Parametric Models and Spectral analysis for Classification in Brain-Computer Interfaces

S Kelly[1], D Burke[1,2], P de Chazal[1], R Reilly[1]

[1]Electronic and Electrical Engineering, National University of Ireland, Dublin, Ireland
[2]National Rehabilitation Hospital, Dublin, Ireland
simon.kelly@ee.ucd.ie

Abstract

Parametric modelling strategies and spectral analysis are explored in conjunction with Linear Discriminant Analysis to facilitate an EEG based direct-brain interface for use by disabled people. A self-paced typing exercise is analysed by employing for feature extraction, respectively, an autoregressive model, an autoregressive with exogenous input model, and a time-frequency decomposition of the data. Modelling both the signal and noise is found to be more effective than modelling the noise alone with the former yielding an accuracy of 70.7% and the latter an accuracy of 57.4%. Using the raw samples of a short-time power spectral density estimate of each trial as features yielded an accuracy of 62.5%.

Introduction

For some people with very severe disabilities (e.g. amyotrophic lateral sclerosis or brainstem stroke), a brain computer interface (BCI) [1] may be the only feasible channel for communicating with others and for environment control. Electroencephalogram (EEG) based BCIs harness electrical signals recorded non-invasively from the scalp. With billions of oscillating communities of neurons as its source, the human EEG potentials are manifested as aperiodic unpredictable oscillations with intermittent bursts of oscillations having spectral peaks in certain preferred bands: 0.1-3.5Hz (delta), 4-7.5Hz (theta), 8-13Hz (alpha), 14-30Hz (beta) and >30Hz (gamma) [2].

Given specific experimental protocols, it is possible to elicit meaningful Event Related Potentials (ERPs) [3] occurring within the background EEG. A common goal of much BCI research is to detect these small ERPs (of order 1μV) from the background EEG (10–50μV) and classify them for subsequent use as inputs for a computer interface. Ensemble averaging of the evoked potential epochs is often performed in detection, but requires a large number of sweeps to obtain a suitable estimate, and the response may vary widely across trials in amplitude, time course, and scalp distribution. The much more favourable but more difficult alternative is single-trial detection of EPs, which if used in a real-time BCI would result in much higher information transfer rates. Several methods for single trial extraction exist e.g. [4], [7], for a review see [5].

Aim

The aim of this paper is to describe the main signal processing stages required for Brain Computer Interfaces. As an example of these stages, the classification of left/right self-paced voluntary finger movement is attempted from single-trial EEG epochs.

BCI Implementation

A BCI can be used for tasks such as making binary decisions, moving a cursor on a screen or selecting icons for control or communication. Figure 1 shows the scheme for signal processing in a BCI. Real-time performance of a BCI typically requires single-trial detection and associated processing. Data is usually acquired at multiple scalp locations, with the remaining stages being artifact removal, feature extraction and pattern classification. Methods for each are proposed with practical results given for comparison of the three cases taken for feature extraction.

In a self-paced voluntary finger movement paradigm, the system must recognise left or right finger movement from EEG. This paradigm is expected to produce an ERP known as the Bereitschaftspotential (BP) – a gradually rising negative potential occurring about 1 second preceding the onset of movement [3]. The occurrence of the BP is largely localised around the sensorimotor area, being most prominent on the contralateral side (electrodes C3 and C4 on the left and right sides respectively, following the American Electroencephalographic Society's 10-20 system [11]).

Methods

A. Artifact Removal:

Artifacts are undesirable potentials from nonneural sources such as eye movements and scalp-recorded Electromyographic (EMG) activity which are regarded as noise. They are usually much more prominent than the EEG, leading to unacceptable negative signal to noise ratios. In a task involving vertical eye movements, Vertical Electrooculographic (VEOG) peaks are a cause for particular concern, as these peaks are picked up to a certain degree at all scalp locations, particularly the frontal sites.

Simpler ways to address this issue include instructing the subject not to blink or move his/her eyes, or to disregard those trials contaminated by artifacts, but these methods can be unreasonable for subjects with disabilities.

0-7803-7503-3/02/$17.00 ©2002 IEEE

Fig. 1 *Processing Stages necessary for BCI analysis and implementation*

A signal processing approach to the removal of artifact can also be applied. EEG recordings $x_1, x_2, \ldots, x_n$ from *n* scalp locations, are grouped into a vector **x,** where $\mathbf{x}= [x_1\ x_2 \ldots x_n]^T$, and for the purposes of this study may be regarded as the linear combination of *m* distinct "source signals" $\mathbf{s} = [s_1\ s_2 \ldots s_m]^T$. Thus one set of signals is a linearly transformed version of the other and an invertible mixing matrix **A** fully describes this transform: $\mathbf{x} = \mathbf{As}$. Some assumption must be made on the number of "sources". This is something that cannot realistically be done for something so complex as the brain, but of concern here is just one particular source signal – the artifact source.

A linear transformation $\mathbf{y} = \mathbf{Wx}$ of the observed data can be used, which separates out the artifact as a source component distinct from the rest of the components in some way, and deleting that component one can work backwards using the inverse transformation to obtain the original observed data with no trace of the artifact peaks at any location. The effectiveness of this method depends on both the validity of the linearity assumption and the particular distinction that is assumed to exist between the artifact component and the rest of the "sources".

In Principle Component Analysis (PCA) the components are assumed to be uncorrelated. The appropriate transformation can be found by performing singular value decomposition on the covariance matrix of the observed **x**, a simple calculation. Figure 2 shows a single trial measured from scalp location FT8, which, although not located close to the eyes, is contaminated with a reasonably large VEOG peak. The thicker trace represents the data at the same location after PCA-based separation and removal of the artifact principle component, which was identified by cross-correlation with a simultaneously recorded VEOG signal.

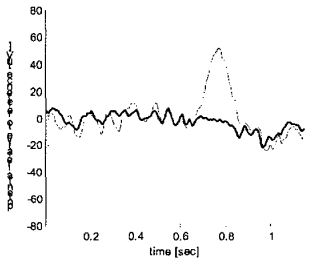

Fig. 2 *A single trial, pre- and post- artifact removal.*

B. Feature Extraction:

The main goal in feature extraction is to characterise an object to be recognised by measurements whose values are very similar for objects in the same category but very different for objects in different categories.

We deal with three cases here. The first two are based on parametric modelling, the technique by which a mathematical model is fitted to a time series. The third is based on spectral analysis.

Case 1: The EEG time series is fitted with an Autoregressive (AR) model. This all-pole model lends itself well to producing the dominant frequencies occurring in the EEG [6]. The AR model can be intuitively rephrased in the frequency domain as a white noise source driving a spectral shaping network, as illustrated in Figure 3.

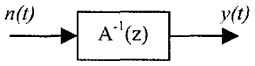

Fig. 3 *AR model*

In this study we use the coefficients $a_1, a_2, \ldots$ of the AR model as features for pattern recognition. Specifically, coefficients from separate AR models for the C3 and C4 electrode time series are combined in a feature vector. When calculating the AR parameters, we solve the Yule-Walker equations [6].

Case 2: Going one step further, we can attempt to model the ERP by using an ensemble-averaged template, filtered by an AutoRegressive Moving Average filter. Using a common denominator results in the AutoRegressive with eXogenous input (ARX) model illustrated in Figure 4.

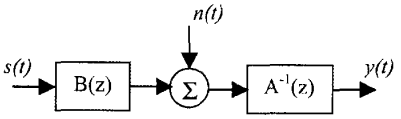

Fig. 4 *ARX model*

For the full ARX model, in terms of the shift operator *q*, and assuming a sampling interval of one time unit, we have

$$A(q)y(t) = B(q)s(t) + n(t)$$

The prediction is written as

$$\hat{y}(t) = -a_1 y(t-1) - \ldots - a_{na} y(t-na) + b_1 s(t-k) + \ldots + b_{nb} s(t-k-nb)$$

where *na* and *nb* are the model orders and k is the delay. Defining θ as the vector of parameters we can write the prediction error

$$e(t,\theta) = y(t) - \hat{y}(t,\theta).$$

The mean square error is

$$E(\theta) = \frac{1}{N}\sum_{t=1}^{N} e^2(t,\theta),$$

where N is the number of samples. We choose $\hat{\theta}$ such that it minimises E resulting in a least squares problem. We define two cost functions as feature vectors:

$$E_L = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left[\hat{s}_3^l(t) - s_3^l(t)\right]^2} + \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left[\hat{s}_4^l(t) - s_4^l(t)\right]^2} \quad \ldots(1a)$$

$$E_R = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left[\hat{s}_3^r(t) - s_3^r(t)\right]^2} + \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left[\hat{s}_4^r(t) - s_4^r(t)\right]^2} \quad \ldots(1b)$$

where the notation s_3^l denotes the template ERP from electrode C3 for left movements and $\hat{s}_3^l$ is the corresponding single trial extracted ERP. Thus for a left movement we might expect to see a low value for E_L and a high value for E_R and vice versa.

Case 3: Oscillatory activity in the brain occurs when populations of neurons form complex networks involving feedback loops. Amplitude increases and decreases in specific frequency bands known as Event Related Synchronisation (ERS) and Event Related Desynchronisation (ERD) respectively, have been observed during, and are believed to be a result of, planning or imagination of limb movement [7]. One such example is an ERD of the mu band (7-13Hz) observed preceding the onset of finger movement, and, like the Bereitschaftspotential, is more pronounced on the contralateral sensorimotor area.
Given this phenomenon, another set of features include those representing the power in this frequency band and how it changes in time over the trial. To this end, a short-time power spectral density (STPSD) estimate can be calculated for each trial – a window of samples is taken in the data and moved along a certain amount, with Welsh's Periodogram PSD estimate performed on each section.

C. *Classification:*

Linear discriminants were used as the classifier model for this study, providing a parametric approximation to Bayes' rule [9]. In response to a set of input features, the output of the classifier is a set of numbers, representing the probability estimate of each class (in our case a left or right movement). The final classification is obtained by choosing the class with the highest probability estimate. Linear discriminants partition the feature space into the different classes using a set of hyper-planes. Optimisation of the model is achieved through direct calculation and is very efficient thus lending itself well to real-time applications.
When implementing a classifier it is important to be able to estimate the expected performance of the classifier on data not used in training. The available data is divided into independent training and testing sets. There are a number of schemes for achieving this and the most suitable for the size of data set used in this study, is *n*-fold cross validation [9]. This scheme randomly divides the available data into *n* approximately equal sized, mutually exclusive "folds". For an *n*-fold cross validation run, *n* classifiers are trained with a different fold used each time as the testing-set, while the other *n-1* folds are used for the training data. The choice of *n* influences the ratio of data used for training/testing with an optimal value of *n* in the range 5-20. Cross validation estimates are generally pessimistically biased, as training is performed using a sub-sample of the available data.

Dataset

The above stages and methods were employed in a left/right classification exercise with the following paradigm:
EEG signals were recorded from one subject in three sessions with a break between each session. The subject was sitting upright, arms relaxed resting on a table, and fingers in the standard typing position at a Qwerty computer keyboard (index fingers at 'F','J' and little fingers at 'A',';'). The task was to press the aforementioned keys with the corresponding fingers in a self-chosen order and timing ('self-paced key typing'). EEG activity was recorded with Ag/AgCl electrodes referenced to the nasion at a sampling rate of 1000 Hz using a band-pass filter from 0.05 to 200 Hz. The data was down-sampled to 100Hz.
The data set consists of 413 single trials of EEG data recorded from 27 channels, each trial labelled left or right. The epochs are 1500 ms long and end 120 ms *before* the keystroke, thus avoiding effects of EMG activity masquerading as control signals.

Results and Discussion

To confirm the expected physiological behaviour, right movement epochs at C3 and C4 were ensemble averaged, and the results are illustrated in Figure 5.

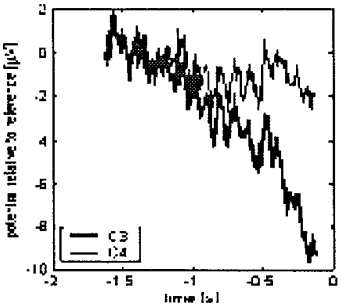

Fig. 5 Ensemble ERPs for a right movement

The Bereitschaftspotential is seen as expected, with the most pronounced ERP being elicited on the contralateral electrode position over the sensorimotor area. This also

shows that during data acquisition EMG/EOG artefacts are successfully avoided, and so artifact removal is unnecessary.
Ten Fold cross validation was used to determine all classification results, and all three cases for feature extraction detailed above were tried, the results of which follow:

Table I illustrates the LDA accuracy results for different AR model orders for case 1; the model order 4, yielding the highest optimum LDA of 57.4%,

1	2	3	4	5	6	7	8
52.1	56.9	57.3	57.4	55.5	53.2	53.5	54.1

Table I: *LDA accuracy (%) vs. AR order*

Table II illustrates the percentage accuracy of the LDA for different model order combinations for case 2. The optimum model orders are found to be na=6 and nb=3, which yield an overall classification accuracy of 70.7%.

nb \ na	1	2	3	4	5	6	7	8
1	64.0	66.7	68.1	67.7	68.6	68.1	67.7	67.0
2	68.7	67.9	69.3	67.9	68.2	69.1	68.1	67.2
3	69.7	70.1	70.1	69.9	70.2	70.7	70.2	69.3
4	68.8	69.5	69.6	69.8	69.4	69.6	69.2	67.5

Table II: *LDA accuracy (%) vs. ARX orders*

Figure 6 shows pseudocolor plots of the mu band section of the STPSD at C4, averaged over all left and right trials respectively. In the finger movement paradigm it is expected to observe at either sensorimotor area a decrease in power (ERD) in the mu band for contralateral (left) finger movement, relative to that for ipsilateral (right) movement. This is observed in the plot, but is not pronounced. A number of features were tried.

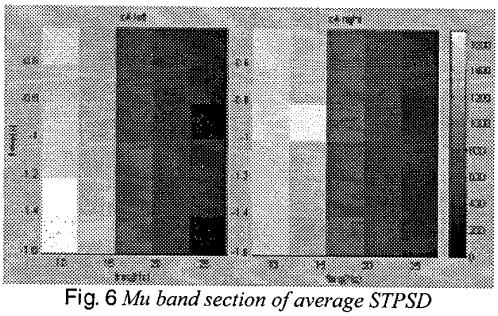

Fig. 6 *Mu band section of average STPSD*

First the values (raw samples) of the first two columns are taken – 12 features. This yields a poor accuracy of 48.7%, although taking the two blocks where the average values differ most out of the 12 yields 51.8%.
The expected ERD can be harnessed by taking the difference between the first and last block in the first and second columns – this yields 52.8%. Looking at the plots we see on average a relatively large difference between the first and fifth block in the first column for left movement, and between the fourth and sixth block in the second column for right movement. Taking these two features yields 55.9%. Taking as features 50-60 raw samples over the entire STPSD where on average they differ most gives ~62.5%.

Conclusions

The BCI methodology incorporating an ARX model in the feature extraction stage yields the best results. Modelling both the signal and noise is found to be more effective than modelling the noise alone, which is an intuitively satisfying result.
In our short time window of data the expected ERD phenomenon is not successfully harnessed as a feature for classification, and taking raw samples as features with no physiological context performs better.
The authors are currently working on applying these results to visual evoked potential and P300 experiments where the evoked potentials are much more well defined (in terms of morphology) than that of the BP. We anticipate that the ARX model should yield comparatively better results for a more complex template. Also these experiments involve prominent artifacts, thus warranting the use of the aforementioned artifact removal techniques.
Possible augmentations to the above feature extraction stage include taking the coefficients of both B(z) and A(z) as the features in the ARX case, and/or combining the spectral features with the ARX features.

Acknowledgements
The authors wish to thank Dr. Benjamin Blankhertz and Prof. Klaus-Robert Müller, Fraunhofer FIRST for the use of their data [10].

References

[1] Wolpaw, J.R., *et al* (2000): "Brain-Computer Interface Technology: A Review of the First International Meeting", *IEEE Trans. Rehab. Eng.*, 8, 2.
[2] Niedermeyer, E., Lopes Da Silva, F. (1999): "Electroencephalography: Basic Principles, Clinical Applications, and Related Fields", 4th edition, Lippincott Williams & Wilkins, Baltimore.
[3] Misulis, K. E. (1994): "Spehlmann's Evoked Potential Primer", Butterworth-Heinemann, MA.
[4] Gharieb, R.R., Cichocki, A., (2001): "Noise Reduction in Brain Evoked Potentials Based on Third Order Correlations" *IEEE Trans.Biomed. Eng., 48, 5.*
[5] Childers, D. G., Perry, N. W., Fischler, I. A., Boaz, T., Arroyo, A., A. (1987): "Event-Related Potentials: A Critical Review of Methods For Single-Trial Detection", *CRC Critical Reviews in Biomedical Engineering*, 14, 3.
[6] Pardey, J., Roberts, S., Tarassenko, L. (1996): "A review of parametric modelling techniques for EEG analysis", *Med. Eng. Phys.*, 18, pp 2-11.
[7] Cerutti, S., Chiarenza, G., Liberati, D., Mascellani, P., Pavesi, G. (1988): "A parametric method of identification of single trial event-related potentials in the brain", *IEEE Trans. Biomed. Eng.*, 35, 9, pp. 701-711.
[8] Pfurtscheller, G., Neuper, C., Schlogl, A., Lugger, K., (1998): "Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters", *IEEE Trans. Rehab. Eng.*, 6, 3, pp 316 – 324.
[9] B.D. Ripley. Pattern Recognition and Neural Networks. Cambridge University Press. 1996.
[10]Blankertz, B., Curio, G., Muller, K. R. (2002), "Classifying Single Trial EEG: Towards Brain Computer Interfacing", in Advances in Neural Information Processing Systems (NIPS 01), ed. Diettrich, G., Becker, S., Ghahramani, Z., vol. 14, MIT Press.
[11]American Electroencephalographic Society "Guidelines for Standard Electrode Position Nomenclature", *J. Clin. Neurophysiol.*, 8(2): 200-202, (1991).