NEURAL NETWORK CLASSIFICATION OF EVENT RELATED POTENTIALS FOR THE DESIGN OF A NEW COMPUTER INTERFACE

Christopher Konger, Jose C. Principe, and Murat Taner

Electrical Engineering Department,
University of Florida, Gainesville, FL 32611

ABSTRACT

In this work we discuss the classification of single event related potentials (ERP) with a neural network classifier. The purpose is to discriminate the YES and NO answers thought by a computer user in response to words flashed in the computer display, and to assess the feasibility of building a new computer interface which we call the cortical mouse.The performance of the neural network is compared with a classifier using an eigenvector based extraction method. In this study the neural network outperformed the other classifier. For some subjects the classification is above 90%, but the introspected variability is still large.

INTRODUCTION

Event related potentials (ERPs) carry information about the intention of the subject response to external stimuli. This features is being explored in our laboratory to discriminate between YES and NO answers thought by the subject in response to questions presented in the computer display. Our goal is to have quadriplegics control the cursor motion in a computer screen, and allow them to use the computer for education, leisure activities or to help them interact with the external world more efficiently. We called this new interface the cortical mouse.

The problem of such project resides in the difficult task of discriminating YES from NO answers using ERPs. ERPs are faint, short transients in the electroencephalogram (EEG). The normal way to cope with the negative signal-to-noise ratio is to average the ERPs. However in our project this has two drawbacks: One is related to the long time involved to make a decision. The other is the subject habituation to a stimulus that may violate the condition of equal ERP shape required in the averaging. We are therefore exploring the possibilities of making the discrimination based on single ERPs.

This preliminary study was designed to show that the computer could distinguish simple YES-NO thoughts by processing the evoked response potentials (ERPs) embedded in the electroencephalogram. This report outlines the experimental design, the results from two different pattern classification schemes (an eigenvector based extraction method and a neural network) and gives directions for further study.

This work was partially funded by RADC contract # F30602-88-D-0027, and Florida High Technology Council 1990.

EXPERIMENTAL PARADIGM

ERP's carry information about a subject's reaction to external stimuli. Previous neuropsychological studies have shown that the signal feature N400 (the negative peak at 400 ms from stimulus onset) is highly correlated to internal cognitive processing in response to the stimulus [1].

In this study, the subject was seated in front of a computer screen on which words (UP, DOWN, LEFT, RIGHT) were flashed in a paired succession. If the two words were the same, the subject was instructed to think YES, otherwise for mismatched pairs the subject thought NO. The EEG's of the subject were sampled synchronously with the word presentation and was digitized with a 12 bit A/D converter (125 Hz sampling frequency) and stored in a data file. These files consisted of 4096 msec of data from three channels: the Pz and Cz channels (10/20 international system), and the electrooculogram (EOG). Along with the raw EEG data, the computer created files indicating the sequence of word pair presentations. The latter was used in training of the classifiers since the input signals could be grouped into the YES-NO classes. For this classification, a single channel was utilized (P_z), provided that no eye blinks were detected. Figure 1 shows the timing diagram of the data collection. A full account of the experimental setup can be found in [2].

Data were taken for six subjects. For each volunteer, three sessions of data collection were made. Each session consisted of 48 presentations of word pairs. These sessions were then used to train the classifiers, as well as testing them on an "unknown" set of trials.

FIGURE 1
DATA COLLECTION TIMING

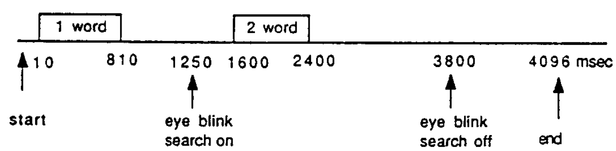
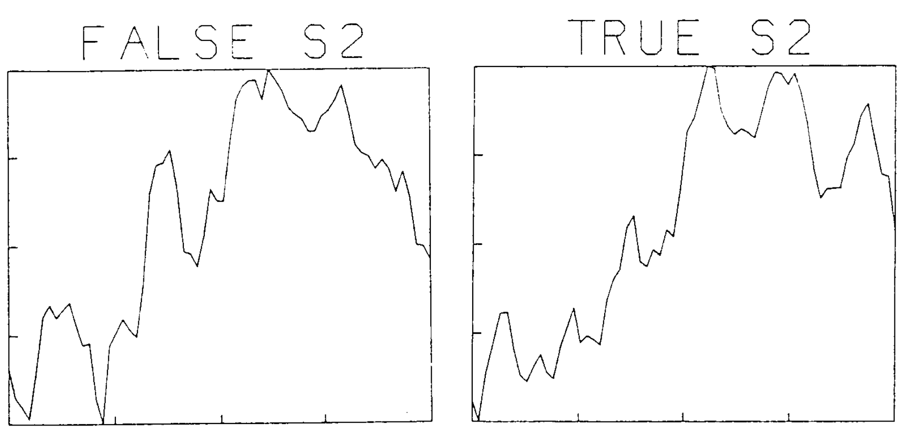


FIGURE 2
ERP ON CHANNEL PZ



RESULTS

Due to the variability of the EEG between subjects, the classifiers were trained for a specific subject by using the data from one or two sessions. A typical example of the ERPs for the YES and NO answers is presented in Figure 2. Once trained, the classifiers where tested on another session not used during training. Subject #1 was bilingual. It was found that there was a difference between sessions when he was requested to think in English and his native language. Due to this variability between sessions, it was decided not to use this data for the training and testing. All other subjects had English as their native language. The data from subject #2 also had to be discarded as the subject was extremely sleepy (frequent alpha bursts), had frequent eye blinks as indicated by the EOG, and confessed frequent mistakes. Subjects #3 through 6 had acceptable data and were used for training and testing of the classifiers.

The Eigenvector Classifier

The eigenvector based classifier is fully described in [3]. Basically the algorithm orders the eigenvectors of data partitions to achieve data reduction. For each partition a feature is selected to represent the data, until only two features are reached to classify the data into two classes. In this procedure the Fisher's ratio is used as a figure of merit. For this study, a 64 input window centered around the P300/N400 features was used (10 partitions of 10 points with 4 points of overlap). This window thus started with sample 225. TABLE 1 summarizes the results for each of the subjects, indicating the success of correctly classifying the training set as well as the test set.

TABLE 1
64 Input Eigenvector Classifier

Subject	Training Hit Rate (%)	Training Session	Testing Hit Rate (%)	Testing Session
3	58/75	1	40/40	3
3	71/69	2	56/54	3
3	63/64	1+2	44/56	3
4	71/65	1	50/44	3
4	65/56	2	50/54	3
4	61/56	1+2	46/56	3
5	69/73	1	29/42	3
5	65/75	2	52/50	3
5	61/58	1+2	38/48	3
6	x/75	1	x/48	3
6	58/73	2	48/50	3
6	x/64	1+2	x/44	3

x: did not converge

These results are very homogeneous, and are not very encouraging. The hit rate for the training sets vary between 55 and 75%, i.e., slightly better than chance. There are no significant differences between the Cz and Pz channels. The performance in the test set is worse than for the training sessions, being at or below chance level. When both sessions are used for training, the performance of the classifier in the test case seems more robust. Interpreting these results, it appears that the classifier is not powerful enough to extract similar features between events belonging to the same class. A scatter plot of the trials in the feature plane shows that the two classes are intermingled. One explanation for these results is the known negative signal-to-noise ratio (SNR) of the ERP with respect to the background EEG. If this is the case, a narrower window around the P300/N400 would yield better results, however this would increase the sensitivity to latency of the subject's decision making.

The Neural Network Classifier

A feedforward multilayer perceptron using the backpropagation learning rule was next trained with the data. The Rumelhart and McClelland software package [4] was used to implement the perceptron. This perceptron consisted of 64 input units, 4 hidden units, and 2 output units. The results for the Pz data are displayed in TABLE 2.

TABLE 2
64 INPUT NEURAL NETWORK

Subject	Training Hit Rate (%)	Training Session	Testing Hit Rate (%)	Testing Session
3	88	1+2	52	3
4	86	1+2	57	3
5	89	1+2	65	3
6	87	1+2	42	3

There are two important points to be made: the classification rate for all of the subjects is above 86%, and the network performed at chance levels in the test condition. These results were quite confusing, for the first condition indicates that the classifier was able to make a reliable distinction between the two classes, however the degradation of performance when testing the third session should not have been so large. It was thought that the problem was insufficient iterations during the training stage, thus giving a convergence to a local minima. After experimenting with a very long training period (above 10,000 iterations) which yielded the same results, it was concluded that the solution was to be found elsewhere.

It is now thought that the difference between training and testing results may be attributed to an insufficient number of training patterns. When 2 sessions are used, only 96 patterns are utilized to train a network of 264 weights. This may be sufficient for learning the original 96 patterns, but is insufficient for extrapolating to the set of new patterns. To test this hypothesis with the already collected data, the obvious solution was to decrease the number of input neurons. While this decreases the number of weights and the number of needed training patterns, it also increases the sensitivity to window placement, as discussed in the Bayesian case above.

A new network of 16 inputs, 4 hidden units, and 2 outputs was used for training and testing. TABLE 3 shows the results

It may be noted that the performance of the perceptron is NOT uniformly good across subjects (ranging from 63% to 91% hit rate), i.e., the perceptron is not always learning the training set.

Subject	Training Hit Rate (%)	Training Session	Testing Hit Rate (%)	Testing Session	Window
3	65	1+2	42	3	224
4	63	1+2	73	3	237
5	91	1+2	85	3	224
6	90	1+2	52	3	248
6	65	1+2	NP		224

TABLE 3
16 INPUT NEURAL NETWORK

NP=Not Performed

When comparing with the 64 input model, one can attribute this result to the limiting window width. This is supported by noting how a varied window placement changed the hit rate for subject #6 from 65% to 90%. Thus it appears that a window of 16 points is too small for the reliable learning of ERP patterns.

It is also noted that the performance between the training and testing did not change as drastically as for the 64 input case. The network makes two classifications at chance level, and two that may be considered reasonable (85% and 73%).

Upon observing the activations of the network during the decision process, it was noted that for large activations (defined as being 0.75 of the maximum output of 1.00) the network was very accurate. See TABLE 4.

TABLE 4
16 INPUT NEURAL NETWORK

Subject	High Activation (%)	Training Hit Rate (%)	Testing Hit Rate (%)
3	71	89	59
4	75	95	97
5	93	94	91
6	90	92	58

Using this criteria for high activation, a real time system could reject trials that would give erroneous results, thereby increasing the hit rate. The price paid for such an option is that sometimes the question must be repeated (on the average 17.5% of the time). Note that this method increases all of the training hit rates to 89% and better - even with the short window effects noted in TABLE 3. For the test case, subjects 4 and 5 exhibited hit rates above 90%, while subjects 3 and 6 were slightly above chance.

CONCLUSIONS

In this work an experiment was designed to make binary decisions via brain waves. It was shown that it may be feasible to develop such a mentally controlled system that would be a powerful new paradigm for interactive computing. In cases where the user's hands are occupied and where voice cannot be used (as in noisy environments) the control of a menu driven program by pure thought is convenient. The interaction involves a dialogue, so it will be slow. However, for disabled subjects such as paraplegics, it is a very appealing solution. For efficient interaction, the decision algorithms must be able to classify an answer given a single ERP. This is an extremely difficult classification problem. This study tested the performance of an eigenvector based classifier, as well as a one hidden layer neural network, for decision making.

The performance of the eigenvector classifier was at chance level, both for training and test sets. This was interpreted as saying that the classifier was not powerful enough to discriminate between the two classes.

The neural network performance was better. It was shown that the network has the capability of distinguishing between the two classes, as shown by a hit rate of above 85%. However, the performance for the test set was poorer. This was attributed to the small number of training patterns relative to the number of weights in the network. By reducing the number of inputs to the network from 64 to 16, this ratio was corrected, however other problems surfaced. These arise because of the sensitivity of placement of the narrow window within the EEG samples. For two of the subjects, the performance was relatively high (73% and 85%), but the others exhibited chance decision behavior. When the level of activation is considered, the performance improved slightly. The price paid is that sometimes the stimulus would have to be repeated (17% of the time). The subjects seem to be divided into two distinct groups: one where the performance is very high (in the 90% bracket), and another with lesser reliability (around 60%). Further work needs to be done to establish the best window location, the window size, network topology and dimension. These results do show that the goal of designing a mentally controlled interface is feasible today.

REFERENCES

1. E. Donchin, *Evoked Potentials in Psychiatry*, Plenum 1980.

2. D. Childers, J. Principe, A. Arroyo, *RADC Tech. Rep.* F30602-88-D-0027, 1989.

3. D. Childers, et al., *IEEE Trans. Biomed. Eng.*, vol. BME-29, #6, 423-438, 1982.

4. J. McClelland, D. Rumelhart, *Exploration in PDP*, MIT Press, 1988.