2012 Ninth International Conference on Wearable and Implantable Body Sensor Networks

Brain-computer interface signal processing algorithms: A Computational Cost vs. Accuracy Analysis for Wearable Computers

Ali Ahmadi, Omid Dehzangi, Roozbeh Jafari
Department of Electrical Engineering
Univeristy of Texas at Dallas, Richardson, TX 75080–3021
Email: {ali.ahmadi, omid.dehzangi, rjafari}@utdallas.edu

***Abstract*—Brain Computer Interface (BCI) is gaining popularity due to recent advances in developing small and compact electronic technology and electrodes. Miniaturization and form factor reduction in particular are the key objectives for Body Sensor Networks (BSNs) and wearable systems that implement BCIs. More complex signal processing techniques have been developed in the past few years for BCI which create further challenges for form factor reduction. In this paper, we perform a computational profiling on signal processing tasks for a typical BCI system. We employ several common feature extraction techniques. We define a cost function based on the computational complexity for each feature dimension and present a sequential feature selection to explore the complexity versus the accuracy. We discuss the trade-offs between the computational cost and the accuracy of the system. This will be useful for emerging mobile, wearable and power-aware BCI systems where the computational complexity, the form factor, the size of the battery and the power consumption are of significant importance. We investigate adaptive algorithms that will adjust the computational complexity of the signal processing based on the amount of energy available, while guaranteeing that the accuracy is minimally compromised. We perform an analysis on a standard inhibition (Go/NoGo) task. We demonstrate while classification accuracy is reduced by 2%, compared to the best classification accuracy obtained, the computational complexity of the system can be reduced by more than 60%. Furthermore, we investigate the performance of our technique on real-time EEG signals provided by an eMotiv® device for a Push/NoPush task.**

Keywords-BCI; EEG classification; Go/NoGo task; real-time application; cost vs. accuracy analysis

I. INTRODUCTION

BODY SENSOR NETWORKS (BSNs) are suitable for continuous monitoring of human body. These sensors are positioned on body and can monitor, for example, human physiological state or movements. Monitoring several modalities using BSNs have been explored including electroencephalography (EEG), electrocardiography (ECG), electromyography (EMG) and body movements. EEG, and applications of BCI possibly have been least explored by the BSN community due to the challenges associated with the implementation of the complex signal processing algorithms on BSN processing units.

The primary aim in BCI systems is to provide communication channels to translate brain rhythms of an individual into application-specific signals for computers. BCI allows a person to use mental processes to communicate with external devices without relying on neuromuscular control [1].

There are several scientific challenges before every-day use of BCI systems become a reality. BCI systems often require complex signal processing algorithms due to the noisy signals that are present in EEG. Therefore, the ability to extend the BCI systems to real-time processing would be an issue. Enhancing wearability, portability and durability are three major objectives in the design and development of wearable and power-aware BCI systems. Lowering the power consumption translates to reducing the size of the battery and the form factor, improving the wearability of the device. Computational complexity of BCI algorithms directly affects the power consumption and the ability of the system for real-time processing. In order to reduce the power consumption of a real-time BCI, special techniques are required to reduce the computational complexity of the BCI based on the amount of energy available in the battery.

There have been many efforts to develop computationally intensive techniques for BCI systems with enhanced accuracy [1, 2, 3]. However, high accuracy is achieved at the expense of increased computational complexity. In real-time applications, it is highly desirable to consider simpler mathematical models to reduce the computational cost while maintaining adequate classification accuracy. In this paper, we setup a classification system to classify single-trial EEG signals for a Go/NoGo task. Some commonly used feature extraction methods in the BCI research such as band power, wavelet coefficients, and Short Time Fourier Transform (STFT) are employed. We present a sequential feature selection algorithm to determine an effective subset of the features while considering the trade-off between the computational complexity and classification accuracy of the system. In order to find a highly discriminative feature subspace as well as to reduce the computational cost, we present a cost vs. accuracy analysis for the feature selection. We chose Support Vector Machine (SVM) classifier to evaluate the performance of the system as it has a solid foundation in statistical learning theory and is shown to be a reliable method for EEG data classification [3, 4]. Finally, we verify our method on a Push/NoPush task with the real-time EEG signals recorded by an eMotiv® device as an empirical case study to illustrate the effectiveness of the proposed method of cost vs. accuracy analysis.

The organization of this paper is as follows. In section 2, we review the previous work and discuss the strengths and weaknesses. In section 3, we describe the Go/NoGo task and data acquisition. Feature extraction and system architecture are covered in section 4. In section 5, we present the feature selection method using the proposed cost vs. accuracy analysis. Experimental results are illustrated and the trade-off between computational complexity and accuracy of the system is discussed in section 6. Finally, section 7 concludes the paper.

978-0-7695-4698-8/12 $26.00 © 2012 IEEE
DOI 10.1109/BSN.2012.19

II. LITERATURE SURVEY

BCI has been gaining much attention as a solution to convert brain signals to usable control commands. BCI can use a variety of electrophysiological sources. However, most of the current BCI implementations rely on three main electrophysiological sources: motor imagery, steady-state visual evoked potential, and P300 potential. In this paper, we focus on P300, which can be observed in the EEG signals during the Go/NoGo task. We attempt to use algorithms with low computational cost for signal processing to justify power requirements in wearable systems. In the following, we provide an overview on the current state-of-the-art signal processing techniques.

In [5], authors analyze changes in EEG power and synchrony between pairs of channels during the Go/NoGo task. They reported that these changes happen at different time windows and frequency bands for the Go and the NoGo trials. Authors in [6] presented statistical analyses that correspond to coherence in the Go/NoGo task for each condition. They calculated the time course of the coherence between the channels F3, F4, C3 and C4 using Fast Fourier Transform (FFT) and provided a comparison between the Go and the NoGo. Khan et al. [7] extracted spatial patterns from the EEG signals in gamma band using CSP method. They used radial basis function to classify four imagery movements.

Feature selection is an effective way to improve the classification performance. By selecting a representative subset of the features, not only performance might be improved but also the computational cost would decrease. Cabrera et al. [9] used an exhaustive search to select a subset of combined features. Their results showed the effectiveness of channel and feature selection for improving the discrimination capabilities. In [10], a genetic based feature selection approach was presented in which a subset of features was generated by applying wavelet to the EEG data, selected using Genetic Algorithm (GA). The authors reported a significant improvement by using a group of features selected by the GA. Francesco et al. [11] used Principal Component Analysis (PCA) to reduce the complexity of classification of cortical neural signals. The authors showed that the cost was reduced significantly by using a small subset of the features while the accuracy was slightly affected.

In order to facilitate the everyday use of BCI systems, the signal processing units must be light-weight, portable, low-power and real-time. Therefore, there has been much research interest in real-time implementation of BCI systems. In [12], a real-time wireless EEG-based BCI with four channel signal acquisition was presented for drowsiness monitoring. The proposed system acquires and analyzes the EEG signals in real-time and provide warning signals when needed. The authors in [13] introduced an embedded BCI system in which they designed and developed a real-time digital signal processing unit with a Bluetooth module to receive the EEG signals from sensors on a headband and process the EEG signals in near real-time.

Although real-time implementation of BCI algorithms has been investigated previously, generalizable techniques that incorporate the computational cost of signal processing with the accuracy requirements have not been proposed. In this work, we begin with features that have low computational cost such as band power, STFT and wavelet coefficients. We incorporated the computational cost in the feature selection algorithm to ensure the accuracy constraint is accommodated while the computational cost is reduced. The details of our approach are illustrated in the following sections.

III. TASK DESCRIPTION AND DATA COLLECTION

In this study, we consider the Go/NoGo inhibition task to develop our BCI signal processing algorithms and evaluate cost vs. accuracy trade-off. The Go/NoGo task is a type of continuous performance neuropsychological test that has been designed to explore complex attention function such as response inhibition. This task requires a level of semantic abstraction to make a correct response. It includes Go item (image of a car) and NoGo item (image of a dog). In this task, an identical image is repeated; therefore the perceptual properties of the item stay identical. When the participant sees a Go object, he or she will have to press a button, and in case of NoGo, he or she will have to refrain from pressing the button. The EEG signals corresponding to this task resembles how well the individual can inhibit some cues and select others. Fig. 1 shows the stimulus for Go and NoGo objects.

Continuous EEG signals were recorded from 64 silver/silver-chloride electrodes mounted within an elastic cap (Neuroscan Quickcap) which are placed according to the International 10–20 electrode placement standard (Compumedics Inc.). The data was collected using a Neuroscan SynAmps2 amplifier and Scan 4.3.2 software sampled at 1 kHz with impedances typically below 10 kΩ. Blinks and eye movement were monitored via two electrodes, one mounted above the left eyebrow and one mounted below the left eye. The EEG data were segmented offline into 2s epochs spanning 500ms before to 1500ms after the presentation of the visual stimuli.

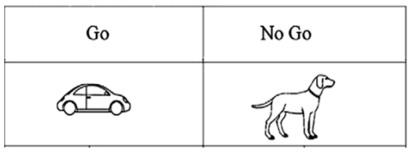

Figure 1. Sample of stimulus for Go/NoGo task.

IV. SIGNAL PROCESSING

A. System Architecture

In this study, we present a lightweight classification method for single trial EEG. First, the baseline is removed, that is, the average of baseline segment (0-500ms) subtracted from all the samples of each trial. Then, the data were re-referenced to the average potential over the entire head. In the next step, we applied a band-pass filter (0.2 – 30 Hz) to eliminate high frequency and very low frequency noise. Then, we used discrete time wavelet, STFT and wavelet packet tree to extract features. Feature selection block selects a subset of features based on the accuracy and the cost constraints. Selected features are applied to the classifier. Fig. 2 shows the signal processing flow of the system.

B. Feature extraction

Band Power

Analyzing frequency spectrum for different frequency bands is a commonly used method for single trial EEG classification [5, 14]. These sub-bands are called delta (δ), theta (θ), alpha (α), beta

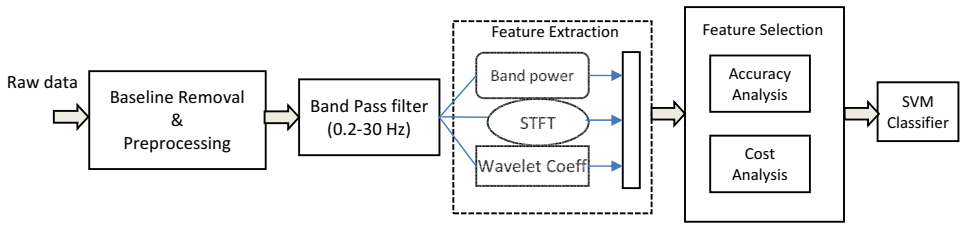

Figure 2. Signal Processing Architecture

(β) and gamma (γ) bands. There are no strict frequency ranges for these different bands. In this paper, ranges are selected as follows: delta (0.5–4), theta (4–8), alpha (8–13), beta (13–25) and gamma (25–40).

Wavelet transformation is a time-scale analysis method and has the capacity of representing local characteristics of the signal in the time and frequency domains. As mentioned before, the sampling frequency used on the EEG data was 1 kHz. We used 7-level wavelet to decompose each trial to its corresponding signals in different frequency bands. After frequency decomposition, power of signal for each frequency band and the original signal was determined according to Eq. (1). Then, relative band powers are extracted by dividing band power values to the power of the original signal. Selection of a suitable mother wavelet for a given application depends on a number of factors, including the type of the signals, the computational requirements of the algorithm and the objectives for the signal processing [15]. Haar wavelet is a suitable choice for mother wavelet in real-time applications because of its low computational cost. In this work, we used Haar wavelet.

$$P_d = \sum_{i=1}^{n} x_i^2 \tag{1}$$

STFT

STFT is an extension of conventional Fourier analysis for non-stationary data. STFT performs FFT on consecutive segments or blocks of data that are assumed stationary, and is equivalent to a sliding window that analyzes the local frequency content of the signal. The STFT for signal $x(\tau)$ windowed by a fixed-length function $w(t - \tau)$ is defined by,

$$STFT_x(t,f) = \int_{-\infty}^{\infty} x(\tau) w(t-\tau) e^{-j2\pi f\tau} d\tau \tag{2}$$

the STFT power or energy, $P_x(t, f)$ is defined by,

$$P_x(t,f)_{STFT} = |STFT_x(t,f)|^2 \tag{3}$$

In our analysis, we use 1s window and 90% overlap for STFT. The length of each trial is 2s spanning 500ms before to 1500ms after representation of the stimulus. Neuroscientists have shown that 300ms after the visual stimulus, the effect of cognitive task appears in the EEG signals [5, 16]. Therefore, we only used STFT windows that contain 500ms to 800ms (evidence of the visual stimulus to 300ms after) in each EEG trial. In each window, we extract four values which represent the power of the signal in the frequency bands $\delta, \theta, \alpha,$ and β.

Wavelet Coefficients

In the discrete wavelet transform, the input signal is decomposed into low and high frequency bands using low-pass and high-pass filter pairs followed by decimation. In wavelet packet transform, this decomposition is done towards lower frequency band. At each level, the input signal decimated by a factor of 2. In this work, each trial has 2000 samples and we applied a 7-level wavelet packet to decompose the signal in different frequency bands with compressed coefficients. Because of decimation at each level, we have a total number of 16 coefficients in the last level.

C. Classifier

In this work, we employ SVM classifier because of its advantages including good generalization properties, insensitivity to overtraining, and robustness to the curse-of dimensionality. The SVM approach, which is successfully used in many different applications, offers an effective classification strategy in separating the input feature vectors in which the input vector X is projected into a scalar value $f(\boldsymbol{x})$ as,

$$f(\boldsymbol{x}) = \sum_{i=1}^{N} a_i y_i K(\boldsymbol{x}_i, \boldsymbol{x}) + b \tag{4}$$

where $y_i = \{-1, 1\}$, the vectors X_i are the support vectors, N is the number of support vectors, α_i are adjustable weights, b is the bias term, and the function $K(\boldsymbol{x}_i, \boldsymbol{x}) = \emptyset(\boldsymbol{x}_i)^t . \emptyset(\boldsymbol{x})$ is the kernel, where $\emptyset(.)$ is a mapping from the input space to a high dimensional space which creates nonlinear decision boundaries.

V. Feature Selection with Complexity Analysis

A. Measure of Complexity

Power consumption is an important challenge in wearable BCI systems. Computational complexity is a measure of power consumption. In order to consider the cost of features, we extracted the required cost for each individual feature in our feature set. There are different measures to extract the complexity of an algorithm. A widely used measure for complexity is the number of FLoating point Operations Per Second (FLOPS). This measure is a precise representative for required hardware resources. In Digital Signal Processing (DSP), the computational effort is mainly made on additions, subtractions, and multiplications. Since FLOPS is actually a measure for the number of additions, subtractions, and multiplications, it is sufficiently accurate to provide a fair comparison. In this paper, we use FLOPS as a measure for complexity of signal processing. We extracted features on 2s segment of data. In the following, we discuss the complexity of features in terms of FLOPS.

In band power feature extraction, each level of the wavelet transform will take two convolution operations on the input signal

with n samples and filter function of length f. We used a 2s window of the EEG data for band power features. Each convolution needs $fn/2$ operations. Therefore, each wavelet level requires fn floating-point operations. Then, the number of operations for L level wavelet is fnL. Eq. (1) which computes the signal power needs n FLOPS. Features 1 to 4 in the feature vector are band power features.

STFT features are extracted from 1s windows, w, with 90% overlapping. The computational complexity of the FFT is $O(nlogn)$ where n is the length of the signal. Suppose $n = 2^m$, there is a total of $m = log_2 n$ stages of computation, each requiring $3n/2$ additions and multiplications. STFT for each individual window of length n needs $3n/2\ logn$ operations.

As described earlier, wavelet features are extracted by using wavelet packet tree. Assuming that the input signal has n samples and the filter has the length f, each convolution takes $fn/2$ operations. The first wavelet level requires fn floating-point operations. In the next level, because of the decimation, $n/2$ samples remain for processing. Therefore, the number of floating-point operations required for multi-level wavelet is:

$$Number\ of\ operations = fn + \frac{fn}{2} + \frac{fn}{4} + \cdots \leq 2fn \qquad (5)$$

Table I demonstrate complexity measures for each category of features.

B. Cost vs. Accuracy Analysis

Feature selection is used to reduce the dimensionality of the feature set by selecting a subset of features. This not only reduces the computational cost, but improves the classification performance, especially when dealing with a small sample size. Feature selection has been a popular method in single trial EEG classification [9, 10]. However, improving the classification accuracy has been the main concern in previous feature selection methods.

In Fig. 3(a), pseudo code of a simple greedy search algorithm for sequential forward selection is illustrated. Starting from an empty set Y_0, we sequentially add feature $\boldsymbol{x}_i^*$ that maximizes the objective function, $Accuracy(.)$ (i.e. cross-validated classification accuracy) when combined with the features in Y_{i-1} that have currently been selected to form Y_i. Then, we search for the best subset, Y_w leading to the maximum accuracy. This feature selection strategy is effective to increase the classification accuracy; however, the low cost features are not of special interest.

TABLE I. COMPUTATIONAL COST IN TERMS OF FLOPS FOR EACH FEATURE IN THE FEATURE SET.

Feature No.	Feature Type	Required FLOPS
1	Band power	(7fn + n)/2
2	Band power	(7fn + n)/2
3	Band power	(6fn + n)/2
4	Band power	(5fn + n)/2
5-20	Wavelet Coeff	~(2fn)/2
21-24	STFT (500-1500)	(1.5*wlog2w)/2
25-28	STFT (600-1600)	(1.5*wlog2w)/2
29-32	STFT (700-1700)	(1.5*wlog2w)/2
33-36	STFT (800-1800)	(1.5*wlog2w)/2

n: Number of sample points in each trial.
f: filter length in mother wavelet
w: STFT window

In this work, we formulate a sequential feature selection algorithm based on the analysis of accuracy vs. computational cost. The objective of the proposed feature selection method is to take into account both accuracy and computational cost of the selected subset of the features. In this way, a candidate feature has to satisfy two objective functions: 1) the classification accuracy, $Accuracy(.)$, and 2) The measure of complexity presented in the previous section, FLOPS(.), in a stepwise manner. The pseudo code shown in Fig. 3(b) summarizes the formulated stepwise feature selection in this paper. The algorithm receives a set of features $\{\boldsymbol{x}_i \mid i=1, ...,m\}$ and a desired threshold which is the maximum acceptable drop in accuracy compared to the best classification accuracy obtained, and returns Y_w as the subset of the features which is the best subset that satisfies the two objective functions. Starting from an empty set, Y_0, in the first step, we select the candidate features that meet the accuracy constraint (Accuracy is in the range defined by the threshold). As the second step, we select the feature $\boldsymbol{x}_i^*$ that minimizes the objective function, FLOPS(). We continue selecting features in the feature pool and pick the feature set resulting in the highest cross validated accuracy.

Input: the set of extracted features, $\{\boldsymbol{x}_i \mid i=1, ...,m\}$
$Y_0 \leftarrow \{\Phi\}$ // Start with the empty set
$Subset\_Accuracy_0 \leftarrow \{\Phi\}$ // initialize accuracy of the selected subsets
for i=1 to m & $\boldsymbol{x}_i.selected = 0$ // where $\boldsymbol{x}_i$ not marked
  $\boldsymbol{x}_i^* \leftarrow argmax\ \{Accuracy\ (Y_{i-1} \cup \boldsymbol{x}_i) \mid i=1, ...,m\ \&\ \boldsymbol{x}_i.selected = 0\}$
  // Select the next best feature which is not selected yet
  $\boldsymbol{x}_i^*.selected \leftarrow 1$ // Mark $\boldsymbol{x}_i^*$ as selected
  $Y_i \leftarrow \{Y_{i-1} \cup \boldsymbol{x}_i^*\}$ // form the i[th] feature subset by adding $\boldsymbol{x}_i^*$
  $Subset\_Accuracy_i \leftarrow Accuracy\ (Y_i)$ // store the accuracy of Y_i
end for
$w \leftarrow argmax\ \{Subset\_Accuracy_i \mid i=1, ...,m\}$ // find index of the best subset
return Y_w

(a) Pseudo code for sequential feature selection

Input: The set of extracted features, $\{\boldsymbol{x}_i \mid i=1, ...,m\}$
  The desired threshold on accuracy tolerance, $thresh$
$Y_0 \leftarrow \{\Phi\}$ // Start with the empty set
$Subset\_Accuracy_0 \leftarrow \{\Phi\}$ // initialize accuracy of the selected subsets
for i=1 to m & $\boldsymbol{x}_i.selected = 0$ // where $\boldsymbol{x}_i$ is not selected
  $Candidates \leftarrow \{\Phi\}$
  $max\_Accuracy \leftarrow max\{Accuracy\ (Y_{i-1} \cup \boldsymbol{x}_i) \mid i=1, ...,m\ \&\ \boldsymbol{x}_i.selected = 0\}$
  k ← 0
  for i=1 to m & $\boldsymbol{x}_i.selected = 0$ // where $\boldsymbol{x}_i$ not marked
    $k \leftarrow k+1$
    $Candidates_k \leftarrow Accuracy\ (Y_{i-1} \cup \boldsymbol{x}_i) \geq max\_Accuracy - thresh$
  end for // Select features that satisfy the Accuracy constraint
  $\boldsymbol{x}_i^* \leftarrow argmin\ \{FLOPS\ (Y_{i-1} \cup \boldsymbol{x}_j) \mid \boldsymbol{x}_j \in candidates\}$ // Select the best feature that satisfies the computational cost constraint
  $\boldsymbol{x}_i^*.selected \leftarrow 1$ // Mark $\boldsymbol{x}_i^*$ as selected
  $Y_i \leftarrow \{Y_{i-1} \cup \boldsymbol{x}_i^*\}$ // form the i[th] feature subset by adding $\boldsymbol{x}_i^*$
  $Subset\_Accuracy_i \leftarrow Accuracy\ (Y_i)$ // store the accuracy of Y_i
end for
$w \leftarrow argmax\ \{Subset\_Accuracy \mid i=1, ...,m\}$ // find index of the best subset
return Y_w

(b) pseudo code for the proposed stepwise feature selection

Figure 3. pseudo codes of the feature selection strategy for (a) simple sequential forward selection, and (b) the proposed feature selection.

VI. EXPERIMENTAL RESULTS

A. Results

We conduct the experiments with the Go/NoGo task, as described in section III. 17 subjects are participated in this study. In [16], authors analyzed the influence of perceptual categorization on inhibitory processing by measuring N2-P3

response in the Go/NoGo task. They mentioned that N2 (a negative peak around 200ms following the visual stimulus) is found over fronto-central areas and P3 (a positive peak around 300ms following the visual stimuli) is a fronto-central component. They considered averaging across participants on frontal channels like Fz for N2 component and central electrodes such as FCz for P3 effect. They showed that these two channels are good candidates to observe the N2-P3 responses in the Go/NoGo task. Therefore, we consider the same channels in this study. We extract the band power, wavelet, and STFT features from the Fz and FCz channels as described in the feature extraction section. To evaluate the performance of the classification system, we use the ten-fold cross validation technique (10-CV) in which 0.9 of the whole training data is used in the training phase and the rest is used for testing. This partitioning is permuted 10 times and 10 full permutation is done and averaged to obtain the results. For the result comparison, a SVM classifier with an RBF kernel function is used. We use a grid search to optimize the two parameters (C and γ) of the kernel for each validation set.

Table II reports the classification results of the system on 17 subjects for the Go/NoGo task. The first three columns in Table II show the classification results of each single feature (i.e. band power, STFT and wavelet) applied to the EEG signal. The forth column shows the results of the best subset of features selected by the sequential feature selection algorithm. The values in the parentheses demonstrate the number of features that are selected. As reported in Table II, the sequential feature selection method considerably improves the classification accuracy of the system compared to each individual feature set. However, the computational cost of the system is not taken into consideration.

Table III illustrates the results of the system with the cost vs. accuracy analysis incorporated in the feature selection algorithm with 0.5%, and 2% thresholds, respectively. For example, 0.5% threshold dictates that the final accuracy can be at most 0.5% below the best achievable accuracy. In Table III, columns two and three are the accuracy and computational cost for the standard forward selection method shown in Figure 3(a). The three columns in the middle of Table III are the accuracy, cost, and percentage cost reduction using the proposed feature selection method in Figure 3(b) for tolerating 0.5% drop of accuracy. The left three columns are results for 2% threshold. It is shown in the last column of Table III that the computational cost in terms of FLOPS decreases significantly by minimally compromising the accuracy. The reduction in FLOPS is more than 60% if we permit 2% drop in accuracy of the system which is a reasonable choice when a real-time system is desirable.

One can observe in Table III that the proposed feature selection method is able to maintain good classification accuracy at a low computational cost. The thresholds that specify the margin of promising accuracy is a tunable parameter. With a smart choice of threshold, considerable reduction in computational cost is achieved.

B. Disscusion

As we mentioned earlier, feature selection is a well-established concept in machine learning theory. Many standard algorithms are proposed to select the best subset of features in a way to solve the problem of dimensionality curse in order to improve the classification performance. Avoiding irrelevant and redundant features is another reason for applying the feature selection methods with the aim to keep only useful information for the subsequent classification task. Another advantage of using the feature selection methods is to decrease the size of feature vectors leading to reduction in the computational cost of the classifier.

Fig. 4 shows the computational cost vs. accuracy curve for the Go/NoGo task. The horizontal axis is the computational cost for feature extraction. The tradeoff curve shows significant potential in the computational cost reduction by a slight compromise in the accuracy. This measure can be adjusted on-the-fly for mobile systems executing BCI algorithms, considering the amount of available energy. For example, in the case of using rechargeable batteries, when the system is low on energy and the user may not have access to recharging supply, the computational load and the power consumption may be decreased by a factor of three (from 30 KFLOPS to 10 KFLOPS – as shown in Fig. 4) at a minor compromise on the accuracy (less than 3%) prolonging the battery life.

TABLE II. CLASSIFICATION RESULTS ON DIFFERENT SUBJECTS

Subjects	**Accuracy (%)**			
	Band Power	***STFT***	***Wavelet***	***Best subset***
Sub 1	88.88	90.74	87.65	91.75 (14)
Sub 2	79.02	82.51	83.21	88.11 (14)
Sub 3	85.0	89.37	90.0	93.17 (11)
Sub 4	78.0	79.33	77.33	84.66 (13)
Sub 5	80.48	84.14	83.53	90.24 (23)
Sub 6	82.12	82.12	81.00	86.59 (18)
Sub 7	80.55	84.06	80.76	86.81 (16)
Sub 8	81.72	84.94	82.25	87.09 (16)
Sub 9	82.01	83.59	81.48	83.59 (20)
Sub 10	80.43	89.85	84.78	91.30 (17)
Sub 11	78.33	82.22	80.55	83.95 (10)
Sub 12	78.02	82.41	82.96	88.46 (11)
Sub 13	82.19	82.19	78.53	87.43 (14)
Sub 14	80.05	80.32	81.82	86.84 (20)
Sub 15	80.43	82.29	83.76	96.50 (19)
Sub 16	81.58	82.27	79.37	85.46 (13)
Sub 17	71.79	83.33	79.48	89.47 (21)
Average	80.56	83.96	82.44	88.32

TABLE III. CLASSIFICATION RESULTS AND THEIR COSTS USING STEPWISE FEATURE SELECTION WITH THE THRESHOLD = 0.5 % AND THRESHOLD = 2 %

Sub	**No cost analysis**		**Threshold = 0.5 %**			**Threshold = 2 %**		
	Acc	***FLOPS***	***Acc***	***FLOPS***	***Cost (%) Red***	***Acc***	***FLOPS***	***Cost (%) Red***
Sub 1	89.47	41,882	88.34	32,407	22.6	85.26	9,984	76.2
Sub 2	85.46	24,407	85.46	7,474	69.4	85.46	7,474	69.4
Sub 3	91.75	26,933	91.67	24,933	74.3	91.35	17,459	35.2
Sub 4	88.11	26,933	88.11	9,459	64.9	88.11	9,459	64.9
Sub 5	93.17	22,933	93.12	16,933	26.2	90.62	14,949	34.8
Sub 6	84.76	24,407	84.56	16,933	30.6	84.56	16,933	30.6
Sub 7	90.24	34,407	89.54	16,933	50.8	89.22	16,933	50.8
Sub 8	86.59	36,407	85.87	28,933	20.5	85.21	16,933	53.5
Sub 9	85.81	41,882	85.71	9,459	77.4	85.71	9,459	77.4
Sub 10	87.09	41,882	86.52	34,407	17.9	85.48	9,459	77.4
Sub 11	83.59	38,882	83.06	7,474	80.8	83.06	7,474	80.8
Sub 12	91.30	37,882	90.45	30,407	19.8	89.13	14,949	60.5
Sub 13	83.95	22,933	83.88	16,933	26.2	83.33	16,933	26.2
Sub 14	88.46	24,407	87.91	24,407	0	86.26	16,933	30.6
Sub 15	87.43	41,882	86.67	32,423	22.6	84.81	15,474	63.1
Sub 16	86.84	41,882	86.38	9,459	77.4	83.55	1,984	95.3
Sub 17	96.50	41,882	94.80	31,882	23.9	93.70	16,933	59.6
Avg	88.27	33,637	87.77	20,639	38.7	86.75	12,925	61.6

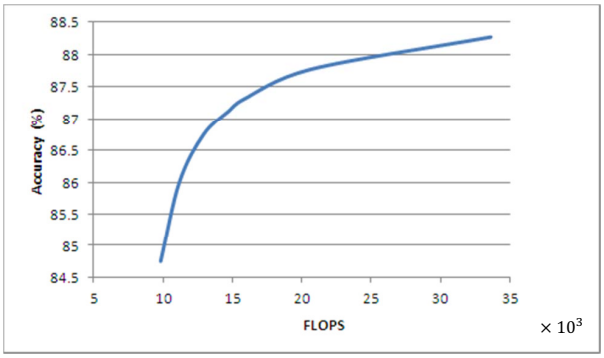

Figure 4. Accuracy versus Computation Cost.

The amount of available energy can be a parameter to tune the cost and accuracy. As a future work, we aim to model the supply power characteristics in the system and set the required threshold for the feature selection method.

C. Online Application (eMotiv® Push/NoPush)

In this section, we perform our analysis on the data obtained from a wireless wearable EEG headset by eMotiv®, which has 14 channels and 2 reference electrodes. The EEG data were recorded for two events: imagine pushing or not pushing an object. In this task, the participant imagines pushing or not pushing (neutral) for 5 seconds. This task contains 100 trails equally distributed between Push and NoPush. The same feature extraction methods were used to extract relevant features. As the sampling frequency changes, the data window and the number of data samples for feature extraction also changes. We use a window of length 2.7s (350 samples) for band power features. For STFT, data is segmented in windows of 64 samples, and then Fourier transform is applied to extract STFT features. Wavelet coefficients are extracted for a window of 2s. Table IV reports the results of the cost vs. accuracy analysis for the online Push/NoPush task. As reported in Table IV, the required FLOPS for this task is much smaller compared to the Go/NoGo task due to smaller frequency of sampling and number of electrodes.

We show that by allowing a slight drop in accuracy, we can reduce the computational cost significantly. The results in Table IV demonstrate that the proposed cost vs. accuracy analysis for feature selection is effective for the online application under consideration.

TABLE IV. RESULTS ON THE COST VS. ACCURACY ANALYSIS FOR FEATURE SELECTION ON THE ONLINE TASK OF PUSH/NOPUSH

Feature selection method	Accuracy	FLOPS	Cost (%) Reduction
no cost measure	78 %	1,763	N/A
Threshold = 1 %	77 %	767	56.5
Threshold = 2 %	76 %	652	63.1
Threshold = 3 %	75 %	442	74.9
Threshold = 6 %	72 %	96	94.6

VII. CONCLUSION

In this paper, we presented a system for single-trial EEG classification, also known as BCI. We investigated the computational cost vs. accuracy of the signal processing at the feature level using two EEG data sets. We successfully incorporated the computational cost of the signal processing in the feature selection mechanism. Results of our evaluation on the Go/NoGo task and the online Push/NoPush task show that with a small tolerance in the accuracy, the computational cost can be reduced significantly. We will use these findings to develop effective power-aware classification systems for real-time BCI applications.

ACKNOWLEDGMENT

The authors would like to thank Drs. Mandy Maguire and John Hart for helpful discussions and providing the inhibition data (Go/NoGo tasks) used in this study.

REFERENCES

[1] J. Wolpaw, N. Birbaumer, W. Heetderks, D. McFarland, P. Peckham, G. Schalk, E. Donchin, L. Quatrano, C. Robinson, and T. Vaughan, "Brain-computer interface technology: a review of the first international meeting," *Rehabilitation Engineering, IEEE Transactions on*, vol. 8, no. 2, pp. 164–173, 2000.

[2] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," *Proceedings of the IEEE*, vol. 89, no. 7, pp. 1123–1134, 2001.

[3] D. McFarland, W. Sarnacki, and J. Wolpaw, "Electroencephalographic (EEG) control of three-dimensional movement," *Journal of Neural Engineering*, vol. 7, p. 036007, 2010.

[4] G. Schalk, K. Miller, N. Anderson, J. Wilson, M. Smyth, J. Ojemann, D. Moran, J. Wolpaw, and E. Leuthardt, "Two-dimensional movement control using electrocorticographic signals in humans," *Journal of neural engineering*, vol. 5, p. 75, 2008.

[5] T. Harmony, A. Alba, J. Marroqun, and B. González-Frankenberger, "Time-frequency-topographic analysis of induced power and synchrony of EEG signals during a Go/No-Go task," *International Journal of Psychophysiology*, vol. 71, no. 1, pp. 9–16, 2009.

[6] I. Paul-Jordanov, M. Bechtold, and C. Gawrilow, "Methylphenidate and if-then plans are comparable in modulating the P300 and increasing response inhibition in children with ADHD," *ADHD Attention Deficit and Hyperactivity Disorders*, pp. 1–12, 2010.

[7] Y. Khan and F. Sepulveda, "Brain-computer interface for single-trial EEG classification for wrist movement imagery using spatial filtering in the gamma band," *Signal Processing, IET*, vol. 4, no. 5, pp. 510 –517, oct. 2010.

[8] K. Sadatnezhad, R. Boostani, and A. Ghanizadeh, "Classification of BMD and ADHD patients using their EEG signals," *Expert Systems with Applications*, 2010.

[9] A. Cabrera, D. Farina, and K. Dremstrup, "Comparison of feature selection and classification methods for a brain–computer interface driven by non-motor imagery," *Medical and Biological Engineering and Computing*, vol. 48, no. 2, pp. 123–132, 2010.

[10] Y. Atum, I. Gareis, G. Gentiletti, R. Acevedo, and L. Rufiner, "Genetic feature selection to optimally detect P300 in brain computer interfaces," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*. IEEE, 2010, pp. 3289–3292.

[11] F. Tenore, V. Aggarwal, J. White, M. Schieber, and N. Thakor, "Computational complexity versus accuracy in classification of cortical neural signals," in *Neural Engineering, 2009. NER'09. 4th International IEEE/EMBS Conference on*. IEEE, 2009, pp. 750–753.

[12] C. Lin, Y. Chen, T. Huang, T. Chiu, L. Ko, S. Liang, H. Hsieh, S. Hsu, and J. Duann, "Development of wireless brain computer interface with embedded multitask scheduling and its application on real-time driver's drowsiness detection and warning," *Biomedical Engineering, IEEE Transactions on*, vol. 55, no. 5, pp. 1582–1591, 2008.

[13] C. Lin, L. Ko, C. Chang, Y. Wang, C. Chung, F. Yang, J. Duann, T. Jung, and J. Chiou, "Wearable and wireless brain-computer interface and its applications," *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*, pp. 741–748, 2009.

[14] D. Allen and C. MacKinnon, "Time-frequency analysis of movement-related spectral power in EEG during repetitive movements: A comparison of methods," *Journal of neuroscience methods*, vol. 186, no. 1, pp. 107–115, 2010.

[15] L. Safavian, W. Kinsner, and H. Turanli, "A quantitative comparison of different mother wavelets for characterizing transients in power systems," in *Electrical and Computer Engineering, 2005. Canadian Conference on*. IEEE, 2006, pp. 1461–1464.

[16] M. Maguire, M. Brier, P. Moore, T. Ferree, D. Ray, S. Mostofsky, J. Hart Jr, and M. Kraut, "The influence of perceptual and semantic categorization on inhibitory processing as measured by the N2-P3 response," *Brain and cognition*, vol. 71, no. 3, pp. 196–203, 2009.