MODEL BASED EEG SIGNAL PURIFICATION TO IMPROVE THE ACCURACY OF THE BCI SYSTEMS

Farid Atry, Amir H. Omidvarnia, and S. Kamaledin Setarehdan

Control and Intelligent Processing Centre of Excellence, ECE Department, Faculty of Engineering, University of Tehran
Tehran, Iran
email: ksetareh@ut.ac.ir

ABSTRACT

Brain-Computer Interfaces are an interesting emerging technology that translates intentional variations in the Electroencephalogram (EEG) into a set of particular commands in order to control a real world machine. For this purpose it is necessary to classify EEG signals correlated with various physical or mental activities. Most of the work in BCI research is devoted to increase the accuracy of the EEG classification. Due to the noisy nature of the EEG including the background brain activity, one of the potential approaches to increase the classification accuracy is to improve the SNR of the EEG signals. In this paper EEG signal denoising in some active channels is investigated using the parametric models developed for relating their signals to the signals of all other channels. The models are used for signal purification in the selected channels. It is shown that the purified signals can improve the classification accuracy of the EEG signals up to 15%.

1. INTRODUCTION

EEG signals as a new communication means between a paralyzed person who is unable to interact physically with his environment has shown promising abilities and applications. There are several different approaches where each method uses a different property of the EEG signal for EEG based communication. For example, the P300 based BCI system uses the P300 component of the EEG signal (a signal peak which appears 300 ms after the time that one's attention is triggered by a subject) [1,2]. In the motion based BCI system, the variation in the EEG signal due to the movement or the imagination of the movement in a particular body organ like hands or feet is used for this purpose [3-5]. Finally, there are some particular mental tasks such as multiplication of two digits, rotating an imaginary object in the three dimensional space and/or writing a letter to a friend which causes a detectable variation in the EEG signal and therefore is used for EEG based communication [6].

Various approaches for EEG signal classification in each of these classes with different degrees of success were reported in the past [1-8]. The main problem with most approaches is the low classification accuracy, a main reason for which is the noisy nature of the EEG signals.

One of the main sources of noise and artefact in the EEG signals is the interferences from other bio-potentials sources like the electro-occulogram (EOG), the electrocardiogram (ECG), the electromyogram (EMG), and most importantly the background activity of the brain itself [9,10]. It is believed that only particular parts of the brain are activated in response to a particular BCI task. This means that the EEG channels, which are closer to the active brain regions (*active channels* for short), have more relevant information with the BCI tasks compared to all other channels. Therefore, it seems to be helpful if one could *purify* the signal of an active channel using the functional relationship between the neighbouring channels.

This paper describes a novel method for signal purification of the active channels by deriving a model between the active channel and all other channels. To show the effectiveness of the proposed algorithm, same feature extraction and classification schemes were applied to the EEG data both before and after purification of the signals of the active channels by the proposed algorithm. The comparative results show an improvement of about 15% in the classification accuracy using a 10 by 10 fold cross validation scheme.

The rest of the paper is organised as follows. Section 2 describes the data set used in this work. In section 3, the proposed algorithm is explained in details first. Then the EEG features and the classification scheme used in this work are described. The comparative results of the application of the proposed algorithm to the data set are demonstrated in Section 4. Finally, Section 4 concludes the paper and summarises the results.

2. THE DATA SET

To demonstrate the effectiveness of the proposed algorithm we used the EEG data obtained from the BBCI group [7]. It was recorded from a normal subject during a no-feedback BCI session. The subject was sitting on a chair with his arms in the resting position on the table and his fingers in the standard typing position on the computer keyboard. The task was to use the index and little fingers of both left and right hands for typing the characters in a self-chosen order and timing. A data set including 316 epochs of 500 ms length each (i.e. typing 316 letters) was recorded by a sampling rate of 100Hz. Each of the recordings was started 630 ms before the physical key press and ended 130 ms before the physical key press. The data where recorded from the standard channels of F3, F1, Fz, F2, F4, FC5, FC3, FC1, FCz, FC2, FC4, FC6, C5, C3, C1, Cz, C2, C4, C6, CP5, CP3, CP1, CPz, CP2, CP4, CP6, O1, O2 [6] (figure

1). In our study only the first 26 channels were considered by omitting the signals from the O1 and O2 channels.
Then, the data set was divided into the two main classes where one class contains EEGs related to typing by the left hand and the other class contains EEGs related to typing by the right hand. The aim of this work was to classify the EEG signals of the two classes.

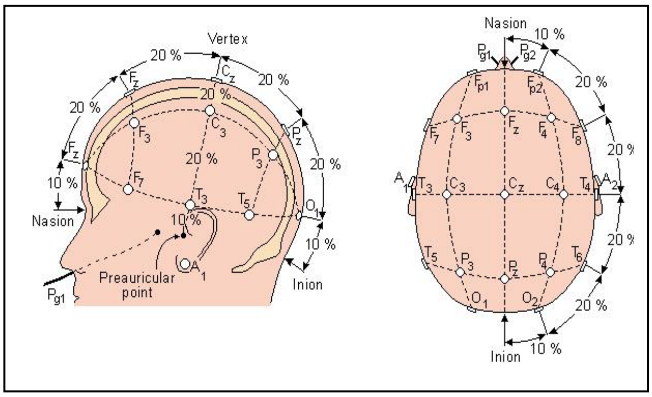

Figure 1: The standard locations of the EEG electrodes (10-20 Standard)

3. METHODOLOGY

3.1. EEG Signal Purification

Background brain activities which are irrelevant to the BCI tasks continuously generate EEG signals that can be recorded almost anywhere over the scalp in all channels.
These signals interfere with the EEG signals triggered by the BCI tasks and generated by only particular regions of the brain. The amplitude of the recorded signal in a given channel depends highly to the distance between the source of the signal and the channel and also to the transfer function of the brain tissue between them (in other words, the spatial filter between the source and channel). In addition to the background EEG, there are other sources of artefact like ECG, EOG, EMG, motion artefacts, eye blinking and the 50Hz from the power line, which are usually affect all EEG channels [8]. Except the power line noise that is almost similar for all channels, other noise and artefacts have different effects on different channels depending on transfer function between the channel and the artefact source. So each channel needs its own estimation of amount of artefact and noise interference. Also, the different noise and artefact sources are not available individually and only the recorded mixture of signals from different sources is available. We believe that the recorded mixture of various signals at different channels can be used for estimating the original artefacts affecting each channel.
For this purpose, the block diagram of Figure 2 is proposed that consists of an “Interference” block and an “ARMA” block. A set of selected channels are fed into the “Interference” block to estimate the amount of the interfering signal from each one of these channels affecting the active output channel represented by the “Target Output” in Figure 2. As a result, the common component between each input channel and the output channel will be estimated by the "Interference" block. The “ARMA” block simultaneously estimates the output by filtering white noise in a manner that sum of its estimated signal and the output signal from the "interference" block reconstruct the “Target Output”. It must be noted that they work simultaneously and they are not independent from each other.

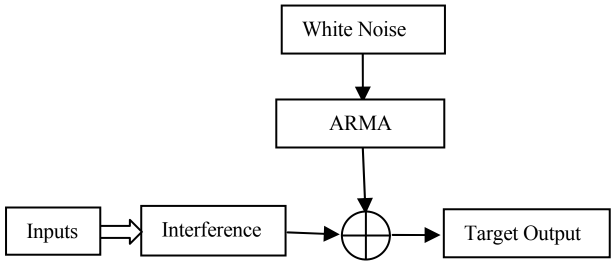

Figure 2: Block diagram used for other channels interference reduction

The “Interference” block can be implemented in several ways, for example linear, nonlinear and/or stochastic. If a linear model is considered for the interference block, the whole scheme will be very similar to the Box-Jenkins structure [11, 12] for which the block diagram is shown in Figure 3 and its mathematical description is as follows:

$$y(t) = [B(q)/A(q)]u(t-n_k) + [C(q)/D(q)]e(t) \qquad (1)$$

For calculating the parameters of each interference block, an iterative algorithm [11] was employed.
Due to the non-stationary nature of both the EEG signals and the artefacts, the coefficients of the filters must be estimated using the short length of the data recorded during each trial (500 ms). This means that the coefficients have to be estimated for each trial separately independent from other trials. Estimating these coefficients provides the “ARMA” and the “Interference” blocks filters for each trial. Interference of the “Target Output” for each trial is computed by applying estimated “Interference” filters to their associate inputs and summing up their outputs. Subtracting this computed interference from target output provides a new signal with a less amount of the interfering signals which is called denoised signal. Applying this to all EEG channels in each trial provides a new denoised data set to be used in feature extraction procedure described in Section 3.2. As it will be shown in Section 3.2 the coefficients of these filters can be used as the signal features.
It is shown in [12] that the EEG signal correlated with the finger movements (like the typing activity in this work) has most of its energy in the 0.5-8Hz frequency range. Therefore, the EEG signals in the data set were initially filtered to reduce the noise effects prior to processing by the proposed algorithm.
Denoising can also be carried out in the frequency domain. In this case the Box-Jenkins estimates the amount of the “Target Output” in each frequency, using the “Interference”

inputs of the nearby frequencies. After deducting the estimated “Interference” from the “Target Output” the resulting signal is returned back to the time domain and used as the denoised data. The advantage of denoising in the frequency domain is that the EEG signals which are correlated with finger movements have frequency patterns that are not easily distinguishable in time domain. Using frequencies out of the frequency band of our interest (.5-8Hz) helps to find a more accurate model by providing more information about the noise.

3.2. Feature Extraction and Classification

In this work, two different groups of features were extracted from the original EEG and the denoised signals to test the proposed algorithm. The first group was the commonly used AR coefficients. The AR coefficients of a signal $x[n]$ satisfy the following equation:

$$x[n] = \sum_{i=1}^{N} a_i x[n-i] + v[n] \qquad (2)$$

where a_i are the AR model coefficients of order N, $x[n]$ is the input signal (the EEG signal) and $v[n]$ is the white noise.

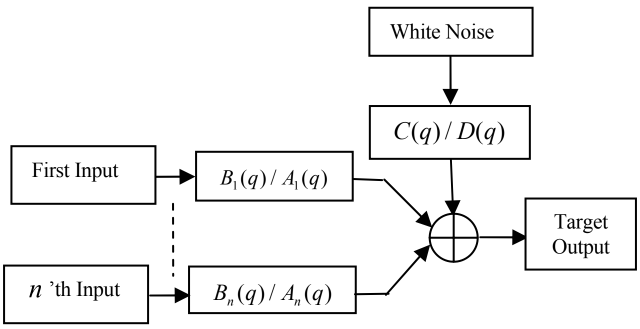

Figure 3: The block diagram of the Box-Jenkins model

As a reference, first the original raw data were considered in the feature extraction and classification procedure. For all of the raw signals recorded from the 26 channels during each typing activity, the coefficients of the AR models of order 5 were computed by the *Burg* method producing a total number of 130 coefficients. These coefficients can be considered as the features for that particular typing activity.
An artificial neural network with the Multilayer Perceptron (MLP) structure was used in this work as the classifier. Ten linear neurons were selected as the first layer, each getting an input from each of the input features. A sigmoid neuron was also selected as the output neuron.
Due to the large number of features, it was necessary to reduce the complexity of the problem by using only a small number of selected channels. For this purpose almost all possible combination of channels were tested using a 10 by 10 fold cross validation procedure. The resulting best set of channels for separating the two classes of “typing by left hand” or “typing by right hand” include the C1, C2, C3 and C4 channels producing only 53% of the classification accuracy. Then the purified EEG signals of the same channels by the proposed method in Section 3.1 were used for feature extraction and classification. A correct classification rate of about %68 for the data purified in the time domain and about %73 for the data purified in the frequency domain was achieved showing an improvement of about %15 and %20 in the classification compared to the initial attempt using the raw data respectively.
The second group of the features that were considered in this work was the coefficients of the Multi-Variable AR (MVAR) model on the EEG signal. Compared to the AR model, the MVAR tries to not only find a relationship between the current amount of a channel and its previous values, but also its relationship to the values of some other neighbouring channels. Mathematically it can be represented as follows:

$$x[n,CH] = \sum_{i=1}^{N} a_i x[n-i,CH] + \sum_{j=1}^{M}\sum_{i=1}^{N_j} b_j x[n-i,Ch_j] \qquad (3)$$

where CH is the channel itself, Ch_i are the other neighbouring channels and $v[n]$ is the white noise.
Again using the original raw data as the reference, the set of the MVAR coefficients $[a_1,\ldots,a_N,b_1,\ldots,b_M]$ that were computed by a least square algorithm, used as the features for that particular typing activity. Due to the large number of features, similar to the previous attempt, these features fed into the classifier in a 10 by 10 fold cross validation framework. Best classification accuracy was obtained by setting C3 and C4 channels as output when input channels were set to (F3 and CPz) and (F4 and CPz) respectively. A correct classification rate of about %55 was achieved this time. Applying the MVAR feature extraction method to the denoised data set provide up to %69 classification rate for both the frequency and time domain denoised data sets.
Finally, the coefficients of Box-Jenkins structure employed as features. It means that each channel considered separately as the “Target Output” and some of the other channels as the “Interference” block inputs. Due to the large number of coefficients (features) an exhaustive search is carried out. First, the features of each channel fed into an ANN and the best possible classification rate is computed. Then the process was repeated for all possible combinations of two, three and four channels. It was found that the combination of two channels produces the best possible classification rate. A similar exhaustive procedure is carried out for selecting the “Interference” inputs. Finally, the best classification rate of 68% and 70% was achieved when C3 and C4 considered as “Target Output” and input channels were set to (F3 and CPz) and (F4 and CPz) for both time and frequency domain denoised data respectively.

4. RESULTS

Classification rates of AR and MVAR features of the original and the denoised data are presented in Table 1. As it can be seen in Table 1, the AR and MVAR features of the denoised data could provide up to 15% better classification rate than AR and MVAR features of original data. Also, it is clear from Table 1 that denoising in the frequency domain could provide more purified signal than denoising in the time domain.

Table 1: classification rate using the features extracted from original and denoised data denoised in the time and or frequency domains.

The features	The data	Correct Classification Rate
AR Coefficients	original	53%
AR Coefficients	denoised in the time domain	68%
AR Coefficients	denoised in the frequency domain	73%
MVAR Coefficients	original	55%
MVAR Coefficients	denoised in the time domain	69%
MVAR Coefficients	denoised in the frequency domain	69%
BJ Coefficients	denoised in the time domain	68%
BJ Coefficients	denoised in the frequency domain	70%

5. CONCLUSIONS

In this paper the subject of Brain-Computer Interfacing was addressed. Brain-Computer Interfacing is an interesting emerging technology that translates intentional variations in the Electroencephalogram (EEG) into a set of particular commands in order to control a real world machine. For this purpose it is necessary to classify EEG signals correlated with various physical or mental activities. Most of the work in BCI research is devoted to increase the accuracy of the EEG classification. Due to the noisy nature of the EEG including the background brain activity, one of the potential approaches to increase the classification accuracy is to improve the SNR of the EEG signals. In this paper EEG signal denoising in some active channels was investigated, using the parametric models developed for relating the signals of the active channels to the signals of all other channels. The models are used for signal purification in the selected channels which are closer to the active brain regions for a particular BCI task.

It was shown that the purified signals can improve the classification accuracy of the EEG signals up to 15% compared to the classification results obtained using the original data.

6. REFERENCES

[1] Farwell L.A., Donchin, E. “Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials,” Electroencephalography & Clinical Neurophysiology. 70(6):510-23, 1988.

[2] Donchin, E., Spencer, K.M., Wijensinghe, R. “The mental prosthesis: Assessing the speed of a P300-based brain-computer interface,” *IEEE Trans. Rehab. Eng*. 8:174-179, 2000.

[3] C. Neuper, A. Schlögl, G. Pfurtscheller. “Enhancement of left-right sensorimotor EEG differences during feedback-regulated motor imagery” *J Clin Neurophysiol*. 16(4):373-82, 1999.

[4] Pfurtscheller G, Neuper C, Schlögl A, Lugger K. “Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters”, *IEEE Trans Rehabil Eng.* 6(3):316-25, 1998.

[5] Gert Pfurtscheller, Christa Neuper. “Motor Imagery and Direct Brain–Computer Communication”, Proceedings of the IEEE, VOL. 89, NO. 7, JULY 2001.

[6] Charles W. Anderson, Saikumar V. Devulapalli, Erik A. Stolz, “EEG as a means of communication: preliminary experiments in EEG analysis using neural networks,” ASSETS 1994: 141-147

[7] http://ida.first.fhg.de/projects/bci/competition

[8] Jonathan R. Wolpaw, Dennis J. McFarland, Theresa M. Vaughan, and Gerwin Schalk, “The Wadsworth Center Brain–Computer Interface (BCI) Research and Development Program”, IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING, VOL. 11, NO. 2, JUNE 2003

[9] Ha¨ma¨la¨inen M., Hari R., Ilmoniemi R. J., Knuutila J. and Lounasmaa O. V. (1993), “Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain,” *Rev. Mod. Phys.* 65**,** 413–497.

[10] Barry D. Van Veen, Wim van Drongelen, Moshe Yuchtman, and Akifumi Suzuki, “Localization of Brain Electrical Activity via Linearly Constrained Minimum Variance Spatial Filtering,” IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 44, NO. 9, SEPTEMBER 1997 867

[11] User manual of MATLAB 6.5, “http://www. mathworks.com”

[12] Roman Krepki, Benjamin Blankertz, Gabriel Curio, and Klaus-Robert Müller, “The Berlin Brain-Computer Interface (BBCI): towards a new communication channel for online control of multimedia applications and computer games,” *9th International Conference on Distributed Multimedia Systems (DMS'03)*, pages 237-244, 2003.