ATTENTIONAL MODULATION OF AUDITORY EVENT-RELATED POTENTIALS IN A BRAIN-COMPUTER INTERFACE

N. Jeremy Hill, [1,*] *Thomas Navin Lal,* [1] *Karin Bierig,* [1] *Niels Birbaumer* [2] *and Bernhard Schölkopf* [1]

[*] jez@tuebingen.mpg.de
[1] Department of Empirical Inference for Machine Learning and Perception,
Max Planck Institute for Biological Cybernetics, Tübingen, Germany
[2] Institute of Medical Psychology and Behavioral Neurobiology,
Eberhard Karls University, Tübingen, Germany

ABSTRACT

Motivated by the particular problems involved in communicating with "locked-in" paralysed patients, we aim to develop a brain-computer interface that uses auditory stimuli. We describe a paradigm that allows a user to make a binary decision by focusing attention on one of two concurrent auditory stimulus sequences. Using Support Vector Machine classification and Recursive Channel Elimination on the independent components of averaged event-related potentials, we show that an untrained user's EEG data can be classified with an encouragingly high level of accuracy. This suggests that it is possible for users to modulate EEG signals in a single trial by the conscious direction of attention, well enough to be useful in BCI.[†]

1. INTRODUCTION

In some cases, it has been possible to allow an entirely paralysed or "locked-in" patient to communicate, using brain signals measured externally by EEG. The patient is trained, usually over the course of several months, to produce two distinguishable types of signal (one representing "yes" and the other "no") which are then classified by computer. Successful approaches to brain-computer interfaces (BCI) include paradigms based on slow cortical potentials [1], signals from the motor and pre-motor areas related to the imagination of voluntary movements [2,3] and evoked potentials in response to visual stimulus events [4]. The different paradigms work to varying degrees, depending on the patient. In some cases, long immobility and the degeneration of the pyramidal cells of the motor cortex may mean the patient can no longer produce classifiable imagined-movement signals. In the most severe cases, where the eyes are completely immobile, the visual modality becomes too limited to be useful for presentation of stimuli and feedback. Thus, there is considerable motivation to explore new BCI paradigms, and in particular to develop systems that rely only on auditory or tactile stimuli, since these modalities usually function fully in paralysed patients.

[†] NB: an extended analysis of these data will be presented at NIPS 2004, Vancouver.

0-7803-8665-5/04/$20.00 ©2004 IEEE

We present the results of experiments on healthy subjects, designed to develop a paradigm for BCI in which a user can make a binary choice by directing his or her attention to one of two concurrent auditory stimulus streams. The paradigm is based on results from the 60's and 70's from Hillyard et al. [5] and others (see [6,7] for a review) which indicate that, when a person listens to two sequences of auditory stimuli from the left and right, the event-related potentials (ERPs) in the EEG signal are modulated by the listener's selective attention to one or the other. Their results, usually measured with just one or two electrodes, are significant when averaged over a large number of ERPs. By using machine-learning techniques to classify signals measured from multiple electrodes, we aim to determine whether *single* trials, each lasting no more than a few seconds and therefore containing only a small number of ERPs, are classifiable enough for the effect to be potentially useful in BCI.

Rather than relying on extensive training of the subjects to meet a fixed criterion, our approach is to train our classifier (offline) to use the signals that an untrained subject can produce in a single two-hour experimental session.

2. EXPERIMENTAL SETUP

EEG signals were recorded from 15 healthy untrained subjects (9 female, 6 male) between the ages of 20 and 38, using 39 silver chloride electrodes, referenced to the ears. An additional EOG electrode was positioned lateral to and slightly below the left eye, to record eye movement artifacts—blinks and horizontal and vertical saccades all produced clearly identifiable signals on the EOG channel. The signals were filtered by an analog band-pass filter between 0.1 and 40 Hz, before being sampled at 256 Hz.

Subjects sat in front of a computer monitor and performed eight 10-minute blocks each consisting of 50 trials. Each trial lasted 6.5 seconds, with a pause of between 2 and 4 seconds in between trials for the subject to relax. For the duration of the trial, a fixation point was visible on the screen, and subjects were asked to keep their gaze fixed on this, minimizing as far as possible eye movements, blinks and swallowing, and not making any voluntary muscle movements, since we wished to ensure that the signals would be as free as possible from artifacts (i.e. signals that could not be obtained from a paralysed patient).

Timing of the events in a single trial was as follows (in milliseconds after the start of the trial):

0	fixation point on
1000	visual cue on
1500	visual cue off
2000	auditory stimuli start
6000	auditory stimuli stop
6500	fixation point off

The auditory stimuli were presented from speakers situated to the subject's left and right. The visual cue on each trial was an arrow pointing left or right, indicating whether the subject should attend to the left or right stimulus sequence. In each block of 50 trials, the left and right stimuli were cued 25 times each, in random order.

The auditory stimuli were two periodic sequences of 50-msec-long square-wave beeps. Each sequence contained "target" and "non-target" beeps: the first three in the sequence were always non-targets, after which they could be targets with independent probability 0.3. The right-hand sequence consisted of eight beeps of frequencies 1500 Hz (non-target) and 1650 Hz (target), repeating with a period of 490 msec. The left-hand sequence consisted of seven beeps of frequencies 800 Hz (non-target) and 880 Hz (target), starting 70 msec after start of the right-hand sequence and repeating with a period of 555 msec.

The subject's task was to count the number of target beeps in the sequence indicated by the arrow, ignoring the other sequence. In the pause between trials, they were instructed to report the number of target beeps using a numeric keypad (in order to keep the measurement period free of movement artifacts, a few practice trials usually had to be run beforehand so that the subject learned not to start the hand movement for response before the fixation point disappeared).

The sequences differed in location and pitch in order to help the subjects focus their attention on one sequence and ignore the other. The regular repetition of the beeps, and the different periodicities of the two sequences, were designed to allow the average ERP to a left-hand beep on a single trial to be examined with minimal contamination by ERPs to right-hand beeps, and vice versa: when the periods of one sequence are averaged, signals correlated with that sequence add in phase, whereas signals correlated with the other sequence spread out, out of phase. Comparison of the average response to a left beep with the average response to a right beep, on a single trial, should thus emphasize any modulating effect of the direction of attention on the ERP, of the kind described by Hillyard et al. [5].

3. ANALYSIS AND RESULTS

Each trial from each subject was first examined by eye, and trials were rejected if they contained obvious large artifact signals caused by blinks or saccades (visible in the EOG and across most of the frontal positions), small periodic eye movements, or other muscle movements (neck and brow, judged from electrode positions O9 and O10, Fp1, Fpz and Fp2). Between 6 and 228 trials had to be rejected out of 400, depending on the subject.

For each subject, trial and electrode, the average ERP following a left-side beep was computed, and then the average ERP following a right-side beep. These two average signals were then concatenated: the result was 142 (left) + 125 (right) = 267 time samples in each of the 40 channels (39 EEG + 1 EOG), for a total of 10680 input dimensions to the classifier.

The classifier used was a linear hard-margin Support Vector Machine (SVM) [8]. To evaluate its performance, the trials from a single subject were split into ten non-overlapping partitions of equal size: each such partition was used in turn as a test set for evaluating the performance of the classifier trained on the other 90% of the trials. Before training, linear Independent Component Analysis (ICA) was carried out on the training set in order to perform blind source separation—this is a common technique in the analysis of EEG data [9,10], since signals measured through the skull, meninges and cerebro-spinal fluid are of low spatial resolution, and the activity measured from neighbouring EEG electrodes can be assumed to be highly correlated mixtures of the underlying sources. For the purposes of the ICA, the concatenation of all the preprocessed signals from one EEG channel, from all trials in the training partition, was treated as a single mixture signal. A 40-by-40 separating matrix was obtained using the stabilized deflation algorithm from version 2.1 of FastICA [11]. This matrix, computed only from the training set, was then used to separate the signals in both the training set and the test set. Then, the signals were centered and normalized: for each averaged (unmixed) ERP in each of the 40 ICs of each trial, the mean was subtracted, and the signal was divided by its 2-norm. Thus the entry K_{ij} in the kernel matrix of the SVM was proportional to the sum of the coefficients of correlation between corresponding epochs in trials i and j. The SVM was then trained and tested. Single-trial error rate was estimated as the mean proportion of misclassified test trials across the ten folds. For comparison, the classification was also performed on the mixture signals without ICA, and with and without the normalizing step.

Results are shown in table 1. For readability, standard error values for the estimated error rates are not shown:

standard error was typically ±0.025, and maximally ±0.04. It can be seen that the best error rate obtainable with a given subject varies according to the subject, between 3% and 37%, in a way that is not explained by the differences in the numbers of good (artifact-free) trials available. ICA generally improved the results, by anything up to 14%. Normalization generally produced a small improvement.

subj.	#good trials	-	-	ICA	ICA
		-	norm'd	-	norm'd
CM	326	0.08	0.06	0.06	**0.04**
CN	250	0.26	0.19	0.28	**0.14**
GH	198	0.34	0.27	0.35	**0.22**
JH	348	0.21	0.19	0.14	**0.08**
KT	380	0.23	0.21	0.15	**0.07**
KW	394	0.18	0.14	0.06	**0.03**
TD	371	0.22	0.18	0.15	**0.10**
TT	367	0.32	**0.31**	0.33	0.32
AH	353	0.22	0.22	0.17	**0.16**
AK	172	0.35	0.31	0.34	**0.22**
CG	271	0.37	0.29	0.31	**0.28**
CH	375	0.31	0.28	0.26	**0.22**
DK	241	0.34	0.34	0.35	**0.30**
KB	363	0.21	0.21	0.15	**0.10**
SK	239	0.47	0.43	0.40	**0.37**

Table 1: error rates for 15 subjects with and without ICA, and with and without normalization (the lowest error rate for each subject is in **bold**)

Thus, promising results can be obtained using the average ERP in response to a small number of auditory stimuli, using ICA followed by per-channel normalization (last results column): error rates of 5–15% for some subjects are comparable with the performance of, for example, well-trained patients in an SCP paradigm [1], and correspond to information transfer rates of 0.4–0.7 bits per trial (say, 4–7 bits per minute).

In order to examine the extent to which the dimensionality of the classification problem could be reduced, recursive feature elimination [12] was performed (limited now to normalized, unmixed data). For each of ten folds, ICA and normalization was performed, then an SVM was trained and tested. Each independent component contributes 267 features (averaged, unmixed time samples) to the representation of a trial, so the elimination score for an IC was equal to the sum of the squares of the 267 elements of the hyperplane normal vector from the trained SVM, corresponding to those features. The IC with the lowest score was deemed to be the least influential for classification, and its set of 267 features was removed from the representation. Then the SVM was re-trained and re-tested, and the elimination process iterated until one channel remained. The removal of batches of features in this way is similar to the Recursive Channel Elimination approach to BCI introduced by Lal et al. [3], except that, here, independent components are removed instead of mixtures. A convenient acronym would therefore be RICE, for Recursive Independent Component Elimination.

The results of feature elimination are plotted in figure 1, showing estimated error rates averaged over ten folds against the number of ICs used for classification. Each subject's initials, together with the number of useable trials that subject performed, are printed to the right of the corresponding curve. It can be seen that a fairly large number of ICs (around 20–25 out of the 40) contribute to the classification: this may indicate that the useful information in the EEG signals is diffused fairly widely between the areas of the brain from which we are detecting signals (indeed, this is in accordance with much auditory-ERP and mismatch negativity research, in which strong signals are often measured at the vertex, quite far from the auditory cortex [5–7]). One of the motivations for reducing the dimensionality of the data is to determine whether performance can be improved as irrelevant noise is eliminated, and as the probability of overfitting decreases. However, these factors do not seem to limit performance on the current data: for most subjects, performance does not improve as features are eliminated, instead remaining roughly constant until fewer than 20–25 ICs remain. A possible exception is KT, whose performance may improve by 2–3% after elimination of 20 components, and a clearer exception is CG, for whom elimination of 25 components yields an improvement of roughly 10%.

The ranking returned by the RICE method is somewhat difficult to interpret, not least because each fold of the procedure can compute a different ICA decomposition, whose independent components are not necessarily readily identifiable with one another. A thorough analysis is not possible here—however, with the mixture weightings for many ICs spread very widely around the electrode array, we found no strong evidence for or against the particular involvement of muscle movement artifact signals in the classification.

RICE was also carried out using the full 400 trials for each subject (results not shown). Despite the (sometimes drastic) reduction in the number of trials, rejection by eye of artifact trials did not raise the classification error rate by an appreciable amount. The one exception was subject SK, for whom the probability of mis-classification increased by about 0.1 when 161 trials containing strong movement signals were removed—clearly this subject's movements were classifiably dependent on whether he was attending to the left or to the right.

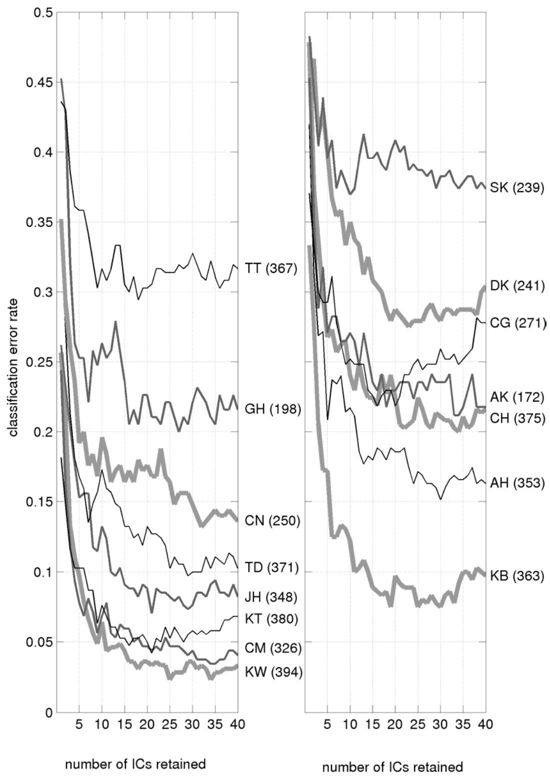

Figure 1: Results of
Recursive Independent Component Elimination

4. CONCLUSION

Despite wide variation in performance between subjects, which is to be expected in the analysis of EEG data, our classification results suggest that it is possible for a user with no previous training to direct conscious attention, and thereby modulate the event-related potentials that occur in response to auditory stimuli reliably enough, on a single trial, to provide a useful basis for a BCI. The information used by the classifier seems to be diffused fairly widely over the scalp. While the ranking from recursive independent component elimination did not reveal any evidence of an overwhelming contribution from artifacts related to muscle activity, it is not possible to rule out completely the involvement of such unwanted signals—possibly the only way to be sure of this is to implement the interface with locked-in patients, preliminary experiments for which are in progress.

5. ACKNOWLEDGMENTS

We would like to thank Prof. Kuno Kirschfeld and Bernd Battes for the use of their laboratory and EEG equipment.

6. REFERENCES

[1] N. Birbaumer, A. Kübler, N. Ghanayim, T. Hinterberger, J. Perelmouter, J. Kaiser, I. Iversen, B. Kotchoubey, N. Neumann and H. Flor, "The Thought Translation Device (TTD) for Completely Paralyzed Patients," *IEEE Transactions on Rehabilitation Engineering,* 8(2):190–193, June 2000.

[2] G. Pfurtscheller, C. Neuper, A. Schlögl, and K. Lugger, "Separability of EEG Signals Recorded during Right and Left Motor Imagery using Adaptive Autoregressive Parameters." *IEEE Transactions on Rehabilitation Engineering*, 6(3):316–325, 1998.

[3] T. N. Lal, M. Schröder, T. Hinterberger, J. Weston, M. Bogdan, N. Birbaumer and B. Schölkopf, "Support Vector Channel Selection in BCI," *IEEE Transactions on Biomedical Engineering*, 51(6):1003–1010, June 2004.

[4] E. Donchin, K. M. Spencer and R. Wijesinghe, "The mental prosthesis: Assessing the speed of a P300-based Brain-Computer Interface," *IEEE Transactions on Rehabilitation Engineering,* 8:174–179, 2000.

[5] S. A. Hillyard, R. F. Hink, V. L. Schwent and T. W. Picton, "Electrical Signs of Selective Attention in the Human Brain," *Science,* 182:177–180, 1973.

[6] R. Näätänen, "Processing Negativity: An Evoked-Potential Reflection of Selective Attention," *Psychological Bulletin,* 92(3):605–640, 1982.

[7] R. Näätänen, "The Role of Attention in Auditory Information Processing as Revealed by Event-Related Potentials and Other Brain Measures of Cognitive Function," *Behavioral and Brain Sciences,* 13:201–288, 1990.

[8] B. Schölkopf and A. Smola, *Learning with Kernels,* MIT Press, Cambridge, USA, 2002.

[9] K. R. Müller, J. Kohlmorgen, A. Ziehe, and B. Blankertz, "Decomposition Algorithms for Analysing Brain Signals," in S. Haykin (ed.) *Adaptive Systems for Signal Processing, Communications and Control,* pp. 105–110, 2000.

[10] A. Delorme and S. Makeig, "EEGLAB: An Open Source Toolbox for Analysis of Single Trial EEG Dynamics Including Independent Component Analysis," *Journal of Neuroscience Methods,* 134:9–21, 2004.

[11] A. Hyvärinen, "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis," *IEEE Transactions on Neural Networks,* 10(3):626–634, 1999.

[12] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene Selection for Cancer Classification using Support Vector Machines," *Journal of Machine Learning Research,* 3:1439–1461, 2003.