifts and electrode movements cause occurrence of linear drifts in EEG trials. This type of artifacts was removed by fitting a straight line to the datapoints of the signal, and if the slope of the calculated line is more than 60 μv, then the epoch's linear trend was removed.

ICA decomposition: Independent component analysis (ICA) algorithm was used for removing the artifacts, eye-movements and blinking using *"runica infomax"* in MATLAB. ICA decomposed the channels' signals into independent signals (based on orthogonality using singular value decomposition algorithm) as a common method for artifact removal and source localization of EEG signals (this method described in [8, 9]).

ICA Reconstruction: After removing the artifactual independent components (IC), remaining signals was used for reconstruction of channels' EEG. Then, ICA decomposition was applied again in order to guarantee the noise removal procedure.

Windowing: Each signal should be synchronized with the stimuli onset timing of the paradigm; therefore, 256 samples (1second) from the onset of the stimuli was selected as a single trial.

Training set labeling: Single-trial P300 detection algorithm is based on a supervised learning classification. therefore 60% of the dataset were selected randomly as training set, and each trial of training set were labeled as target or non-target.

C. Feature Extraction

The major challenge in optimizing the performance of the P300-BCI is to enhance the real-time detection of P300. The process of real-time detection consists of an optimal P300 features extraction in order to employ a simpler classification algorithm to increase the processing speed in real-time applications. P300 like other event related potentials has a very low amplitude (μv) compared to baseline EEG signals (mv); this per se compromise the detection accuracy and cause a trade-off between detection speed and accuracy. Additionally, high signal-to-noise ratio (SNR) of EEG signals make this very challenging [10]. Various studies have been employed orthogonal linear transformation [11], blind source separation, wavelet transform [12, 7, 2, 13-15] and other advanced techniques to overcome these challenges. The key point of established feature extraction techniques is to utilize most distinctive features to reduce the computational time.

In this study, two feature extraction algorithms were designed based on wavelet transform (WT) properties:

CWT: In the first approach, continuous WT (CWT) were applied on the trials using *Mexican-hat* wavelet (scales 30-100). It was shown that this wavelet has a good similarity on scales of 30-100 with P300 component (this range of scale is associated with 0.6-2 Hz in frequency domain) [7]. In this method, each trial was swept by different scales of wavelet and their correlation was calculated for each time-scale as similarity coefficients. Wavelet coefficients of a signal x (t) at time point p are defined as:

$$\int_0^{256} x(t)\varphi(s,p,t)dt \tag{2}$$

Where s is the wavelet scale, t is trial data-points, and φ is the chosen wavelet (*Mexican-hat*).

Then, CWT coefficients were averaged over different scales and extermum values of obtained vector was stored. It is assumed that the maximum of averaged curve has the amplitude of A_0 that happens at time T_0. The goal is to find the two local minimums, i.e., one just after A_0 and another just before A_0 with amplitudes of B_1 and B_2 respectively. Using "equations (3) and (4)", and the obtained extermum properties, two heuristic features over averaged scales can be defined. "A" as the similarity amplitude and "T" as ratio of latency.

$$A = (A_0 - B_1) + (A_0 - B_2) \tag{3}$$

$$T = {|T_0 - 300|}/{300} \tag{4}$$

For detecting P300 wave the amplitude of the peak feature should have "large" value and time difference feature should be as "small" as possible (zero is considered ideal). Therefor, *A* and *T* are two features to be extracted by applying CWT. This method was confirmed to provide robust features for single-trial P300 detection, although calculation of correlation and sweeping the signals for all scales increase the processing time [7, 13].

(DC)WT: In the second approach, a combination of discrete WT (DWT) and CWT was applied for providing more robust features. In this method, discrete wavelet transform was used for multi-resolution decomposition of signal into 'details' and 'approximation' (high frequency and low frequency) components. B-Splines wavelets were chosen as mother function in this study due to their high resemblance with brain evoked responses. Five levels of DWT transformed the signals into 64–128 Hz, 32–64 Hz (gamma), 16–32 Hz (beta), 8–16 Hz (alpha), 4–8 Hz (theta) and the last approximation giving the activity in the 0.1–4 Hz (delta). Since, each step of DWT decomposition divide the signal into two components by down-sampling, delta and theta band (0.1-8 Hz) contains 16 data points. These 16 data points were stored as DWT features.

Thanks to low frequency of evoked potentials, delta and theta decompositions were used for reconstructing the signal and up-sampling. The reconstructed signal from delta and theta band results in a smooth signal 0-8 Hz. CWT as it explained earlier was applied on reconstructed signal and CWT and DWT features was stored (18 features).

Final number of features in this method is these features multiple by the number of selected channels. Since, dealing with large number of features leads more computational time for classification; selected features should be reduced into lower dimensions using principle component analysis (PCA). PCA reduced the features dimensionality into lower orthogonal dimensions using Eigen vectors of features. The number of principle component should be chosen based on the percentage of variance that retained. The optimal percentage of retained variance is supposed to be 99%. The whole procedure of this approach is depicted in Figure 2.

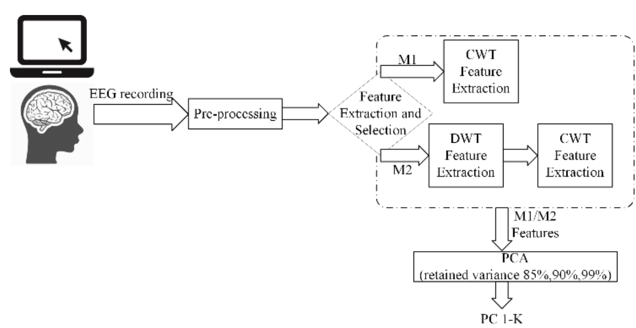

Fig. 2 Feature extraction methods designed based on two approaches. CWT and DWT features of each channel should transformed to a lower dimensions using PCA algorithm. The controller decide on the number of principle components (K) to be selected as the inputs of classifier.

D. Classification

Single-trial P300 detection requires accurate classification of extracted features. Numerous studies have attempted to enhance the classification algorithm by utilizing linear and nonlinear methods [16-18]. However, avoiding complex but reliable classifiers benefits the enhancement of computational time. Logistic regression and support vector machine (SVM) (both linear and Gaussian kernels) were used as fast binary classifiers in this study.

Selected principle components and trials label of training dataset would be used in these three supervised classification algorithms (logistic regression, SVM with linear kernel, and SVM with Gaussian kernel).

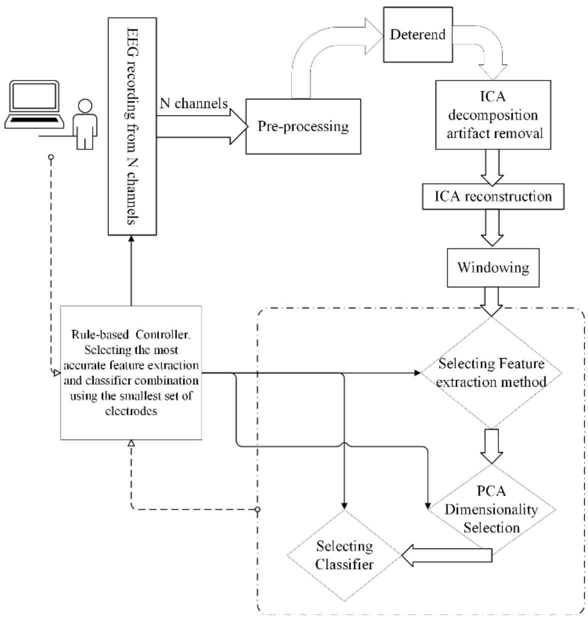

Fig. 3 Flowchart of rule-based controller of P300-BCI system. The user will choose the criteria for accuracy and computational time based on the application and system would be modified.

E. Controller

The aim of controlling the single-trial P300 detection algorithm is to choose the best procedure for detecting P300 accurately in minimum time. In this study, 60% of the data was selected randomly for training, 20 % for cross validation, the rest of data (20 %) as test dataset, and a rule-based controller was chosen for selecting the best combination of the algorithms in order to fulfill the accuracy and computational time conditions. This controller can decide on number of selected channels (NC), feature extraction algorithm (FE), number of selected features based on PCA output (FS), and selecting the type of classifier (CS). In this step, training and cross-validation set were used to evaluate each possible combination of parameters. However, there are huge number of different possibility for NC and calculation of all possible combinations is impossible. Therefore, a certain set of well-studied channels was used (Table 1 shows some of the outputs of this part).

There are five different level for each method, namely very low, low, normal, high, and very high. Computational time was normalized between 0-1, and 0.2 as threshold of each level (0-0.2 very low, 0.2-0.4 low, 0.4-0.6 normal, 0.6-0.8 high, and 0.8-1 very high). Accuracy percentage of cross validation set was divided to 50-65% very low, 65-75% low, 75-85% normal, 85-90% high, and 90-100% as very high. The controller aims to maintain the system to perform in a condition which user can define (e.g. very low computational time and high or standard accuracy percentage), shown as user condition (UC).

The system start with using Cz channel signals, CWT method of feature extraction and using just one principle component and logistic regression (the fastest method). The rule-based controller works as follow:

Step 1. If the accuracy is lower than UC, then other classifiers would be evaluated, and the one with highest accuracy is selected.

Step 2. If the accuracy is lower than UC, increase the number of K (number of principle components) that 99% variance retained.

Step 3. If the accuracy is lower than UC, evaluate both feature extraction methods and select the most accurate algorithm.

Step 4. If the accuracy is lower than UC, evaluate the second set of channels.

Step 5. Repeat steps 1-4 until the UC condition for accuracy percentage is provided.

In this system, accuracy percentage condition has priority over the computational time; therefore, after fulfilling the accuracy criterion, the computational time would be evaluated using all possible methods without changing the number of channels and the fastest method would be chosen; then, the UC condition for accuracy percentage should be reassessed. By defining the method, system use test dataset to evaluate the system again and results would be shown to the user and waiting for confirmation or new set of criteria. The flowchart of this system is depicted in Figure 3.

Table 2 Accuracy (A) and normalized computational time (NCT) for some of the possible combinations.

NC	FE	FS	CS	A%	NCT
Cz	2	3	LR	68.23	0.21
Cz, Pz	1	3	SVML	66.12	0.16
Pz	2	3	SVMG	73.29	0.19
C1	1	3	SVMG	61.19	0.18
Cpz	1	3	SVML	54.49	0.11
Fcz	2	3	LR	60.10	0.16
C1,Cz	2	2	SVMG	65.12	0.21
Cz, Pz, Fcz	2	2	SVMG	75.45	0.25
5	1	2	SVML	86.28	0.35
8	1	2	SVML	89.36	0.54
19	2	3	LR	96.38	0.96

III. CONCLUSIONS

The P300-BCI appears to be the most commonly used BCI system. Despite its popularity among researchers, it is apparent that many P300-BCI systems must be improved before they can be considered as an alternative communication device for individuals. In this paper, a rule-based controller system was applied to optimize the accuracy and processing time of single-trial P300-BCI system.

The accuracy of each set of data would be compared to select the smallest set of channels providing fastest computation and highest possible accuracy. Once the sub-group of channels selected, the system can be work in test condition using less number of channels. This controlling system can provide a better performance of a typical BCI system in various applications. Future work is focus on testing this system in various applications and compare it with current systems using fuzzy controllers and other soft computing methods for a better modification.

ACKNOWLEDGMENT

We would like to acknowledge and thank the University of Malaya and Ministry of Higher Education for providing High Impact Research Grant, account codes E000007-20001 to fund this project. This research was also supported by Special Prime Minister's project (Project No: 66-02-03-0061/oracle 8150061).

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

1. Burns A, Adeli H, Buford JA. Brain–Computer Interface after Nervous System Injury. The Neuroscientist. 2014:1073858414549015.
2. Motlagh FE. P300 detection of brain signals using a combination of wavelet transform techniques: Universiti Putra Malaysia; 2012.
3. Sanei S, Chambers JA. EEG signal processing. John Wiley & Sons; 2013.
4. Farwell LA, Donchin E. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroencephalography and clinical Neurophysiology. 1988;70(6):510-23.
5. Donchin E. Surprise!… surprise? Psychophysiology. 1981;18(5): 493-513.
6. Krusienski DJ, Sellers EW, McFarland DJ, Vaughan TM, Wolpaw JR. Toward enhanced P300 speller performance. Journal of neuroscience methods. 2008;167(1):15-21.
7. Motlagh F, Tang S, Motlagh O, editors. Optimal accuracy and runtime trade-off in wavelet based single-trial P300 detection. Signal and Image Processing Applications (ICSIPA), 2013 IEEE International Conference on; 2013: IEEE.
8. Schlögl A, Keinrath C, Zimmermann D, Scherer R, Leeb R, Pfurtscheller G. A fully automated correction method of EOG artifacts in EEG recordings. Clinical neurophysiology. 2007;118(1):98-104.
9. Li Y, Ma Z, Lu W, Li Y. Automatic removal of the eye blink artifact from EEG using an ICA-based template matching approach. Physiological measurement. 2006;27(4):425.
10. Li K, Raju VN, Sankar R, Arbel Y, Donchin E. Advances and challenges in signal analysis for single trial P300-BCI. Foundations of Augmented Cognition. Directing the Future of Adaptive Systems. Springer; 2011. p. 87-94.
11. Dien J, Spencer KM, Donchin E. Localization of the event-related potential novelty response as defined by principal components analysis. Cognitive Brain Research. 2003;17(3):637-50.
12. Xu N, Gao X, Hong B, Miao X, Gao S, Yang F. BCI competition 2003-data set IIb: enhancing P 300 wave detection using ICA-based subspace projections for BCI applications. IEEE transactions on biomedical engineering. 2004;51(6):1067-72.
13. Motlagh F, Tang S, Motlagh O, editors. Combination of continuous and discrete wavelet coefficients in single-trial P300 detection. Biomedical Engineering and Sciences (IECBES), 2012 IEEE EMBS Conference on; 2012: IEEE.
14. Quiroga RQ, Garcia H. Single-trial event-related potentials with wavelet denoising. Clinical neurophysiology. 2003;114(2):376-90.
15. Bostanov V, Kotchoubey B. The t-CWT: a new ERP detection and quantification method based on the continuous wavelet transform and Student's t-statistics. Clinical neurophysiology. 2006;117(12):2627-44.
16. Kaper M, Meinicke P, Grossekathoefer U, Lingner T, Ritter H. BCI competition 2003-data set IIb: support vector machines for the P300 speller paradigm. Biomedical Engineering, IEEE Transactions on. 2004;51(6):1073-6.
17. Krusienski DJ, Sellers EW, Cabestaing F, Bayoudh S, McFarland DJ, Vaughan TM et al. A comparison of classification techniques for the P300 Speller. Journal of neural engineering. 2006;3(4):299.
18. Cecotti H, Gräser A. Convolutional neural networks for P300 detection with application to brain-computer interfaces. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 2011;33(3):433-45.

Author:	Farid Esmaeili Motlagh
Institute:	Centre for Innovation in Medical Engineering (CIME)
City:	Kuala lumpur
Country:	MALAYSIA
Email:	farid_motlaq@hotmail.com