International Journal of Medical, Medicine and Health Sciences
ISSN: 2517-9969
Vol:1, No:12, 2007

Discrimination of Alcoholic Subjects using Second Order Autoregressive Modelling of Brain Signals Evoked during Visual Stimulus Perception

Ramaswamy Palaniappan

***Abstract*—**In this paper, a second order autoregressive (AR) model is proposed to discriminate alcoholics using single trial gamma band Visual Evoked Potential (VEP) signals using 3 different classifiers: Simplified Fuzzy ARTMAP (SFA) neural network (NN), Multilayer-perceptron-backpropagation (MLP-BP) NN and Linear Discriminant (LD). Electroencephalogram (EEG) signals were recorded from alcoholic and control subjects during the presentation of visuals from Snodgrass and Vanderwart picture set. Single trial VEP signals were extracted from EEG signals using Elliptic filtering in the gamma band spectral range. A second order AR model was used as gamma band VEP exhibits pseudo-periodic behaviour and second order AR is optimal to represent this behaviour. This circumvents the requirement of having to use some criteria to choose the correct order. The averaged discrimination errors of 2.6%, 2.8% and 11.9% were given by LD, MLP-BP and SFA classifiers. The high LD discrimination results show the validity of the proposed method to discriminate between alcoholic subjects.

***Keywords*—**Linear Discriminant, Neural Network, Visual Evoked Potential.

I. Introduction

DIGITAL spectral analysis using autoregressive (AR) models have proven to be superior to classical Fourier transform techniques due to the ability of AR models to handle short segments of data, while giving better frequency resolution and smoother power spectra than Fourier methods. Furthermore, AR methods need only one or more cycles of sinusoidal-type activity to be present in the segment to produce good spectral peaks and they also provide the ability to observe small shifts in peak frequencies, which are not easily observed with Fourier derived spectra [1].

AR models are more popular than the other linear parametric models like moving average (MA) and autoregressive moving average (ARMA) due to their inherent computational efficiency [2]. The AR model coefficients can be easily estimated using recursive methods like Levinson-Durbin [3] or Burg [4]. In addition, AR coefficients can be efficiently updated when new data becomes available through the use of Kalman filter equations. On the other hand, MA and ARMA require complicated procedures to estimate the model coefficients [2].

AR models have been used in a broad spectrum of applications ranging from identification, prediction and control of dynamical systems and digital spectral analysis including analysis of biomedical signals like electroencephalogram (EEG) [5], [6] and Visual Evoked Potential (VEP) [7], [8].

VEP is typically generated in response to external visual stimulus. This electrical signal consists of the activity of an ensemble of neuronal generators producing rhythmic activity in several frequency ranges. These activities are normally random, however with the application of sensory stimulus like visually seeing a set of pictures, these generators are coupled and act in a coherent manner. Synchronisation of this activity gives rise to VEP and its analysis has become very useful for neuropsychological studies and clinical purposes [7]-[9].

In this paper, the goal is to use optimal AR model (i.e. the second order) to model single trial gamma band VEP signals for discrimination of alcoholic subjects using three different classifiers: Simplified Fuzzy ARTMAP (SFA) neural network (NN), Multilayer-perceptron-backpropagation (MLP-BP) NN and Linear Discriminant (LD). Previous works in classification of alcoholics and controls have used VEP signal energy after some filtering [7], [8].

A second order is proposed as the optimal order for AR modelling because of the pseudo-periodic property exhibited by the VEP signals in gamma band. Optimal here means a model with low order that does not compromise the discrimination performance. Lower order means faster computation and simpler design solutions. The suitability of the second order is proved by the experiments, which show that VEP patterns exhibit pseudo-periodic behaviour, which can be optimally modelled using second order AR.

II. Autoregressive Systems

A real valued, zero mean, stationary, non-deterministic, AR model of order p is given by

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

Manuscript received June 16, 2005.

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

where p is the model order, $x(n)$ is the data of the signal at sampled point n, a_k are the real valued AR coefficients and $e(n)$ represents the error term independent of past samples. The error term is assumed to be a zero mean white noise with finite variance, σ_e^2. In applications, the values of a_k and σ_e^2 have to be estimated from finite samples of data $x(1), x(2), x(3), \ldots\ldots, x(N)$.

Many different techniques have been proposed to estimate a_k [10]. The most common method is to use the autocorrelation technique of solving the Yule-Walker equations [3] but a shortcoming of this approach lies in its huge computational time. Thus, recursive algorithms have been developed which are based on the concept of estimating the parameters of a model of order p from the parameters of a model of order $p-1$. Some of these methods are like Burg's algorithm [4] and Levinson–Durbin algorithm [3].

Burg's method is more accurate than Levinson-Durbin since it uses the data points directly unlike the latter method, which relies on the estimation of the autocorrelation function, which is generally erroneous for small data segments. The earlier method also uses more data points simultaneously by minimising not only a forward error (as in the Levinson-Durbin case) but also a backward error.

Burg's method is common is AR literatures and as such, only a brief discussion of the algorithm will be given here. The steps are

Step 1. Calculate initial values

- Error variance, $\hat{\sigma}_e^2(0) = \frac{1}{N}\sum_{0}^{N-1}[x(n)]^2$ where $x(n)$ is the n^{th} sampled data with mean value subtracted
- Forward error, $e_n(0) = x(n)$
- Backward error, $b_{n-1}(0) = x(n-1)$

Step 2. Calculate reflection coefficient and error variance

- Reflection coefficient,

$$\pi_m = -2\frac{\sum_{n=m}^{N-1} b_{n-1}(m-1)e_n(m-1)}{\sum_{n=m}^{N-1}[e_n^2(m-1) + b_{n-1}^2(m-1)]}$$

- Error variance,

$$\sigma_e^2(m) = [1-|\pi_m|^2]\sigma_e^2(m-1)$$

Step 3. Update Error and AR coefficients

- AR coefficients,

$$\left.\begin{aligned} a_k(m) &= a_k(m-1) + \pi_m a_{m-k}(m-1) \\ a_m(m) &= \pi_1 \end{aligned}\right\} \begin{aligned} m &> 1 \\ m &= 1 \end{aligned}$$

- Forward Error Update,

$$e_n(m) = e_n(m-1) + \pi_m b_{n-1}(m-1)$$

- Backward Error Update,

$$b_n(m) = b_{n-1}(m-1) + \pi_m e_n(m-1)$$

Step 4. Repeat steps 2 and 3 (with m incremented by one) until the selected model order p is reached.

Proofs and details of this algorithm can be found in [4], [10].

These AR coefficients are then used to obtain the power spectral density (PSD) values by using the equation [10]

$$S(f) = \frac{\hat{\sigma}_p^2 T}{|\sum_{k=0}^{p} a_k e^{-i2\pi fkT}|^2}, \quad (2)$$

where $S(f)$ represents the PSD function, T is the sampling period and $\hat{\sigma}_e^2(p)$ is the unbiased estimated variance of the residuals.

III. Second-order AR Model for Pseudo-Periodic VEP

A second-order AR process may be written as

$$x(n) = a_1 x(n-1) + a_2 x(n-2) + e(n). \quad (3)$$

The AR characteristic equation is given by

$$x(B) = 1 - a_1 B + a_2 B^2, \quad (4)$$

where the backshift operator, $B^t = x(n-t)$. The roots of (4) can be found by setting $x(B)=0$. The reciprocals of these roots, G_1 and G_2 are

$$G_{1,2} = -2a_1\left[[a_1 \pm \sqrt{a_1^2 + 4a_2}\,]\right]^{-1}. \quad (5)$$

The autocorrelation (AC) function is now given as

$$R(k) = a_1 R(k-1) + a_2 R(k-2), \text{ where k=1,2,3……} \quad (6)$$

If the roots of (6) are real and different, it can be shown that the AC function of the second-order AR model consists of a mixture of damped exponentials. The positive root causes AC to remain positive as its decays exponentially and the negative root causes AC to alternate sign while decaying.

If the roots are complex, the AC function displays damped sinusoidal behavior, which denotes that the AR represented time series exhibits periodic behavior. The AC function is now given by

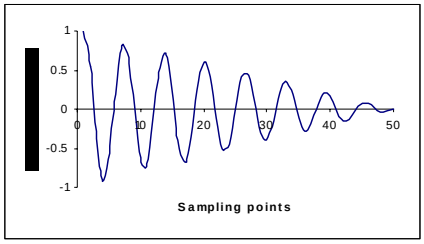
$$R(k) = \frac{A^k \sin(2\pi f_0 k + \theta)}{\sin\theta}, \text{ where k=0,1,2,3……} \quad (7)$$

The parameters in (7) are damping factor, $A = \sqrt{|a_2|}$; dominant frequency, $f_0 = \tan^{-1}[\sqrt{4a^2 - (a_1)^2}\,/\,a_1] * f_S$ with f_s as the sampling frequency and phase, $\theta = \tan^{-1}[(1+a_2)/(1-a_2)]\tan 2\pi(f_0 / f_S)$.

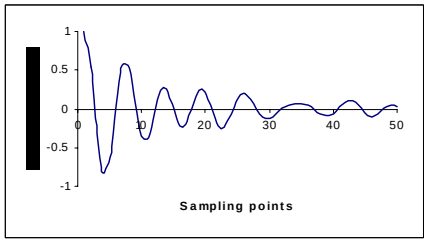
Figure 1(a) shows an example of AC damped sinusoidal behavior given by (7). Notice that the period for the plot is approximately 6.4 data points, which corresponds to the actual value of f_o=40 Hz with f_s=256 Hz. Calculation of f_o from (7) gives 39.9 Hz, which is close to the actual value. This analysis shows that periodic time signals can be suitably represented by second-order AR model. However, for pseudo-periodic

signals, calculation of f_o from (7) is not very accurate and therefore power spectral density (PSD) analysis is required.

The term pseudo-periodic is used here because in practical applications, the AC periodicity is only approximate as shown in Figure 1(b) for an extracted VEP segment.



(a)



(b)

Fig. 1 Damped sinusoidal autocorrelation function (a) theoretical (b) extracted VEP segment

IV. Single Trial VEP Extraction

There is a major problem encountered in analysing VEP signals, which comes from the contamination of spontaneous background EEG brain activity, which is many times higher in amplitude as compared to VEP signals. The predominant method of extracting VEP signal is to use signal averaging from a certain number of VEP signals [11]. However, there are numerous problems associated with this method like the variation in latency and amplitude for a similar stimulus across different sessions even for the same subject and the difficulty in analysing single trial VEP cannot be addressed by signal averaging alone.

In this paper, EEG contamination is avoided by using VEP signals in the gamma band range. In this method, the requirement of having to increase the signal-to-noise ratio (SNR) of VEP to background EEG by signal averaging is removed. The method relies on the assumption that gamma band spectrum is evoked during visual stimulus [9]. Since EEG activity is band-limited from 0 to 30 Hz, a high-pass filter that cuts off signals with frequencies below this range will suffice to separate VEP from EEG. That is because gamma band (>30 Hz) is beyond the normal EEG spectral range, technique like high-pass filtering is sufficient to obtain the gamma band VEP component from the EEG signal.

Here, these VEP signals were high pass filtered using a 5^{th} order Elliptic digital filter with a 3-dB cut-off frequency at 30 Hz. Order 5 was used since it is sufficient to give a minimum attenuation of 30dB in the stop band with a transition band from 30 to 35 Hz. Elliptic filter was selected as this filter requires lower order than other IIR filters like Butterworth. Forward and reverse filterings were performed to achieve zero phase response i.e. to avoid any phase distortion because Elliptic filter is a non-linear filter. First, the filtering was done in the forward direction, then the filtered sequence was reversed and run back through the filter. The result has precisely zero phase distortion and magnitude modified by the square of the filter's magnitude response. Care was taken to minimise startup and ending transients by matching initial conditions. The ripple in the passband was kept below 0.5 dB.

V. Classifiers

A. LD

LD classifier [12] is a linear classification method that is computationally attractive as compared to other classifiers like artificial neural network. It could be used to classify two or more groups of data. Here, LD was used to discriminate the VEP feature vectors into one of the two categories (alcoholic and control). The classify function in MATLAB (Mathworks Inc.) with linear distance measure was used as it was assumed that the distribution of the VEP feature vectors to be multivariate normal density with similar covariance structure.

In principle, any mathematical function may be used as a discriminating function. In case of the LD, the VEP training feature vectors were used to derive the linear discriminant functions as

$$F = \sum_{i=1}^{N} x_i w_i + a \ , \tag{8}$$

where x_i was the set of AR coefficients from the VEP feature vectors, N was the number of features, w_i and a were the coefficients and constant, respectively. The discriminating function was formed in such a way that the separation (i.e. distance) between the groups was maximised, and the distance within the groups was minimised i.e. the parameters w_i and a have to be determined in such a way that the discrimination between the groups was best. In other words, in the feature space with the dimensions equal to the number of features, linear planes were introduced to divide the data into different groups. Using these discriminant functions, the discriminant scores of each test VEP feature vector occurring in each of the groups were computed. The test VEP feature vector was then assigned to the group with the highest score and then compared with the actual group to determine the classification error.

B. MLP-BP NN

MLP-BP NN [13] was used in addition to LD classifier to compare the discrimination 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 (alcoholic and control). The number of hidden nodes was set at 20, which gave the best results after some preliminary simulations.

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 represented by the VEP feature vector, while for the other category, it was set to 0.

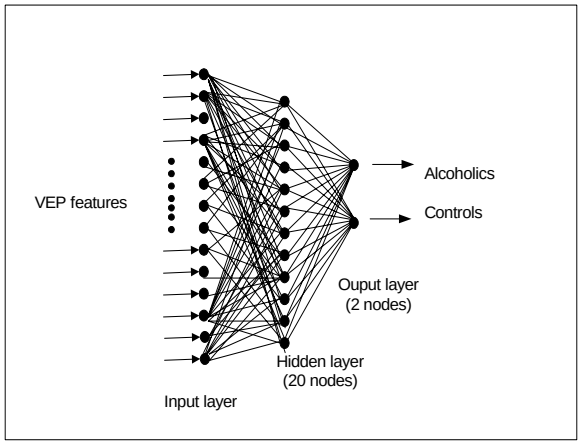


Fig. 2 MLP-BP network as used in the study

C. SFA

These VEP feature vectors were also discriminated by SFA. SFA was chosen for comparison due to its high speed training ability in fast learning mode. SFA is a type of neural network that performs incremental supervised learning [14]. It consists of a Fuzzy ART module linked to the category layer through an Inter ART module.

During training (supervised learning), Fuzzy ART receives a stream of input features representing the pattern and the output classes in the category layer are represented by a binary string with a value of 1 for the particular target class and values of 0 for all the rest of the classes.

Inter ART module will create mappings between the Fuzzy ART output to either the alcoholic or control category. For all the input patterns presented, it creates a dynamic weight link that consists of a many to one or one to one mapping between the output layer F_2 of Fuzzy ART and category layer.

Inter ART module works by increasing the vigilance parameter (VP), ρ of Fuzzy ART by a minimal amount to correct a predictive error at the category layer. Parameter ρ calibrates the minimum confidence that Fuzzy ART must have in an input vector in order for Fuzzy ART to accept that category, rather than search for a better one through an automatically controlled process of hypothesis testing. Lower values of ρ enable larger categories to form and lead to a broader generalisation and higher code compression.

The testing stage works similar to the training stage except that there will be no match tracking. This is because the input presented to Fuzzy ART will output a category in layer F_2, which will be used by the Inter ART module to trigger the corresponding category layer node that refers to the predicted class. Figure 3 shows the SFA network architecture as used in the experimental study. For further details on SFA, refer to [14].

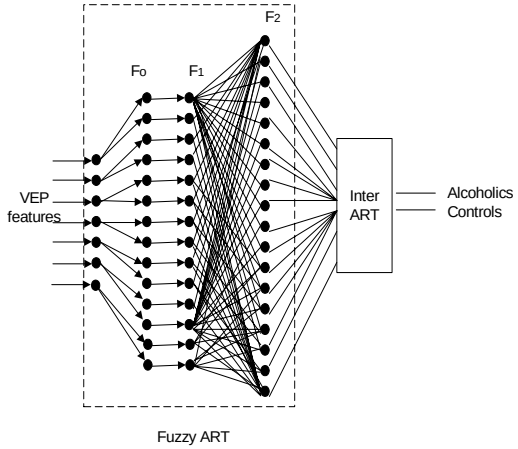


Fig. 3 SFA network as used in the study

VI. Experimental Study

As mentioned earlier, the objective of the study was to discriminate alcoholics from controls using gamma band VEP signals represented by second order AR model.

The alcoholics were significantly older than the controls [t(118.9)=12.64, p=0.0001]. The mean age for the control group was 25.81 years old (SD=3.38) ranging from 19.4 to 38.6 years of age. The mean age of alcoholic group was 35.83 (SD=5.33), ranging from 22.3 – 49.8 years. The alcoholics tested had been abstinent for a minimum period of one month (through closed ward detention). Therefore, all alcoholics were fully detoxified and had no alcohol available for that period of hospitalisation. Alcoholic individuals were excluded from the study if they had history of drug dependence, major psychiatric illness, or other diseases related to overt liver, metabolic, vascular and neurological. Most of the alcoholics had been drinking heavily for a minimum of 15 years. The diagnosis of alcohol abuse was made by the intake psychiatrist of the Addictive Disease Hospital in Brooklyn according to DSM-III criteria. The alcoholics were non-amnesics. The controls were carefully matched for age and were not alcoholics or substance abusers. They were also matched for socioeconomic status.

Measurements were taken for one second from 64 electrodes placed on the subject's scalp, which were sampled at 256 Hz. The electrode positions were located at standard sites (Standard Electrode Position Nomenclature, American Encephalographic Association). The electrode positions are as shown in Figure 4. These sites are extension to the 10-20 electrode positioning system [15]. The VEP data was extracted from subjects while being exposed to a single stimulus, which are pictures of objects chosen from the 1980 Snodgrass and Vanderwart picture set [16]. These pictures were common black and white line drawings like aeroplane, hand, banana, bicycle, ball, etc. executed according to a set of rules that provide consistency of pictorial representation. The

pictures have definite verbal labels i.e. they are easily named. This fact is important as some amnesics may perform differently on recognition tasks using complex (abstract) pictures [17].

VEP signals with eye blink artifact contaminations were removed in the pre-processing stage using the fact that VEP signals above 70μV denotes occurrence of eye blinks. VEP data were extracted using Elliptic filter to remove contamination from overlapping EEG. Figure 5 (a) shows an example of a recorded EEG signal, while Figure 5 (b) shows an example of the extracted gamma band VEP signal. Figure 6 shows some examples of the Snodgrass and Vanderwart pictures.

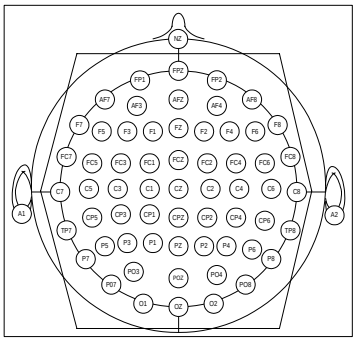
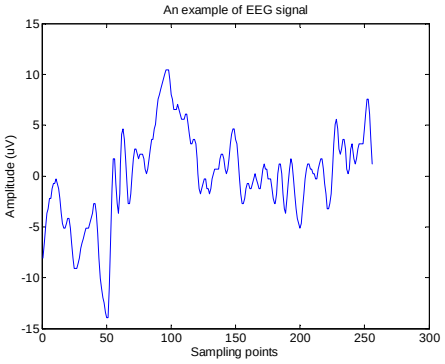
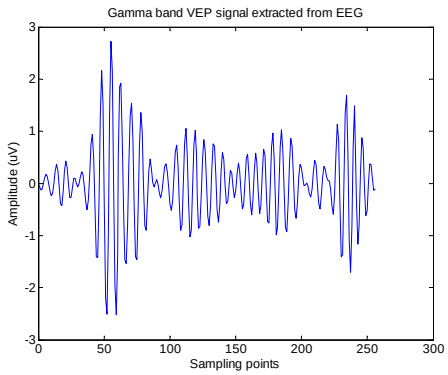


Fig. 4 64 channel electrode system



(a)



(b)

Fig. 5 (a) An example of recorded EEG signal (b) extracted gamma band VEP signal from EEG

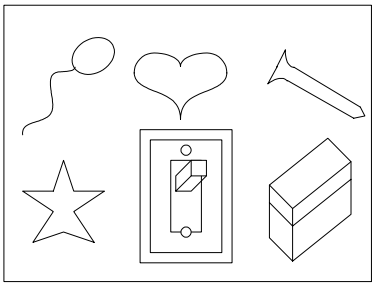


Fig. 6 Snodgrass and Vanderwart pictures

In the experimental study, EEG signals were recorded from 20 subjects: 10 alcoholics and 10 controls, with each subject completing 40 trial sessions giving a total of 800 EEG patterns. As mentioned earlier, single trial gamma band VEP were extracted from the EEG signals using the Elliptic filter. Next, Burg algorithm was used to derive the second order AR coefficients. After this, the PSD for each channel was derived and the peak value of the PSD values for all the 64 channels were concatenated into one feature vector. These vectors were used in training the classifiers for discriminating alcoholic subjects.

The inputs to the classifiers will be the peak PSD values from 64 channels. A total of 800 VEP feature vectors (20 subjects x 40 trials) were used in the experimental study. Half of the feature vectors were used in training and the remaining half in testing. The selection of the feature vectors for training and testing were chosen randomly. A modified four fold cross validation procedure was used to increase the reliability of the results. In this procedure, the entire data for an experiment (i.e. 800 VEP feature vectors) were split into four parts, with equal number of feature vectors from each subject. Training and testing were repeated for four times where for each time, two different parts were used for training and the remaining two parts for testing. This was done to increase the reliability of the discrimination results.

VII. Results

Table I shows the discrimination error using the 3 different classifiers where the four different datasets (from modified four fold cross validation) were used. It could be seen that the best discrimination of alcoholics and controls was given by LD, followed by MLP-BP and SFA. LD discrimination gave the averaged false positive (FP) error of 2.8%, with a false negative (FN) error of 2.5%, i.e. with an overall averaged error of 2.6%. FP occurs when a control VEP feature vector is detected as from alcoholic category, while FN occurs when an alcoholic VEP feature vector is detected as from control category.

VIII. Discussion

This paper proposed a second order AR model to discriminate alcoholics using single trial gamma band VEP signals classified using SFA, MLP-BP and LD classifiers, where LD gave the best discrimination performance. The results also showed that using second order for AR model is

suitable for classification purposes because of the pseudo-periodic nature of the gamma band VEP signals. Also, using a fixed second order circumvents the requirement of having to find the suitable order and has the advantage of lower computation time and smaller system design due to its low order. Conclusively, the high discrimination accuracy obtained in the experimental study showed that the proposed method of using single trial gamma band VEP signals modelled with second order AR model could be used to discriminate between alcoholics and control subjects. This would be useful in applications to screen alcoholics for certain purposes.

TABLE I
DISCRIMINATION ERROR USING DIFFERENT CLASSIFIERS

	SFA			MLP-BP			LDA		
Validation dataset	FP	FN	Overall	FP	FN	Overall	FP	FN	Overall
1	13	8	10.5	2	1	1.5	3	0	1.5
2	13	13	13	2	2	2	3	3	3
3	15	17	16	4	2	3	3	2	2.5
4	11	5	8	2	7	4.5	2	5	3.5
Average	13.0	10.8	11.9	2.5	3.0	2.8	2.8	2.5	2.6

ACKNOWLEDGMENT

We thank Prof. Henri Begleiter at the Neurodynamics Laboratory at the State University of New York Health Center at Brooklyn, USA who recorded the raw EEG data and Mr. Paul Conlon, of Sasco Hill Research, USA for sending the data to us.

REFERENCES

[1] Jansen, B.H., Bourne, J.R., and Ward, J.W., "Autoregressive Estimation of Short Segment Spectra for Computerized EEG Analysis," *IEEE Transactions on Biomedical Engineering*, vol. 28, no. 9, pp. 630-638, September 1981.

[2] Ning, T., and Bronzino, J.D., "Autoregressive and Bispectral Analysis techniques: EEG Applications," *IEEE Engineering in Medicine and Biology Magazine*, vol. 9 no.1, pp. 47-50, March 1990.

[3] Box, G.E.P., and Jenkins, G.M., *Time Series Analysis: Forecasting and Control* Holden Day, San Francisco, 1976.

[4] Burg, J.P., "A new analysis technique for time series data," In Childers, D.G. (ed.), *Modern Spectrum Analysis,* New York, IEEE Press, pp. 42-28, 1978.

[5] 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.

[6] 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.

[7] Palaniappan, R., and Raveendran, P., "Using Genetic Algorithm to Identify The Discriminatory Subset of Multi-channel Spectral Bands for Visual Response", *Applied Soft Computing*, vol. 2, Issue 1F, pp. 48-60, August 2002.

[8] Palaniappan, R., Raveendran, P., and Omatu, S., "VEP Optimal Channel Selection Using Genetic Algorithm for Neural Network Classification of Alcoholics," *IEEE Transactions on Neural Network*, vol. 13, issue 2, pp.486-491, March 2002.

[9] Basar, E., Eroglu, C.B., Demiralp, T., and Schurman, M., "Time and Frequency Analysis of the Brain's Distributed Gamma-Band System," *IEEE Engineering in Medicine and Biology Magazine,* pp. 400-410, July/Aug. 1995.

[10] Shiavi, R., *Introduction to Applied Statistical Signal Analysis, 2nd edition,* Academic Press, 1999.

[11] Aunon, J.I., McGillem C.D., and Childers, D.G., "Signal Processing in Event Potential Research: Averaging and Modelling," *CRC Crit. Rev. Bioeng.,* vol 5, pp. 323-367, 1981.

[12] Fukunaga, K., *Introduction to Statistical Pattern Recognition, 2nd edition*, Academic Press, 1990.

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

[14] Kasuba, T., "Simplified Fuzzy ARTMAP," *AI Expert*, vol. 8, no. 11, pp. 19-25, 1993.

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

[16] Snodgrass, J.G., and Vanderwart, M., "A Standardized Set of 260 Pictures: Norms for Name Agreement, Image Agreement, Familiarity, and Visual Complexity," *Journal of Experimental Psychology: Human Learning and Memory,* vol. 6, no. 2, pp. 174-215,1980.

[17] Starr, A., and Philips, L., "Verbal and motor memory in the amnesic syndrome," *Neuropsychologia*, vol. 8, pp.75-88, 1970.