EEG-based communication via dynamic neural network models

William D. Penny and Stephen J. Roberts
{w.penny, s.j.roberts}@ic.ac.uk
Department of Electrical and Electronic Engineering, Imperial College, London SW7 2BT, U.K.

Abstract— *The overall aim of this research is to develop an EEG-based computer interface. In this paper we report on an offline analysis of EEG data recorded from 7 subjects performing two different pairs of cognitive tasks; motor imagery versus a baseline task and motor imagery versus a maths task. For the imagery versus baseline pairing, discrimination was good in three subjects, marginal in two and not possible in the other two. For the imagery versus maths pairing, discrimination was very good in two subjects, good in 4 and marginal in one. The data was analysed using lagged-AR feature vectors and a Bayesian logistic regression classifier with temporal smoothing. Enhanced spectra are shown highlighting differential spectral activity for each task pairing. The results suggest that combinations of different task pairings and dynamic neural network models have the potential to drastically reduce the time it takes for a new user to learn to use an EEG-based computer interface.*

I. Introduction

Thanks to recent advances in computing and in the understanding of the electroencephalogram (EEG) many researchers now believe it is possible to build an EEG-based communication device, and a number of prototype systems have appeared in research laboratories around the world. For a review of recent approaches and discussion of key issues see Vaughan *et al.* [13] and Seabrook [11].

A typical system converts fixed features of the spontaneous EEG (8-13Hz activity [14], or slow potentials [2]) into cursor movement on a computer screen. The user controls the direction of movement (up or down) by performing one of two mental tasks. A common task pairing is motor imagery ('imagine moving your hand') versus a baseline task ('just look at the cursor'). These tasks form a starting point for a biofeedback training process where the user modifies their thought patterns in response to observed cursor movements and thereby slowly gains some degree of control over the cursor. Though successful [14], this process is rather long, taking up to several weeks before subjects can achieve sufficient accuracy to interface to, say, a spelling device [2].

Informal experimentation with an EEG-based computer interface developed at Imperial College led us to contend that the process could be considerably speeded up if (i) different initial task pairings were used (allowing better initial discrimination) and (ii) if, instead of using fixed features of the EEG, adaptive pattern recognition methods were used.

In one pilot subject, for example, it was noticed that an imagery versus maths pairing resulted in better discrimination than imagery versus baseline. Is this peculiar to our pilot subject or is this generally the case ? To answer this question we carried out formal experiments on a group of seven subjects.

One aspect of these experiments was concerned with the biofeedback process itself, where EEG signals were actually used to drive cursor movements. Analysis of this data is reported elsewhere [9], [10]. In this paper we report only on an offline analysis of sections of the data where no cursor movement took place. The purpose of this analysis is to find out if the imagery versus maths pairing is generally easier to discriminate than imagery versus baseline.

Our data analysis method uses lagged-AR features and a Bayesian logistic regression model with temporal smoothing. These components are described in the next section along with a detailed account of the experimental protocol.

II. Methods

A. Data collection

We analyse data from a pilot subject and from a group of seven subjects involved in a more formal experiment. The pilot subject was a 31 year old man and the other seven subjects consisted of four men and three women aged between 24 and 34. Each experimental session lasted 1 hour and a half.

A single channel of EEG was recorded between electrodes placed over the left and right sensorimotor cortex at positions C3' and C4' in the 10/20 system[1]. The

This work was supported in part by the UK Engineering and Physical Sciences Research Council (EPSRC), grant number GR/K79062.

[1] C3' and C4' are 3cm behind C3 and C4. For the pilot subject

signals were recorded with respect to a reference electrode placed over the right mastoid and the impedance between each electrode and the reference was less than 5KΩ. We also recorded surface EMG from the subjects' right forearm extensor in order to monitor hand movements (if the subjects were left-handed we took EMG from the left forearm). Inspection of the EMG trace allows us to check that, when the subject is imagining movement, no significant muscle activity is present.

The EEG and EMG signals were amplified using two ISODAM amplifiers. Two on-board analog filters were used to remove activity below 0.1Hz and above 100Hz. The signals were then digitised to 11-bit accuracy at a sampling rate of 384 Hz (by recording segments of 128 samples every 1/3 second). All data was stored on computer disk for later analysis.

For each subject two runs of recordings were made for each pair of cognitive tasks. In each run, the subject performed one of two task pairings: (i) motor imagery versus baseline and (ii) motor imagery versus maths. For the motor imagery task, subjects were asked to imagine opening and closing their right or left hand (depending on handedness). For the maths task subjects were instructed to serially subtract seven from a large number [12], [3]. Ten seconds of EEG were recorded during each task. Subjects repeated each task $N = 6$ times per run (making 12 'trials' per run) and there was a one minute rest period between runs.

B. Signal Processing

A pth-order 'lagged-autoregressive' (LAR) model was applied to a half-second window of data which moved along 1/5 second at a time. This produced 50 LAR vectors per trial and thus 600 vectors per run.

The LAR model had a lag of L by which we mean that the taps of the AR model were separated by L samples. We note that the LAR model is able to detect frequencies up to $1/L * (384/2)$Hz and a pth order model can pick up $p/2$ peaks in the associated power spectrum. This spectrum may be calculated directly from the AR coefficients [7].

The reason for using a lagged model rather than a non-lagged one is that lagged models are less sensitive to noise. The approach is used for embedding dynamical systems where the lag is chosen to be at the first zero crossing of the autocorrelation function [4]. The efficacy of the lagged-AR versus non-lagged AR approach is discussed in [10]. In this paper we choose L and p by cross-validation.

C. Bayesian Logistic Regression

The resulting LAR features, $\{\boldsymbol{x}_t\}$, were classified using a logistic regression model where the predicted class label, $\hat{Z}_t$, is generated according to

$$y_t = P(\hat{Z}_t = 1 \mid \boldsymbol{w}_t) = g(\boldsymbol{w}_t^T \boldsymbol{x}_t) \qquad (1)$$

where $g(a_t)$ is the logistic function

$$g(a_t) = \frac{\exp(a_t)}{1 + \exp(a_t)} \qquad (2)$$

a_t is the model 'activation' at time t and $\boldsymbol{w}_t$ is a column vector of weights.

The models were trained using the Bayesian evidence framework [6]. This results in both an estimate for the weights, $\hat{\boldsymbol{w}}_t$, and an estimate for the distribution of the weights. This distribution is approximated by a Gaussian with mean $\hat{\boldsymbol{w}}_t$ and covariance $\boldsymbol{\Sigma}_t$. The distribution on the weights implies a distribution on the activation, the mean and variance of which are given by

$$\bar{a}_t = \hat{\boldsymbol{w}}_t^T \boldsymbol{x}_t \qquad (3)$$

and

$$s_t^2 = \boldsymbol{x}_t^T \boldsymbol{\Sigma}_t \boldsymbol{x}_t \qquad (4)$$

If the uncertainty in the weights is taken into account when making a prediction (as it should be) then the correct predictive output to use is the 'moderated' output

$$\tilde{y}_t = P(\hat{Z}_t = 1) = \int P(\hat{Z}_t = 1 \mid a_t) p(a_t) da_t \qquad (5)$$

This integral can be accurately approximated by [6]

$$\tilde{y}_t = g(K(s_t)\bar{a}_t) \qquad (6)$$

where

$$K(s_t) = \left(1 + \frac{\pi s_t^2}{8}\right)^{-1/2} \qquad (7)$$

The moderation changes the actual output, y_t, to a moderated output, $\tilde{y}_t$, which is nearer to 0.5 by an amount which is dependent on the uncertainty on the weights. Moderated outputs are typically better than unmoderated outputs in terms of the likelihood of predictions [6].

D. Temporal smoothing

For EEG data, averaging classifier outputs over a number of consecutive data segments is known to significantly increase classification accuracy [1]. In this paper, averaging is performed, not in the output space, but in the space of activations, or the 'latent' space. This scheme arises from considerations of how to make optimal decisions in a 'committee' of classifiers [8].

To this end, we construct a committee of logistic regression classifiers formed by combining model predictions from T previous time steps. The distribution of

positions C3 and C4 were used.

activations is now viewed as a Gaussian *mixture* where the distribution of classifier activations at time t (as described in the previous section) corresponds to a single component in that mixture. The mean and variance of the Gaussian mixture are

$$a_{com}(t) = \frac{1}{T} \sum_{t}^{t-T+1} \bar{a}_t \tag{8}$$

and

$$s_{com}^2(t) = \frac{1}{T} \sum_{t}^{t-T+1} s_t^2 + \frac{1}{T} \sum_{t}^{t-T+1} (\bar{a}_t - a_{com}(t))^2 \tag{9}$$

where the second term arises because the variance of a Gaussian mixture is the mean of the component variances plus the variance of the component means. The moderate committee output is then given by [8]

$$\tilde{y}_{com}(t) = g(K(s_{com}[t])a_{com}[t]) \tag{10}$$

This is the final output which is used to make predictions and classifications. The scheme requires a stored buffer of the previous T activations and variances and classifications are not considered for $t < T$, ie. before the buffer is full. The efficacy of the above approach is discussed in [10].

A smoothing value of $T = 10$ was used in all experiments. This corresponds to averaging over two seconds (as the window is moved on 1/5 second between LAR vectors).

E. Cross-validation

We analyse each data set on a run-by-run basis, where $N = 6$ pairs of tasks are performed in each run. To assess the accuracy of classification we use a cross-validation (CV) process where classifiers are trained on $N-1$ task pairs and tested on the remaining task pair. This is repeated in the usual CV manner and average results are computed.

Discrimination accuracy is assessed on a segment-by-segment basis by reporting correct classification rates (CCRs).

Optimal values of L and p are chosen on the basis of minimum CCR on the CV test sets. So as not to bias the results, only one of the two CV runs was used for selecting L and p. Overall classification results are reported on both CV runs. Also, we restrict the permissible values of L and p to $L = 4, 8$ and $p = 4, 8, 12$.

F. Enhanced spectra

Once we have trained a classifier to discriminate between two different cognitive tasks on the basis of LAR features it is interesting to then go back and look at what are the typical LAR features and corresponding spectra associated with each task. This can be achieved by taking class averages of LAR feature vectors after discarding those that were incorrectly classified (we do this on the CV test sets). This is equivalent to the enhanced averaging method described by Gevins and Morgan [5].

TABLE I

Classification rates: percentage correct ± one standard deviation

Subject	Imagery vs. Baseline	Imagery vs. Maths
D	56 ± 16	60 ± 11
I	75 ± 13	81 ± 6
K	79 ± 15	76 ± 7
N	63 ± 12	67 ± 14
ST	71 ± 15	68 ± 10
SU	61 ± 12	66 ± 14
Z	53 ± 21	84 ± 9

III. Results

The parameters L and p were tuned to each subject and task pair using the cross-validation process described above and applied to a single run. In table 1 we report the cross-validation accuracies as measured over both runs. The tuning process, however, is unlikely to have biased the results as the accuracy on the 'tuned' run was not consistently higher than that of the other run. This is understandable as the data sets are quite large (600 data points) and the classifiers have very few parameters (4, 8 or 12).

For the pilot subject recordings were only made for the imagery versus maths task pair. Analysis of the data showed that the tasks could be correctly discriminated with an accuracy of 80 ± 12 per cent. The corresponding enhanced spectra are shown in Figure 2(P).

Classification rates are reported for the seven subjects involved in the formal experiment in Table 1. For the motor imagery versus baseline pairing, discrimination is good for three subjects, marginal for two and not possible for the other two. For the motor imagery versus maths pairing, discrimination is very good for two subjects, good for four subjects and marginal for one.

The corresponding enhanced spectra are shown in Figures 1 and 2. For motor imagery versus baseline the differential activity is in the mu-band (8-13Hz) and for subject I there is also differential activity in the beta band (14-20Hz). For motor imagery versus maths differential activity is in the mu-band with subjects I and P also showing differences in the theta-band (4-7Hz).

IV. Discussion

Our experiments show that we can discriminate between a motor imagery task, a maths task and a baseline task and that discrimination is better for the imagery versus maths task pairing than for the imagery versus baseline pairing. Importantly, this is achievable with a single channel of EEG (only three electrodes) and requires no significant subject training. The results were achieved by using lagged-AR features and a Bayesian logistic classifier with temporal smoothing. The results suggest that combinations of different task pairings and dynamic neural network models have the potential to drastically reduce the time it takes for a new user to learn to use an EEG-based computer interface.

References

[1] C. Anderson and Z. Sijercic. Classification of EEG Signals from four subjects during five mental tasks. In *Proceedings of the conference on engineering applications in neural networks (EANN'96)*, pages 407–414. Systems Engineering Association, 1996.

[2] N. Birbaumer et. al. A spelling device for the paralysed. *Nature*, 398:297–298, 1999.

[3] T. Fernandez et. al. Eeg activitation patterns during the performance of tasks involving different components of mental calculation. *Electroencephalography and Clinical Neurophysiology*, 94:175–182, 1995.

[4] N.A. Gershenfeld. *Directions in Chaos*, volume 2, chapter An Experimentalist's Introduction to the Observation of Dynamical Systems, pages 310–384. World Scientific, Singapore, 1989.

[5] A.S. Gevins and N.H. Morgan. Classifier-directed signal processing in brain research. *IEEE Transactions on Biomedical Engineering*, 33(12):1054–1068, 1986.

[6] D.J.C. Mackay. The evidence framework applied to classification networks. *Neural Computation*, 4(5):720–736, 1992.

[7] J. Pardey, S. Roberts, and L. Tarassenko. A Review of Parametric Modelling Techniques for EEG Analysis. *Med. Eng. Phys.*, 18(1):2–11, 1996.

[8] W.D. Penny, D. Husmeier, and S.J. Roberts. *Artifical Neural Networks in Biomedicine*, chapter The Bayesian Paradigm: second generation neural computing. Springer, 1999.

[9] W.D. Penny and S.J. Roberts. Experiments with an eeg-based computer interface. Technical report, BCI Workshop, Albany, USA., June 1999.

[10] S.J. Roberts and W.D. Penny. Real-time Brain-Computer Interfacing. Technical report, Department of Electrical Engineering, Imperial College, 1999. Also submitted to Engineering in Medicine and Biology.

[11] R.H.C. Seabrook. The Brain-Computer Interface: Techniques for Controlling machines. Technical report, University of Maryland, 1994. Available from ftp.clark.net in /pub/seabrook.

[12] C.J. Stam, T.C.A.M. van Woerkom, and W.S. Pritchard. Use of non-linear eeg measures to characterize eeg changes during mental activity. *Electroencephalography and Clinical Neurophysiology*, 99:214–224, 1996.

[13] J.R. Wolpaw T.M. Vaughan and E. Donchin. EEG-Based Communication: Prospects and Problems. *IEEE Transactions on Rehabilitation Engineering*, 4(4), 1996.

[14] J.R. Wolpaw, D.J. McFarland, D.J. Neat, and C.A. Forneris. An EEG-based Brain-Computer Interface for Cursor Control. *Electroencephalography and Clinical Neurophysiology*, 78:252–259, 1991.

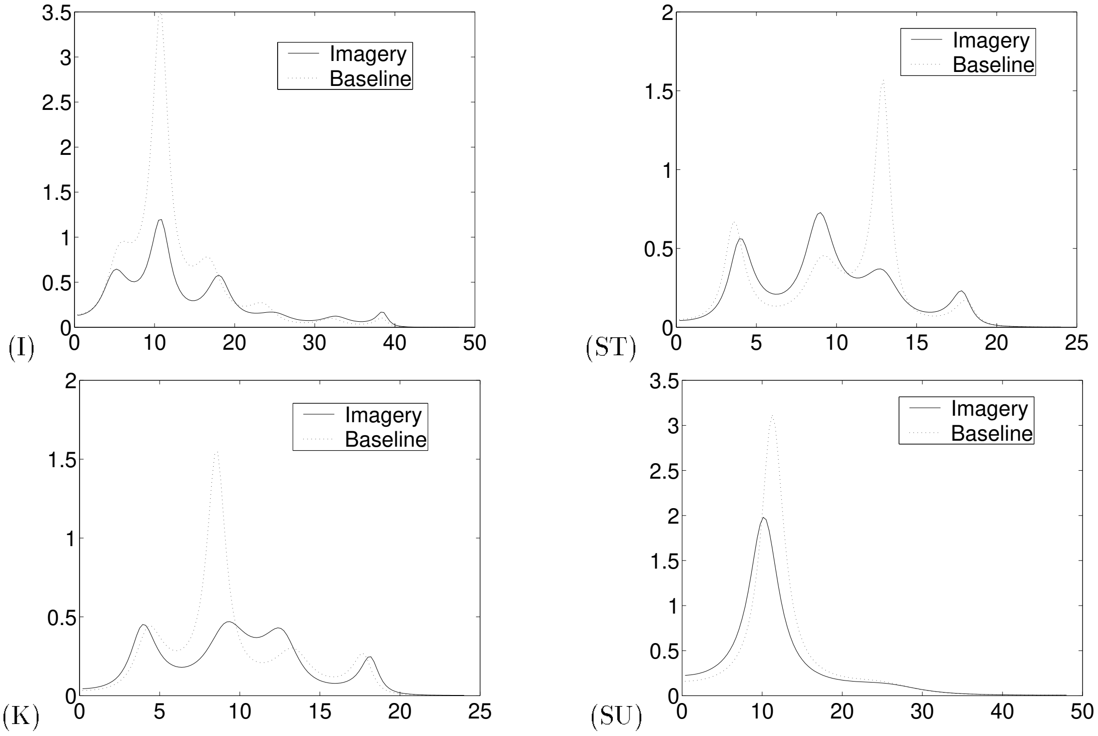

Fig. 1. Enhanced spectra for motor imagery versus baseline tasks for four of the subjects. Each subject shows differential activity in the mu-band (8-13 Hz) and subject I also shows differences in the beta band (14-20 Hz).

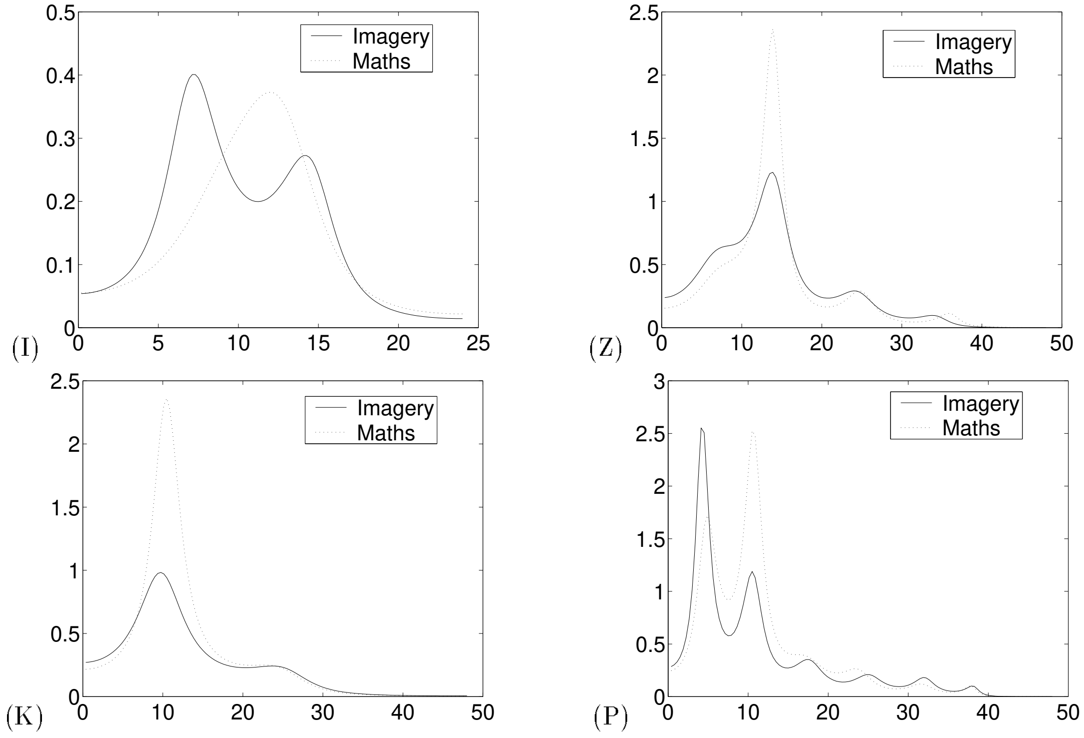

Fig. 2. Enhanced spectra for motor imagery versus maths tasks for three of the subjects plus the pilot subject (P). Each shows differential activity in the mu-band (8-13 Hz) with subjects I and P also showing differences in the theta-band (4-7 Hz).