Bayesian neural networks for detection of imagined finger movements from single-trial EEG

Will Penny & Steve Roberts
w.penny@ic.ac.uk s.j.roberts@ic.ac.uk
Department of Electrical Engineering
Inperial College
Exhibition Road
London SW7 2BT, UK

Abstract

Imagined finger movements can be discriminated from background EEG with a typical accuracy of 75%. This shows that in principal it is possible to build a 'brain-computer interface' whereby a person can control a device by imagining a movement. The EEG corresponding to imagined movement is recognised by a committee of Bayesian neural network classifiers trained on complexity features.

1 Introduction

There is a growing interest in computer interfaces which use information derived from non-invasive physiological measures such as the electroencephalogram (EEG).

The majority of existing interfaces use biofeedback which involves a training period in which the user modifies his EEG in response to some feedback on how well he is doing. A typical method is that of Wolpaw *et. al.* (1994) where the movement of a cursor on a computer screen is determined by some fixed feature of the EEG such as alpha-band power. The user then learns to modify his alpha-band power, by entering into various levels of relaxation, and after several days practice is able to use the interface.

In this paper, however, we look at an approach which does not use biofeedback. The basic idea is that the computer adapts rather than the human. This requires a phase in which the prospective user performs a particular fixed task, such as imagining making finger movements, at regular self-paced or externally paced intervals whilst the EEG is recorded. A computer then adapts its recognition algorithm (e.g. neural network) in order to discriminate between imagined movements and the background EEG. If this can be done reliably then the future occurence of such imagined movements can be used as a switch to control whatever device the user wishes to interface with. This is the protocol we investigate in the paper. We present results from an off-line analysis of EEG data from normal subjects performing imagined hand movements. A neural network is then trained on the EEG data to see if it is possible to discriminate between imagined movement and background EEG.

2 Methods

2.1 Experimental set-up

EEG data was collected from 2 healthy right-handed volunteers (subjects A and B) who participated in recording sessions made up of blocks of approximately 10 minutes each.

In the first block, the subject made externally-paced finger movements in response to a computerised stimulus sequence. This stimulus is identical to that described in Pfurtscheller *et. al.* (1994) and consists of circles appearing on the right or left side of a computer screen accompanied by an auditory stimulus and presented in a known pseudo-random order. The stimuli appear at a frequency of one every 12 seconds. Subjects are asked to look at a cross in the centre of the screen in order to keep eye movements to a minimum and to press a switch with their left or right finger depending on the stimulus.

In the second block the stimulus was the same but the subject was asked to just imagine making the the finger movements.

Two further blocks of recordings were made whilst the subject made voluntary self-paced real finger movements at intervals of about 10 seconds. In the first of these blocks the subject had their eyes closed and in the second they had their eyes open.

A third subject (subject C) also took part in externally-paced recordings. For this subject, an additional 'control' block was recorded in which the subject was asked to view the stimulus but not to make any movements or imagine doing so.

The EEG was recorded from a grid of electrodes shown in Figure 1 using a sampling frequency of 125Hz. The electrodes are positioned over the right and left motor cortex in a subset of the positions used in Pfurtscheller *et. al.* (1994). All signals were recorded with respect to the ground electrode and the impedance between each electrode and the reference was less than 5K. The electromyogram (EMG) was recorded from the left and right finger muscle and the electroculagram (EOG) was recorded from the right eye to monitor eye blinks.

2.2 Signal processing

The EEG signals were transformed by the Laplacian operator which refers electrode voltages to a local average thus localising the information received from each

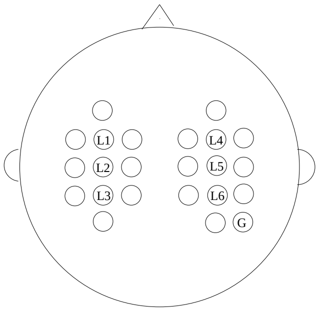

Figure 1: *EEG montage: Laplacian channels L2 and L5 correspond to the C3 and C4 positions in the standard 10-20 system. All other electrodes are placed 3cm apart in a grid surrounding C3 and C4. G denotes the ground electrode.*

channel. In this way, the 22 electrode voltages were reduced to the six Laplacian channels (L1,L2,L3,L4,L5 and L6) shown in Figure 1.

The EEG signals were then processed to extract features which discriminate between movement and non-movement. Pfurtscheller *et. al.* (1994) has shown that the mu-rhythm, an 8-12 Hz oscillation in the motor cortex, disappears 1-2 seconds prior to real finger movement. In the same study Pfurtscheller also found an increase in 40 Hz EEG accompanying movement.

A preliminary analysis of our data, however, revealed that neither of these features were particularly discriminative of movement on a single trial basis. For this reason, we decided to use complexity features as described in Rezek and Roberts (1996).

To this end, we used the method of delays to embed each Laplacian signal in a d-dimensional embedding space (d=20). Samples are embedded over a 2 second window. The singular value decomposition (SVD) of the windowed samples is then made. We then define the complexity of the signal as one minus the proportion of signal variance explained by the first singular value. Thus, signals which are dominated by a single mode (eg. the mu-rhythm) will have low complexity. We therefore expect an increase in EEG signal complexity when the subject makes a movement.

Each of the complexity signals is then smoothed with a 2 second Hanning window before the pattern recognition stage.

2.3 Pattern recognition

The time-series of complexity values are then compared at times around the movement versus baseline times 6 seconds before each movement. This results in two groups of complexity values (one for movements and one for non-movements) which are to be discriminated by a neural network. This reduces the problem to one of static pattern recognition.

The resulting patterns are then classified by Multilayer Perceptrons (MLPs) trained with a regularised cost function according to the Bayesian framework introduced by Mackay (1992). We use three regularisers: one for the hidden-layer weights, one for the hidden-layer biases and one for the output unit's weights and biases - because each of these weight groups will develop on different scales. Consequently, the networks are trained on the cost function

$$E = G + \frac{1}{2}\sum_k \alpha_k E_w^k \tag{1}$$

where G is the usual cross-entropy term, α_k is the kth regularizer, and E_w^k is the total size of the kth weight group which we can write this as

$$E_w^k = \frac{1}{2}\boldsymbol{w}^T \boldsymbol{I}_k \boldsymbol{w} \tag{2}$$

where $\boldsymbol{w}$ is the weight vector and $\boldsymbol{I}_k$ is a diagonal matrix having ones along the diagonal that pick off weights in the kth group.

The regularizers are initially set to small arbitrary values (10^{-6}) and the networks are trained to a certain tolerance (relative change in cost function). The regularizers are then re-estimated according to Mackay's formula

$$\alpha_k = \frac{\gamma_k}{2E_w^k} \tag{3}$$

where γ_k is the number of well determined parameters in group k which is given by

$$\gamma_k = \sum_j^W \frac{\lambda_j - \alpha_k}{\lambda_j}(\boldsymbol{V}^T \boldsymbol{I}_k \boldsymbol{V})_{jj} \tag{4}$$

where the sum runs over the W network weights and λ_j and the jth column of $\boldsymbol{V}$ are respectively the jth eigenvalue and eigenvector of the Hessian matrix

$$\boldsymbol{A} = \nabla\nabla G + \sum_k \alpha_k \boldsymbol{I}_k \tag{5}$$

The Hessian $\nabla\nabla G$ is computed by the outer product approximation as described in Bishop (1995).

Once the regularisers have been re-estimated the network is then trained from where it left off until a specified lower tolerance value is achieved. Re-estimation and further training continue according to some tolerance regime until a minumum tolerance is reached.

The above re-estimation scheme means that the network is automatically regularized and so will not overfit the data. This eliminates the need for a separate validation set for choosing optimal regularizer values.

All networks are trained with the Broyden-Fletcher-Goldfarb-Shanno (BFGS) Pseudo-Newton algorithm. All results are compared with linear discriminant analysis.

3 Results

3.1 Signal processing

Figure 2 shows a typical time series of complexity values for a subject imagining making right and left movements. For right movements we clearly observe a dip in complexity at movement times. This shows that we are not solely picking up the desychnronisation of the mu-rhythm as this would cause an increase in complexity. For left movements the trend is not so clear.

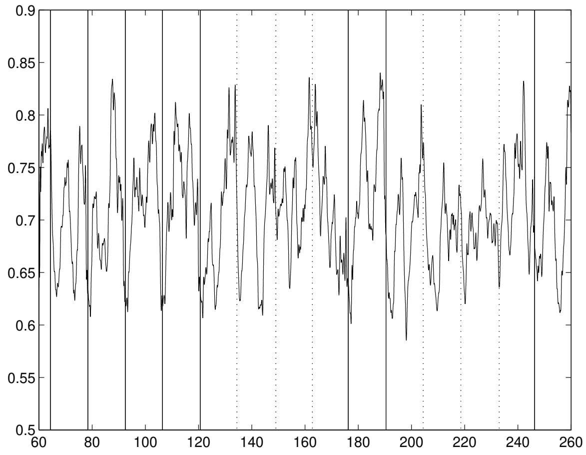

Figure 2: *Time-series of complexity values from Laplacian channel L2 for subject B imagining making right and left movements in response to an external cue. The solid vertical lines indicate right cues and the broken lines indicate left cues.*

3.2 Pattern recognition

The time-series of complexity values are then compared at movement times and at baseline times 6 seconds before each movement. This results in two groups of 6-dimensional complexity values (one dimension from each Laplacian channel).

Columns 1 and 2 of Table 1 shows the classification accuracy of linear and neural net classifiers on data from 3 subjects performing real and imagined movements under different conditions. Each data set consists of about 80 patterns (40 movements and 40 non-movements).

The results are from a 10-fold cross-validation of each data set. In each fold, 60 networks were trained having 1,2,3,4,5 or 6 hidden units. For each structure 10 networks were trained from different random initializations. The 10 out of 60 networks having the highest evidence (the evidence is a Bayesian model selection criterion described in Mackay (1992)) were then used in a uniformly weighted committee to classify the test data in that fold. The results show only a modest improvement over linear discriminant analysis which suggests that the decision boundaries are linear in most of these data sets. This is also indicated by the predominance of single hidden unit networks in the committees.

Table 1: *Classification accuracy, *indicates the p-value from a* χ^2 *test being greater than 0.05.*

Subject/Condition	Linear Discriminant with 6 inputs	Neural Network with 6 inputs	Feature selection from 30 inputs
A/self-paced, eyes-closed	69	70	67
A/self-paced, eyes-open	73	73	90
A/ext-paced, real	71	72	75
A/ext-paced, imagined	66	66	70
B/self-paced, eyes-closed	89	90	99
B/self-paced, eyes-open	68	73	62
B/ext-paced, real	88	94	90
B/ext-paced, imagined	71	71	89
C/ext-paced, imagined	75	75	79
C/ext-paced, control	63*	63*	56*

The non-significant results on the control data set tell us that, for externally-paced movements, the classifiers are not picking up features in the EEG caused by the evoked response of the stimuli.

3.3 Feature selection

We have also looked at sampling the time-series of complexity values at 0, 0.5, 1, 1.5 and 2 seconds prior to each movement. Over the 6 channels this gives a total of 30 input features. This is too many inputs to put directly into a neural network given that we only have 80 data points. We have therfore looked at univariate rank correlation measures and multivariate linear forward stepwise selection to reduce the number of features. Figure 3 shows the features selected for one of the data sets. The selected features vary considerably between data sets. Feature selection from a richer pool of features also yields better classification results as shown by the results in column 3 of Table 1.

4 Discussion

It is possible to detect imagined finger movements from the background EEG on a single trial basis with a typical accuracy of 75%. This does vary between subjects and has yet to be demonstrated in a larger population. The accuracy may be made higher still by using spatial coherence measures or some other measure of the spatial complexity of the EEG. It also seems that feature selection will play a major role in the pattern recognition stage. To this end, we have looked at Automatic Relevance Determination (Mackay (1995)). Apart from performing a soft feature selection the method also allows vague prior knowledge to be embedded into classifiers, such as 'features concurrent with movement cues will be more discriminative' or 'electrodes 2 and 5 contain the most information'. Early results show a marginal improvement on the data sets we have looked at. Our next point of focus is to classify left movements versus right movements. Preliminary results suggest that this can be performed with about 80% accuracy.

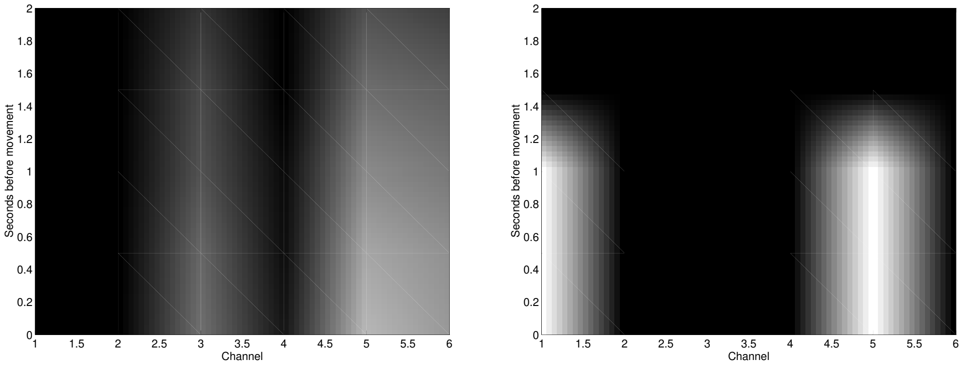

Figure 3: *The left figure shows the value of the Wilcoxon statistic for each feature rescaled to lie between 0.5 (black) and 1 (white). This is a measure of the rank-correlation between feature value and class label. The right figure shows the features selected by multivariate linear forward stepwise selection. Feature selection took place in each fold - if a feature was selected in every fold it is displayed white, if it is never selected it is displayed black. These features are from subject B making self-paced movements with eyes-closed.*

Acknowledgements

Will Penny is a research fellow funded by the UK Engineering and Physical Sciences Research Council (EPSRC) (GR/K79062), whose support we gratefully acknowledge. The authors would also like to thank Dr. Maria Stokes of the Royal Hospital for Neurodisability who is a co-grant-holder on the project.

References

Wolpaw, J.R. & McFarland, D.J. (1994) Multichannel EEG-based brain-computer communication. *Electroencephalography and Clinical Neurophysiology* **90**:444-449.

Pfurtscheller, G., Flotzinger, D. & Neuper, C. (1994) Differentiation between finger, toe and tongue in man based on 40 Hz EEG. *Electroencephalography and Clinical Neurophysiology* **90**:456-460.

Rezek, I.A. & Roberts, S.J. (1996) Stochastic complexity measures for physiological signal analysis. Submitted to *IEEE transactions on Biomedical Engineering.* Available from http://www.ee.ic.ac.uk/hp/staff/sjrob/pubs.html.

Mackay, D.J.C. (1992) The evidence framework applied to classification problems. *Neural Computation* **4**(5):720-736.

Bishop, C.M. (1995) *Neural Networks for Pattern Recognition.* Oxford:Oxford University Press.

Mackay, D.J.C. (1995) Bayesian non-linear modelling for the 1993 energy prediction competition. In G. Heidbreder (Ed.), *Maximum Entropy and Bayesian Methods, Santa Barbara 1993*, Dordrecht: Kluwer.