Discriminating Mental States Using EEG Represented by Power Spectral Density

Jack Culpepper

Department of Computer Science
Harvey Mudd College
Claremont, CA 91711
jack@cs.hmc.edu

Abstract

Artificial neural networks were trained to classify segments of 12 channel EEG data into one of five classes corresponding to five cognitive tasks performed by one subject. Three-layer feedforward neural networks were trained using a validation set to control over-fitting. Independent Component Analysis (ICA) was used to segregate obvious artifactual EEG components from other sources, and a frequency-band representation was used to represent the sources computed by ICA. The most notable result is an 85% accuracy rate on differentiation between two tasks, using a segment of EEG 1/20th of a second long.

1 Introduction

For quite some time now, we have been using our physical actions to control computers. The most primitive interfaces (e.g. punch cards) were designed to accomodate the requirements of the computer, rather than to provide the fastest and most intuitive link between the human and the machine. However, as interface technology progressed, efforts to improve this link became more concentrated on making things easier for the human: the keyboard was invented, and then the mouse. Although there are many other, less widely used techniques for controlling computers, these tools have become the modern standard. The others include eye tracking devices that perform the same basic function as a mouse, and speech recognition, which can be used to perform the functions of both the mouse and keyboard. Different types of input devices appeal to people with different needs (for example, the physically handicapped), and some are more appropriate for certain applications than others. Yet, they all share one similarity which is particularly relevant to this study: the requirement that the user encode his intentions into a stream of physical movements, which are then decoded by the computer into a set of signals to be used for control.

It seems natural to wonder whether the physical encoding process is necesary. Can we not convey our intentions to the computer in a more direct fashion by skipping the physical layer entirely? Assuming we had some method of sensing and differentiating between two or more of the mental processes inside our brains, we could use these few states as a simple alphabet, from which a more complicated language of control could be built. The process of encoding intentions would take place entirely inside a person's head. For example, instead of thinking "I want that window to move to the foreground, so I will move the mouse to it and click," a user would think, "I want that window to move to the foreground, so I will think A,B,C,A," where A, B, C, A is a sequence of mental states recognized as signals by the computer, and which trigger a series of menu choices on a screen. A further question to ask is whether this mental encoding process is necessary, either. Could we not use the power of the computer to decode our intentions as we think them? Unfortunately, this question is beyond the scope of this paper. At present, we do not know enough about how the underlying brain processes encode thoughts to perform this task well.

Electroencephalogram (EEG) signals are an important source of information for studying the underlying brain processes that make up our thoughts and actions. EEG recorders with up to 256 electrodes are currently in use, and experiments with them produce large amounts of raw data. Thus, the question naturally arises: How much can we find out about the brain's activities from all that data? Can we read the mind of a person by properly deciphering their EEG? It has been known for some time that this is possible, to a modest extent. Recently, more than a few researchers have investigated using EEG signals as a new way of conveying intentions to a computer. EEGs produced during a very limited set of mental tasks can be classified according to tasks. Off-line analysis of multi-channel EEG has been performed by the groups of Wolpaw and Pfurtscheller since 1992, and more recently by Anderson's group.

In this article, I describe the methods and results of my experiments with EEG signals recorded from one subject while the subject performed five mental tasks. The EEG signals were processed by Independent Component Analysis to remove artifacts, and presented to three layer artificial neural networks using a frequency-band representation.

2 Related Work

Prior to acquiring my own EEG signals, I spent four months reviewing the available literature, and experimenting with the data used by Anderson, Devulapalli, and Stolz in [1]. This data was originally taken by Keirn and Aunon in [5], and, for this and other reasons, my work is heavily based on both [1] and [5]. Both [1] and [5] investigated the classification of five different mental tasks: a baseline resting task, mental multiplication, geometric figure rotation, mental letter composition, and visual counting.

In [5], Keirn and Aunon recorded data from seven subjects using six channels. They transformed the data into features based on spectral estimates calculated from both the Fourier transform of the windowed autocorrelation function and a scalar autoregressive (AR) model. Their features included asymmetry ratios and power values for each channel from four standard frequency bands: delta (1-3 Hz), theta (4-7 Hz), alpha (8-13 Hz), and beta (14-20 Hz). Asymmetry ratios given by

$$\frac{(R-L)}{(R+L)}$$

were calculated for all combinations of left to right electrodes and frequency bands, where R and L are power values from right and left hemisphere electrodes. A second set of features was made up of the AR coefficients from all channels, concatenated together. These features were extracted from a single quarter-second window per trial. Classification was performed with a quadratic Bayesian classifier, and it was found that all pairs of tasks could be discriminated 84.6% of the time using the AR coefficients as features.

In [1], Anderson classified EEG from one subject who performed five mental tasks. He used the 6-channel data sampled at 250 Hz from [5], and modeled signals in a half-second window as 6th order autoregressive processes. Through a cross-validation study of a variety of neural network topologies, he found that a single hidden layer of 20 input units performed best. Attempting to differentiate between five tasks resulted in a 31-54% classification accuracy. He also found that averaging the output of the network over consecutive inputs improved performance, and that by averaging over long periods of time, one could achieve of 100% classification accuracy. By averaging the output of the best performing network over 20 consecutive, overlapping time windows which amounted to 5 seconds of actual EEG data, Anderson was able to get a 33-71% classification accuracy.

In [2], Anderson, Devulapalli, and Stolz evaluated different signal representations. They used the data from [5], and presented 1,580 patterns of quarter-second data windows to neural networks with a single hidden layer, the size of which they varied from 1 to 40 nodes. Four different representations were tried: unprocessed, low-pass filtered, Karhunen-Loeve (K-L), and frequency-band. An unprocessed feature vector was formed by concatenating one quarter second (62 samples) of data from each of the six channels (372 values per pattern). The low-pass filtered representation was similar to this, except that an FIR low-pass filter with a cut-off frequency of 40 Hz was applied to the samples, followed by down-sampling by a factor of 2 (every other sample was removed). Thus, there were half as many features in the lowpass filtered vectors as there were in the unprocessed vectors. When classifying high-dimensional data, equivalent or better generalization accuracy is often achieved by classifying data obtained by projecting the original data onto the first *n* eigenvectors, where *n* is much smaller than the dimensionality of the original data [2]. The K-L representation was formed by projecting each 372-dimensional pattern onto the first 50 eigenvalues. The frequency-band representation used here was the same as that used in [5]. Anderson et al found that the frequency-band representation yielded the best results: 73.9% classification accuracy with 40 hidden nodes. The other representations had classification accuracies which hovered around 50%.

Although the network with 40 hidden nodes did outperform those with lesser numbers of node, it was not by much. In general, it helped very little to increase the number of hidden nodes, although it helped the frequency-band representation the most. This seems to indicate that the network which was trained using the frequency-band representation was learning the most.

In [17], Wolpaw and McFarland let four subjects teach themselves to control electrical potentials from two bipolar electrodes located around positions C3 and C4. The potentials were used to control a cursor on a screen, and the subjects had to move the cursor into predefined areas. After 20-30 training sessions, subjects could move the cursor into given corners of the screen with 41-70% success rate. For comparison, a random cursor movement would have a 25% success rate.

In [13], Peters, Pfurtscheller, and Flyvbjerg classified brain states corresponding to the intention of movement in the left and right index finger and right foot. This classification was done using a "committee" of artificial neural networks each processing an individual channel of a 56-electrode EEG. The size of the committee was chosen for optimal accuracy, and was found to be 9-12, confirming a previous result of Wolpaw and Pfurtscheller which was obtained differently. A classification rate of 83% was obtained when differentiating between two intented movements using a one second window of EEG. When the number of movements was increased to three and four, the classification accuracy dropped to around 70%.

For the experiment presented here, I used the mental states chosen in [5], and took my own 12-channel data. I used Independent Component Analysis to remove the artifactual components found in my EEG, and trained three-layer neural networks of different sizes on the Fourier transform of a varied window length of EEG. Using a one half second window of EEG, I obtained a classification accuracy of 94% when differentiating between the geometry and multiplication tasks. Using a 1/20th second window of EEG, I obtained a classification accuracy of 85% when differentiating between the geometry and multiplication tasks.

3 Method

3.1 EEG Signal Recording

The data for this experiment were taken using the neuroscience facilities at Pomona College, which are located in a small two story house that has been converted into a laboratory. The subject was seated in a closet with dim lighting, a comfortable chair, and a computer running the program used to present visual stimulae to the subject, NeuroScan Inc.'s STIM 2.0. The computer had internal fans which produced some slight noise, but other than that, the closet was fairly silent. A QuikCap-64 was used to record from positions FPZ, F3, FZ, F4, FCZ, C3, CZ, C4, PZ, P3, POZ, and P4, shown in Figure 1 and defined by the 10-20 system of electrode placement [19]. These twelve channels were referenced to electrically linked mastoids at M1 and M2. The impedance of all electrodes was kept below 20 Kohms. The data were recorded at a sampling rate of 250 Hz with a SynAmps Model 5038 EEG amplifier, which uses a 16 bit A/D converter. A serial cable connected the computer used to present the stimulae to the SynAmps EEG amplifier, and was used to signal when a stimulus was presented. The SynAmps was programmed to do analog bandpass filtering from 0.15-30Hz, and was calibrated with a known voltage before the recording session. Eye blinks were detected by means of a separate channel of data recorded from an electrode placed below the subject's left eye (VEOG).

Data were recorded from one subject, a 21-year-old, right-handed, male, college student. Five different programs of stimulus were presented using STIM, each displaying a total of ten images. The subject was given written instructions at the beginning of each program of ten images. In general, the instructions were to view an image related to a particular mental task, and to concentrate on the task, after hearing an audible tone, until the next image was presented. The programs of stimulus took the following format, shown in Figure 2:

- An image was presented for 5 seconds.
- A blank (dark) screen was presented for 5 seconds.
- A 1 KHz tone (beep) sounded.
- The blank screen continued for another 5 seconds.
- The next image in the program was presented.

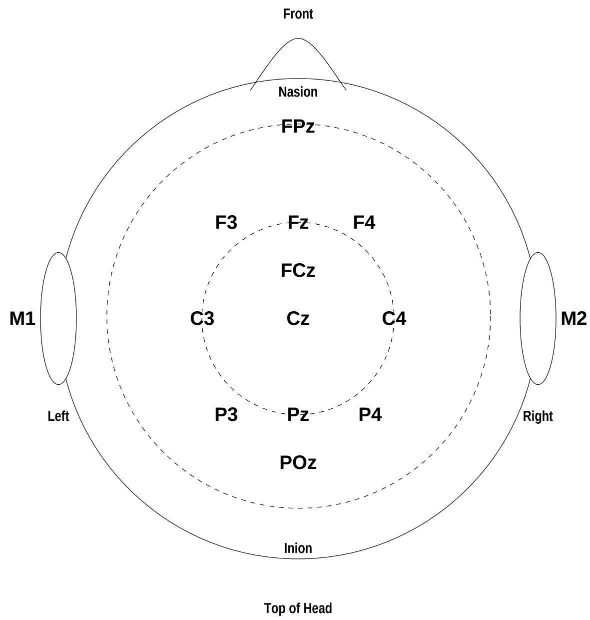

Figure 1: Electrodes were placed at positions FPZ, F3, FZ, F4, FCZ, C3, CZ, C4, PZ, P3, POZ, and P4.

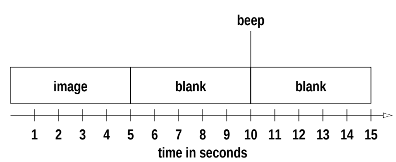

Figure 2: The format of the stimulus programs.

All tasks were performed with the subject's eyes open. The tasks used in this experiment are the same as those chosen by Keirn and Aunon in [5] to invoke hemispheric brainwave asymmetry. The five tasks were:

- **Baseline Task:** The instructions given to the subject preceding the stimulus program were not to perform a specific mental task, but to relax as much as possible, make as few movements as possible, and think of nothing in particular. This task is considered a baseline task for alpha wave production and was used as a control measure of the EEG. The ten images presented were all exactly the same, and consisted of a white screen.
- **Letter Task:** The subject was shown images consisting of a black word on a white background. Each word was indicative of a friend or family member (e.g., "father", "mother", "aunt", "uncle", etc.), and the subject was asked to mentally compose a letter to that person without vocalizing or making any physical movements.
- **Math Task:** The subject was shown images consisting of nontrivial multiplication problems, such as 89 times 67, and was asked to solve them without vocalizing or making any physical movements. The problems were designed so that they could not be solved in the time allowed. Although they repeated, the subject did not solve any of them to completion.
- **Geometric Figure Rotation:** The subject was shown images of three-dimensional figures (rendered and shaded), and asked to visualize them rotating about an axis. The figures were all three dimensional extrusions of randomly drawn two dimensional shapes.
- **Visual Counting:** The subject was shown an image of black arabic numerals on a white background, and asked to visualize similar numerals being written onto a blackboard, one after another, sequentially in ascending order, the previous numeral being erased before the next being written.

Data were recorded for 15 seconds per image, or 150 seconds per program. Thus, one recording of all five programs resulted in 750 seconds of data. The five programs were each recorded twice, giving us a total of 1500 seconds of data containing 100 beeps, each indicating the start of a five second period during which the subject was to be concentrating on a particular brain state.

After having some time to examine the data, it became apparent that CPZ, channel 14 on the Quik-Cap, had not been recording properly. However, all the other channels were quite clear, and since CPZ is along the center line, I noted that disgarding it would not have an adverse effect on the assymmetry ratios if I wanted to use them later on in the experiment. This left me with 11 channels of scalp data, and a single channel for detecting eye movements.

3.2 Artifact Removal

Contamination of EEG activity by eye movements, blinks, cardiac signals, and muscle and line noise is a serious problem for EEG interpretation and analysis. One way of dealing with this problem is to simply reject segments of EEG with unacceptable amounts of noise. However, this may result in an unacceptable amount of data loss. Fortunately, there are other, algorithmic alternatives to disgarding data. One algorithm in particular stands out from the rest: Independent Component Analysis. To understand what it does and why it serves our purpose, it is of use to gain some context with regard to the type of data we are dealing with when we analyze EEG.

The signals we hope to measure when we record the voltage potentials from a subject's scalp are those that result from the activity of neurons some significant distance away from the electrode we are using to take the measurement. Each electrode is "hearing" a aummation of all the neural activity in the vicinity. The signals differ from one another by virtue of the fact that they are located in different geographical areas of the scalp. Neural activity in an area which is close in proximity to one electrode will be "louder" in the recording produced by that electrode than one further from the source. Thus, in an ideal situation, each electrode would detect a unique linear mixture of all the neural activity happening in a subject's brain.

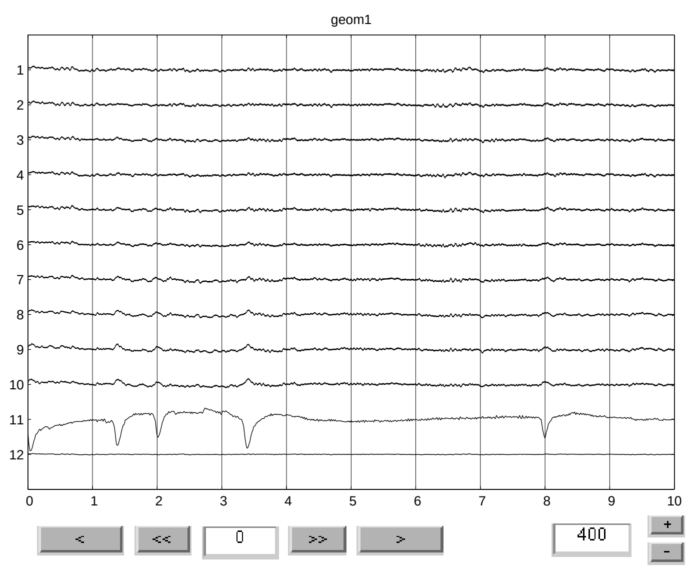

Figure 3: The first 10 seconds of the geometry task, before ICA processing.

Unforunately, the ideal linear mixture is augmented by other electrical activity which does not pertiain to neurons being fired. Typically, these noise signals are much greater in amplitude than the signals of interest, and have the effect of obliterating a good amount of useful information. Some of the noise signals, such as those resulting from eye blinks and other muscle movements, are infrequent enough such that the segments of data in which they appear can be simply disgarded without losing too much. Others, such as cardiac signals and eye movements, are regular enough to make obtaining useful data a cumbersome task. The problem of removing this noise form the interesting signal can be stated as follows: from N unique linear mixtures of an undetermined number of sources, can we somehow separate out N statistically independent mixtures? In other words, can we “unmix” the statistically unrelated noise onto a separate channel from the interesting signals? In fact, it has been known for some time that this is possible.

Figure 4: The first 10 seconds of the geometry task, after ICA processing.

Independent Component Analysis, proposed by Bell and Sejnowski in [10], is a simple neural algorithm that blindly separates mixtures of independent sources using infomax. In [10], they show that maximizing the joint entropy of the output of a neural processor minimizes the mutual information among the output components. Bell and Sejnowski offer the following two reasons for why ICA is suitable for performing blind source separation on EEG data: (1) it is plausible that EEG data recorded at multiple scalp electrodes are linear sums of temporally independent components arising from spatially fixed, distinct, or overlapping brain or extra-brain networks, and, (2) spatial smearing of EEG data by volume conduction does not involve significant time delays [21].

3.3 Representation of EEG Signals

The key to training a neural network to do a reliable discrimination is finding a suitable representation of the EEG signals. Since the early days of automatic EEG processing in the medical community, representations based on a Fourier transform have been most commonly applied to the problem of discriminating and classifying EEG patterns. This approach builds upon earlier observations that there are some characteristic waveforms that fall primarily within four frequency bands: delta (1-3 Hz), theta (4-7 Hz), alpha (8-13 Hz), and beta (14-20 Hz).

In related work, Anderson, Devulapalli, and Stolz [2] found that a frequency-band representation yielded the best results of four methods that they tried. Others [13] have had success with similar representations as well. I wanted to use ICA for artifact removal, but I realized that if I did so, I would not be able to use the asymmetry ratios, which had shown very positive results in the past, in my signal representation. This is because the sources computed by ICA do not have the same spatial relationship to the skull as do the signals derrived from the electrodes. However, I reasoned that although the asymmetry ratios certainly emphasized the differences between mental states in certain frequency bands, these differences would still be present in any frequency-band representation of the data, regardless of whether or not I precomputed them and presented them to the network explicitly.

Thus, I decided to use a representation based on the power spectral densities of the sources computed by ICA. With a sample rate of 250 Hz and 12 channels of data, each five second window of time during which the subject was to be concentrating on a particular brain state contained 15,000 data points. After computing the ICA sources and disgarding the one which was representative of eye and muscle movements, I was left with 11 channels of data.

Inside the period of concentration, I took ten windows of 11 channel EEG data, each offset by 50 samples from the one before it. For example, the first of the ten started at the beep, the second started 50 samples after the beep, etc. Since there were 100 windows per session, and 2 sessions of each mental state, each mental state was represented by 200 feature vectors. Of these features, half were used for training and half for validation.

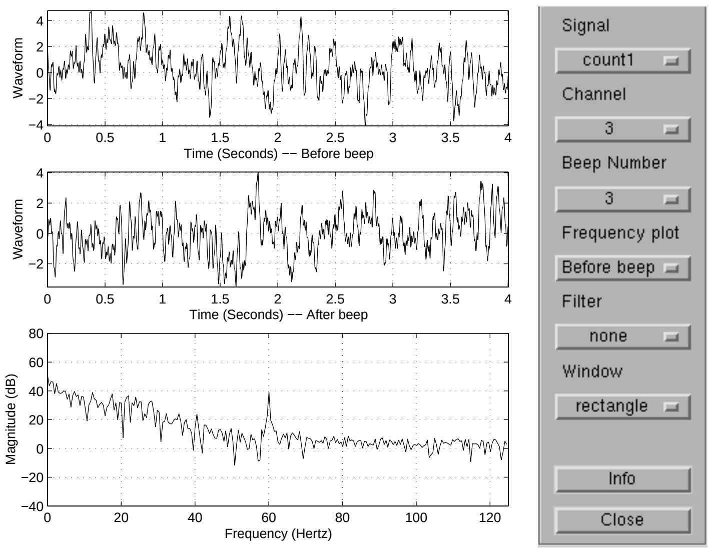

Figure 5: The third ICA source and its Fourier Transform around beep 3 of the counting task.

The length of the window was varied from one half second (125 samples) to one nineteenth of a second (12 samples). To point out what may not be obvious, windows longer than 50 samples overlapped other windows, while those shorter did not. For each window, I computed the Discrete Fourier Transform of each channel, which left me with 11 vectors containing a number of power values equal to the number of samples in the time domain. The power spectral density was computed by taking the dot product of the Fourier transform with its conjugate, and dividing the resulting vector by the number of power values in the Fourier transform. These 11 power spectral density vectors were concatenated together to form a presentation vector.

3.4 Pattern Classification

Three-layer feedforward artificial neural networks were trained using Matlab with the standard backprop algorithm (tbpx). The learning rate was dynamically adjusted as the network trained, such that as the standard squared error (SSE) of the network decreased, the learning rate increased, and when the SSE increased, the learning rate fell back down to a preset minimum. With regard to the number of hidden nodes in each hidden layer, a variety of thoughtfully selected configurations were tried. Some of the best performing configurations were 40-5, 50-10, 100-10, 250-50, 1000-100, where 40-5 indicates that there were 40 hidden nodes in the first layer, and 5 in the second. All networks were trained on Turing, the Harvey Mudd College Computer Science Department's six processor Sun Ultra Enterprise 3000 with 1.5 GB of memory.

Networks were trained to differentiate between all pairs of mental states, and all triples of mental states. One-hot encoding was used to enumerate mental states, and a "correct" classification was one in which the correct output was larger than all other outputs. Half of the total number of feature vectors were used as a validation set to prevent over-fitting, and training was stopped when the SSE of the validation set did not decrease for 200 epochs. Once this occurred, the network weights were reset to their values at the point when the network received the highest "score" upon evaluation of the validation set, where the “score” of the network was the number of validation patterns classified correctly. Each recording donated 100 patterns to the pool, and there were two recordings made of each brain state. Thus, for two-way differentiation 200 patterns were used for training, and 200 different patterns were used for validation. For three-way differentiation, 300 patterns were used for training, and 300 different patterns were used for validation.

4 Results

Table 1 shows the results of differentiation between two mental tasks. The notation used for the “Best Network” column indicates the number of hidden nodes in each hidden layer. For example, “40-5” indicates that there were 40 nodes in the first hidden layer, and 5 in the second. In the “Best Classification” column, the percentage indicates the number of test patterns classified correctly. Below the percentage is an indication of which mental tasks were used in the differentiation which resulted in that particular percentage. Clearly, some pairs of tasks are more easily differentiated between than others.

Table 1: Results for Two-way Differentiation

Window Length (seconds)	Best Network	Best Classification	Worst Classification	Training Time (minutes)
0.5 125	40-5	94% geom, mult	77% count, mult	7
0.25 62	40-5	90% base, letter	69% count, letter	6
0.125 31	100-10	82% base, count	69% count, letter	3
0.05 13	1000-100	85% geom, mult	67% count, letter	10

Table 2 shows the results of differentiation between three mental tasks. The classification accuracies indicate that this was a much harder problem for the network to learn. Another indication of the difficulty of the problem is that networks with a greater number of hidden nodes performed better. When only differentiating between two mental tasks, it was sufficient to use a lesser number of hidden nodes: after a certain threshold, classification accuracy did not increase if the number of hidden nodes was increased. After finishing with the results for three-way differentiation, it became clear that it would not be worth the training time to compute them for four-way differentiation.

Table 2: Results for Three-way Differentiation

Window Length (sec, samples)	Best Network	Best Classification	Worst Classification	Training Time (min)
0.5 125	100-10	86% base, letter, mult	71% count, letter, mult	30
0.25 62	250-50	77% count, geom, letter	63% count, geom, letter	16
0.125 31	250-50	74% geom, letter, mult	56% count, letter, mult	8
0.05 13	250-50	66% geom, letter, mult	51% count, letter, mult	6

5 Conclusion

Accurate, two-way differentiation can be done using a short window of EEG data. This is probably the most significant result of the experiment, because most applications for control systems have real-time requirements. For example, with a recognition rate of one symbol per second, it would be very difficult to steer a wheelchair or compose a letter on the computer. On the other hand, with a rate of sixteen symbols per second it might be possible to accomplish something.

Increasing the number of hidden nodes increases the accuracy of classification, although the increase in accuracy is very gradual after a point. Unfortunately, since larger networks take longer to train, there is a threshold at which the return in accuracy does not justify the investment in training time.

ICA is fast and useful for removing artifacts without discarding useful data. This experiment is one more verification that the ICA algorithm works as a method of removing artifacts from EEG data. Furthermore, the processing time required to run the ICA algorithm was insignificant compared to the time required to train the neural networks using backprop.

6 References

[1] C. W. Anderson, S. V. Devulapalli, and E. A. Stolz. Determining Mental State from EEG Signals Using Neural Networks. *Scientific Programming*, 4(3):171-183, Fall 1995.

[2] C. W. Anderson, S. V. Devulapalli, and E. A. Stolz. EEG Signal Classification with Different Signal Representations. *Neural Networks for Signal Processing V*, pages 475-483. IEEE Service Center, Piscataway, NJ, 1995.

[3] T. Fernandez, T. Harmony, M. Rodriguez, J. Bernal, J. Silva, A Reyes, and E. Marosi. EEG Activation Patterns During the Performance of Tasks Involving Different Components of Mental Calculation. *Electroencephalography and Clnical Neurophysiology*, 94:175-182, 1995.

[4] W. Finnoff, F. Hergert, and H. G. Zimmerman. Improving Model Selection by Nonconvergent Methods. *Neural Networks*, 6(6):771-783, 1993.

[5] Zachary A. Keirn and Jorge I. Aunon. A New Mode of Communication Between Man and His Surroundings. *IEEE Transactions on Miomedical Engineering*, 37(12):1209-1214, December 1990.

[6] Zachary A. Keirn. Alternative Modes of Communication Between Man and Machine. Master's thesis, Purdue University, 1988.

[7] E. Stolz. Multivariate Autoregressive Models for Classification of Spontaneous Electroencephalogram During Mental Tasks. Master's thesis, Electrical Engineering Department, Colorado State University, Fort Collins, CO, 1995.

[8] Patricia F. Harner and Theda Sannit. A Review of the International Ten-Twenty System of Electrode Placement. Grass Instrument Company, Quincy, MA, 1974.

[9] S. Amari, A. Cichocki, and H. Yang. A New Learning Algorithm for Blind Signal Separation. *Advances in Neural Information Processing Systems*, 8:757-763, 1996.

[10] A. J. Bell and T. J. Sejnowski. An Information-Maximization Approach to Blind Signal Separation and Blind Deconvolution. *Neural Computation*, 7:1129-1159, 1995.

[11] G. Pfurtscheller, D. Flotzinger, and J. Kalcher. Brain-Computer Interface: A New Communication Device for Handicapped Persons. *Journal of Microcomputer Applications*, 16(3):293-299, 1993.

[12] S-Y. Tseng, R-C. Chen, F-C. Chong, and T-S. Kuo. Evaluation of Parametric Methods in EEG Signal Ananlysis. *Med. Eng. Phys.*, 17:71-78, January 1995.

[13] Bjorn O. Peters, Gert Pfurtscheller and Henrik Flyvbjerg. Prompt Recognition of Brain States by their EEG Signals. 1997.

[14] Pfurtscheller, G., D. Flotzinger and C. Neuper. Differentiation Between Finger, Toe, and Tongue Movement in Man Based on 40 Hz EEG. Electroenc. Clin. Neurophys. 90, 456-460.

[15] Pfurtscheller, G., M. Pregenzer, and C. Neuper. Visualization of Sensorimotor Areas Involved in Preparation for Hand Movement Based on Classification of m and Central b Rhythms in Single EEG Trials in Man. Neurosc. Lett. 181, 43-46. 1994.

[16] Pregenzer, M., G. Pfurtscheller, and D. Flotzinger. Selection of Electrode Positions for an EEG-based Brain-Computer Interface. Biomed. Technik 39, 264-269. 1994.

[17] E. Wolpaw, J. R. and D. J. McFarland. Multichannel EEG-based Brain-Computer Communication. Electroenc. Clin. Neurophys. 90, 444-449. 1994.
[18] Nunez, P.L. and R. Katznelson. Electrical Fields of the Brain: The Neurophysics of EEG. New York: Oxford University Press. 1981.
[19] Harner, Patricia F. A Review of the International Ten-Twenty System of Electrode Placement. Quincy: Grass Instrument Company. 1974.
[20] T-P. Jung, C. Humphries, T-W. Lee, S. Makeig, M. J. McKeown, V. Iragui, and T. Sejnowski. Removing Electroencephalographic Artifacts: Comparison between ICA and PCA. Neural Networks for Signal Processing, VIII.
[21] S. Makeig, A. J. Bell, T-P. Jung, and T. J. Sejnowski. Independent Component Analysis of Electroenephalographic Data. Advances in Neural Information Processing Systems 8.
[22] A. J. Bell and T. J. Sejnowski. An Information-Maximization Approach to Blind Separation and Blind Deconvolution. Neural Computation, 7:1129-1159, 1995.