Improving the Performance of Two-state Mental Task Brain-Computer Interface Design Using Linear Discriminant Classifier

Ramaswamy Palaniappan, *Member, IEEE* and Nai-Jen Huan

***Abstract* — The purpose of this study is to motivate the use of the simpler Linear Discriminant (LD) classifier as compared to the commonly used Multilayer-perceptron-backpropagation (MLP-BP) neural network for Brain Computer Interface (BCI) design. We investigated the performances of MLP-BP and LD classifiers for mental task based BCI design. In the experimental study, EEG signals from five mental tasks were recorded from four subjects and the classification performances of different combinations of two mental tasks were studied for each subject. Two different AR models were used to compute the features from the electroencephalogram signals: Burg's algorithm (ARB) and Least Square algorithm (ARLS). The results showed that in most cases, LD classifier gave superior classification performance as compared to MLP-BP, with reduced computational complexity. However, the best mental tasks for each subject were the same using both classifiers. ARLS gave the best performance (93.10%) using MLP-BP and (97.00%) using LD. As the best mental task combinations varied between subjects, we conclude that for different subjects, proper selection of mental tasks and feature extraction methods would be essential for a BCI design.**

***Keywords* — Autoregressive, Electroencephalogram, Neural Network.**

I. Introduction

BRAIN Computer Interface (BCI) designs might give new lives to those individuals who are completely paralysed. With this motive, for the last ten years, the volume and pace of BCI research have grown tremendously [1]. In 1995, there were no more than six active BCI research groups, and in the year 2000, there were more than 20 [1]. BCI designs are very useful for completely paralysed individuals[1] to communicate with their external surroundings using their brain thoughts. These individuals could have become completely paralysed after being involved in an accident or caused by cerebral palsy, spinal cord injuries or due to some diseases such as amyotrophic lateral sclerosis, brainstem stroke and muscular dystrophies. BCI design is also suitable for use in simple hands off menu selection on the screen.

There are a few non-invasive methods for obtaining these brain signals to be utilised in a BCI design. Electroencephalogram (EEG) signals recorded at the scalp during some mental tasks have been used by some of the research groups [2-4]. Some others utilise single-trial visual evoked potential (VEP) signals where the subjects gaze at a screen full of alphabets or menus [5]. Synchronisation and desynchronisation of μ-rhythm extracted during sensory motor tasks is another method for BCI design [6]. Reviews of some of these technologies and developments in this area are given by Wolpaw *et al* [1].

Previous studies [3-4] on mental task BCI used Multilayer-perceptron-backpropagation (MLP-BP) neural network (NN), which is computationally expensive, so as an alternative, we propose the use of Linear Discriminant (LD) classifier in this paper.

We investigate the classification performances of a two-state BCI design using LD and MLP-BP NN classifiers. The classification performances using two autoregressive AR models: Burg (ARB) and Least Square (ARLS), used to compute the features from the EEG signals are also studied. The EEG signals were recorded during five different mental tasks from four different healthy subjects. These mental tasks were: geometrical figure rotation, mathematical multiplication, mental letter composing, visual counting and baseline-resting task. The designed BCI is individual BCI, which is suitable for use by a particular individual. Our results showed that we cannot expect to use a particular set of mental tasks for all the subjects due to different thought patterns that are obtained from different subjects.

The AR methods used order six. After these features are computed, LD and MLP-BP classifier is used to classify ten different combinations of two mental tasks. The output of the two-state BCI design could be used to control the movement of a cursor to select a target on a computer screen, to move a wheel chair or to be used with some translation schemes like Morse Code [4], which would provide a channel for paralysed individuals to communicate with their external surroundings.

As for application of the BCI design, a Morse Code scheme could possible be used to translate the outputs of BCI design into letter/words like 'water', 'tv', etc. For example, the letter 'A' in Morse Code is represented by a dot followed by a dash, so the user need to think of different mental tasks to 'spell' out the alphabet 'A'. For instance, if mental task 'letter' represents dot, 'maths'

R. Palaniappan is with the Dept. of Computer Science, University of Essex, Colchester, CO4 3SQ, United Kingdom. (tel: +44(0)-1206-872773; fax: +44(0)-1206-872788; e-mail: rpalan@essex.ac.uk, palani@iee.org).

N. Huan is with the Faculty of Information Science and Technology, Multimedia University, Melaka, 75450, Malaysia (e-mail: njhuan@mmu.edu.my).

[1] Individuals who have lost all forms of control over their peripheral nerves and muscles.

represents dash, the user thinks of letter mental task, followed by maths. By using the same concept, any BCI user could communicate with external world through 'thinking' of spelling out an alphabet instead of verbally spell it out.

II. DATA

The EEG data used in this study were collected by Keirn and Aunon [2]. Here, data from 4 subjects were used. The subjects were seated in an Industrial Acoustics Company sound controlled booth with dim lighting and noise-less fan (for ventilation). An Electro-Cap elastic electrode cap was used to record EEG signals from positions C3, C4, P3, P4, O1 and O2 (shown in Figure 1), defined by the 10-20 system [7] of electrode placement. The impedance of all electrodes were kept below 5 KΩ. Measurements were made with reference to electrically linked mastoids, A1 and A2. The electrodes were connected through a bank of amplifiers (Grass7P511), whose band-pass analog filters were set at 0.1 to 100 Hz. The data were sampled at 250 Hz with a Lab Master 12-bit A/D converter mounted on a computer. Before each recording session, the system was calibrated with a known voltage. Signals were recorded for 10s during each task and each task was repeated for 10 sessions where the sessions were held on different weeks. The EEG signal for each mental task was segmented into 20 segments with length 0.5 s. The sampling rate was 250 Hz, so each EEG segment was 125 data points in length.

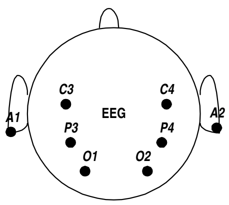

Fig. 1. Electrode placement.

In this paper, EEG signals from four subjects performing five different mental tasks were used. The data is available online at http://www.cs.colostate.edu/~anderson. These mental tasks were:

a) Baseline task (Bas). The subjects were asked to relax and think of nothing in particular. This task was used as a control and as a baseline measure of the EEG signals.
b) Math task (Mat). The subjects were given nontrivial multiplication problems, such as 42 times 18 and were asked to solve them without vocalising or making any other physical movements. The tasks were non-repeating and designed so that an immediate answer was not apparent. The subjects verified at the end of the task whether or not he/she arrived at the solution and no subject completed the task before the end of the 10 s recording session.
c) Geometric figure rotation task (Rot). The subjects were given 30 s to study a particular three-dimensional block object, after which the drawing was removed and the subjects were asked to visualise the object being rotated about an axis. The EEG signals were recorded during the mental rotation period.
d) Mental letter composing task (Let). The subjects were asked to mentally compose a letter to a friend or a relative without vocalising. Since the task was repeated for several times the subjects were told to continue with the letter from where they left off.
e) Visual counting task (Cnt). The subjects were asked to imagine a blackboard and to visualise numbers being written on the board sequentially, with the previous number being erased before the next number was written. The subjects were instructed not to verbalise the numbers but to visualise them. They were also told to resume counting from the previous task rather than starting over each time.

Keirn and Aunon [2] specifically chose these tasks since they involve hemispheric brainwave asymmetry (except for the baseline task). For example, it was shown by Osaka [8] that arithmetic tasks exhibit a higher power spectrum in the right hemisphere whereas visual tasks do so in the left hemisphere. As such, Keirn and Aunon [2] and later, Anderson *et al* [3] proposed that these tasks are suitable for brain-computer interfacing.

Keirn and Aunon [2] in their paper used 2 s segment that comprised of 512 data points (artifact-free EEG signals) from a total of 2500 sample points recorded in 10s for one session. Furthermore, first quarter second (64 data points) of each 2 s segment was also used for analysis. However, in this paper, we used all the 2500 sample points recorded for 10s without any form of filtering. This saves computational time and design cost.

The studies in [3, 4] were more similar to the work here in terms of EEG data but used MLP-BP NN only for classification and also differed in some of the feature extraction methods.

III. FEATURE EXTRACTION

A real valued, zero mean, stationary, autoregressive process of order p is given by

$$x(n) = -\sum_{k=1}^{p} a_k x(n-k) + e(n), \tag{1}$$

where p is the model order, $x(n)$ is the signal at the sampled point n, a_k are the real valued AR coefficients and $e(n)$ represents the error term independent of past samples. The term autoregressive implies that the process $x(n)$ is seen to be regressed upon previous samples of itself. The error term is assumed to be a zero mean noise with finite variance. In applications, the values of a_k have to be estimated from finite samples of data $x(1), x(2), \ldots, x(N)$.

In this paper, we used Burg's method [9] to estimate the AR coefficients. The method is more accurate as compared to other methods like Levinson-Durbin as it uses the data point directly. Furthermore, Burg algorithm uses more data points by minimising both forward error and backward error.

In computing AR coefficients, order six was used because other researchers [2-4] have suggested the use of order six for AR process for mental task classification. Therefore, we had six AR coefficients for each channel, giving a total of 36 features for each EEG segment for a mental task. Next, we used the LS algorithm as proposed in [10] in addition to the Burg's method to estimate the AR coefficients and we computed the 36 features.

IV. LD CLASSIFIER

LD classifier is one of the linear classification methods that require fewer examples in order to obtain a reliable classifier output [11]. It is also a simpler and computationally attractive as compared to other classifiers. LD was used to classify different combinations of mental task pairs represented by the different EEG feature vectors.

A brief description of the LD method used in this paper is as follows [11]. Assume that each data element S_i has m features and the number of examples is n where each example is assigned to one of the two classes $C= \{0, 1\}$. Then S is a matrix of size n x m, and C is a vector of size n. Assume N_0 and N_1 are the number of elements for class 0 and 1, respectively.

The mean μ_c of each class C is the mean over all S_i with i being all elements with in class C. The total mean μ of the data is calculated as

$$\mu = \frac{N_0\mu_0 + N_1\mu_1}{N_0 + N_1} \tag{2}$$

The covariance matrix, Cv of the data is the expectation value computed as

$$Cv = E < (s-\mu)^T (s-\mu) > \tag{3}$$

Then, the weight vector w and the offset w_0 are

$$w = Cv^{-1}(\mu_1 - \mu_0)^T \quad \text{and} \quad w_0 = -\mu w \tag{4}$$

The weight vector w determines a separating hyperplane in the m-dimensional feature space. The normal distance $D(x)$ of an element x is computed as

$$\begin{aligned} D(x) &= xw + w0 \\ &= (x - \mu)w \\ &= (x-\mu)C^{-1}(\mu_1 - \mu_0)^T \end{aligned} \tag{5}$$

If $D(x)$ is bigger than 0, x is assigned to class 1 while if $D(x)$ is smaller than 0, x is assigned to class 0. However, if $D(x)$ = 0, it means all elements of x are part of the separating hyperplane.

A total of 200 EEG patterns (20 segments for each EEG signal x 10 sessions) were used for each subject for each mental task in this experimental study. Therefore, for each simulation, there were 400 EEG patterns from two mental tasks, where half of the patterns were used in training and the remaining half in testing. The selection of the patterns for training and testing were chosen randomly.

V. MLP-BP NN

In this paper, MLP-BP NN [12] was used in addition to LD classifier to compare the classification performances. Figure 2 shows the architecture of the MLP-BP NN used in this study. The output nodes were set at two so that the NN could classify into one of the two categories representing the mental task. The hidden layer nodes were varied from 20 to 100 in steps of 20 but only the maximums of these are reported here due to space constraints. The maximum value is used to ensure a lower bias with comparing the classification performances with LD.

Training was conducted until the average error fell below 0.01 or reached a maximum iteration limit of 2000. The average error denotes the error limit to stop NN training. The average error is the average of NN target output subtracted by the desired target output from all the training patterns. The desired target output was set to 1.0 for the particular category representing the mental task of the EEG pattern being trained, while for the other category, it was set to 0.

VI. RESULTS AND DISCUSSION

Tables 1- 4 show the classification results for different combinations of mental task pairs for subjects 1,2,3,4, respectively. The coefficient variation (CV) i.e. the mean/standard deviation measures the spread of variation in the classification accuracies.

TABLE 1: CLASSIFICATION RESULTS FOR SUBJECT 1

Mental Task Combination	MLP-BP NN AR Method ARB	MLP-BP NN AR Method ARLS	MLP-BP NN Best AR Method	LD AR Method ARB	LD AR Method ARLS	LD Best AR Method
Bas, Cnt	78.40	80.60	ARLS	79.50	82.50	ARLS
Bas, Let	74.50	72.90	ARLS	76.50	75.50	ARB
Bas, Mat	75.50	74.70	ARB	81.00	83.00	ARLS
Bas, Rot	76.40	76.20	ARLS	80.50	84.00	ARLS
Let, Cnt	64.20	64.80	ARLS	65.50	68.00	ARLS
Let, Rot	67.40	71.60	ARLS	78.00	74.00	ARB
Mat, Cnt	81.50	79.70	ARLS	86.50	86.00	ARB
Mat, Let	79.00	81.00	ARLS	86.50	85.50	ARB
Mat, Rot	82.30	86.50	ARLS	87.00	89.00	ARLS
Rot, Cnt	71.40	69.40	ARB	76.00	74.00	ARB
Ave	75.06	75.74	ARLS	79.70	80.15	ARLS/ARB
CV	0.08	0.08		0.08	0.08	
Max	82.30	86.50		87.00	89.00	
Best Combination	Mat, Rot	Mat, Rot		Mat, Rot	Mat, Rot	

TABLE 2: CLASSIFICATION RESULTS FOR SUBJECT 2

Mental Task Combination	MLP-BP NN AR Method ARB	MLP-BP NN AR Method ARLS	MLP-BP NN Best AR Method	LD AR Method ARB	LD AR Method ARLS	LD Best AR Method
Bas, Cnt	71.30	68.80	ARB	68.00	65.50	ARB
Bas, Let	80.60	79.90	ARB	81.50	80.50	ARB
Bas, Mat	74.50	75.30	ARLS	81.00	80.50	ARB
Bas, Rot	70.90	70.50	ARB	74.00	67.50	ARB
Let, Cnt	88.80	88.00	ARB	87.00	83.00	ARB
Let, Rot	73.90	75.30	ARLS	73.50	72.50	ARB
Mat, Cnt	72.95	71.70	ARB	74.50	70.50	ARB
Mat, Let	74.30	73.00	ARB	72.00	75.50	ARLS
Mat, Rot	68.85	70.60	ARLS	66.50	68.00	ARLS
Rot, Cnt	68.00	65.60	ARB	72.00	72.00	ARB
Ave	74.41	73.87	ARB	75.00	73.55	ARB
CV	0.08	0.09		0.08	0.08	
Max	88.80	88.00		87.00	83.00	
Best Combination	Let, Cnt	Let, Cnt		Let, Cnt	Let, Cnt	

TABLE 3: CLASSIFICATION RESULTS FOR SUBJECT 3

Mental Task Combination	MLP-BP NN AR Method ARB	MLP-BP NN AR Method ARLS	MLP-BP NN Best AR Method	LD AR Method ARB	LD AR Method ARLS	LD Best AR Method
Bas, Cnt	57.60	59.90	ARLS	64.50	67.00	ARLS
Bas, Let	59.85	61.30	ARLS	71.00	68.00	ARB
Bas, Mat	67.85	67.90	ARLS	70.50	68.00	ARB
Bas, Rot	74.25	73.40	ARB	77.00	75.00	ARB
Let, Cnt	61.85	64.50	ARLS	63.50	64.50	ARLS
Let, Rot	77.30	74.90	ARB	81.00	82.00	ARLS
Mat, Cnt	63.80	62.50	ARB	65.50	63.50	ARB
Mat, Let	71.15	68.20	ARB	73.50	76.00	ARLS
Mat, Rot	71.55	73.00	ARLS	76.50	76.00	ARB
Rot, Cnt	69.40	69.90	ARLS	73.50	74.00	ARLS
Ave	67.46	67.55	ARLS	71.65	71.40	ARB
CV	0.10	0.08		0.08	0.08	
Max	77.30	74.90		81.00	82.00	
Best Combination	Let, Rot	Let, Rot		Let, Rot	Let, Rot	

TABLE 4: CLASSIFICATION RESULTS FOR SUBJECT 4

Mental Task Combination	MLP-BP NN AR Method ARB	MLP-BP NN AR Method ARLS	MLP-BP NN Best AR Method	LD AR Method ARB	LD AR Method ARLS	LD Best AR Method
Bas, Cnt	76.40	76.40	ARLS	79.50	82.50	ARLS
Bas, Let	63.30	65.30	ARLS	68.50	67.50	ARB
Bas, Mat	86.10	84.90	ARB	89.50	91.50	ARLS
Bas, Rot	80.95	81.80	ARLS	85.00	86.50	ARLS
Let, Cnt	71.35	71.80	ARLS	81.50	81.00	ARB
Let, Rot	72.55	74.10	ARLS	77.50	79.50	ARLS
Mat, Cnt	92.70	93.10	ARLS	95.50	97.00	ARLS
Mat, Let	78.10	78.20	ARLS	85.00	86.00	ARLS
Mat, Rot	76.20	72.70	ARB	84.50	81.50	ARB
Rot, Cnt	73.75	77.50	ARLS	83.50	89.50	ARLS
Ave	77.14	77.58	ARLS	83.00	84.25	ARLS
CV	0.11	0.10		0.09	0.09	
Max	92.70	93.10		95.50	97.00	
Best Combination	Mat, Cnt	Mat, Cnt		Mat, Cnt	Mat, Cnt	

It could be seen that for each subject, the best mental task pair was different. However, this best mental task pair was independent of the classifier or AR feature extraction method. In terms of AR feature extraction method, it is difficult to generalise as ARLS gave the best performance for some subjects, while for some others, it was ARB.

For subjects 1, 3 and 4, LD gave better performance that MLP-BP when considering the best classification performance, while for subject 2, MLP-BP gave slightly better performance. Therefore, we propose that LD be used for future work involving classification of mental task pairs as LD is computationally less complex. As a simple and convenient comparison of complexity, we decided to use the computation time as the experiments were run on the same computer and programming software. LD required 40 μs on average to train and test an EEG pattern, while MLP-BP required more than 120.8 ms on average to train and 80 μs on average to test an EEG pattern.

VII. CONCLUSION

In this paper, we have used LD and MLP-BP NN to classify mental tasks using AR features from EEG signals. We compared the LD and MLP-BP NN classification performances for each subject. Our results showed that ARLS gave the best classification performances for some subjects, while it was ARB for some others. Besides this, the results indicated that the best mental task pair was invariant to classifier and feature extraction method. In general, LD classifier gave better classification performances as compared to MLP-BP NN. Therefore, we suggest the use LD classifier due to its often superior performance, lower computational complexity and faster response. Finally, we draw the conclusion that different subjects have different combination of best mental tasks and proper selection of the best mental tasks is essential for a successful BCI design.

ACKNOWLEDGMENT

The authors would like to acknowledge the assistance of Dr. C. Anderson of Colorado State University, USA for giving permission to use the EEG data.

REFERENCES

[1] Wolpaw, J.R., Birbaumer, N., Hectderks, W.J., McFarland, D.J., Pecleham, P.H., Schalk, G., Donchin, E., Quatrano, L.A., Robinson, C.J., Vaughan, T.M. "Brain-Computer Interface Technology: A Review of the First International Meeting," *IEEE Transactions on Rehabilitation Engineering*, vol. 8 no. 2, pp. 164-173, June 2000.

[2] Keirn, Z.A., and Aunon, J.I., "A new mode of communication between man and his surroundings," *IEEE Transactions on Biomedical Engineering*, vol. 37, no.12, pp. 1209-1214, December 1990.

[3] Anderson, C.W., Stolz, E.A., and Shamsunder, S., "Multivariate autoregressive models for classification of spontaneous electroencephalogram during mental tasks," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 3, pp. 277-286, 1998.

[4] Palaniappan, R., Paramesran, P., Nishida, S., and Saiwaki, N., "A New Brain-Computer Interface Design Using Fuzzy ARTMAP," *IEEE Transactions on Neural System and Rehabilitation Engineering*, vol. 10, pp. 140-148, Sept. 2002.

[5] Donchin, E., Spencer, K.M., and Wijesinghe, R., "The mental prosthesis: assessing the speed of a P300-based brain-computer interface," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 174-179, June 2000.

[6] Pfurtscheller, G., Neuper, C., Guger, C., Harkam, W., Ramoses, H., Schlogl, A., Obermaier, B., Pregenzer, M., "Current trends in Graz brain-computer interface (BCI) research," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 216-219, June 2000.

[7] Jasper, H., "The ten twenty electrode system of the international federation," *Electroencephalographic and Clinical Neurophysiology*, vol.10, pp. 371-375, 1958.

[8] Osaka, M., "Peak alpha frequency of EEG during a mental task: task difficulty and hemispheric differences," Psychophysiology, vol. 21, pp. 101-105, 1984.

[9] Shiavi, R., *Introduction to Applied Statistical Signal Analysis*, 2^{nd} *edition*, Academic Press, 1999.

[10] Neumaier, A., Schneider, T., "Estimation of parameters and eigenmodes of multivariate autoregressive models," *ACM Transactions on Mathematical Software*, vol. 27, no.1, pp. 27-57, 2001.

[11] Schlogl, A., "A New Linear Classification Method for an EEG-based Brain-Computer Interface," *Technical report MDBC*, Austria, 2001.

[12] Rumelhart, D.E., and McCelland, J.L., *Parallel Distributed Processing: Exploration in the Microstructure of Cognition*, MIT Press, vol. 1, 1986.