Experiments with an EEG-based computer interface

William D. Penny and Stephen J. Roberts

Department of Electrical Engineering
Imperial College,
London SW7 2BT, UK

July 27, 1999

Abstract

We report on experiments with an EEG-based computer interface developed at Imperial College. The experiments are inspired by a programme of research in Albany by Wolpaw *et al.* We first briefly review their programme by describing what has been achieved and what remains to be done. This then informs the design of experiments at Imperial which are intended to improve EEG-based computer interfaces by reducing the time taken by a user to achieve cursor control. A combination of approaches are taken to achieve this end; (i) using different cognitive tasks, (ii) using different EEG features and (iii) using more sophisticated pattern recognition algorithms.

1 Introduction

Thanks to recent advances in computing and cognitive electrophysiology, 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.* [32] and Seabrook [29]. Moreover, such interfaces are beginning to be used by people with severe physical disabilities [3, 11].

In short, current approaches fall into two categories - (i) those using evoked responses (ER), ie. EEG signals elicited by specific evoking stimuli and (ii) those using spontaneous EEG components, that is, spontaneous in the sense that they occur in the course of normal brain function and are not strongly linked to evoking stimuli. The ER approaches have the advantage that the signals can be readily detected by simple processing methods but have the disadvantage of requiring the relevant sensory modality to be wholly or in part devoted to EEG-based communication eg. in some ER systems attention is paid (or not) to a visual stimulus occupying a major part of the visual field [16, 31]. The 'spontaneous' approaches, however, do not demand the constant committment of a sensory modality and therefore may prove more convenient to use, more compatible with simultaneous performance of other activities and more practical for individuals with impaired modalities.

In this report we focus on the latter approach, in particular the methods pioneered by Wolpaw *et al.* based on mu-rhythm training [35, 15, 34].

1.1 Mu-rhythm training

Wolpaw *et al.* have shown that people can learn to use spontaneous EEG activity to control the movements of a cursor on a computer screen. Cursor movement was chosen because it is objective, easily implemented and readily quantified and can serve as a prototype for control of a wide variety of prosthetic devices. Their system is based on the mu-rhythm which is known to 'desynchronise' (reduce in power) during movement planning, imagination or execution. Unlike occipital alpha, the mu rhythm can be modulated over a period of seconds (rather than minutes [7]).

In their 1991 experiment [35] subjects were trained to perform one-dimensional movement (ie. up or down). Five subjects were chosen from a pool of 60 on the basis of having a strong mu-rhythm. The subjects' task was to move a cursor in order to hit targets which appeared at the top or bottom of the screen - ie. ballistic cursor control. After several weeks of training, four out of the five subjects achieved significant control (typically 90% of targets were hit at a rate of about 20 per minute).

The system used two electrodes placed just in front of and behind C3 and upward cursor movement was proportional to the signal power in the 8-12Hz band as measured by a Fourier transform method.

In later work, subjects were asked to perform a more difficult task - to move the cursor to a specific location and to keep it there in order to intercept a target - ie. graded cursor control [15].

In a 1994 experiment subjects attempted 2D cursor control (ie. up, down, left and right) [34]. After 6 to 8 weeks of training, 4 out of 5 subjects achieved significant control (typically 65% of targets were hit at a rate of 20 per minute).

1.2 The training process

The first notable feature of the training process is that it took several weeks. In the Wolpaw 1991 experiment, for example, each subject undertook 15-20 half hour training sessions over a period of at least 3 weeks. Each session consisted of 6 or 7 runs where each run lasted about 3 minutes and approximately 100 trials were made in each session.

Another feature of the training process is that the tasks were graded so that the feedback was always positive eg. in the first two or three sessions of the Wolpaw 1991 experiment, the cursor always moved up and the subject just had to increase the speed of reaching the target. Only in later sessions did the cursor move in both directions.

A third important feature of the training process is that, although subjects reported using motor imagery in the earlier sessions, towards the end of training such imagery was often found to be no longer necessary. Furthermore, in the later sessions, cursor control was found to be compatible with the simultaneous performace of other activities such as speaking or mental calculation [35]. Thus it appears that, like conventional skills, EEG-based cursor control, once learned, no longer requires intense concentration.

1.3 Open problems

Despite the considerable achievements of the Wolpaw research program a number of open problems and areas for improvement remain:

1. **Accuracy**. Currently, the accuracy of 1D ballistic cursor control is about 90% and of 2D control is about 65%. Can this be improved upon ?
2. **Speed**. Currently, both 1D and 2D systems operate at about 10 to 15 bits per minute. This corresponds to approximately 2 to 3 characters per minute. Can this be improved upon ?
3. **Training time**. To achieve the above bit-rates takes between 3 and 8 weeks of biofeedback training. Can this process be speeded up ?
4. **Subject variability**. EEG-based interfaces seem to work well for some people and not at all for others. One out of the five of Wolpaw's subjects, for example, failed to achieve cursor control whilst another

did so with 95% accuracy. Furthermore, these 5 subjects were selected from a pool of 60 ! Is it possible to develop an interface that can be used by anyone ?

5. **Artifact elimination**. Accuracy, speed, subject variability and learning time are all affected by muscle artifacts appearing in the EEG trace. So this is really a sub-problem. Can automatic artifact removal methods be improved upon ?
6. **Graded control**. A mouse-driven cursor offers the user graded control - ie. not just control of the direction of movement (ballistic control) but also the amount. Research in this area is at an early stage.
7. **On-Off switch**. Because the EEG components occur in the course of normal brain function some extra signal is needed to flag when EEG communication is intended and when it is not. For example, though a subject may be physically connected to a BCI, he may wish to be involved in other activities eg. resting, and not want to be communicating via the EEG all the time. Therefore, the interface needs an 'on/off switch' (see eg. Birch and Mason [4]).
8. **Application as an assistive technology**. Work in this area is at an early stage. Miner *et al.* [11] report on an individual with Amyotrophic Lateral Sclerosis (ALS) who achieved cursor control via mu-rhythm training and Birbaumer *et al.* [3] reported on two individuals with ALS who were trained to use a spelling device via operant conditioning of slow cortical potentials. Research with other patient groups has been reported more recently by the Graz group [17].

Of all these problems, the one we address in this report is (3) training time. We attempt to reduce the training time by (i) trying an additional cognitive task (ie. not just imagined hand movement) (ii) using a variety of EEG features (ie. not just the mu rhythm) and (iii) using more sophisticated signal processing and pattern recognition methods.

These changes shift the onus of communication somewhat from the user to the computer. We do, however, recognise that some adaptation by the user may be necessary and for this reason also investigate the effect of biofeedback training, albeit over a much shorter period.

We now briefly review other work which is relevant to each of the three possible methods of improvement.

1.4 Other EEG features

A number of features of the EEG other than the mu-rhythm have also been found to be correlated with movement planning and imagery. Pfurtscheller *et al.*, for example, have shown that motor planning signals can be picked up in other frequency bands [26] and by autoregressive models [19] (they have also developed EEG-based interfaces based on these and other features [25]). More recently, McFarland *et al.* have shown that the beta rhythm is also correlated with imagined movement [14]. Because these signals correspond to a reduction in power in particular frequency bands they are called Event-Related Desynchronizations (ERDs).

Real and imagined hand movements are also accompanied by DC changes in the EEG (the 'readiness' or 'bereitschaft' potential - RP or BP) recorded at C3, C4 and Cz (and referenced to linked mastoids) [6, 3]. For movement planning and imagery these changes are bilateral, but for actual movement they are lateralized (this is opposite to mu rhythm changes which are contralateral for planning and imagery but bilateral for actual movement). In a study of 20 subjects, Toro *et al.* [33] found that, across the population, RPs were not correlated with ERDs. RPs therefore provide an independent source of information which may be very useful for the subsequent classification stage.

Following Pfurtscheller *et al.* we also intend to use AR features. The features will be derived from a single bipolar channel of EEG derived from electrodes placed at C3' and C4' (ie. 3cm posterior to C3 and C4). Our earlier study on detecting imagined movements (which used an array of electrodes over sensorimotor cortex) suggests that these are the optimal positions [23]. These positions were also found to be optimal in a study of movement planning [24].

We use only a single channel as, ultimately, any EEG-based interface must be easy to wear and putting on lots of electrodes is both cumbersome and time-consuming. Further, by looking at bipolar C3'-C4' we can immediately detect differences in hemispheric activity eg. the contralateral mu-rhythm associated with movement planning [26] and imagery [23] and, hopefully, detection of cognitive components associated with arithmetic tasks [9][1] (see next section). A further advantage of using C3'-C4' is resistance to eye movement artifact (this resistance is due to the fact that eye movement artifact reduces in proportion to the distance posterior to the plane of the eyes and C3' and C4', being equidistant from this plane, will pick up a common amount of contamination which will not therefore appear in the bipolar signal C3'-C4'). A disadvantage of using bipolar C3'-C4' is that the relevant DC changes will not be picked up (because they are bilateral).

1.5 Other cognitive tasks

Other work in spontaneuos EEG analysis by Anderson *et al.* has looked at discrimination of different cognitive tasks [9, 2]. Using EEGs recorded from six channels the following tasks could be discriminated with an accuracy of over 80 per cent; baseline, mental arithmetic, geometric figure rotation, mental letter composing and visual counting. Importantly, as little as 0.25s of data is sufficient to discriminate between the tasks, making rapid communication a possibility.

However, as these tasks are well defined (more so than motor imagery) it may not be the case that, after a period of feedback training à la Wolpaw, cursor control could become unconscious (see section 2.2). But this is an open question (as these cognitive tasks may just define a starting point).

In this paper we consider a single additional cognitive task: mental arithmetic. In particular, we use the 'serial subtraction of sevens from a large number' task as it is known to be detectable on the basis of hemispheric differences in the EEG [30].

1.6 Other analysis methods

Like Pfurtscheller *et al.* and Anderson *et al.* we also intend to use neural networks to classify the AR features. There are, however, two novel aspects to our analysis.

Firstly, we intend to use a dynamic models (see section 2.3.3), which constrains the cursor movement at a given time step to be dependent on the cursor movements at previous time steps. This, according to our preliminary investigations, results in smoother cursor control.

Secondly, we infer not just the parameters of our models (eg. weights in a neural net) but also the uncertainty on the parameters. This allows us to estimate the confidence associated with each subsequent classification. If the cursor is then allowed to move only for high confidence classifications then the system should, in principle, perform automatic artifact rejection ie. the quality of cursor control should not be seriously degraded by eye or muscle movement. The system should also reject signals containing irrelevant cognitive components (eg. during lapses of concentration).

[1] This work used electrodes at C3, C4, P3, P4, O1 and O2 - as we are using a subset of their positions it may still be possible to recognise cognitive components associated with mental arithmetic - this possibility is supported by our preliminary study [22].

1.7 Aims of present study

1. To find out how accurately different cognitive tasks (imagined hand movement, mental arithmetic and a baseline task) can be discriminated by analysis of single-channel EEG.
2. To see if biofeedback, in the form of cursor movements, can improve the discrimination.

2 Methods and Materials

2.1 Data collection

EEG data were collected from 4 male and 3 female healthy volunteers aged between 24 and 34. Each subject participated in a recording session lasting between 1 hour and 1 hour 30 minutes. The EEG was recorded from electrodes positioned over the left and right sensorimotor cortex in positions C3' and C4' (3cm posterior to the C3 and C4 positions in the 10/20 system).

The 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. The surface electromyogram (EMG) was recorded from the left or right forearm extensor (depending on the handedness of the subject). The EEG and EMG signals were amplified using two ISODAM amplifiers. Two analog filters on board the ISODAMs 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 and, during the biofeedback tasks, was also analysed on-line in order to drive the cursor movements.

2.2 Experimental tasks

During each session, 9 runs of recordings were made. There was a one minute rest period between runs.

At the beginning of each session recordings were made whilst the subject performed each of the following tasks; 30 seconds of continuous movement of eyes from left edge to right edge of computer screen, 30 seconds of continuous movement of eyes from top to bottom of computer screen, 30 seconds of frowning and 15 seconds of eye blinks. The rest of the session consisted of 9 runs of 3 minute recordings. The first run was a 'warm-up run' so that subjects could familiarise themselves with the experimental set-up. In the next 8 runs subjects performed one of two task pairings; (1) imagined movement (IM) versus baseline (B) or (2) IM versus mental arithmetic (MA), and one of two biofeedback conditions; (1) none (stationary cursor) or (2) moving cursor.

The ordering of runs was chosen so that the no-feedback trials for a given task pair preceded the feedback trials for that task pair - as it was thought necessary to give the user some practice at performing the tasks before they are given the added task of controlling cursor movements. The ordering of the task pairs was randomized for each subject to eliminate any possible bias due to fatigue.

In each run the subject was asked to perform 6 of each of the tasks and the tasks were interleaved eg. IM followed by B followed by IM etc. Subjects were prompted to perform the tasks by following instructions on the computer screen. There was a 2 second delay between tasks and a 3 second delay after each task instruction, before recordings began. The instructions for each of the tasks were as follows:

1. IM:Please look at the cursor and imagine opening and closing your hand.
2. MA:Please look at the cursor and serially subtract 7 from X.

3. B:Please just look at the cursor.

For the IM task, the subject was previously told that the hand they should imagine moving was to be their right or left hand (depending on handedness). For the MA task X was a randomly chosen 3 digit number whose last two digits were not divisible by 7.

For all of the tasks a window appears on the screen containing instructions. When the instructions disappear a diamond shaped 'cursor' appears. For the no-feedback trials this cursor remained stationary. For the feedback trials the subjects were told 'to look at the cursor and to perform the tasks in such a way as to make the cursor hit the targets as quickly as possible'. The subjects were presented with a target, a square, which was located at the top or bottom of the screen and the cursor moved towards the target in proportion to how well the subject was performing the associated task; upward movements were always associated with IM and downward movements with B or MA. The cursor movements were driven by the online analysis method described below.

2.3 Online signal analysis

2.3.1 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 10 second 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[2]. We note that the LAR model is able to detect frequencies up to $f_{max} = 1/L * (384/2)$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 [18].

The reason for introducing a lag into the AR model is to change the effective sampling frequency so that the model fits to the frequency range of interest $(0..f_{max})$ rather than, due to the high sampling rate, to a much broader frequency range (0..384/2). The benefit of this approach is demonstrated for a synthetic signal in Figure 1.

2.3.2 Bayesian Logistic Classifier

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

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

where $g(a_t)$ is the logistic function

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

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 [13]. 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

$$\bar{a}_t = \hat{\boldsymbol{w}}_t^T \boldsymbol{x}_t \tag{3}$$

and

$$s_t^2 = \boldsymbol{x}_t^T \boldsymbol{\Sigma}_t \boldsymbol{x}_t \tag{4}$$

[2] The LAR model was implemented by first downsampling the signal by a factor of L. The downsampling included a low-pass filtering so that high frequency components would not fold down into low frequency components. A standard AR model was then applied to the resulting signal.

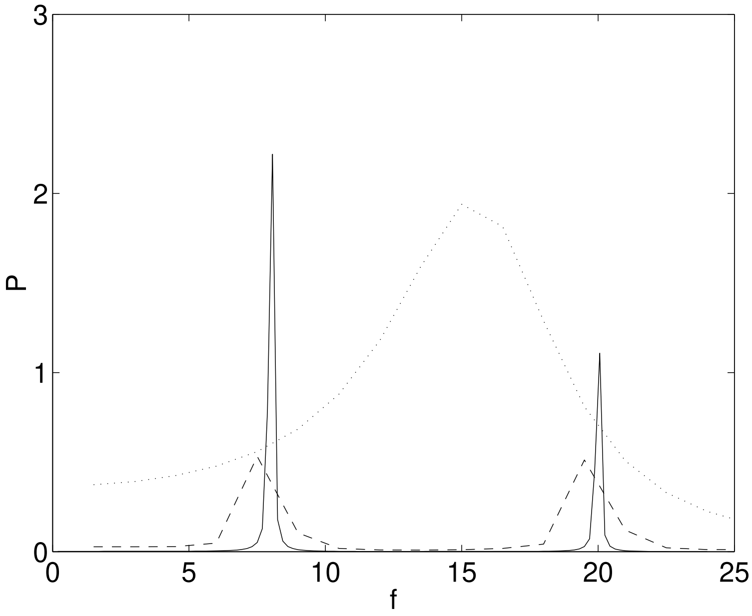


Figure 1: The figure shows spectral estimates of data consisting of an 8Hz sine wave, a 20Hz sine wave and additive Gaussian noise sampled at 384Hz for an AR (8) model (dotted line), an AR(48) model (dashed line) and a lagged-AR(8) model with a lag of 6 (solid line).

If the uncertainty in the weights is taken into account when making a prediction (ie. cursor movement) then the correct predictive output to use is the 'moderated' output

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

This integral can be accurately approximated by [13]

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

where

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

The moderation changes the actual output, y_t, to a moderated output, $\tilde{y}_t$, which is nearer to 0.5 (the prior probability for a two-class problem) 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 [13].

2.3.3 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 [20].

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

$$s^2_{com}[t] = \frac{1}{T} \sum_{t}^{t-T+1} s^2_t + \frac{1}{T} \sum_{t}^{t-T+1} (\bar{a}_t - a_{com}[t])^2 \tag{9}$$

where the second equation arises because the variance of a Gaussian mixture is the mean of the component variances plus the variance of the component means. The moderated committee output is then given by [20]

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

This is the final output which is used to make 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.

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 AR samples).

2.3.4 Generation of cursor movements

The window containing the cursor had 79 cursor positions in the vertical direction. In each feedback trial the cursor starts at offset 0 (in the middle) and has 38 positions to go to reach either of the targets.

Given classifier output, y_t, a number between 0 and 1, the cursor position is increased d_t pixels up the screen where d_t was given by

$$d_t = k(y_t - \pi) \tag{11}$$

and k is a 'speed' constant which was set so that the minimum time to reach the target was 5 seconds (given that the cursor can move once every third of a second this works out as $k = 3.8$). The cursor position

$$c_t = \sum_{t=1}^{T} d_t \tag{12}$$

is stored as a real number and plotted discretely to the nearest pixel position. The variable π introduces a bias into the cursor movements such that *a priori* the cursor is twice as likely to go in the target direction as the other direction (this is achieved by setting $\pi = 1/3$ for up trials and $\pi = 2/3$ for down trials)[3]. This ensures that the feedback given to the subject is, psychologically, positive. A maximum trial time of 10 seconds was also imposed.

3 Results

We report results from online experiments by analysing the EEG data on a segment-by-segment basis (we could measure the proportion of targets hit but this would tie the results rather strongly to the details of the particular communication protocol which is yet to be optimised). Figure 2 shows classification accuracies on seven subjects for (a) stationary cursor recordings and (b) moving cursor recordings for the two pairs of cognitive tasks.

[3] Note that this bias is used only for the purpose of providing feedback to the subject. For the purpose of evaluating each subject's performance, the bias is removed.

3.1 Effect of cognitive task pairing and feedback

The effect of cognitive task pairing and the use of biofeedback on the accuracy of cursor movement can be quantified from an analysis of variance (ANOVA).

From the data set as a whole one can derive a regression model in which the accuracy of cursor movement, A, is modelled as a linear function of task pairing, T, feedback condition, F and subject identity, S_i, as follows

$$A = w_0 + w_t T + w_f F + \sum_{i=1}^{7} w_i S_i \tag{13}$$

where $T = 1$ for imagery versus maths, $T = 0$ for imagery versus baseline, $F = 1$ for feedback and $F = 0$ for no-feedback, and the variable S_i is unity if the data point is from subject i and zero otherwise. Application of this 'full' model to our experimental data yields the following coefficients: $w_0 = 54.0$, $w_t = 5.9$ and $w_f = -3.4$. That is, on average, performing a maths task leads to a 5.9% increase in accuracy and receiving feedback leads to a 3.4% decrease in accuracy. But are these differences significant ?

By removing the task variable and re-computing the regression equation for this 'reduced' model and then comparing the predictive power of the reduced and full models using an F-test [10], one can conclude that the effect of cognitive task pairing is significant; imagery versus maths gives significantly ($p < 0.003$) more accurate cursor movement than imagery versus baseline.

A similar analysis for the feedback variable shows that, overall, the effect of feedback is not significant ($p = 0.09$). If we just look at those experiments involving imagery versus baseline tasks, we also see that the effect of feedback is not significant ($p = 0.94$). If we then look at just the imagery versus maths tasks, we see that the effect of feedback is significant ($p < 0.012$) and the effect is to reduce the accuracy of cursor movement.

If we then look at just the stationary cursor trials, we see that, again, the imagery versus maths tasks result in significantly more accurate cursor movement ($p < 0.0005$). For the moving cursor trials, however, there is no significant difference ($p = 0.33$).

3.2 Accuracy and Communication Rate

Overall, in the moving cursor trials, four out of seven subjects achieved at least 75% accuracy. The accuracy, averaged over all subjects was, however, only 61.3%. This corresponds to a communication rate measured in bits per second (or bit-rate) given by

$$C = r(1 - H[p]) \tag{14}$$

where p is the probability of a correct classification, $H[p]$ is the entropy of the classification, and r is the number of classifications per second [5]. As we are classifying 2 seconds of EEG data at a time we have, $r = 0.5$. For $p = 0.613$, we have a bit-rate of $C = 0.02$.

In a companion paper [28] we show that, by including a reject option in which no cursor movements are made for low confidence classifications, the overal classification rates can be greatly increased. The cost of this increased accuracy is slower cursor movement, but this nevertheless represents a significant improvement. For a 'hard-rejection' option [28], for example, by rejecting 43% of the data (giving $r = 0.285$), 86% average classification rates are obtainable. This corresponds to a bit-rate of $C = 0.12$. This represents a 6-fold increase in bit-rate over the classification scheme which doesn't use a reject option.

This bit-rate corresponds to 1 bit communicated approximately every 8 seconds. Allowing 3 to 4 bits for represeting each letter of the English language [5], gives a potential communication rate of 2 to 3 letters per minute. Remember, this is for the average subject. For the best subjects, this increases to 5 to 10 letters per minute.

A further benefit of the 'rejection' approach is that sections of EEG containing irrelevant cognitive components (eg. during lapses of concentration) do not result in spurious cursor movement. This greatly enhances the robustness of the system.

3.3 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 is equivalent to the enhanced averaging method described by Gevins and Morgan [8]. Figure 3 shows enhanced spectra for subject 5 performing the imagery versus maths pairing. This is quite representative of all the subjects with the majority of differential activity in the mu-band (8-13 Hz). Some subjects also showed differences in the theta (4-7 Hz) and beta (14-20 Hz) bands. Enhanced spectra are shown for many of the subjects in a companion paper [22].

4 Discussion

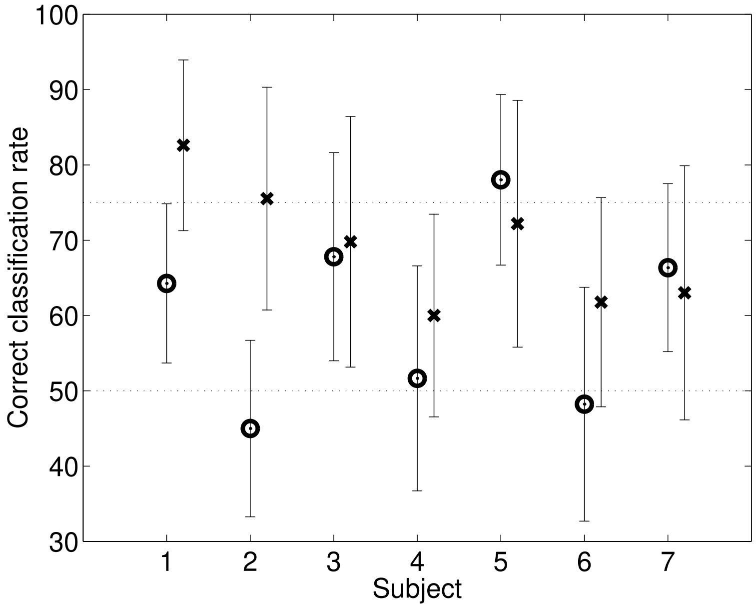
We note that four out of the seven subjects achieved at least 75% accuracy in the moving cursor trials. They would would therefore be able to immediately use a wordprocessor package via our EEG-based interface using, for example, the protocol developed by Birbaumer *et al.* [3]. Two of these subjects would use an imagery versus baseline strategy and the other two imagery versus maths. With the use of a reject option at the pattern classification stage, the average subject could communicate at a rate of 2 to 3 letters per minute (note that this result is obtained from averaging over all seven subjects, not just the four best ones).

We can also see the technology improving in other ways, by following one of two distinct research paths.

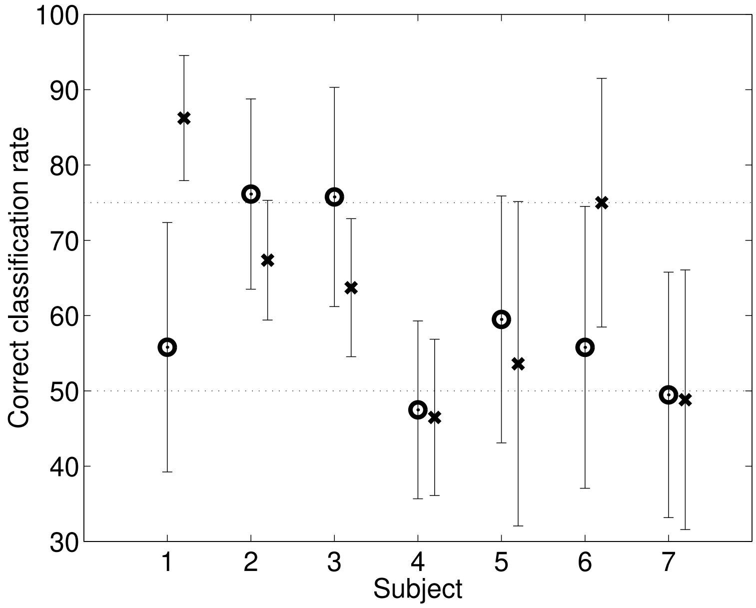
Firstly, we could persevere with methods involving little or no biofeedback training and focus on other ways of improving classification accuracy. One method might be to also use 'readiness' potentials (RPs) [6] or 'slow cortical potentials' [3]. Recent research shows that RPs and spectral changes (mu-rhythm desynchronization) are uncorrelated [33]. The RPs would therefore provide an additional source of information. A second method is to look at using additional features such as complexity [27] or to pre-process the signals in more adventurous ways eg. using an Independent Component Analysis embedding [12]. A third method is to use HMMs. HMMs have not, as yet, been applied to the whole database but qualitative results on one subject suggest the approach is promising. The main strength of HMMs is in analysing *sequences* of data (via Viterbi decoding). We therefore envisage that HMMs can best be utilised by retrospectively analysing sequences of intended cursor movements in order to better decode the transmitted message. Use can also be made of error-correcting codes.

Secondly, we could train subjects using a biofeedback approach. This involves the interaction of two adaptive controllers; the user and the computer. One promising approach for handling this is to use nonstationary classification algorithms [21] which acknowledge that the statistics of each class (ie. how the user moves the cursor up or down) can change with time.

We plan to incorporate all of the more recent technical developments into a new version of the recognition software and to test the resulting system on a large number of normal subjects. The interface will then be tested on patients having severe neurological disabilities but known to be cognitively aware.



(a)



(b)

Figure 2: The figure shows classification accuracies on seven subjects for the imagery versus baseline pairing (marked as circles) and the imagery versus maths pairing (marked as crosses). The error bars are plotted at one standard deviation. Each result is from approximately 1200 data segments derived from 12 repetitions of each task where each task lasts up to ten seconds. Plot (a) is for **stationary cursor** recordings: for the imagery versus baseline pairing, discrimination is good in one subject, significant in three and not possible in the other three. For the imagery versus maths pairing, discrimination is good in four subjects and significant in three. Plot (b) is for **moving cursor** recordings: for the imagery versus baseline pairing, discrimination is good in two subjects and not possible in the other five. For the imagery versus maths pairing, discrimination is good in two subjects, significant in two subjects and not possible in the other three.

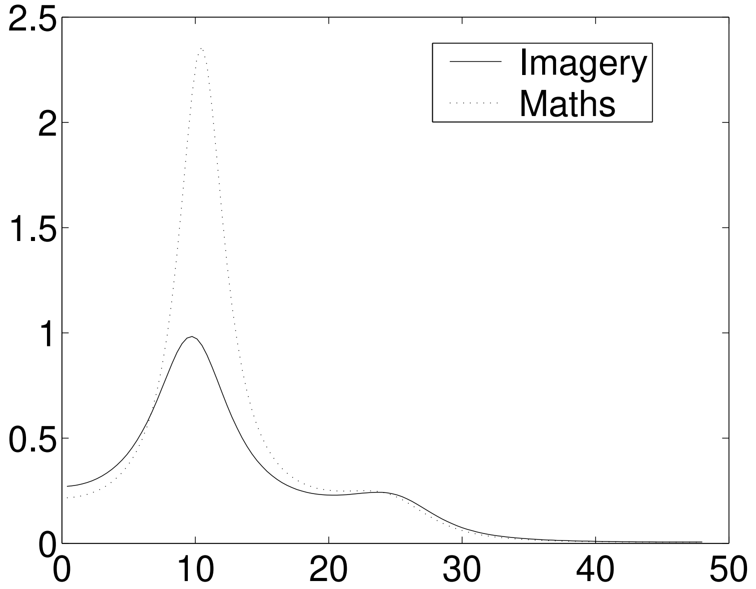


Figure 3: Enhanced spectra for subject five performing motor imagery and maths tasks whilst the cursor was stationary.

Acknowledgements

The authors would like to thank the subjects who took part in the experiment. W. D. Penny was supported by the UK Engineering and Physical Sciences Research Council (EPSRC), grant number GR/K79062.

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] C. Anderson, E. Stolz, and S. Shamsunder. Discriminating Mental Tasks Using EEG Represented by AR Models. In *Proceedings of the 1995 IEEE Engineering in Medicine and Biology Annual Conference, Sept 20–23, 1995, Montreal, Canada*, 1995.

[3] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kubler, J. Perelmouter, E. Taub, and H. Flor. A spelling device for the paralysed. *Nature*, 398:297–298, 1999.

[4] G.E. Birch and S.G. Mason. Brain Computer Interface Research at the Neil Squire Foundation. In *Brain-Computer Interface Technology: Theory and Practice, First International Meeting*, pages 50–53, Rensselaerville Institute, New York, June 1999.

[5] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* John Wiley, 1991.

[6] R. Cunnington. Movement-related potentials in Parkinson's disease: motor imagery and movement preparation. *Brain*, 120:1339–1353, 1997.

[7] S.T. Elder, J.K. Lashley, N. Kedouri, D. Regenbogen, S. Martyn, G. Roundtree, and C. Grenier. Can subjects be trained to communicate through the use of EEG biofeedback ? *Clinical Biofeedback and Health*, 9:42–47, 1986.

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

[9] Z.A. Keirn and J.I. Aunon. A new mode of communication between man and his surroundings. *IEEE Transactions on Biomedical Engineering*, 37(12):1209–1214, 1990.

[10] D.G. Kleinbaum, L.L. Kupper, and K.E. Muller. *Applied Regression Analysis and Other Multivariable Methods.* PWS-Kent, Boston, 1988.

[11] D.J. McFarland L.A. Miner and J.R. Wolpaw. Answering Questions with an Electroencephalogram-based Brain-Computer Interface. *Archives of Physical and Medical Rehabilitation*, 79:1029–1033, 1998.

[12] D. Lowe. Feature space embeddings for extracting structure from single channel wake EEG using RBF networks. In *Neural Networks for Signal Processing: Proceedings of the IEEE workshop*, pages 428–437, 1998.

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

[14] D.J. McFarland, L. McCane, L.A. Miner, T.M. Vaughan, and J.R. Wolpaw. EEG Mu and Beta Rhythm Topographies with Movement Imagery and Actual Movement. In *Society for Neuroscience Abstracts*, page 1277, 1997.

[15] D.J. McFarland, G.W. Neat, R.F. Read, and J.R. Wolpaw. An EEG-based method for graded cursor control. *Psychobiology*, 21(1):77–81, 1993.

[16] G.R. McMillan and G.L. Calhoun et al. Direct brain interface utilizing self-regulation of steady-state visual evoke response (SSVER). In *Proceedings of RESNA, June 9-14*, pages 693–695, 1995.

[17] C. Neuper, C. Guger, E. Haselsteiner, B. Obermaier, M. Pregenzer, H. Ramoser, A. Schlogl, and G. Pfurtscheller. Current Trends in Graz Brain-Computer Interface (BCI) Research. In *Brain-Computer Interface Technology: Theory and Practice, First International Meeting*, pages 84–89, Rensselaerville Institute, New York, June 1999.

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

[19] M. Peltoranta and G. Pfurtscheller. Neural network based classification of non-averaged event-related EEG responses. *Medical and Biological Engineering and Computing*, 32:189–196, 1994.

[20] W.D. Penny, D. Husmeier, and S.J. Roberts. The Bayesian Paradigm: second generation neural computing. In P.J. Lisboa, editor, *Artificial Neural Networks in Biomedicine.* Springer, 1999.

[21] W.D. Penny and S.J. Roberts. Dynamic Logistic Regression. In *International Joint Conference on Neural Networks (IJCNN)*, 1999.

[22] W.D. Penny and S.J. Roberts. EEG-based communication via dynamic neural network models. In *International Joint Conference on Neural Networks (IJCNN)*, 1999.

[23] W.D. Penny, S.J. Roberts, and M.J. Stokes. Imagined hand movements identified from the EEG mu-rhythm. *Journal of Neuroscience Methods*, 1998. Submitted.

[24] G. Pfurtscheller and A. Berghold. Patterns of cortical activation during planning of voluntary movement. *Electroencephalography and Clinical Neurophysiology*, 72:250–258, 1989.

[25] G. Pfurtscheller, D. Flotzinger, and J. Kalcher. Brain-Computer Interface - a new communication device for handicapped persons. *Journal of Microcomputer Applications*, 16:293–299, 1993.

[26] G. Pfurtscheller, D. Flotzinger, and C. Neuper. Differentiation between finger, toe and tongue movement in man based on 40 Hz EEG. *Electroencephalography and Clinical Neurophysiology*, 90:456–460, 1994.

[27] S.J. Roberts, W. Penny, and I. Rezek. Temporal and spatial complexity measures for EEG-based brain-computer interfacing. *Medical and Biological Engineering and Computing*, 37(1):93–99, 199.

[28] S.J. Roberts and W.D. Penny. Real-Time Brain-Computer Interfacing: A Preliminary Study Using Bayesian Learning. *Medical and Biological Engineering and Computing*, 1999. Submitted.

[29] 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.

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

[31] E.E. Sutter. The Brain Response Interface: communication through visually-induced electrical brain responses. *Journal of Microcomputer Applications*, 15:31–45, 1992.

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

[33] C. Toro, G. Deuschl, R. Thatcher, S. Sato, C. Kufta, and M. Hallett. Event-related desynchronization and movement-related cortical potentials on the ECOG and EEG. *Electroencephalography and Clinical Neurophysiology*, 93(5):380–389, 1994.

[34] J.R. Wolpaw and D.J. McFarland. Multichannel EEG-based brain-computer communication. *Electroencephalography and Clinical Neurophysiology*, 90:444–449, 1994.

[35] 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.